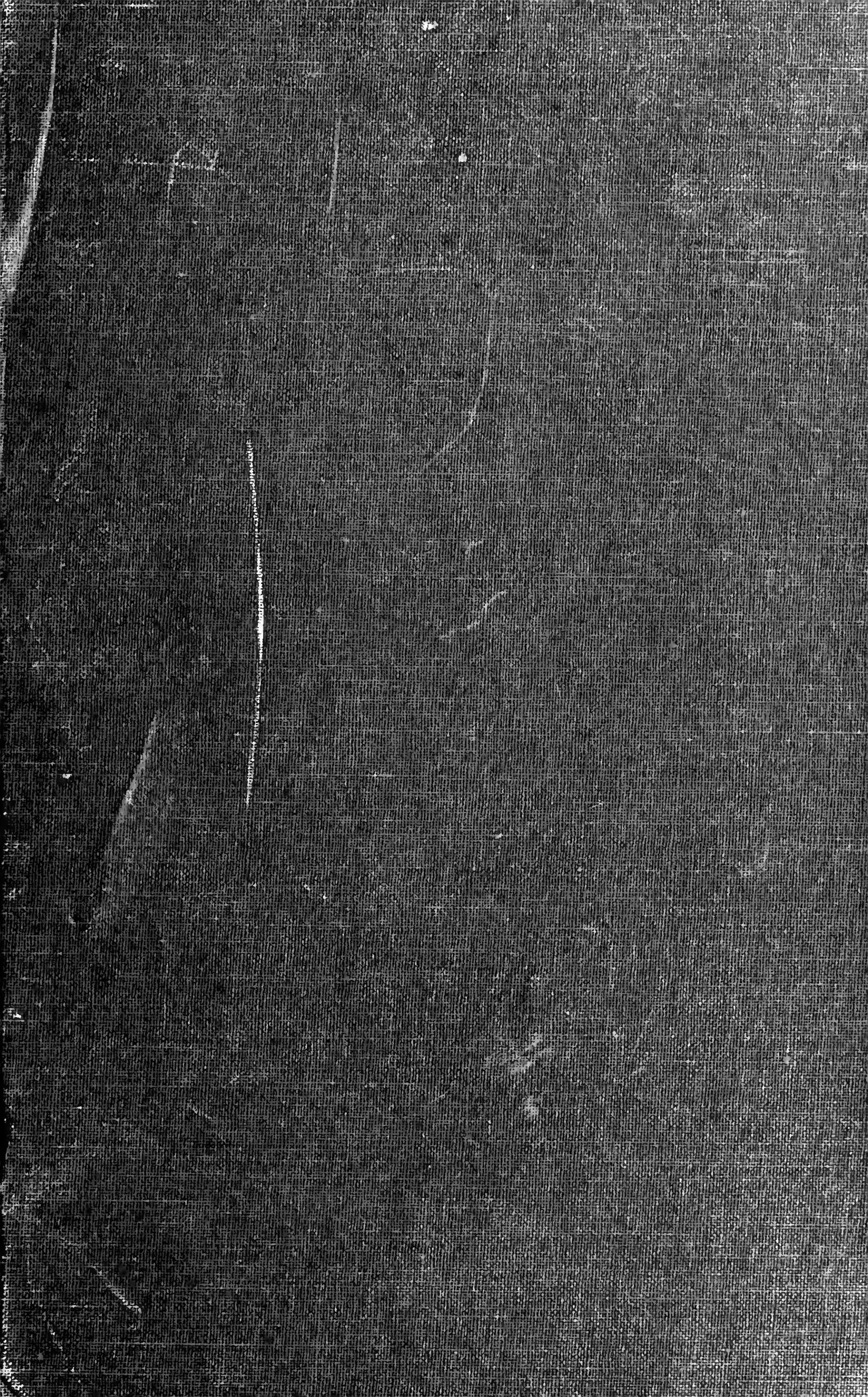


MBL/WHOI
0 0301 0010506 0

SYMPOSIUM ON
INFORMATION THEORY IN BIOLOGY

SYMPOSIUM ON

INFORMATION THEORY IN BIOLOGY

Gatlinburg, Tennessee, October 29–31, 1956

Edited by

HUBERT P. YOCKEY
Oak Ridge National Laboratory

With the assistance of

ROBERT L. PLATZMAN
Purdue University

HENRY QUASTLER
Brookhaven National Laboratory

SYMPOSIUM PUBLICATIONS DIVISION
PERGAMON PRESS
NEW YORK · LONDON · PARIS · LOS ANGELES

PERGAMON PRESS INC.
122 East 55th Street, New York 22, N.Y.
P.O. Box 47715, Los Angeles, California

PERGAMON PRESS LTD.
4 & 5 Fitzroy Square, London W.1.

PERGAMON PRESS, S.A.R.L.
24 Rue des Écoles, Paris, Ve

First published 1958

Library of Congress Card No. 58-9687

Printed in Northern Ireland at The Universities Press, Belfast

CONTENTS

PAGE

Foreword — ix
A. M. Weinberg

Preface — xi

PART I. INTRODUCTION

A Primer on Information Theory — 3
Henry Quastler

Some Introductory Ideas Concerning the Application of Information Theory in Biology — 50
Hubert P. Yockey

PART II. STORAGE AND TRANSFER OF INFORMATION

Editorial Introduction — 61

The Cryptographic Approach to the Problem of Protein Synthesis — 63
George Gamow and Martynas Yčas

The Protein Text — 70
Martynas Yčas
Discussion — 101

Protein Structure and Information Content — 103
L. G. Augenstine
Discussion — 123

Specific Mechanisms of Protein Synthesis and Information Transfer in the Developing Chick Embryo — 124
H. R. Mahler, H. Walter, A. Bulbenko and D. W. Allmann
Discussion — 135

The Mechanism of Action of Methyl Xanthines in Mutagenesis — 136
Arthur L. Koch

Evidence for a Negative Feedback System Controlling Liver Regeneration — 148
André D. Glinos

Fluctuations in Neural Thresholds — 153
Lawrence S. Frishkopf and Walter A. Rosenblith

PART III. DETERMINATION OF INFORMATION MEASURES

Editorial Introduction — 169

Chemistry and Biochemistry at Low Temperatures and Discrimination of States and Reactivities — 171
Simon Freed
Discussion — 180

75193

Information Content of Tracer Data With Respect to Steady-state Systems 181
MONES BERMAN and ROBERT L. SCHOENFELD

The Domain of Information Theory in Biology 187
HENRY QUASTLER

Discussion 196

Some Membrane Phenomena from the Point of View of Information Theory 197
HERMAN BRANSON

Efficiency of Information Transmission by Biochemical Co-factors 204
PETER D. KLEIN

Discussion 209

Antigenic Specificity 211
BERNARD N. JAROSLOW and HENRY QUASTLER

Information Content and Biotopology of the Cell in Terms of Cell Organelles 218
CHARLES F. EHRET

Quantification of Performance in a Logical Task With Uncertainty 230
A. RAPOPORT

PART IV. DESTRUCTION OF INFORMATION BY IONIZING RADIATION

Editorial Introduction 239

Electron Spin Resonance in the Study of Radiation Damage 241
WALTER GORDY

A Physical Mechanism for the Inactivation of Proteins by Ionizing Radiation 262
ROBERT PLATZMAN and JAMES FRANCK

Information and Inactivation of Biological Material 276
HAROLD J. MOROWITZ

Discussion 281

The Absence of Radiation-Induced Disulfide Interchanges 283
ARTHUR L. KOCH

A Proposed Mechanism of Protein Inactivation 287
L. G. AUGENSTINE

Discussion 291

PART V. AGING AND RADIATION DAMAGE

Editorial Introduction 293

A Study of Aging, Thermal Killing, and Radiation Damage by Information Theory 297
HUBERT P. YOCKEY

Entropic Contributions to Mortality and Aging 317
GEORGE A. SACHER

A Quantitative Description of Latent Injury from Ionizing Radiation 331
H. A. Blair

Some Notes on Aging 341
Hardin B. Jones

Cancer as a Special Case of a General Degenerative Process 347
Harry Auerbach

Discussion 351

Free Radicals as a Possible Cause of Mutations and Cancer 353
Walter Gordy

PART VI. INFORMATION NETWORKS

A Probabilistic Model for Morphogenesis 359
Murray Eden

Functional Geometry and the Determination of Pattern in Mosaic Receptors 371
John R. Platt

PART VII. THE STATUS OF INFORMATION THEORY IN BIOLOGY

The Status of Information Theory in Biology: A Round Table Discussion 399
Edited by Henry Quastler

Author Index 403

Subject Index 411

FOREWORD

Alvin M. Weinberg
Director, Oak Ridge National Laboratory

The reader of this book may wonder why it is that an institution such as the Oak Ridge National Laboratory, which is primarily interested in the control and release of nuclear energy, should also be interested in sponsoring a meeting on Information Theory in Health Physics and Radiobiology.

The answer rests in the fact that among the activities that are pursued at this Laboratory there are two which bear very directly on general problems of growth and of the impairment of growth by radiation and allied agents. Broad programs in fundamental research in the basic physical mechanisms and in the basic biological manifestations of radiation damage have been established in the Health Physics Division and in the Biology Division. In the Biology Division there is a great deal of experimental work being done on protein synthesis, on the mechanism of action of the nucleic acids, and on problems of the characterization of the nucleic acids. In the Health Physics Division there is a lively interest in the problems of dosimetry and the basic mechanisms of the interaction of radiation and matter. It is in establishing a tie-up between the physical and biological aspects of radiation damage that information theory may play an important role. We hope that this conference will help to assess the value of information theory to phenomena involved in the interaction of radiation and living matter.

PREFACE

BIOLOGY has made extensive use of many instruments and methods developed in the physical sciences. In recent years certain biological fields have been able to make increasing use of and have developed mathematical methods for their special purposes. Perhaps the reason this has not occurred earlier is that sufficiently *simple and important* systems or situations do not present themselves in biology. The life sciences, therefore, have developed most of their theoretical structure without mathematics playing a leading role. Nevertheless, the need for mathematical methods in biology has long been felt, as the pioneering work of Fisher, Haldane, Wright, and others, has emphasized.

The possibility that the life sciences could develop mathematical systems suitably their own, so that this form of research could be added to the already powerful research tools available, was the common denominator in this symposium. In order to address ourselves to a single task, the principal emphasis was on information theory. The reader will note that in several papers there is, willy-nilly, a reference or two to the ideas of cybernetics. Perhaps this presages a greater influence in biology of this mathematical sibling of information theory.

Our symposium and this book owe a debt to the pioneering effort oi Henry Quastler and the book he edited in 1952 entitled *Information Theory in Biology*. Among the newcomers to the fields of biology in which information theory has found an application since 1952 is radiobiology. Since radiation is an excellent way of introducing noise, the force of information theoretic ideas may well be effective in achieving a better understanding of radiobiologic problems in the future. By the same token, health physics will benefit by an appreciation of the relation between radiation damage and aging.

Our book is about a mathematical theory but it is also a book about experimental biology. This is properly so, for the development of clear ideas about nature is as much a part of science as any activity carried on in the laboratory. Experiment and theory do their best work in double harness. It should be understood that, although something has been done here to bring earlier work up to date and to present new material, much remains to be done. The information theory point of view suggests many problems of both a theoretical and an experimental character. We hope that good advantage will be taken of this fact.

The conference was entitled *A Symposium on Information Theory in Health Physics and Radiobiology* and was held in Gatlinburg, Tennessee, 29–31 October 1956. The articles composing this volume were written in the ensuing months and so represent the authors' results and opinions after the contact with his confreres. As a result of this collocution some of the authors contributed additional papers not given at the symposium.

The symposium could not have been carried through without the labor and good judgment of the other editors, Professor Robert L. Platzman and Dr Henry Quastler. It is a pleasure to acknowledge the encouragement of

Dr Alvin M. Weinberg and Dr Robert A. Charpie, Director and Assistant Director of the Oak Ridge National Laboratory. Appreciation is due Dr Karl Z. Morgan, Director, Health Physics Division and to Dr Alexander Hollaender, Director, Biology Division. Any success that has been achieved is due to the contributors and to those named above. Responsibility for errors or omissions is that of the undersigned.

H. P. Y.

Health Physics Division
Oak Ridge National Laboratory
Oak Ridge, Tennessee

PART I

INTRODUCTION

A PRIMER ON INFORMATION THEORY*

HENRY QUASTLER
Biology Department, Brookhaven National Laboratory, Upton, New York

SYNOPSIS

I. *Introduction:* Historic development of information theory; reason for its present popularity. System theories in general; specific role of information theory.

II. *The representation of information:* Paul Revere's code; essential features of representation of intelligence. Possibilities of representing information; variety of means; data and operations; conscious and non-conscious acts of representation; generalized meaning of 'information'. 'Real' and 'symbolic' events; abstraction preceding representation. Symbol, alphabet, 'words': units of representation. Binary representation, or standard method of symbolization; (a) simplest case: number of 'real' categories an integral power of two, words of equal length; (b) any number of categories: words of unequal length, FANO's 'confusion-proof' code, minimum-bulk code; (c) groups of events represented by single words; (d) unequal probabilities: general rule to obtain a minimum-bulk code; (e) any probabilities: general formula of minimum bulk in standard representation; (f) representation theorem. Exercises.

III. *The measure of information or uncertainty:* Information acquired and uncertainty abolished. The amount of uncertainty a function of probabilities of events, not their nature, causes and consequences; amount of information and representability; the H-function; the 'bit'. Some properties of the Shannon-Wiener information function: independence, continuity, additivity, naturalness of scale; values for probabilities zero and one; effects of averaging and of pooling. Exercises.

IV. *Information measurements pertaining to two related variables:* Generalized meaning of 'communication'. An example of two related variables: heights of father and daughters; joint uncertainty, internal constraints: the T-function; effects of scale on H and T. Two-part systems in general: the six information functions for two variables. Communication systems: nomenclature. Noise: the height correlation as example of a noisy channel; channel capacity; manipulation of information does not increase its amount. Error detection and correction: redundant information; theorem of the noisy channel; economics of error checking. Actual communication systems: signals and channels as physical entities. Exercises.

V. *Organization (systems, structures, pattern):* Organization, communication, redundancy. Systems analysis: informational analysis, informational challenge and performance; general limitations on information-processing. Multi-part systems. Unitization. Conclusion.

Appendix I*: The evaluation of information content:* Typical difficulties. The relativity of information measures; arbitrariness of selections which determine actual values. Approximation methods. Examples: rate of information transmission in conversation; information content per printed letter.

Appendix II*: Answers to exercises.*

* This paper is based on a report of the same name issued as Office of Ordnance Research Technical Memorandum 56–1, in January 1956. The memorandum was written at the suggestion of Dr SHERWOOD GITHENS, Jr, Director of the Physical Sciences Division of the Office of Ordnance Research; he and his staff were very helpful at all stages of the execution of this project, and it is a pleasure to extend them thanks. The present revised version was prepared at Brookhaven National Laboratory, under the auspices of the Atomic Energy Commission.

I. INTRODUCTION

THERE appears to be a gap in the literature on information theory. One can find several articles explaining what information theory is; several of those are understandable to a reader with little knowledge of mathematics; in fact, some are at the general magazine level. One can also find a number of books and articles explaining how information theory is to be used, but all of these are on a highly technical level. I am not aware of any presentation which is not highly technical and rigorous yet sufficiently explicit and pragmatic to enable a reader to make some practical use of information theory. This paper is intended to serve as a stopgap to fill a temporary current need. It is designed to have some of the aspects of an elementary textbook, including a few exercises. The examples are largely drawn from communication engineering and engineering psychology, these being the most convenient ways of interpreting information theory; however, the whole theory could be expounded without any reference to conscious communication.

Information theory is based on the concept that information is measurable. This idea is not new. In physics, the notion of a measurable relation between information and degree of orderliness (entropy) dates back to BOLTZMANN's work in 1872 and its development in 1929 by SZILARD (1). In 1918, the statistician R. A. FISHER (2) needed a criterion to assess the degree to which the information contained in experimental data is utilized by a given statistical procedure; he worked out a measure of information which has been used in statistics ever since. Later, the need arose for a measure of information-carrying potential as a consequence of the tremendous development of telecommunication, and in 1928, R. V. L. HARTLEY (3) published such a measure. In 1948, WIENER (4) observed that a measure of information content is a basic ingredient to the study of communication, which itself is a basic ingredient to the study of control in its broadest sense. In the same year, the communication engineer C. E. SHANNON (5) published an article on the mathematical theory of communication which in several respects went beyond previous studies. This article is highly technical; it is very difficult reading; it appeared in a specialized journal (*The Bell Systems Technical Journal*) and it pertained to no other field than telecommunication. It certainly did not look like an article destined to reach wide popularity among psychologists, linguists, mathematicians, biologists, economists, estheticists, historians, physicists . . . yet this is what happened. In 1949, the University of Illinois Press issued a book (6) which consisted of a reprint of SHANNON's earlier article and a paper by WARREN WEAVER; in this paper, the generality of the concept of 'amount of information' was forcefully expounded. The literature on 'information' has been increasing ever since at an almost explosive rate.

What are the reasons for these startling consequences of such a highly specialized article? One reason is, of course, that it is a very good article. The other is that the concept of a measure of information fulfills a general and deep need of our time. The sheer bulk of the information now available increases at a rapid rate. Accordingly, the representation of information becomes a more and more critical problem, and information theory offers general principles concerning representation. Also, we are developing organizations

which are more and more complex; they depend for their functioning on successful and efficient communication, and information theory offers general principles about communication.

Information theory is related to a group of specialties which are either new or have recently increased tremendously in popularity, as has information theory itself. To name but a few: operations analysis; the theory of experimental design; decision theory; theory of linear programming; cybernetics; game theory; theories of administration; group dynamics; and others.

A little pondering over this list suggests a few generalizations:

(i) all of these sciences are mathematical, the fields of probability and statistics being referred to most frequently;
(ii) each employs a system of evaluation* of something;
(iii) each deals with complex situations; in every one, there is a multiplicity of possible choices as arrangements of some sort, and in most of them, a large number of interrelated factors affect the choices. Thus, they all can properly be called system theories;
(iv) none of these sciences is primarily concerned either with the physical nature of the system considered or with the mechanisms by which its parts are interrelated.

These are the common features. However, each of the endeavors named is a special science since each deals with a different aspect of systems. It is an open question whether because of their similarities these different sciences can be gathered under one common discipline that could be called General Systems Theory (and is the basis of a society formed in 1954).

Information theory, thus, is only one of several system theories. The particular concern which characterizes it in contrast to the related specialities is measurement of the degree to which a thing (or a condition, or an event) is *specified*; that is, of the degree to which it differs from other possible things (conditions, events). Communication and organization are treated in terms of a mutual specification. One way to illustrate the essence of information theory is to compare it with statistics. Statistics and information theory both deal with the diversity among the elements of a set, but in different ways. Statistics treats diversity as a nuisance, and tries to establish what can be stated or done *in spite* of it. Information theory treats diversity as an asset without which operations such as selection, communication, representation, specification, would not be possible; it tries to establish what can be achieved because of a certain degree of diversity. The 'information' evaluated in Information Theory is thus not the every-day information. The 'information' in a message, for example, as a type of event, is the measure of the amount of knowledge (intelligence) which a message of this sort ideally can convey through the medium of symbolic representation.

To many people, information theory looks highly promising on first contact; to some people, it still looks promising after serious study. Information theory has become well established among engineers. In psychology, it has achieved

* In some cases values are frankly imposed, in others they are inferred from observed behavior. This does not necessarily mean that the values must be consciously imposed; goal-directed behavior can occur without any conscious act of fixing values.

a certain status. In biology, economics, political science, esthetics, linguistics, information theory has interested many people, but the active users are only a handful.

II. THE REPRESENTATION OF INFORMATION

'One if by land, two if by sea.' Paul Revere and his fellow citizens did not know information theory, but they knew and utilized what is at the basis of information theory, namely, the principle of the representation of intelligence. The Paul Revere code is not quite up to modern standards, but it has the essential properties:

(i) The news concerning the road of approach of the enemy was translated into another kind of intelligence, namely, lights hoisted on a steeple; this translation is useful because it transforms a hard-to-broadcast piece of intelligence into one which is easily broadcast;

(ii) the range of all possible events was subdivided into categories of interest. The code could have been reduced to one which indicated only the enemy's arrival. It could have been expanded into one conveying more accurately the direction of approach, or signalling the enemy's strength and other details of possible interest. In this case, a more complicated code would have been necessary, and this would have increased the possibility of misunderstandings. Proper economy of categorization is an important feature in representing information;

(iii) the representation employed a code previously agreed upon. No light meant no enemy approaching, one indicated land, two indicated sea. It seems that no agreement was made concerning simultaneous approach by land and by sea, but the message 'three lights' would have been correctly interpreted by all concerned.

Possibilities of Representing Information

There is no limit to the number of possibilities of representing one kind of information by another. (It may be observed that the term 'information' covers more ground than the word 'intelligence'. Generally, the word intelligence is restricted to conscious information.) The only condition for representation is that a complete system of translation, a code, be agreed upon. The limitations are set only by the ability to discriminate information to be represented, by the ability to produce accurately a desired representation, and by the range of the code.

Representation is not restricted to discrete categories of intelligence. A continuum of information or state of affairs can be represented by a physical continuum such as a range of voltages or the rotation of a shaft. Any kind of information, discrete or continuum, can be represented by the charges in an electron tube, by the magnetization of a spot on a metallic surface, by the deflection of the beam of a cathode ray tube, by a light falling upon a photographic emulsion, and so forth (7).

It is possible to represent not only data, but also operations on data. On a slide rule, numbers, a kind of intelligence, are represented on scales by marks, which is a form of encoding. The operation of multiplication is encoded by positioning the slide and the indicator, and decoded by the act of reading from

a scale. The slide rule, thus, is an early and modest example of an information-handling device. At present, machines exist which accept and store considerable symbolic information concerning data and operations, and which will execute a wide range of manipulation with this information. There is good reason to believe that machines will actually be built which can compute any number that can be computed, and which, even more generally, can arrive at the results of any thinking which can be described by explicitly-defined operations.

In the situations so far mentioned the relation between the original and the translated (coded) intelligence was based on a mutual agreement. Translation of information is not restricted to such a situation; it may be based on a one-sided choice of code, one not explained to the receiver. This occurs in the case of conditioned reflexes: each time a dog is fed, a bell rings; after some time, the information 'the bell is ringing' comes to represent the information that 'food is about to be served', and leads to preparations for eating. The code is established by the experimenter; the dog is not consulted. In fact, the representation of information does not have to be at the level of conscious awareness in any way. For instance: the system which regulates one's breathing and thereby his oxygen intake has come to depend for its regulation not on the oxygen content of the blood itself, but on the concentration of carbon dioxide. Ordinarily, the CO_2 level in the blood is a reliable representation of the O_2 level; under certain conditions the representation ceases to be correct, and then difficulties can occur. In all of these cases, just as in those with arbitrarily-fixed codes, information theory is concerned with the general laws which govern the possibility of translating one kind of information into another; it will be obvious by now to the reader that the term 'information' in the technical sense covers a good deal more than in everyday language.

'Real' and 'Symbolic' Events

We now turn to a more formal and general discussion of the principles of representation of information. We will deal here only with information that occurs in discrete units; however, the transition to treatment of the continuous-function type of information would not be difficult.

In our discussions, the terms *real* and *symbolic* will replace the cumbersome expressions 'something to be represented' and 'something representing'. It will be remembered that 'real' and 'symbolic' refer not to properties of things, but to their functions in a given situation, and that the term 'symbolic' may but does not necessarily imply that a conscious act of symbolization has occurred. The 'symbolization' of the need for oxygen by the carbon dioxide level illustrates that 'symbolic' is here used in a wider sense than is customary.

'Information' is not a disembodied something; it is always related to some actual carrier—a thing or an event. We will use whichever word is appropriate in a given situation; the word *event* is most frequently used as a generic term. However, it must always be remembered that other terms could be substituted; information theory applies equally to all kinds of carriers of information. In formal language, one refers to the information carriers as elements of discourse, or points in sample space, or configurations of properties.

A concrete event, in all its richness of detail, is not amenable to complete representation. The only complete representation of a particular man, at a

particular moment, is that man, at that moment. Amenable to symbolic representation are only certain aspects of the concrete event, for instance, the fact that this man, at that time, belonged to a category labelled 'male' student junior year, college of engineering, grade average 4.32, etc. This kind of information lends itself to representation, e.g., by the position of certain slots on a Hollerith card. In general, when we speak of representing events, we mean not concrete events in their whole individuality, but only their abstractions as instances of a category of events. In formal language, the aspect of informational interest of an event is its 'class membership', or the name of the 'set of points' to which it belongs.

The first steps, then, in representing information, are (1) the decision of what to consider as elementary carriers of information, or elementary events, (2) the decision as to what features of these events are to be considered as relevant, and (3) a comprehensive listing of all classes of events corresponding to the various features or combinations of features. Ideally, this analysis of the real situation should be completed before the task of symbolic representation is started; in practice, it will often be convenient to base a temporary system of representation on an incomplete analysis, and introduce subsequent refinements and adaptations as needed. A library catalogue, arranged by subject matter, is a good example of a system of representation which must remain flexible and capable of growing.

Symbol, Alphabet, 'Word'

The basic unit of symbolization is called a *symbol.* The set of all available symbols constitutes an *alphabet.* In the simplest form of representation, each individual event is translated into a single symbol that represents that kind of event. This can be done if the alphabet is large enough to have a separate symbol for each of the categories required to classify real events. The Paul Revere code is an example.

The one-by-one method of representation lacks flexibility and is cumbersome. Often a small alphabet is used to express a wide range of possibilities. In this case, single events must be represented by combinations of symbols; these are called code groups, or 'words'. For instance, an alphabet of twenty-six letters is used to represent several hundred thousands of English words, and is flexible enough to accommodate any number of new words; this is achieved with an average of 4.5 letters per word (where the letters are not used with greatest economy!). In turn, an alphabet consisting of only two symbols, the dots and dashes of the Morse code, is sufficient to account for all the twenty-six letters, plus the ten digits, plus punctuation marks and a few standard concepts, without ever employing more than six symbols in a single code group, or 'word' (as defined above, not in the ordinary sense!). According to GAMOW and YČAS (this volume), each of twenty amino acids in a protein can be represented in the RNA molecule by a 'word' of three nucleotides; in this case, the sequence of letters in the word is considered irrelevant.

Binary Representation

Simplest Case—For developing a general theory of the representation of information, it will be convenient to reduce all representations to some standard

form. Any standard form would be acceptable; it has become customary to employ the simplest of all possible alphabets, the binary alphabet. The two symbols commonly used are '1' and '0'; it must be emphasized that '0' does not necessarily imply the absence of some physical action; e.g., '1' and '0' might stand for right-left, positive-negative, dash-dot, etc. The standard symbolization of any event will be a binary number, such as 1001101 . . . , where the symbolic meaning of each digit and combination of digits is fixed by some law of association.

It must be pointed out that the Morse code is not a strictly binary representation if one thinks of whole messages. The Morse code is really a quaternary code. This is so because, in addition to the 'blacknesses', the dots and dashes, it uses two 'whitenesses' of different length, namely, an inter-letter space and an inter-word space. Both of these are integral parts of the code system, because otherwise we could not know whether this:

. . . . . — .

means 'hen', 'sue', 'sin' or 'site'.

How many events can be represented by words made up of a certain number of binary symbols? There are two different 'words' ('1' and '0') consisting of a single symbol; they can represent a partition of a set of real events into two classes. There are four different two-symbol 'words' (11, 10, 01, 00), and, in general, 2^n code groups consisting of n binary symbols. Accordingly:

A sequence of n binary tests will discriminate between 2^n possibilities;
A sequence of n binary choices will select any one of 2^n alternatives;
A sequence of n binary statements will identify any one of a set of 2^n items, etc.

Conversely, if a code book with r distinct representations is to be made up in standard binary code, then each word will have to be a binary number with about $\log_2 r$ digits*. For instance, eight categories of events can be represented by code groups consisting of three binary symbols ($3 = \log_2 8$):

Category A	1 1 1
Category B	1 1 0
Category C	1 0 1
Category D	1 0 0
Category E	0 1 1
Category F	0 1 0
Category G	0 0 1
Category H	0 0 0

Observe that the meaning of each binary symbol depends on its position and on the nature of the other symbols in the word. For example, a 1 in the second position means 'A or B or E or F'; if preceded by a 0, then it can mean only 'E or F'; if also followed by a 1, then it designates 'E', unequivocally. This

* Logarithms to the base 2 can be found in published tables, or read on a slide rule with a log-log scale, or obtained by multiplying the base-10 logarithms by 3.322.

implies that any digit of the symbolic word, considered by itself, does not necessarily represent a given operation. For instance, a set of recognition operations, such as a naturalist's key, could well be arranged as a sequence of binary tests. In such a case one will not always apply the same sequence of tests; in general, the choice of the second test will depend on the outcome of the first, the choice of the third on the outcome of the previous two, etc. Accordingly, the code book will have to specify what operations and what outcomes a particular symbol designates.

Sequences of code words will represent series of 'events'. Suppose our message were to represent the sequence of events 'G C A'. Using the code given above, we get:

001101111.

Observe that the 'words' in the message are not separated by spaces; a space with a defined symbolic meaning (such as a 'whiteness' in the Morse code) would make the alphabet ternary rather than binary. The receiver will not miss the spaces; he is expected to know the code and, accordingly, to read the message in groups of three digits, beginning with the first one at the left.

Any Number of Categories—In general, the number of categories to be encoded is not an integral power of two (such as 2,4,8,16,32 . . .). We could always use the nearest higher power of two as the basis of the coding scheme. For instance, if we had five categories to represent, then we could simply use a portion of our three-digit code for eight categories:

Category A 1 1 1
Category B 1 1 0
Category C 1 0 1
Category D 1 0 0
Category E 0 1 1

In this example, the symbolization possibilities of the three digits are not fully utilized. This is not economical; one will suspect that it is possible to achieve greater economy in number of digits. This can mean only that some of the five categories will be represented by two digits only. Some words will have two, and some three digits; the decoder will not have the benefit of being able to cut up the message into pieces of equal length. Therefore, it becomes imperative that the words themselves indicate unequivocally the correct partition. This will be the case if no combination of code groups is identical with any other combination of code groups; otherwise, confusion may arise. For instance, the following:

Category A 1 1
Category B 1 0
Category C 0 1
Category D 0 0
Category E 1 1 1

is useless because the message '11111' could be read as 'A E' (11 111) or 'E A' (111 11).

R. M. FANO (8) has devised a simple method for establishing a confusion-proof code. It goes as follows: all the categories to be encoded are divided up into two groups; the symbols '1' and '0' are assigned to these. Each of the two groups is, in turn, subdivided into subgroups; these are designated '1' and '0' in the *second* digit. The procedure of subdividing is continued until no subgroup contains more than a single category. At any stage of partitioning, the subgroups may contain unequal numbers of categories; accordingly, the number of steps to complete the coding does not have to be the same for all categories. This results in words of unequal length. In spite of this, messages composed of code groups formed according to this rule will be perfectly unequivocal.

Fano's method will be illustrated by three ways of making up a code for five categories:

(*a*): separate category 'A' from the others in the first step; use two more steps to subdivide the remaining four categories.

Category	1st step	2nd step	3rd step	Final code
A	1	—	—	1
B	0	0 1	0 1 1	0 1 1
C	0	0 1	0 1 0	0 1 0
D	0	0 0	0 0 1	0 0 1
E	0	0 0	0 0 0	0 0 0

(*b*): Use the first step to separate 'A or B' from 'C or D or E':

Category	1st step	2nd step	3rd step	Final code
A	1	1 1	—	1 1
B	1	1 0	—	1 0
C	0	0 1	—	0 1
D	0	0 0	0 0 1	0 0 1
E	0	0 0	0 0 0	0 0 0

(*c*): First step as in (*a*); the second step is used to separate 'B' from 'C or D or E'; the third separates 'C' from 'D or E'; and the fourth separates 'D' from 'E':

Category	1st step	2nd step	3rd step	4th step	Final code
A	1	—	—	—	1
B	0	1	1	—	0 1
C	0	0	1	—	0 0 1
D	0	0	0	1	0 0 0 1
E	0	0	0	0	0 0 0 0

These are the three Fano codes with five words; all other codes can be reduced to one of these three by rearranging the names of the events. All three codes are confusion-proof. In decoding, one retraces the steps of encoding; the code book shows, unequivocally, whether any symbol in a given sequence is a terminal one, or whether it does not yet identify a single category. For instance, suppose code (*c*) has been used, and the message received was:

000000101.

The first zero indicates 'B or C or D or E'; the second, 'C or D or E'; the third, 'D or E'; the fourth designates 'E', unequivocally, and is a terminal symbol. We mark off the four zeros and proceed. The first symbol of the second code group is a zero, as is the second; this indicated 'C or D or E'; the next symbol is a one, which is a terminal symbol and designates 'C'. The remaining code group, '01', means 'B', and the whole message is decoded unequivocally: 'E C B'.

Code (*b*) has the minimum bulk, or lowest average number of digits per word (2.4 digits, against 2.6 digits for code (*a*) and 2.8 for code (*c*)). The rule to obtain the minimum bulk code with any number of categories is as follows: *all divisions and subdivisions must be between groups of categories of as nearly as possible equal sizes.* To find the word length in this code, determine the largest integer k compatible with the condition that

$$2^k \leqslant r < 2^{k+1}$$

$$(k \leqslant n < k + 1).$$

Then, using equipartition as nearly as possible, each word will be of length k or $k + 1$, and the average number of binary symbols per category encoded will be somewhat larger than $\log_2 r$. In the example just given:

$$r = 5, \log_2 r = 2.33$$

$$k = 2, k + 1 = 3, \text{average length of word} = 2.4.$$

The worst discrepancy between $\log_2 r$ and average word length occurs for $r = 3$. We have:

Category A 1
Category B 0 1
Category C 0 0

$$\log_2 3 = 1.58$$

$$k = 1, k + 1 = 2, \text{average length of word} = 1.67 \text{ symbol}$$

$$\text{excess digits per word} = 1.67 - 1.58 = 0.09 \text{ or } 5.7 \text{ per cent of } 1.58$$

Groups of Events—The excess of average word length over $\log_2 r$ is due to some partition (especially an early one) dividing the set of categories into

portions of unequal size. In the case of three categories the first partition separates category 'A' or 33 per cent of all categories, from 'B or C', representing 67 per cent. This situation can be improved when it is allowed to represent pairs of events, instead of single events. There are nine pairs of the events A, B, and C, designated AA, AB, AC, BA, etc. We use the first symbol to subdivide them into two groups of four and five, respectively; each of these groups is subdivided by the second symbol, etc. We obtain:

Pairs of real events	Symbolic representation
AA	1 1 1
AB	1 1 0
AC	1 0 1
BA	1 0 0
BB	0 1 1
BC	0 1 0
CA	0 0 1
CB	0 0 0 1
CC	0 0 0 0

$$\text{Average: } \frac{29}{9} = 3.22 \text{ symbols per pair of events, or } 1.61 \text{ per single event.}$$

$$\text{Excess digits} = .03 < 2 \text{ per cent}$$

By going from pairs to triplets, the limiting value can be approached still closer. In general, if the *group* of events to be represented can be made as large as desired, then the limiting value can be approached as closely as desired.

Unequal Probabilities—In general, categories occur with unequal probabilities. In this case, subdividing the categories into sub-sets containing equal numbers of categories will not result in a minimum-bulk code. Consider, for instance, the three Fano codes for a set of five categories. Suppose category A accounts for 80 per cent of all occurrences, and the other four for 5 per cent each. In this case, code (*a*) will yield minimum bulk with an average of 1.4 digits per word, followed by code (*c*) with 1.45 and code (*b*) with 2.1 digits. The general rule to obtain a minimum bulk code, with any number of categories and equal or unequal probabilities, is as follows: *all divisions and subdivisions should be between groups of categories of as nearly as possible equal aggregate probabilities.*

The average number of digits in a minimum bulk code is found by the following consideration: let $p(i)$ be the probability of an event falling into i'th category, where i may stand for A, B, C, . . . , if the categories are designated by letters, or for 1, 2, 3, . . . , if the categories are numbered. For the time being, we consider only probabilities which are integral powers of 1/2, i.e., 1/2, $(1/2)^2 = 1/4$, $(1/2)^3 = 1/8$, $(1/2)^4 = 1/16$, etc.; i.e. we set $p(i) = (1/2)^{z_i}$ where z_i is a positive integer. In such a case, each step in the coding procedure can be a partition into groups of equal aggregate probability; then, the code word for

each category will have exactly z_i binary digits. This is illustrated in the following example:

Category	Probability, $p(i)$	z_i	Code representation
A	1/2	1	1
B	1/8	3	0 1 1
C	1/8	3	0 1 0
D	1/8	3	0 0 1
E	1/16	4	0 0 0 1
F	1/32	5	0 0 0 0 1
G	1/32	5	0 0 0 0 0
	32/32 = 1		

The first step separates 'A' ($p = 1/2$) from all other categories (aggregate probability $= 1/2$); the second separates 'B or C' (aggregate $p = 1/4$) from 'D or E or F or G' (aggregate $p = 1/8 + 1/16 + 1/32 + 1/32 = 1/4$); the third separates 'B' from 'C' (p, 1/8 each) and 'D' ($p = 1/8$) from 'E or F or G' (aggregate $p = 1/8$), etc.

The average number of digits per code word is the sum of the z_i's, weighted by the probabilities $p(i)$; in our example:

$$\sum_i p(i) \cdot z_i = 1/2 + 3/8 + 3/8 + 3/8 + 4/16 + 5/32 + 5/32 = 70/32 = 2.19$$

From $\qquad p(i) = (1/2)^{z_i}$

we get: $\qquad \log_2 p(i) = z_i \cdot \log_2 (1/2)$

and, because: $\qquad \log_2 (1/2) = -1$

we have: $\qquad z_i = -\log_2 p(i).$

We get (for $p(i)$'s which are integral powers of 1/2!) the following result:

$$\text{Average number of binary symbols per event} = -\sum_i p(i) \log_2 p(i).$$

We will check this result for the case of equiprobable categories. For r categories, the probability of every one will be $1/r$; so:

$$-\sum_i p(i) \log_2 p(i) = -r \cdot \frac{1}{r} \cdot \log_2 \frac{1}{r} = \log_2 r$$

This is the expression previously obtained for equiprobable categories.

Any Probabilities—What if probabilities are not limited to the values 1/2, 1/4, 1/8, etc.? In this case, it will—in general—not be possible to make divisions into exactly equiprobable groups. We would suspect that in this case the coding will be less than optimally efficient; accordingly, the average length of a code word will be somewhat higher than $-\sum_i p(i) \log_2 p(i)$. The approximation is usually not bad. This is illustrated in the following example which shows the construction of a binary code for the letters of the English alphabet, taking into account their relative frequencies. As expected, it turns out that

each category, i, is represented by a code word of approximately $-\log_2 p(i)$ digits; accordingly, its contribution to the weighted average is not far from the ideal value of $-p(i) \log_2 p(i)$, and the mean code length is only very slightly greater than the limiting value of $-\sum_i p(i) \log_2 p(i)$.

Table I. Fano Code for English Letters

1	2	3	4	5	6	7
i	$p(i)$	Code	No. of digits in code word	$-\log_2 p(i)$	Contribution to weighted average 2 × 4	$-p(i) \times \log_2 p(i)$ 2 × 5
E	.132	111	3	2.92	.393	.384139
T	.105	110	3	3.25	.315	.341411
A	.086	101	3	3.54	.258	.304398
O	.080	1001	4	3.64	.320	.291508
N	.071	1000	4	3.82	.284	.270938
R	.068	0111	4	3.88	.272	.263725
I	.063	0110	4	3.99	.252	.251275
S	.061	0101	4	4.04	.244	.246137
H	.053	0100	4	4.24	.212	.224606
D	.038	00111	5	4.72	.190	.179278
L	.034	00110	5	4.88	.170	.165862
F	.029	00101	5	5.11	.145	.148126
C	.028	00100	5	5.16	.140	.144436
M	.025	000111	6	5.32	.150	.133048
U	.020	000110	6	5.64	.120	.112877
G	.020	000101	6	5.64	.120	.112877
Y	.020	000100	6	5.64	.120	.112877
P	.020	000011	6	5.64	.120	.112877
W	.015	000010	6	6.06	.090	.090883
B	.014	000001	6	6.16	.084	.086218
V	.009	0000001	7	6.80	.063	.061162
K	.004	00000001	8	7.97	.032	.031863
X	.002	0000000011	10	8.97	.020	.017931
J	.001	0000000010	10	9.97	.010	.009965
Q	.001	0000000001	10	9.97	.010	.009965
Z	.001	0000000000	10	9.97	.010	.009965
	1.000				4.144	4.118347

We have already met a situation where a binary code was less than optimally efficient (in the sense of minimum length of code words); that was the case of r equiprobable categories, when r was not an integral power of 2. In this

instance, it was possible to approximate optimal efficiency by symbolizing groups of events instead of single events. The same principle works in the case of probabilities which are not integral powers of (1/2). We will illustrate the method in the case of a situation with two alternatives.

Example: Let there be two categories of events, 'A' and 'B', with associated probabilities, p(A) and p(B):

$$p(\mathrm{A}) = .7$$
$$p(\mathrm{B}) = .3$$

The limiting value of symbols per event is:

$$-\sum_i p(i) \log_2 p(i) = -(0.7 \log_2 0.7 + 0.3 \log_2 0.3) = 0.881291 \ldots$$

If this situation is to be represented on the basis of single events, then one needs one binary digit per event.

Event	Probability	Representation
A	0.7	1
B	0.3	0

Average number: 1.0 symbol per event; excess 12 per cent.

The following two-event clusters are possible: AA, AB, BA, BB. If the two events are independent, then the probability that *both* occur is the product of their individual probabilities:

$$p(\mathrm{AA}) = p(\mathrm{A}) \cdot p(\mathrm{A}),\ \ p(\mathrm{BA}) = p(\mathrm{B}) \cdot p(\mathrm{A}),\ \text{etc.}$$

Setting up a Fano code, we get:

Event	Probability	Representation
AA	.49	1
AB	.21	0 1
BA	.21	0 0 1
BB	.09	0 0 0

Average 1.81, or 0.905 symbols per event; excess 3 per cent.

If we can encode groups of three real events, then we get still closer to optimum economy:

Event	Probability	Representation
AAA	.343	1 1
AAB	.147	1 0
ABA	.147	0 1 1
BAA	.147	0 1 0
ABB	.063	0 0 1 0
BAB	.063	0 0 1 1
BBA	.063	0 0 0 0
BBB	.027	0 0 0 1

Average: 2.686, or 0.895 digit per event; excess $1\frac{1}{2}$ per cent.

Even with more pronounced unbalance of frequencies, the minimum value of binary digits per word is soon approximated. For $p(A) = .89$ and $p(B) = .11$, the limiting value is .50. In single-event-code, one needs one digit per event; for two-event-sequences, .66 digits; for three-event-sequences, .55; and for four-event-sequences, .52.

We have begun our discussion of binary representation with the case of 2, 4, 8, 16, . . . , equiprobable categories. We then generalized to cases with any number of categories, and proceeded from the representation of single events to clusters of events. Next, we introduced unequal probabilities, of value 1/2, 1/4, 1/8, ... Finally, we dropped all restrictions. We can now state, with full generality:

If a real situation is categorized into r categories, with associated probabilities $p(i)$, (*where* $i = 1, 2, \ldots, r$), *then it is possible to represent each event with an average of no more than* $-\sum_{i=1}^{r} p(i) \log_2 p(i)$ *binary symbols.*

Representation Theorem—In general, the closer we want to approximate the minimum bulk of representation, the larger the groups of sequences which must be encoded. This entails the following penalties:

1. There will be a delay in waiting for a whole group of events to occur or to be registered, and

2. The encoding and decoding procedures, and the code book itself, will become the more elaborate the larger the groups coded.

It is obvious that the code which is most economical in terms of bulk of representation is not necessarily optimum in over-all performance. There will be cases where it might be worthwhile to sacrifice economy in word length for ease in decoding. If the reader will work through exercise 4, then he surely will appreciate this possibility. Whether or not minimum bulk of coding is favorable, in a given case, cannot be derived from informational analysis. What information theory does is to establish a limiting value of the number of symbols, of a given kind, which are needed to represent the information in a given factual situation; in some cases, like those here discussed, information theory will also show how such coding economy can be achieved; but it can never prescribe that this is what should be done.

It would be quite legitimate to inquire, at this point, why we have gone to so much trouble to find out how to achieve binary representation with minimum bulk? Is not the result of doubtful value, in view of the fact that a tolerable approximation to minimum bulk can usually be achieved with the simplest means, and that a close approximation often entails prohibitive costs in encoding and decoding? The answer is this: by establishing the minimum length of code words in standard binary representation, we have implicitly established a *general* condition of representability:

If an event can be represented by (*on the average*) *n binary digits, then it can symbolically represent, or be represented by, any other event that can also be coded into n binary digits.*

This can be immediately generalized to groups of events: Let S_x and S_y be the number of real and symbolic events in a group, and n_x and n_y the average

binary representation per event and per symbol. Then, the general condition of representability can be stated as follows:

$$S_y \cdot n_y \geqslant S_x \cdot n_x$$

EXERCISES

1. A weakness of the Paul Revere code is that there is no positive signal for "peace and quiet". Hence, the colonists could not be sure whether the absence of a warning signal meant "peace and quiet" or a disturbance in the communication system. Show how two lights could be used to indicate the four situations by positive signals.

2. Any integer can be written as a sum of powers of 2(1,2,4,8,16, · · ·). For instance:

$$\begin{aligned} 27 &= 16 + 8 + 2 + 1 \\ &= 2^4 + 2^3 + 2^1 + 2^0 \end{aligned}$$

In binary notation, one indicates the power by position, and writes a '1' in appropriate position if this power does enter the sum, a '0' if it does not. Thus, '27' becomes 11011.

(a) Write the following numbers in binary notation: 0,1,2,3,4,5,6,7,8,9,10,12,16,1955.

(b) Write the following binary numbers in decimal notation: 1001, 1011, 10010011, 100000.

Any proper fraction can be written as a sum of powers of 1/2, (1/2, $(1/2)^2 = 1/4$, $(1/2)^3 = 1/8$, etc.). For instance: .75 = 1/2 + 1/4, or, in binary notation, .11.

(c) Translate into decimal notation: .001, .1001001

3. (a) encode the message 'ABCDE' in code (a) of the five-word codes described earlier.
   (b) decode the message: '000001011011' in code (b).

4. This assignment is coded in the Fano code for English letters given earlier.
   0010010011000010111001110001100010011010100100100110000010100011001010
   1101001100000101011111111100100100100111111000110010101101000000101001
   0101100000001111000001011010001010111000100001110110000101011011001010
   0101100101111111010010001000011011111011011101111011000001110010010010
   0011100001110101111111100100111000011111011100101100101110001111011000
   0010011110010010111001000110010100100100100111111000010011011001001100
   1001011101001001011100100111101000001100100000001111000001000100111000
   0010010010011101101000000101101100000111001111111001001001001110110100 0
   000 101101000111111010101011010001011110011001100000000111111001000110
   010110011000 111
   (This assignment is very tedious but it is good practice.)

5. Given a real situation with three categories and probabilities $p(\mathrm{A}) = .8$, $p(\mathrm{B}) = .15$, $p(\mathrm{C}) = .05$. Construct a binary code which comes within 10 per cent of the minimum bulk.

6. A protein is thought to be a linear arrangement of amino acids of which there are (about) twenty kinds in each cell. The specificity of a protein depends mostly on the sequence of amino acids, i.e. a protein can be considered as a 'message' written in a twenty-letter alphabet. It is known that, in the living cell, protein specificity is determined by nucleic acids. These are linear arrangements of nucleotides, of which there are four different kinds.

Question: what is the minimum number of nucleotides needed, on average, to specify each amino acid? Assume all amino acids to be equiprobable.

III. THE MEASURE OF INFORMATION OR UNCERTAINTY

It seems reasonable to equate the amount of information acquired, as a result of an event, to the amount of uncertainty which its occurrence has

abolished*. The prior uncertainty does not depend on the event that has actually happened, but, rather, on the whole set of events which could have happened at this particular occasion. For instance, if one wishes to compute how much information is acquired, on the average, by a glance at the speedometer, one proceeds to estimate how uncertain a motorist is *before* he glances. The amount of this uncertainty must depend on the number of needle positions which the motorist thinks he can distinguish. Suppose his speedometer scale reaches from zero to one hundred and he can read the position to the nearest mile per hour; then, he will be able to distinguish 101 positions, and the amount of his uncertainty will be somehow related to this number. However, it wouldn't be realistic to relate his uncertainty only to this number, 101. Because, suppose his speedometer scale ranges up to 150 instead of 100 miles per hour; yet, when he is driving along the highway at a moderate speed, this extra portion of scale does not contribute in any way to his uncertainty; he will be quite sure that his needle will not be in this interval. In fact, he will expect to find his needle somewhere within a range of about 10 m.p.h., and he will be almost certain to find it within a somewhat larger range of, say, 20 m.p.h. Thus, to describe his uncertainty realistically, we must not only state every possible result of his reading, but will have to qualify each by a statement of expectation or probability.

The Amount of Uncertainty

As before, we turn to a binary situation to obtain a simple perspective of the problem. Suppose somebody has made a record of 100 tosses of a coin; he has registered only whether the coin fell 'head up' or 'tail up', but neglected all other features such as on what spot the coin came down, which direction the head faced, etc. What is the average amount of information in the record of any one toss? In other words, what is the amount of uncertainty before the record is seen?

The uncertainty must be a function of 'two', the number of alternatives; it must be modified by their relative frequencies. If it is known that the record is that of a coin so thoroughly biassed that 'head' always turns up, then there will be no uncertainty at all; if the coin is moderately biassed, then the outcome of a toss will be uncertain but not quite as much as with an unbiassed coin. If we don't know the bias of a particular coin, then we do not know exactly how uncertain we should feel about the outcome of a toss. If we know that the record contains 60 'heads' and 40 'tails', then a record of 'head' will show up with a probability of .60, a record of 'tail' with a probability .40. The uncertainty can be described by a statement of these probabilities:

Probability of head up 0.60

Probability of tail up 0.40

In the same way we can describe any number of binary uncertainties with a 60–40 choice between any class 'A' and its complement 'non-A'—where 'A' and 'non-A' may be males and females, hits and misses, friends and foes.

* At some time there was some discussion whether uncertainty and information should be given opposite signs. Present usage prescribes the same sign for both.

These uncertainties differ in any number of respects from each other. They will be of interest in very different situations; the kind of information needed to produce certainty is not the same; neither is the usefulness of this information, and so on. However, there is something in common between *all* uncertainties which can be characterized by the probabilities:

Probability of 'A' 60
Probability of 'non-A' . . . 1 — .60

One aspect of this 'something-in-common' is that an arrangement of *any* 60 A's and 40 non-A's can be coded to represent any other 60 A's and 40 non-A's —heads or tails, males or females, hits or misses, friends or foes. Once such representation has been established, then the uncertainty concerning one event will be abolished by information concerning the other. We have previously equated the amount of information with the amount of uncertainty it removes. Accordingly, it can be said that the amounts of uncertainty and information must be equal in all situations characterized by a binary alternative with probabilities .60 and .40.

The foregoing consideration exposes the fundamental features of the measure of information:

(1) Information is a measurable abstract quantity; its value does not depend on what the information is about, just as length, or weight, or temperature have values which do not depend on the nature of the thing which is long, heavy, or hot;

(2) Information is related to the ensemble of possible outcomes of an event; its value depends on the probabilities associated with these outcomes, but not on their causes, and not on their consequences.

What remains is the development of a measure which complies with this concept of 'amount of information'; this is merely a technical problem. An obvious generalization states that whenever two events have the same number of possible outcomes, and identical sets of probabilities are associated with the two ensembles of possible outcomes, then these two events have identical information contents. However, we wish to be able to compare events with quite different probability sets; for instance, we wish to be able to say which uncertainty is greater, that associated with a situation with three equiprobable alternatives, or that where there are four possibilities with probabilities .8, .1, .05 and .05. To answer such questions, we have to derive a measure which is a single number, whatever the number of possible categories and their associated probabilities.

Such a measure is readily derived from the equivalence of uncertainty with the information which removes it. We may represent the information content of an uncertainty-removing piece of intelligence in any manner we wish. We stipulate that this information should be represented in a standard fashion, namely, by using a binary alphabet. In addition we stipulate that the binary representation be coded in such a manner that the expected number of symbols is minimized. We thus obtain a unique number; namely, the minimum average number of binary symbols needed to abolish the uncertainty associated with a given situation. This number will be called the amount of uncertainty or information of this situation.

The function here needed has already been derived as the condition of representability. If two situations can be made to represent each other, then information on one can abolish uncertainty concerning the other. Thus, mutual representability implies equal information content, and representation in the standard binary system yields a general measure of information content. This measure is the 'amount of selective information' as defined by SHANNON and WIENER (4, 5). It is expressed as follows:

Let x be a classification with categories i and associated probabilities $p(i)$; then the information content of x is designated $H(x)$ and given by:*

$$H(x) = -\sum_{i} p(i) \log_2 p(i)$$

The units of this function are the binary digits needed for representation of a given event, and are called *bits*. It must be remembered that the 'bit' is a technical unit of amount of information and not a small piece of information. A single chunk of information may contain many bits or a fraction of a bit.

Some Properties of the Shannon-Wiener Information Function

The Shannon-Wiener information function has been derived (admittedly, in a loose fashion) from a consideration of standard representation of information. We will now consider a number of its properties and see that they correspond closely to the behavior which one would intuitively expect from a good measure of information.

(1) *Independence*—Let i be one of the possible categories of an event x, $p(i)$ the associated probability, and $F(i)$ the contribution of the ith category to the uncertainty. It is desirable that $F(i)$ be a function of and only of $p(i)$. The function

$$F(i) = -p(i) \log_2 p(i)$$

fulfills this requirement.

(2) *Continuity*—A small change of $p(i)$ should result in a small change in $F(i)$; in other words, $F(i)$ should be a continuous function of $p(i)$. The function $p(i) \log_2 p(i)$ is continuous.

(3) *Additivity*—It is desirable that the total information derived from two independent sources should be the sum of the individual information; in other

* The information function looks (except for a scale factor) like Boltzmann's entropy-function; this is not a mere coincidence. The physical entropy is the amount of uncertainty associated with a state of a system, provided all states which are physically distinguishable are considered as different, that is, if the categorization is taken with the finest grain possible. In most situations dealt with in information theory, large numbers of states which are physically distinguishable are lumped into equivalent classes. The category "one light on the steeple" is a good example; an enormous number of physically distinct states are compatible with this definition, but they are all lumped into one class. The distinctions upon which categorizations are based are usually a very small percentage of the distinctions one could make. Thus, physical entropy is an upper bound of the information functions which can be associated with a given situation, but it is a very high upper bound, usually very far from the actual value. For this reason, I prefer not to use the word 'entropy' as synonymous with 'information'.

A very thorough discussion of the relation between information and entropy has been given by BRILLOUIN (9).

words, the uncertainty concerning independent events should be the sum of the individual uncertainties.

Let y be an event with categories j and associated probabilities $p(j)$. Let $p(i, j)$ be the probability of the event pair that x falls into category i and y into category j. Then, the function

$$H(x, y) = -\sum_{i,j} p(i, j) \log_2 p(i, j)$$

will measure the uncertainty associated with the event pair.

If x and y are independent events, then

$$p(i, j) = p(i) \cdot p(j)$$

As a matter of fact, this relation is often used to define independence. In this case, we have

$$\begin{aligned} H(x, y) &= -\sum_{i,j} p(i, j) \log_2 p(i) \cdot p(j) \\ &= -\sum_{i,j} p(i, j) \log_2 p(i) - \sum_{i,j} p(i, j) \log_2 p(j) \end{aligned}$$

It is known that

$$\sum_{j} p(i, j) = p(i)$$

$$\sum_{i} p(i, j) = p(j)$$

Substituting these expressions, we obtain

$$\begin{aligned} H(x, y) &= -\sum_{i} p(i) \log_2 p(i) - \sum_{j} p(j) \log_2 p(j) \\ &= H(x) + H(y). \end{aligned}$$

Thus, the Shannon-Wiener function fulfills the postulate of additivity.

(4) *Natural Scale*—the prototype of uncertainty is that associated with a 50–50 choice. So, the unit of uncertainty should be the uncertainty associated with this situation. In this case, both p's have the value 1/2, and

$$H(x) = -(1/2 \log_2 1/2 + 1/2 \log_2 1/2) = 1$$

Thus, the Shannon-Wiener function is seen to have an appropriate scale factor.

We have derived the information function from the postulate of efficient binary representation, and have found that the function so defined has the desirable properties of independence, continuity, additivity, and natural scale. We could have started differently, setting up these four properties as *postulates*. It can be shown that these four postulates (or other sets of four similar postulates) define uniquely the Shannon-Wiener function. Working it this way, we would have derived the fact that the function so defined has the desirable property of efficient binary representation.

The function $F(p)$ is plotted against p in Fig. 1. The graph shows a curve which originates and terminates at $F = 0$, and has a flat top with a maximum

of $F = 0.53$ for $p = 0.37$. Inspection of the graph reveals some more important properties of the function $F(p)$:

(5) $$F(0) = 0:$$

When a particular class of events is certain not to occur ($p = 0$), then it does not contribute to the measure of uncertainty.

(6) $$F(1) = 0:$$

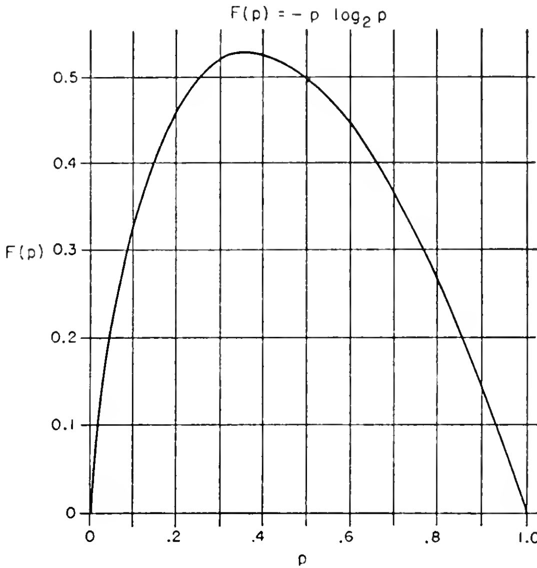


FIG. 1. Graph of $F(p)$ as a function of p

When a particular class of events is certain to occur ($p = 1$), i.e. excludes all other classes, then there is no uncertainty about the outcome.

(7) *Effect of Averaging:*

$$F\left(\frac{p_1 + p_2}{2}\right) \geqslant \tfrac{1}{2}[F(p_1) + F(p_2)]$$

The function of the average is greater than or at least as large as the average of the function. When the probabilities associated with two disjoint categories are averaged, then the uncertainty becomes larger. Figure 2 is a graphical demonstration of this effect.

The extreme case of averaging occurs if all r categories in a classification are considered equiprobable. Then,

$$p(i) = \frac{1}{r}$$

$$\text{max. of } H(x) = -\sum_{i=1}^{r} \frac{1}{r} \log_2 \frac{1}{r} = -r \cdot \frac{1}{r} \log_2 \frac{1}{r}$$

$$\text{max. of } H(x) = \log_2 r$$

In particular in a binary classification,

$$r = 2$$

$$\text{max. of } H(x) = 1$$

Thus, the maximum uncertainty associated with two alternatives is one bit; it occurs if both alternatives are equally probable (this is the case of the unbiassed coin!).

(8) *Effect of Pooling:*

$$F(p_1 + p_2) \leqslant F(p_1) + F(p_2)$$

The function of the sum is smaller than the sum of the functions. That is, pooling of two classes in one equivalence class reduces uncertainty (exactly

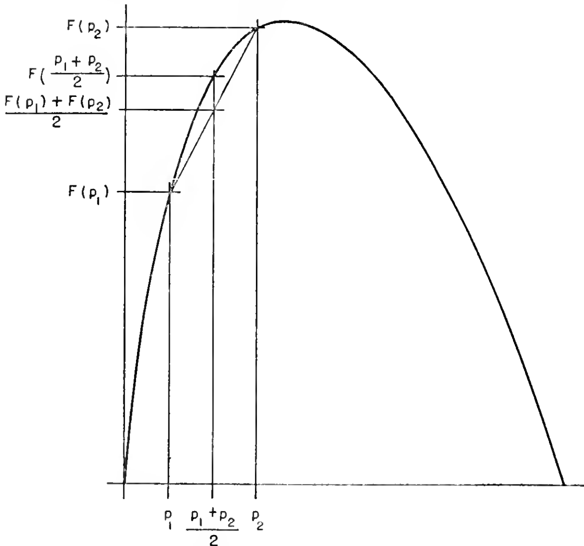


FIG. 2. Graphical demonstration of the effect of averaging

by that uncertainty which is associated with the distinction between the two pooled classes). Extreme pooling results in a single category with probability 1; this means uncertainty 0. Figure 3 demonstrates the effect of pooling.

The function $F(p) = -p \log_2 p$ has been tabulated. The reader is advised to use Fig. 1 to obtain approximate values for use in working the exercises below. For more precise values, one of the existing tables may be consulted (10, 11).

EXERCISES

7. Compute the uncertainty associated with:

$$p(\text{A}) = .60$$
$$P(\text{non-A}) = .40$$

8. Compute $H(x)$ for two alternatives, and plot the value against p(A).

9. Answer the question posed previously: which uncertainty is greater, that associated with a situation (x) with three equiprobable alternatives, or that (y) where there are 4 possibilities with probabilities .8, .1, .05 and .05.

10. Estimate the uncertainty of a motorist like the one described at the beginning of this section.

11. Certain languages have considerably fewer letters than English (that is, about 18 to 20), yet the information content per letter is nearly the same. How is this possible?

12. A situation has an unlimited number of alternatives, with probabilities of 1/2, 1/4, 1/8, 1/16, etc. in geometric progression. What is the measure of uncertainty?

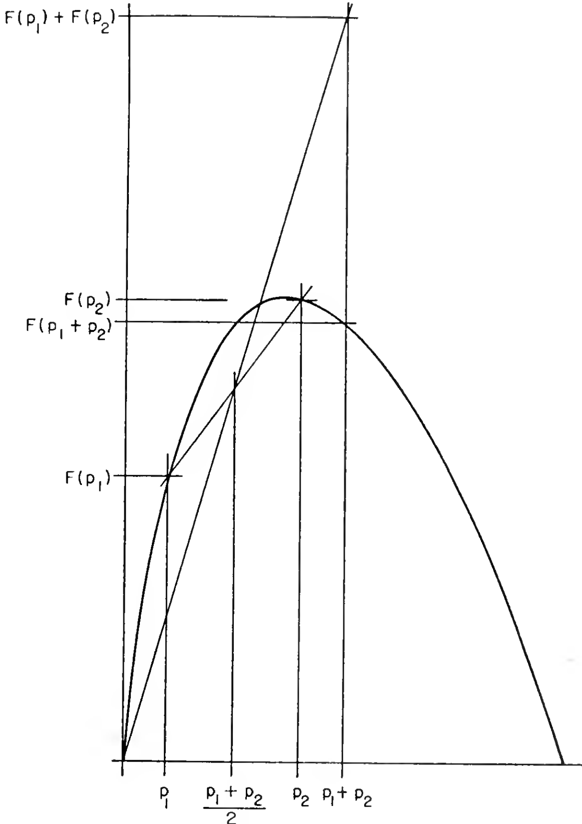


Fig. 3. Graphical demonstration of the effect of pooling

The function of the sum is on the intersection between the curve and the ordinate over the sum; the sum of the functions is on the intersection of the same ordinate with a straight line through the origin and the midpoint of the straight line which connects the intersections of the curve with the ordinates over p_1 and p_2, hence:

$$F(p_1 + p_2) \leq F(p_1) + F(p_2)$$

IV. INFORMATION MEASUREMENTS PERTAINING TO TWO RELATED VARIABLES

In the two preceding sections we have discussed how to represent information, and how to measure amounts of information. Both procedures become important if information is to be manipulated. The manipulation most commonly used is *communication.*

In information theory, we use the word 'communication' in a wider sense than usual—just as the word 'information' is used in a wider sense than usual. We understand by 'communication' any relation between variables, accomplished by any means whatsoever, conscious or otherwise, provided that it results in a mutual reduction of uncertainty. For instance: if one watches one of two tennis players, without looking at the other, he derives a considerable amount of information about the unseen player's action. Thus, the seen player *transmits information* about the unseen player—although in this case, the transmission of information is incidental and not normally utilized, as one ordinarily looks at both players.

An Example of Two Related Variables

The following example is purposely selected to represent an instance of unintentional communication. The table below is based on Pearson and Lee's measurements of heights on 1376 father–daughter pairs. To simplify the analysis, we have grouped the data in coarse intervals of 3 in. each, and converted all frequencies into percentages.

Table II. Heights of Fathers and Daughters; Probabilities and Information Measures

Joint probabilities of heights, $p(i,j)$
(Pearson and Lee's data, 1376 father–daughter pairs)

	y^*: j‡ = 59.5	62.5	65.5	68.5	71.5	74.5	$p(i)$	$-p \log_2 p$
i‡ = 53.5	—	.001	—	—	—	—	.001	.01
56.5	.001	.007	.006	.001	—	—	.015	.09
59.5	.005	.022	.060	.027	.005	—	.119	.37
x† 62.5	.004	.042	.156	.152	.039	.001	.394	.53
65.5	—	.009	.075	.175	.095	.010	.364	.53
68.5	—	.001	.011	.035	.039	.010	.096	.32
71.5	—	—	—	.003	.006	.002	.011	.07
$p(j)$	.010	.082	.308	.393	.184	.023	1.000	1.92
$-p \log_2 p$	.07	.30	.52	.53	.45	.13		
					2.00			

* height of fathers, in 3 in. intervals
† height of daughters, in 3 in. intervals
‡ center of intervals

Information Functions:

$$H(x) = -\sum_i p(i) \log_2 p(i) = 1.92 \text{ bits}$$

$$H(y) = -\sum_j p(j) \log_2 p(j) = 2.00 \text{ bits}$$

$$H(x) + H(y) = 3.92 \text{ bits}$$

$$H(x,y) = -\sum_{ij} p(i,j) \log_2 p(i,j) = 3.70 \text{ bits}$$

$$T(x;y) = H(x) + H(y) - H(x,y) = 0.22 \text{ bits}$$

From the marginal sums, the uncertainties concerning the height of daughters, $H(x)$, and of fathers, $H(y)$, are computed as described in the preceding section. The uncertainty concerning *both* heights in a father–daughter pair is computed in similar fashion from the joint probabilities, $p(i,j)$. This function is properly called the *joint uncertainty*, or uncertainty of the two-part system; its symbol is $H(x,y)$. It is compared to the sum of the two individual uncertainties. If the two heights were completely independent of each other, then the joint uncertainty should be equal to the sum of the individual uncertainties. In our case, it is smaller by 0.22 bits. The deficit is a *measure of the internal constraints* in the system, which lead to an association between heights of fathers and daughters. The function is designated by the symbol $T(x;y)$. Its defining equation is:

$$T(x;y) = H(x) + H(y) - H(x,y)$$

This information function is germane to other statistics which measure the relatedness of two variables, such as the coefficients of correlation and of contingency. The T-measure is of very general applicability; the values of the variables do not have to be quantitative, not even ordered—they must only be distinguishable. For instance, one can compute a T-measure for a relation between color and shape.

The two functions, H and T, differ in the way in which they are affected by change of scale. Let us consider what would have happened if he had chosen one-inch intervals instead of three-inch intervals. It could be the case that only one one-inch interval out of any group of three is occupied at all. Then, the information that a certain height falls into a given three-inch interval would automatically locate it in some one-inch interval; hence, the uncertainty is not increased by the subdivision of intervals. However, this is an extremely unlikely situation. It is much more likely that the three one-inch intervals are populated with approximately equal frequencies. In this case, additional information of $\log_2 3 = 1.58$ bits is needed to specify the proper one-inch interval. Then, the uncertainty concerning the height of fathers with regard to a one-inch scale will be $2.00 + 1.58 = 3.58$ bits, and the uncertainty concerning the height of daughters $1.92 + 1.58 = 3.50$ bits. The joint uncertainty will be increased by a factor of $\log_2 9 = 3.17$, because each cell in the table will be replaced by nine cells as one goes from three-inch intervals to one-inch intervals. If one uses a still finer grain, going from inches to millimetres, then the individual uncertainties can be increased by another 4.7 bits, the joint uncertainty by 9.3 bits. This is quite the expected behavior. The more categories are recognized, the greater the uncertainty of classification. The uncertainty can become infinite for a continuous function. However, it will always remain finite for any set of real observations.

T, on the other hand, depends very little on the scale interval used. With very coarse grouping, T tends to be less. In the extreme cases, where all heights are pooled into one single class, all individual and joint uncertainties vanish, and with them their differences. In the other extreme case, where measurements are taken and registered to so many digits that no two results are alike, we must get $H(x) = H(y) = H(x, y) = T(x; y) = \log_2 1376$. But, between these unreasonable extremes, the measure of constraints is characteristic of the system and not of the scale which is used in measuring it.

Two-part Systems in General

We proceed to a general treatment of a two-part system x, y. Let i and j be the categories of x and y, respectively, and $p(i)$ and $p(j)$ the associated probabilities. Further, let $p(i, j)$ be the probability of the joint occurrence $[(x = i)$ and $(y = j)]$.
Then:

$$H(x) = -\sum_i p(i) \log_2 p(i)$$

$$H(y) = -\sum_j p(j) \log_2 p(j)$$

$$H(x, y) = -\sum_{ij} p(i, j) \log_2 p(i, j)$$

We introduce the conditional probabilities,

$$p_i(j) \ldots \text{Prob } \{y = j \text{ if } x = i\}$$

$$p_j(i) \ldots \text{Prob } \{x = i \text{ if } y = j\}$$

When $x = i$ then y must have *some* value j with certainty (or probability 1.0), that is

$$\sum_j p_i(j) = 1$$

Equally,

$$\sum_i p_j(i) = 1$$

Furthermore, the probability of the joint occurrence $[x = i$ and $y = j]$ can be factored into the product of the probability that x equals i, times the conditional probability that $y = j$ if $x = i$; equally, it can be factored into the product of $p(j)$ times $p_j(i)$. So:

$$p(i, j) = p(i) \cdot p_i(j)$$

$$= p(j) \cdot p_j(i)$$

The conditional probabilities yield naturally conditional uncertainties. For instance, the uncertainty of y, if it is known that $x = i$, will be

$$H_i(y) = -\sum_j p_i(j) \log_2 p_i(i)$$

The average uncertainty of y, under the condition that x is known, is designated by $H_x(y)$. It is obtained as the weighted average of the $H_i(y)$'s:

$$H_x(y) = \sum_i p(i) \cdot H_i(y)$$

Substituting the value of $H_i(y)$, we get

$$H_x(y) = -\sum_i p(i) \sum_j p_i(j) \log_2 p_i(j)$$

and remembering that

$$p_i(j) = \frac{p(i, j)}{p(i)}$$

we get

$$H_x(y) = -\sum_{ij} p(i, j) \log_2 \frac{p(i, j)}{p(i)} .$$

Expanding the logarithm gives

$$H_x(y) = -\sum_{ij} p(i,j)\log_2 p(i,j) + \sum_{ij} p(i,j)\log_2 p(i).$$

Noting that

$$\sum_j p(i,j) = p(i)$$

we get

$$H_x(y) = -\sum_{ij} p(i,j)\log_2 p(i,j) + \sum_i p(i)\log_2 p(i).$$

We have seen that the first term on the right side is $H(x, y)$ and the second $-H(x)$. So:

$$H_x(y) = H(x, y) - H(x) \qquad \text{and} \qquad H(x, y) = H(x) + H_x(y)$$

A parallel development shows that

$$H(x, y) = H(y) + H_y(x)$$

This relation is quite obvious if put into words: the joint uncertainty concerning two variables is equal to the sum of the uncertainty concerning either one variable plus the conditional uncertainty concerning the second variable if the first one is given.

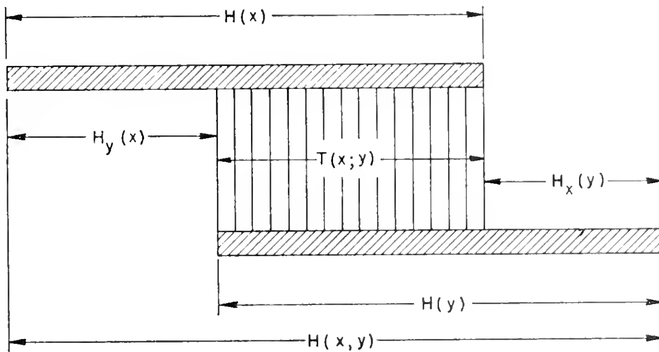


FIG. 4. The relation between information functions shown graphically

The difference in uncertainty concerning y, depending on whether or not x is known,

$$H(y) - H_x(y),$$

is the *gain* in certainty about y derived from observing x. Substituting for $H_x(y)$, we get:

$$H(y) - H_x(y) = H(y) + H(x) - H(x, y)$$

The expression on the right side is the defining equation for $T(x; y)$:

$$H(y) + H(x) - H(x, y) = T(x; y).$$

It follows from this derivation that T is a symmetrical function:

$$T(x; y) = T(y; x) = H(x) - H_y(x) = H(y) - H_x(y)$$

and it becomes clear why T is a measure of the mutual reduction of uncertainty. The relations between the six information functions, $H(x)$, $H(y)$, $H(x, y)$, $H_x(y)$, $H_y(x)$ and $T(x; y)$, can be demonstrated graphically as in Fig. 4.

In normal code representation, i.e. reduced to efficient binary operations, the information functions have the following meaning:

$H(x)$ ....number of operations which specify x

$H_y(x)$ ....no. of operations which specify x if y is given

$T(x; y)$....no. of operations which apply to the specification of both x and y

$H(x, y)$....no. of operations which specify the whole system.

Inspection of the graph shows that:

$$H(x) \geqslant H_y(x)$$

$$H(y) \geqslant H_x(y),$$

that is, the conditional uncertainty cannot be greater than the unconditional uncertainty.*

Communication Systems

When a system not only transmits information but exists primarily for that purpose, then it is called a communication system. No class of two-part systems has received as much attention as that of the communication system. In a simple communication system, the two parts are called the *source* and the *destination* of information. The distinction between source and destination must be based on external grounds; the informational relations between the two are perfectly symmetrical. The relevant states of the source are called the *inputs*, or *signals sent*, and the relevant states of the destination are the *outputs*, or *signals received*. A single state is called a *symbol*, and a higher unit composed of several symbols, a *message*. The conditional probabilities for each pair of signals sent and received form a matrix called the *channel*. Note that the word 'channel' is again used in a sense wider than customary. A 'channel' may but does not have to be a means of physically conveying information. For instance, if two variables x and y do not affect each other but are both affected by a third variable z, then knowledge of the state of x is likely to reduce the uncertainty concerning the state of y, and vice versa; hence, information is transmitted between the two variables, and they are connected by a 'channel' in the sense of information theory—although they do not communicate with each other directly.

* However, this is true only for an *average* conditional uncertainty, and does not apply to every *particular* condition. The following example will help to fix the ideas: Consider a diagnostic test for a certain disease; suppose the nature of the test and the occurrence of the disease are such that in 98 per cent of the patients the test is negative; that of the positive tests, 50 per cent are spurious; and that virtually every case of the disease will give a positive test. Then, if the test is not performed at all, the diagnostician's uncertainty as to the presence of the disease in any given patient, is

$$-(.99 \log_2 0.99 + .01 \log_2 0.01) = .081 \text{ bits/patient.}$$

If the test was negative then the uncertainty is zero. But, if the test is positive, the chances are equal that it is or is not spurious; hence, the uncertainty is 1.0 bit, and the diagnostician is more in doubt than he was before. However, the average uncertainty, conditional upon his performing the test, is reduced to

$$.98 \times 0 + .02 \times 1.0 = 0.020 \text{ bits/patient.}$$

The information functions in a communication system are designated as follows:

$H(x)$ ....uncertainty of source
$H(y)$ ....uncertainty of destination
$H_x(y)$ ....ambiguity
$H_y(x)$ ....equivocation
$T(x; y)$....information transmitted, or communicated

Amounts of information transmitted must be referred to some unit of action. In particular, it is customary to compute transmissions *per symbol* or *per unit time*.

A channel which associates one and only one output with each input, and no output with more than one input, is called a *noise-free* channel or transducer; in this case,

$$H(x) = H(y) = H(x, y) = T(x; y);$$

$$H_x(y) = H_y(x) = 0.$$

We can think of a noise-free channel as a means by which information at the source is represented at the destination. Physically, this involves two acts of representation: first, states of the channel are selected so as to represent the inputs, according to some agreed-upon code; this is called *encoding*. Next, the states of the channel are translated into meaningful states at the destination; this is called *decoding*. All we have stated about representation, representability and amounts of information could now be restated in terms of encoding and decoding operations. In this sense, the relation which we introduced as the 'condition of representability' is also known as the *Theorem of the Noise-free Channel*; and all the examples and exercises of representing information could be re-interpreted as coding operations.

Noise—Few real channels are noise-free; in general, more than one output can follow a particular input. For instance, the 'channel' which links a daughter's height to her father's is far from noise-free; the following table gives the conditional probabilities:

Table III. Data of Table II in Form of a Communication Channel

	Conditional probabilities, $p_j(i)$							$H_j(x)$
	$i = 53.5$	56.5	59.5	62.5	65.5	68.5	71.5	
$j = 59.5$	—	.10	.50	.40	—	—	—	1.36
62.5	.01	.09	.27	.51	.11	.01	—	1.80
65.5	—	.02	.19	.51	.24	.04	—	1.74
68.5	—	—	.07	.39	.45	.09	.01	1.70
71.5	—	—	.03	.21	.52	.21	.03	1.74
74.5	—	—	—	.04	.45	.43	.09	1.55

The last column, $H_j(x)$, is the uncertainty concerning the height of the daughter if the height of the father is known; it is not too surprising to find this uncertainty smallest in the extreme cases, and always smaller than the unconditional uncertainty of 1.92 bits.

The father's height 'communicates' some information about the daughter's height; the amount communicated is 0.22 bits. It is not more than that for a number of reasons. Some of the deficit in information about the daughter's height is undoubtedly due to ignorance, and could be reduced by taking proper account of various concomitant factors. Some of the uncertainty may be irreducible, due to a truly random process—possibly the selection of the particular chromosomes which go into determining the daughter's height. In the strict sense, the term 'noise' is reserved for the effects of random disturbances, and not to the effects of ignorance. However, the problem of the final distinction between uncertainty due to randomness and uncertainty due to ignorance is an extremely delicate one; the practical information analyst will usually be satisfied to treat any uncertainty as due to noise, which results in the greatest reduction of certainty. This interpretation will be subject to revision in the light of additional knowledge.

The two-part system 'father's height–daughter's height' is not a communication system, and this is one reason why so little information is transmitted. Suppose the numbers which define the 'father's heights' categories were not observed in a given population but could be chosen arbitrarily; for instance, they might be input voltages applied to a system. Accordingly, the 'daughters' heights' might be output voltages, and the table of conditional probabilities becomes a statement of the transfer function of the system. It is obvious that this system can be made to transmit more than 0.22 bits per symbol. For instance, using only $j = 59.5$ and $j = 74.5$, with equal frequencies, one would transmit about .90 bits per signal. In general: for each channel, $p_i(j)$, there exists a set of input probabilities, $p(i)$, which maximizes the transmission rate. The rate so obtained is called the *channel capacity*.

Even with best utilization of the possibilities of a channel, it can do no more than transmit all the input information, and in general it will not transmit quite all of it. This leads to an important generalization: *Manipulation of information cannot increase its amount; it can at best preserve it, and it is likely to reduce it.*

This important statement will be clarified by the discussion of an apparent exception. Suppose A wishes to send a message to B over the channel C; conditions being very good, B picks up not only almost perfectly the message sent by A but acquires, in the course of doing so, considerable amount of information about conditions in the channel. His total information received might be more than that contained in A's message; still, he has lost some of the information contained in the message. In general: as a result of manipulating information, there can be more output information than there was input information—but the contribution of the input information to the total cannot be more than the amount of input information.

Error Detection and Correction

A codebook states which output should be associated with any given input. A noise-free channel fulfills these requirements perfectly. In a noisy channel

other outputs than the required ones appear; in other words, a noisy channel produces errors. Errors lead to loss of information, and a reduction in the rate of transmission; in a noisy channel,

$$T(x; y) < H(x)$$

$$H_y(x) > 0.$$

This loss is unavoidable. However, it is at least possible to spot and correct the errors which have occurred. It is one of the main endeavours of information theory to devise methods to do this efficiently.

An error in a message can never be found unless the message contains some extra information which can be used for this purpose. For instance, if the message consists of a string of four digits chosen without any constraint:

5 3 8 7,

one has absolutely no possibility of knowing whether or not it contains any errors. If it has been agreed upon that the message will be repeated, then one can detect errors:

5 4 8 7
5 3 7 7,

and if the message is repeated several times, these errors can be detected and corrected, with arbitrary certainty if the number of replications can be made sufficiently large:

5 3 8 7
5 3 7 7
5 3 8 7
5 4 8 7
5 3 8 1.

In the second case, the possibility of error detection was bought at the price of making two digits do the work of one; the message is said to be 50 per cent *redundant*. In the last case, the price of error correction is the use of five digits to transmit a single one, or a redundancy of 80 per cent.

Introducing redundant information in the form of a simple replication is straight-forward and effective, but not very economical. Error detection could be achieved more efficiently by simply adding the sum of the digits to the message:
be achieved more efficiently by simply adding the sum of the digits to the message:

5 3 8 7 2 3.

Here, the redundant information is only one-third of the total. In fact, giving only the last digit of the sum as 'signature' is almost as effective, and requires only 1 digit in 5, or 20 per cent redundant information. The signature check illustrates a general principle: a given amount of redundant information in a

message can be used for error checking the more effectively the more evenly it is related to all parts of the message.

It is always possible to achieve reliability, in the presence of noise, by the use of redundant information; in fact, one can approach perfect reliability arbitrarily closely if one is willing to provide enough redundant information. The amount of redundant information needed, for a given noise level and a given desired reliability, will depend on the efficiency of coding. The ideal relation between noise level and redundant information needed is formulated in Shannon's fundamental *Theorem of the Noisy Channel.* This theorem can be stated as follows: if a certain amount of information is to be transmitted with perfect reliability in the presence of noise, then it is necessary to provide at least as much redundant information as the amount of equivocation introduced by the noise; furthermore, this amount will be sufficient if the coding is maximally efficient.

There exist several proofs of this theorem; none of them is easy to follow, and all are existence proofs—that is, they prove that an error-checking code exists which will fulfill the requirements, but they do not say how to construct it. In fact, perfectly efficient error-checking codes seem to be realizable only in a few special cases; however, close approximations to ideal efficiency are easily obtained if it is permissible to use message blocks of great length (12).

The economics of error-checking are dominated by three factors:

(I) the frequency and costliness of errors

(II) the cost of adding redundant information

(III) the availability and costliness of checking procedure (encoding and decoding).

The work of Shannon and his followers has dealt with one particular situation: encoding and decoding procedures are supposed to be reliable and gratis, the error frequency is to be reduced to almost zero, and redundant information is supposed to be used as sparingly as possible. As long as the theory is not completed even for this case, one cannot expect to develop a more general theory. Some qualitative notions of what it will entail can be gathered from a consideration of a much-used, and presumably well developed communication system, namely, printed language. Symbols are gathered into various checking units (words, sentences, paragraphs, chapters); on each level, there operate constraints which will help to locate and correct errors. For instance, this sentence will be read corretly even though one letter has been onitted and one word misspelled. It seems that the redundancy per letter, in a coherent English text, is about 60 per cent. Paragraphs are constructed in such a way that the sense can be grasped even if whole words or even sentences are missing or perturbed, and the essence of a whole chapter is, in general, understandable even if a whole paragraph should be left out.

Actual Communications System

So far we have dealt with two-part systems in a purely abstract way. 'Sources' and 'destinations' are defined simply by the states which they can assume. 'Channels' are tables of conditional probabilities; in the simplest case, the channel is a kind of telephone book which associates every input to some

particular output. If the association is not unequivocal, then the channel is said to be noisy. 'Noise' is defined as a random perturbation of the input-output link. Those are nice, clean concepts, not to be confused with realities. The 'channel' exists on paper only, and is not the same as the mechanism which links two parts of a system. The informational relation between heights of fathers and daughters does not reveal the nature of the mechanisms involved; whether fathers affect their daughters' heights by means of their genes, or of the food they provide, or of the mother they select for them, cannot be decided on grounds of informational relations. Indeed, I believe that Buddhist tradition would explain the correlation on the grounds that daughters select their fathers; as far as information theory is concerned, this is perfectly acceptable.

The scheme shown in Fig. 5 is a somewhat closer approximation to reality:

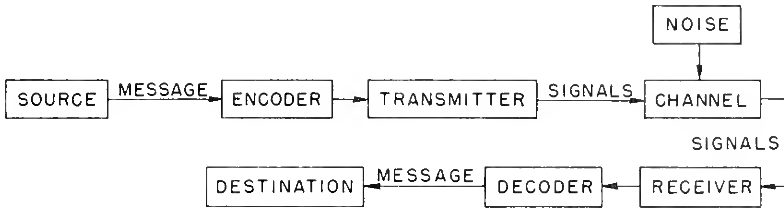


FIG. 5. A diagrammatic representation of a communication system

It is customary to treat all links but the channel as noise-free. If need be, one can introduce noise into the other links of the model by some straight-forward adaptations.

If signals and channels are physical entities, then it is relevant to investigate their physical capacity of carrying information. Suppose the nature of a unit of action and the physical constraints are such that the channel can assume any one of m states during one unit of action; then, these states can be made to represent $\log_2 m$ bits of information. It is the function of the encoder-transmitter system to match the diversity of messages generated by the source to the diversity of states which can be assumed by the channel; those, in turn, are matched to the diversity of messages intelligible at the destination by the receiver-decoder system.

As long as the demands on the channel are light, the matching process is not much of a problem. However, it may become very difficult if the channel is to be driven at capacity, and if the various states of the channel are not of equal value; some may be more subject to noise effects than others, some may need more time than others, some may necessitate more effort than others. In general, one will tend to favor the safest, shortest, and easiest states. However, this must not go too far; if one goes to the extreme of using the very 'best' state, then the channel does not transmit any information at all. To find optimum compromises between informational needs of source and destination and physical capacities of the channel, between amount of information used to carry messages and amount of information needed for noise reduction, is one of the fundamental problems of the theory of information and communication.

EXERCISES

13. The following table gives the number of times the four possible combinations of two flower colors with two pollen shapes were found:

Pollen shape	Flower color: Purple	Red
Long	296	27
Round	19	85

Is there information transmission between these two characters?

14. Define the following functions, and derive their values (in terms of H-functions)

$$T(x, y; z)$$
$$T(x; y, z)$$
$$T(x; y; z)$$

15. A diagnostic test gives the following results:

true negatives ..	85%
false negatives ..	5%
true positives ..	3%
false positives ..	7%

What is the informational value of the test?
What is the maximum informational value that any test could give in this situation?

16. A teletype machine sends 2.3 groups of five binary symbols per second. What is the maximum possible rate of information transmission?

17. Same machine as in Exercise (16). All code groups are equiprobable. Error probabilities are as follows: symbols nos. 1 and 4 are always received correctly, nos. 2 and 3 are wrong 11 per cent of the time, no. 5 is wrong 1 per cent of the time. All errors are equiprobable. Compute equivocation and amount of information transmitted.

18. You are to send 2-bit messages through a channel which has the property that one in five binary symbols is bound to be in error. Construct four sequences of five binary messages which will allow the reconstruction of the original message. What is the efficiency of the code?

V. ORGANIZATION

Systems, Structures, Pattern

A *system* is an organized whole made up of interrelated parts. *Organization* is based upon the interrelations between parts. The parts may be strongly or weakly coupled; their effect on each other may be quantitative or qualitative.

FIG. 6. A simple communication network

If two parts are coupled in any fashion, then knowledge of the state of one must imply some information about the state of the other. Accordingly, any interrelation can be technically represented as a channel. So, two components of a system can be symbolically represented by a simple communication network of two parts, referred to as two nodes and one channel:

Let $H(x)$ be the amount of information needed to know what state x is in. If y is known, some of this information becomes unnecessary, or redundant. This amount, $T(x; y)$, is an index of the degree of coherence, constraint, integration, or organization which prevails in the system.

Consider the pair of words 'green valley'. These two words form a small system—a whole made up of interrelated parts. The whole has a meaning which neither part alone has. The price for this feature is elimination of many other possible connotations of 'green' and 'valley'. As a result, the information content of the word combination is smaller than the combined information contents of the two words. The difference must show up as redundant information. The presence of redundancy implies that each word contains some information about the other. This is best demonstrated by successful error checking. The errors 'preen 'for 'green', and 'volley' for 'valley' would not be found in isolated words, but can be spotted in the pair.

System Analysis—There seem to be three general viewpoints under which relations within a system are assessed: (a) the amount of information transmitted—on the technical, semantic and pragmatic level; (b) the degree of control or cause-effect relations, dominance; and (c) the utility, or value, of the relation to one or both of the related parts. Information theory deals only with the first viewpoint. It does not concern cause-effect relations, or what causes the information to flow, and it is not concerned either with the utility of the flow of information.

Informational analysis of a system will be of interest if and only if the informational challenge is serious, that is, when a system has to process information at a rate which crowds its capabilities. The informational challenge is the result of:

(1) The diversity which is characteristic of the tasks; this can be expressed as H/task. A system which is faced with the same task all the time or most of the time may be working very hard but the difficulty is not an informational one.

(2) The precision which is required; this can be expressed as the ratio T/H. That is, the diversity of tasks is informationally challenging only insofar as it is expressed in a diversity of responses. A system with a small response repertoire may be working very hard, but not in the informational domain.

(3) The time which is allotted for the fulfillment of each task. A system with very modest informational equipment can solve many tasks if given ample time. For instance, the extremely simple logical machine devised by Turing (13) will solve any solvable problem if given very much time.

The time rate of informational challenge of the system is the product

$$\frac{H}{\text{task}} \times \frac{T}{H} \times \frac{\text{tasks}}{\text{unit time}} = T/\text{unit time}.$$

The informational output of the system will be measured in H-measures but the effective output, or informational performance, in terms of T-measures, as T per task or T per unit time. The limits of the informational performance of a system can be found by systematically varying the informational challenge and observing the resulting performance. In such studies it is important to

make sure that the system's performance is limited informationally, and not by difficulties of sensing inputs or generating outputs.

It is possible to vary the informational challenge in a number of modes; e.g. one can vary the number of sources of information, or the amount of information per source. Challenging in various modes reveals whether or not there exist several modes of limitation. It seems that the informational performance which a system can produce in single tasks may be limited by the following factors, singly or in conjunction:

(1) the amount of information which can be processed effectively in a single task,
(2) the number of independent information-carrying components which can be involved in a single act of information-processing,
(3) the informational contribution from each independent component,
(4) all information-carrying components must be assembled within a certain length of time;
(5) in addition, there seem to be two general limitations on time rates: there is a minimum time for each act of information processing, and
(6) the over-all rate of information-processing is limited (only this last limitation has the character of a channel capacity).

This list of limitations is based on psychological experiments (14) but is believed to apply to all types of systems.

Multi-part Systems—The informational system analysis is not restricted to two-part systems. A system of three components can be represented as a three-node network with a connecting channel:

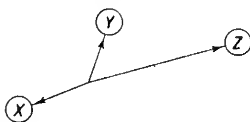


FIG. 7. A simple three-node network

Again, it is merely a matter of convenience which node, or set of nodes, one treats as the input, or independent variate.

The treatment can be extended to any number of components. Thus, a nine-node network is equivalent to one man receiving information from eight sources, or feeding information into eight sinks; or, to four men watching two sources, communicating with each other, and feeding information into three sinks; to a sentence of nine words; to a decision based upon eight factors.

The more parts there are to a system, the more difficult becomes the informational analysis (15, 16). This is territory that has been but recently opened, and we are still largely concerned with the formulation and highly tentative application of concepts. It will be helpful to consider a parallel effort, namely, the study of organization by game theory (17). One result of this study is that each time a new player is added, the organization (the 'game') acquires a new qualitative feature. One-person games deal with problems of *maximum*; the addition of a second person introduces *competition*; of a third person, *coalition*; of a fourth person, an *asymmetric role* of one player in relation to the group of the other three. VON NEUMANN (17) points out that it is at this junction that the most remarkable problems begin to appear; also at this junction,

there occurs a change from a rigorous and complete exposition to a heuristic and incomplete one.

The situation is similar in the study of organization by information theory. Each time a new part is added to a system, a qualitatively new information function appears. As long as one deals with a single variable, the problem is one of *efficient use of existing variations*. A two-part system introduces *relations between parts*; a three-part system, *relations between relations*; a four-part system, *relations between a part and a complex of relations*.

Unitization—It is an empirical fact that when a system is complex enough to require very many components, the phenomenon of *unitization* occurs. That is, some components get organized in such a way that they interact strongly among each other, and act as a unit with respect to the remainder of the system and the external world. Unitization seems to be a necessary evil; it might be an important key for the study of complex organization and complex mental activities. The phènomenon has never been really explained; it is possible that a quantitative treatment will be made possible through the use of information theory (18).

Unitization is always coupled with the phenomenon of limited span. Any real part has a limited information content. In any single act of communication, the capacity for non-redundant transmission of a part is limited by its own information content. This amount must somehow be partitioned into interaction with the external world, and interaction with the other members of the unit. If each of these interactions is to be of significant size, then only a limited number is possible. The interaction of a unit with the outside may be only a fraction of the information traffic within the unit. Hence, several units can be organized into a secondary structure of greater versatility, and this process can be repeated on successive levels of organization.

There appears, thus, a possibility that information theory can be helpful in formulating both the causes and the effects of unitization, and in establishing rational interpretations of the size of the units. This would be a very important contribution to any theory of organization.

Conclusion—We have proceeded from simple processes of representation to discussions of communication and, finally, organization. It was attempted to treat in a heuristic and perspicuous manner the basic principles of Information Theory: there exists a generalized concept of 'information' which includes communication and organization and is so general that every real event or structure has its informational aspects; this general concept is related to a measurable quantity; the operation of taking a measurement of this quantity is done by means of symbolization in a standard language. The functions as defined obey two fundamental theorems: the Representation Theorem, and the Theorem of the Noisy Channel. Both theorems impose a limit on the amount of information which can be effectively processed in a given situation; both also state that it is possible to reach this limit.

APPENDIX I

THE EVALUATION OF INFORMATION CONTENT

The examples and exercises should have familiarized the reader with the techniques of taking information measurements. However, the investigator

who wishes to use this knowledge in his field is bound to run into some difficulties. A typical difficulty is that a natural situation does not present itself neatly classified with a complete set of categories and probability measures. It often takes considerable ingenuity to supplement the missing components of the picture. Wherever ingenuity must be used, the result will not be unequivocal. Hence it becomes important to estimate not individual information measures but rather whole ranges compatible with reasonable assumptions.

The Relativity of Information Measures

'Information content' is a measurable quantity, just as length; and, just as length, it is a function and not a property of a particular set of events. The theory of relativity asserts that the measured length of an object depends on certain relations between the object and the measuring system. However, under everyday conditions these relations will not produce any significant effect and, most of the time, lengths behave as if they were properties of objects. The information content of an event depends on the manner in which this event is related to the frame of reference of the evaluating system. Unlike with length, these relations are not fixed under everyday conditions. Therefore, information content behaves only rarely as if it were a property of an event.

The amount of information, $H(x)$, associated with an event, x, is defined as the expectation of the logarithm of the probability that x will fall into some category, i. Thus, the measure of information depends on three decisions:

(1) the choice of a unit event,

(2) the establishment of categories,

(3) the selection of a set of probability measures.

In general, each of these decisions involves a degree of arbitrariness. Accordingly, a considerable range of information measures will be compatible with a given real situation.

The question of an appropriate selection of a unit event cannot be solved by mechanical application of hard and fast rules. There is a lower limit to the size of elements, imposed by limits of observability. In general, selection of these lower limits will force one to take cognizance of a tremendous amount of detail, most of which is bound to be irrelevant. Thus, one will try to select a unit event broad enough that all irrelevant details are submerged in its internal structure, yet narrow enough so that no relevant relations get lost within the unit event. In practice, one has to make a guess, subject to revision by later experience. This difficulty occurs with all kinds of analyses, and is not specific to informational analysis.

The situation is quite similar with respect to categories. There, too, exists a bound, imposed by the capabilities of discrimination. In general a large number of discriminations can be made which are irrelevant to the problem at hand. For instance, if one deals with the semantic content of a printed message, it will be quite irrelevant to categorize by shapes of letters, quality of paper, type of printing ink, etc. The decision is not always so easy. For instance, in categorizing the atoms found in living matter it will, by and large, not be necessary to distinguish between isotopes; in the overwhelming majority of occasions, differences between isotopes will have no effect. Occasionally, of

course, a particular isotope located in a sensitive spot and decaying at a critical moment can have very large effects. In a case like this, the selection of a set of categories becomes a matter of compromise.

The probabilities, finally, are never actually known. We have to estimate them, on more or less sound bases. In many situations where generalized information theory is used, the bases for estimating probabilities are rather uncertain. Therefore, it becomes important to assess the dependence of information functions on fluctuations of probabilities.

The contingent nature of information measures has not always been obvious. All early applications of information theory dealt with telecommunication systems. In all of these, all informational characteristics are perfectly well defined. In Morse code, all we have to know is whether a particular information-carrying element is a blackness or a whiteness, and whether it is long or short. In pulse code modulation, the only thing that counts is presence or absence of a pulse within a stated interval of time. In pulse amplitude modulation, all information is vested into the amplitude of pulses. In all these cases, there is no question about the informational characteristics of the process under consideration.

The situation is radically different in the larger domain of applied information theory. For instance, take the case of two people transmitting information to each other by talking. The information-carrying element is a clause; to simplify our analysis, let us consider just words (remembering that the information content of a clause cannot be greater than that of its constituent words). Now, each person culls his words from a reservoir which is known to be large, but its actual size is not exactly known. The information content of a single word depends on the probability of its use, and these probabilities are not exactly known either. Furthermore, they will hardly be the same for both persons involved in a conversation. Also, each word can have several meanings, one of which may be more or less determined by the context. The relations between words, meanings, and context, again, are not the same for any two people. This is not all. Information is conveyed not only by the choice of words but also by inflection of voice, loudness, timing, and accompanying gestures. In such a situation we have obviously no hope ever to obtain a precise, unequivocal, and incontestable measure of information content. We are, thus, confronted with two alternatives. These are: not to use information theory, or to try to devise ways of producing usable approximate estimates. Obviously, our choice is the latter alternative (19).

Approximation Methods

It appears that the approximation methods to estimate information functions are based on the following rules:

1. Averaging increases uncertainty;
2. Pooling decreases uncertainty;
3. Disregarding constraints increases uncertainty;
4. Rare events have small effects on uncertainty measures;
5. Small variations in probability have small effects on uncertainty measures;
6. In systems, information functions can be estimated in different ways, and care should be taken to select the most appropriate one;

7. If it is not possible to measure the actual information functions desired, then one can try to substitute closely related measurable quantities.

In the following paragraphs, these rules will be amplified and illustrated.

1. *Averaging Increases Uncertainty*—The fact was demonstrated in Section III. It suggests a simple bracketing procedure: obtain a lower and upper bound of uncertainty by using probabilities which are certainly more and less unbalanced than they actually are. In particular, if the number of categories is known but their respective probabilities are not, then one can follow Laplace's procedure and set all probabilities equal which maximizes uncertainty.

2. *Pooling Decreases Uncertainty*—This, too, has been proven in the third section. It is equally of value in bracketing procedures: using only categories actually discriminated puts a lower bound on uncertainty; assuming more categories than could be of interest establishes an upper bound.

3. *Disregarding Constraints Increases Uncertainty*—Let x and y be different events, where y may differ from x only in time or place of occurrence or in any other respects. If $H(x)$ is the uncertainty of x, and $H_y(x)$ the uncertainty of x if y is known, then:

$$H_y(x) \leq H(x).$$

That is, knowing some other event, y, cannot increase the average uncertainty concerning x; it will leave it unchanged if there is no association between x and y; it will reduce it if constraints exist which are manifested in a statistical association between x and y.

Rule 3 can be used for a bracketing procedure. Disregarding constraints yields an overestimate of $H(x)$; introducing constraints known to be too strong, an underestimate.

Constraints have to be very marked to cause large changes in $H(x)$. For instance, the large inequalities of letter frequency in English texts reduce H from a possible maximum of 4.7 bits per letter to 4.1 bits; the strong constraints between successive letters and words result in an additional reduction to 1.5–2.0 bits per letter.

Formally, rule 3 is a special case of rule 1.

4. *Small Effects of Rare Events*—The information functional is a sum of terms of the form $(-p \log p)$. This function rises steeply between zero and .10, hence, small probabilities contribute little to the total sum. For instance, ten equiprobable alternatives correspond to an H of 3.32. If one of these alternatives is replaced by ten separate sub-categories, each of probability .01, then the resulting H is 3.65. If instead of ten, one introduces 100 equiprobable sub categories, each with probability .001, the resulting H is 3.99, or equivalent to sixteen equiprobable categories.

A good example turned up in a study by A. A. Blank. He calculated the information content of single English words. For particular reasons, the sample was restricted to four letter words. Thorndyke's list contains 1550 such words. H, based on the observed frequency of these words, is 8.13 bits per word. Of these words, 119 occur with the greatest frequencies. Computing H on the basis of these words alone gives a value of 6.34 bits per word. Thus, taking into consideration only about one tenth of all categories already yields about four-fifths of the final information function.

This means that information functions can be estimated successfully as soon as the more common occurrences are categorized. The remaining infrequent occurrences will not contribute very much, and that contribution can be easily bracketed between values based on numbers of categories which are certainly too small and too large.

5. *Small Effects of Small Variations in Probability*—The curve of the function $F(p) = -p \log p$ has a flat top. Small changes in probability in this region have small effects.

Consider the simplest case, of two categories. If their probabilities are equal, then $H = 1$. If the ratio of the probabilities is 1:2, then $H = .92$. If the ratio is 1:3, a very considerable deviation from equality, H is still .81.

For a larger number of categories, the insensitivity of H against probability distortion is still more pronounced. If one replaces equiprobable alternatives by probabilities staggered arithmetically or geometrically, stipulating only that the span between the extreme value should be not more than one order of magnitude, then the resulting changes in H are quite small.

This implies that the assumption of equiprobability, which gives an upper bound as stated in rule 1, will not go very far from the true value unless probabilities are radically unbalanced. The stretch bracketed between an upper bound based on equiprobability, and a lower bound based on a distortion undoubtedly stronger than the real one, will not be very large.

6. *Alternative Ways of Estimating Information Functions*—In systems with several nodes, the compound information functions can always be estimated in several ways. For instance, in a two-node communication system, the quantity which is the function of greatest interest, the amount of information transmitted, $T(x; y)$, can be computed in three alternative ways: as differences between input uncertainty and equivocation, as difference between output uncertainty and ambiguity, or as difference between the sum of uncertainties of input and output and the uncertainty of their union. It usually is worthwhile to inspect the data very carefully to establish which of the set of functions can be most easily and most accurately computed. In many cases, the quantities most readily computed are not those which result directly from the plan of observation or experimentation. For instance, in most experiments it would be natural to measure output uncertainty and ambiguity, but it is easier to measure input uncertainty and equivocation.

7. *Substitution of Related Quantities*—In many cases where it is not practical to compute the proper information measures, one can compute information measures associated with related quantities. Take the case of estimating the amount of information which an individual can transmit after a single glance at a display. This quantity is very difficult to determine; but, it is fairly easy to determine the amount of information which can be elicited from an individual by a short interrogation procedure after he has had a glance at the display. This function is not quite the one we want, but presumably closely related to it. Another example: in the case of mental arithmetic, we have no way of estimating the actual amount of information processed, but we can readily estimate the amount of information which must be processed if computations are done in the way in which the subject claims he computes. In cases of this kind one will use the measurable quantity instead of the desired one. Of

course, results so obtained have to be used with a certain amount of restraint.

Example: Rate of Information Transmission in Conversation—The working of the approximation methods can be shown by two examples. The first example is that which we used to illustrate the need for approximation methods; namely, that of estimating the amount of information in conversation.

We consider first the information carried in words. To establish an upper bound, we ask how much information must be transmitted so that the receiver can recognize every single word spoken.

This upper bound, in bits per second, is the product of the rate of words per second times bits per word. A rate of 2.1 words per second is typical for lively discussions. The number of bits per word in English context has been estimated as 6.5 bits ($\pm$25 per cent). This yields 11 to 17 bits per second.

Words are not the only method of communication between two persons conversing face to face. It can be shown, however, that all other means of communication add little to the total transmission rate.

We will now try to establish a lower bound. Of course, no general lower bound exists; it is easy to find examples where information is transmitted at the rate of 1 millibit per second, or less. What we want is an 'upper lower bound' a lower bound of the amount of information transmitted between people who try to communicate at some speed, and under reasonably favorable conditions. Such a bound is obtained by analysis of pragmatic communication. We look at situations where the verbal messages elicit or control actions. We make an informational analysis of the relations between actions and verbal messages. This will yield an amount of information demonstrably transmitted, and it certainly represents a lower bound to the amount of information communicated.

At this time, we have a single case where pragmatic communication has been evaluated accurately in informational terms. FELTON, FRITZ and GRIER (20) measured the amount of pragmatic communication between an airplane pilot coming in for a landing and the control tower operator. They found an average rate of 2 bits per second, computed in terms of actual effects of the messages. Both pilot and control tower operator have all interest to communicate as fast as they can. On the other hand, they do so in the presence of a very high level of noise which reduces verbal communication to probably about one third of its optimum rate.

We conclude, thus, that information transmitted through verbal communication is certainly not less than 2 bits per second nor more than 17 bits per second, and very likely within the range between 6 and 12 bits per second. This estimate is rough but not at all unrealistic.

Example: Information Content per Printed Letter—A very elegant way of computing an information measure under unfavorable conditions was used by SHANNON in his analysis of the 'entropy' of printed English (21). The information content of a single letter is easily determined as a function of relative letter frequencies. However, constraints between neighboring letters lead to a reduction of information content, and in order to estimate this reduction exactly one would have to investigate the probability distributions for long sequences of letters. This is manifestly impossible. SHANNON, therefore, proceeded to estimate a related quantity; namely, the amount of information

concerning language constraints which can be elicited from a person familiar with printed English by a carefully planned interrogation. The subject is given a text which is truncated at some point; he is asked to guess the next letter. If he is successful, then he is told to go on; if not, he is told to try again. Records are taken of the number of times a letter is correctly identified at the first, second, third, . . . statement. In this setup, the experimenter acts as source of auxiliary information, emitting sequences of the type 'wrong . . . wrong right', with an 'alphabet' of twenty-six different sequences (if repetitions are excluded, the letter must be identified after no more than twenty-five wrong guesses). The informational output of the auxiliary source depends on the relative probabilities of the various sequences. These probabilities are very unequally distributed. In a large percentage of the cases, the first statement is correct; the most frequent message from the auxiliary source is 'right'. The next highest probability is for the sequence 'wrong-right'. Messages with up to three 'wrongs' make up the vast majority of cases; the remaining categories, with from 4 to 25 'wrongs', have low probabilities. As was pointed out before, they contribute little to the estimated value of H. This means that we arrive at an estimate of the information furnished by the auxiliary source essentially as a function of two to four probabilities.

The amount of information per single letter is known to be about 4.1 bits (on the basis of relative frequency of letters in English texts). This is the amount of information per letter which the subject needs to reconstruct the whole text. Of this amount of information, a certain measurable fraction is furnished by the auxiliary source. The remainder must come out of the subject's head, and is based on his knowledge of language constraints. The amount of information so elicited will not be quite as high as the information content of language constraints, but it is a closely related quantity. By the ingenious trick of effectively reducing the size of the alphabet, this quantity has been made easily measurable.

APPENDIX II
ANSWERS TO EXERCISES

1. One light—peace and quiet
   two lights, vertically—enemy approaches by land
   two lights, horizontally—enemy approaches by sea
   two lights, diagonally—enemy approaches by land and sea
   (This is not the only possible solution)
2. (a) 0, 1, 10,11,100,101,110,111, 10000, 1001,1010,1100,10000, 11110100011
   (b) 9, 11, 147, 32
   (c) .125, .6703125
3. (a) 10110100010000
   (b) EDCBA

4. 'Construct a confusion-free code using five binary digits for each letter and compare the performance of this code with that of the above by encoding and decoding a message like this one'.

Use part of the 32 code words made up of 5 binary digits, such as: 11111, 11110, 11101, 11100, etc. The message will be, on average, 21 per cent longer than with the most efficient code (5 is 121 per cent of 4.14), but it is much easier to decode. Some of the unused code words can be used for punctuation, etc. The teletype works on this principle.

5. Limiting value:

$$-(.8 \log_2 .8 + .15 \log_2 .15 + .05 \log_2 .05) = .883$$

Single event code:

A	1	.8
B	0 1	.3
C	0 0	.1

$$\overline{1.20}, \text{ excess is } \frac{1.20 - 0.883}{0.883} = 36 \text{ per cent.}$$

Two-event code:

Event pair	Prob.	Code	
AA	.64	1	.64
AB	.12	0 1 1	.72
BA	.12	0 1 0	
AC	.04	0 0 1 1	.32
CA	.04	0 0 1 0	
BB	.0225	0 0 0 1	.09
BC	.0075	0 0 0 0 1	.0375
CB	.0075	0 0 0 0 0 1	.06
CC	.0025	0 0 0 0 0 0	
	1.0000		1.8675

.934 digits per event

excess = 5½% < 10%

6. Let x designate amino acids, and y nucleotides.

$$n_x = \log_2 20 = 4.322$$

$$s_x = 1$$

$$n_y = \log_2 4 = 2.0$$

$$s_y = \frac{1 \times 4.322}{2.0} = 2.161$$

7.

p	$-p \log_2 p$
.60	.44
.40	.53

$H(x) = .97$

8. The curve looks similar to $F(p)$, but has a flatter top and is symmetrical, with a maximum of 1.0 at $p(1) = .50$.

9. $H(x) = \log_2 3 = 1.58$

	p	$-p \log_2 p$
y:	.8	.26
	.1	.33
	.05	.22
	.05	.22

$H(y) = 1.03$

$H(x) \quad H(y)$

10. A realistic description of his uncertainty might be:

prob (55–64)	.95
prob (55–54)	.02
prob (65–70)	.02
prob (any other speed) =	.01

Within each range, all speeds are considered equiprobable.

We will derive the answer in two steps, obtaining first the uncertainty as to the speed range:

Range	p	$-p \log_2 p$
55–64	.95	.07
50–54	.02	.11
65–70	.02	.11
any other speed	.01	.07
		.36 bits

Next, we observe that the range from 55 to 64 miles per hour contains ten speeds (determined to the nearest mile) which are equiprobable. The uncertainty measure for ten equiprobable categories has been found to be $\log_2 10 = 3.32$. This uncertainty will arise 95 times out of 100; its expected contribution to the total uncertainty is $3.32 \cdot 0.95 = 3.15$. The other ranges are treated equally:

Range	No. of sub-classes (r)	$\log_2 r$	$p \cdot \log_2 r$
55–64	10	3.32	3.15
50–54	5	2.32	.05
65–70	5	2.32	.06
all other	81	6.35	.06
			3.31 bits

We thus need (on average) .36 bits to determine the range of speeds, and an additional 3.31 bits (on average) to identify the speed to the nearest mile, within the range. The total uncertainty is $0.36 + 3.31 = 3.67$ bits.

Of course, different expectations would yield different uncertainties.

11. The letters occur with more nearly equal frequencies.

12. Two bits.

13.

$$H(\text{shape}) = -\left(\frac{323}{427}\log_2\frac{323}{427} + \frac{104}{427}\log_2\frac{104}{427}\right) = .80 \text{ bits}$$

$$H(\text{color}) = -\left(\frac{315}{427}\log_2\frac{315}{427} + \frac{112}{427}\log_2\frac{112}{427}\right) = .83 \text{ bits}$$

$$H(\text{color, shape}) = -\left(\frac{296}{427}\log_2\frac{296}{427} + \frac{27}{427}\log_2\frac{27}{427} + \frac{19}{427}\log_2\frac{19}{427} + \frac{85}{427}\log_2\frac{85}{427}\right) = 1.26 \text{ bits}$$

$$T(\text{color; shape}) = .80 + .83 - 1.26 = .39 \text{ bits}$$

14. $T(x, y; z)$ = mutual reduction of uncertainty between x and y on one hand, z on the other

$$= H(x, y) + H(z) - H(x, y, z)$$

$$T(x; y, z) = H(x) + H(y, z) = H(x, y, z)$$

$T(x; y; z)$ = total constraint in a tri-variate system

$$= H(x) + H(y) + H(z) - H(x, y, z)$$

15.

		Actual: pos	Actual: neg	
Test	pos	3	7	10
	neg	5	85	90
		8	92	

$H(x) = .47$

$H(y) = .40$

$H(x, y) = .84$

$T(x; y) = .03$

The informational value of the test is .03 bits.

Its maximum possible informational value equals the amount of uncertainty before the test, viz. .40 bits.

16. $2.3 \times 5 \times 60 = 690$ bits/minute

17. Begin by computing the output uncertainty. The probabilities of receiving each signal are obtained as the sum of receiving it correctly (0.2 for Nos. 1 and 4, .178 for 2 and 3, .198 for 5) plus the addition due to errors (1/4 of the errors, for each erroneous transmission). This procedure yields $H(\text{out}) = 2.32$ bits. Next, compute the ambiguities. These are zero for symbols no. 1 and 4. For 2 and 3, the ambiguity can be computed as the sum of the information needed to ascertain that an error has occurred ($-0.11 \log_2 0.11 - 0.89 \log_2 0.89$) plus the information needed to find out which of the possible and equiprobable four errors has occurred, which is 0.11×2.0 bits/symbol. Symbol no. 5 is treated similarly. The average of the ambiguities is 0.31 bits, hence T equals $2.32 - 0.31$ or 2.01 bits—a loss of about one-sixth of the input information.

18. One solution is the following:

11000
10101
01110
00011

A single error will result in the reception of a word which is not in the code book. If one follows the rule of substituting that message in the code book which differs from the received one by one digit only, then every error (provided there is only one!) will be corrected.

A five-digit binary message can carry five bits of information. If it is known that one error has occurred somewhere in a group of five symbols, then the information needed to locate the error is $\log_2 5 = 2.33$ bits. With maximum efficiency, one should use only 2.33/5 or 46.5 per cent of redundant information (which could be achieved by coding large sequences of five-digit words!). In our case, the redundant information is 3/5 or 60 per cent, and we transmit with an efficiency of $40/53.5 = 75$ per cent. (Observe that there is less uncertainty if it is known that there is one error in every five-symbol word, than when it is only known that the error rate is 20 per cent!)

REFERENCES

1. L. Szilard: Über die Entropieverminderung einem thermodynamischen System bei Eingriffen intelligenter Wesen. *Z. Phys.* **53**, 840–856 (1929).
2. R. A. Fisher: On the mathematical foundations of theoretical statistics. *Phil. Trans.* (*A*) **222**, 309–368 (1922).
3. R. V. L. Hartley: Transmission of information. *Bell Syst. Tech. J.* **7**, 535–563 (1928).
4. N. Wiener: *Cybernetics*, J. Wiley and Sons, New York (1948).
5. C. E. Shannon: A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423, 623–656 (1948).
6. C. E. Shannon and W. Weaver: *The Mathematical Theory of Communication*, University of Illinois Press, Urbana (1949).
7. L. N. Ridenour: Computer memories. *Sci. Amer.* **192**, 92–100 (1955).
8. R. M. Fano: The transmission of information. *Tech. Rep. Mass. Inst. Tech. Res. Lab. Electron.*, no. 65 (1949)
9. L. Brillouin: *Science and Information Theory*, Academic Press, New York (1956).
10. L. Dolansky and M. Dolansky: *Tables of* $\log_2 1/p$, *etc.*, *Tech. Rep. Mass. Inst. Tech. Res. Lab. Electron.*, no. 227 (1952).
11. E. Klemmer: Tables for computing informational measures. Tech. Rep. A. F. Cambridge Research Center, ARDC.
12. Articles by J. E. Golay, P. Elias, I. S. Reed, R. A. Silverman, and M. Balser in: *Transactions of the I.R.E. Professional Group on Information Theory* (1954).
13. A. M. Turing: On computable numbers, with an application to the Entscheidungsproblem. *Proc. Lond. Math. Soc.* **42**, 230–265 (1937).
14. H. Quastler: Studies of human channel capacity. In: *Information Theory*, ed. by C. Cherry, Academic Press, New York (1956).
15. Wm. McGill and H. Quastler: Standardized nomenclature: an attempt. In: *Information Theory in Psychology*, ed. by H. Quastler, Free Press, Glencoe, Ill. (1955).
16. H. Quastler: Information theory terms and their psychological correlates. *ibid.*
17. J. von Neumann and O. Morgenstern: *Theory of Games and Economic Behavior*, Princeton University Press, Princeton (1947).
18. H. Quastler, H. H. Chase, W. Montagna, M. V. Edds, Jr., P. F. Fenton, and P. B. Weisz: Essays on biological unitization. Rep. Control Systems Laboratory, Univ. Ill., No. R-52 (1953).
19. A. A. Blank and H. Quastler: Notes on the estimation of information measures. Rep. Control Systems Laboratory, Univ. Ill., no. R-56 (1954).
20. F. Fritz and G. W. Grier, Jr.: Pragmatic communication. In: *Information Theory in Psychology*, ed. by H. Quastler, 232–243, Free Press, Glencoe, Ill. (1955).
21. C. E. Shannon: Prediction and entropy of printed English. *Bell Syst. Tech. J.* **30**, 50–64 (1951).

SOME INTRODUCTORY IDEAS CONCERNING THE APPLICATION OF INFORMATION THEORY IN BIOLOGY

HUBERT P. YOCKEY
Oak Ridge National Laboratory, Oak Ridge, Tennessee

Abstract—The model of protein synthesis in the cell which has been built up as the result of the work of many researchers has been used as a basis for applying the principles of information theory in biology. The main line of the argument has been the role of noise in the genome. The discussion has been kept as independent as possible of special models.

It was shown that in a real organism noise must exist in the genome and that an ensemble of organisms may be represented by a probability distribution in H, $\rho(H, \lambda)$. Individuality is thus incorporated in a very natural way. Dancoff's principle requires that there be a lower limit for viability for this distribution, H_d.

The action of a deleterious agent which induces errors in the genome by acting on nucleotide pairs is assumed to be represented by an equation of the first order:

$$\frac{dp_i(j)}{d\lambda} = -J(\lambda)\, p_i(j) + \tfrac{1}{4}J(\lambda)$$

where $J(\lambda)$ measures the effectiveness of the deleterious agent, of which λ is a measure, in producing defects. A differential equation for $H(\lambda)$ is derived and it is shown that $(dH/d\lambda)_{H_d}$ as a function of λ behaves like $J(\lambda)$.

I. INTRODUCTION

INFORMATION theory finds its place in biological thought through its ability to deal quantitatively with organization and specificity. The importance of these concepts has long been recognized in biology, but this realization is rather sterile unless a quantitative form of expression can be found. One is reminded of a quotation from Lord KELVIN, 'When you can measure what you are speaking about and express it in numbers, you know something about it, but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind.'

The need for expressing biological quantities in numbers is clear but solving the problem of how to do it is very much like belling the cat. Biology doesn't seem to have any problems *both really simple* and *terribly important* such as some which occur in the physical sciences. The application of first principles has come much more slowly in biology for perhaps this reason. That ideas of great general application do exist in biology is exemplified by Mendel's laws and by the theory of evolution.

One of the purposes of this article, and indeed one of the purposes of this book, is to explore the practical and theoretical consequences that may be found in the discovery that biochemical specificity of proteins is carried, largely at least, by the exact order of twenty amino-acid residues. The suggestion of

WATSON and CRICK (1) that genetical information is carried by the exact order of four kinds of nucleotide pairs provides a molecular vehicle for the genetic control of protein specificity. GAMOW (2) was the first to see that this control implied the existence of a four-letter to twenty-letter code. Thus by following the logical consequences of purely biological, or perhaps biochemical, problems one is lead directly to a problem purely mathematical in character.

This notion of the role of order, which is basic to information theory, is worth pursuing in biology since it provides a way of measuring what we are speaking about and expressing it in numbers. Furthermore, from the results of applying the theory to specific problems, we may obtain an experimental check on the validity of these ideas as first principles. In this article we shall apply these considerations to the storage and transfer of biochemical specificity. We shall explore, in particular, the role of noise in the genetical message. In my article in Part V the theory is applied to the practical problem of calculating and understanding survivorship curves.

The present status of the means of storage and transfer of specificity is given by GAMOW, by YČAS and by AUGENSTINE in their respective articles in this volume. The question of the exact way in which information is destroyed by read-off error, radiation damage, aging, thermal fluctuations, biochemical side reactions, and so forth, is of equal importance. This problem is also discussed in this volume but no final and detailed account can be given at this writing. Nevertheless, since there is virtue in attempt, we shall attempt the development of a mathematical formalism which is information theoretic in character.

Most animals and plants exist at one time, at least, in the form of a single cell; we can consider that cell to contain a substantial part of the directions for the development of the organism. Since information is conserved unless lost due to noise, it shall be assumed that the mature organism is characterized by substantially the same information content as the fertilized egg or seed. In order to fix the idea we shall develop the formalism on the basis of WATSON and CRICK's suggestion concerning the role of DNA. It should be remembered that the central ideas of this paper are independent of much of the detail embodied in WATSON and CRICK's papers and are dependent only on the possibility of genetical endowment being conveyed by a series of structures composing an information bearing molecule.

Suppose we imagine the symbols A, B, C, D (GAMOW's predilection is to the less prosaic spades, clubs, hearts, diamonds!) arranged in one-to-one correspondence with the nucleotide pairs of the DNA found in a particular given cell. The cell will have been selected from a number of similar but not identical cells in a colony under study. This colony may be thought of as being indefinitely large, so that in principle we may consider the ensemble of all possible organisms identifiable as being members of the colony. Since the number of nucleotides in DNA is finite, the number of elements in this ensemble is also finite. Because of this one-to-one correspondence it will be seen that the set of symbol sequences, which is the mathematical model of the ensemble of organisms, will contain the informational or specificity properties of the ensemble of organisms.

The importance or value of a theory lies, among other things, in its capability of treating a wide variety of phenomena from a single point of view. It is well to think, at the start, of the field of validity this theory may have and, if it should fail, the significance of its failure. If it should be discovered that WATSON and CRICK's suggestion has very little bearing or applicability then this development, while negative, is still a valuable result. One would then perforce search for another explanation for the great detail and specificity characteristic of any biological phenomenon. At present it is the most detailed proposal based specifically on molecular chemistry. The theory here developed is essentially statistical and may be expected to express its results in the form of expectation values, probability distributions, and their functions. The statistical character of the theory is directly in the line of thinking of both modern biology and modern physics. It should be kept clearly in mind that information theory deals with organizational problems and so some aspects of organisms will be outside its scope. In this sense it may be that the role information theory will play in biology will parallel that played by thermodynamics in physics and chemistry.

II. NOISE IN THE GENETICAL INFORMATION

The Instability of a Perfect System

Let us consider an ensemble of organisms and discuss the communication of information from the DNA to protein. There is evidence discussed by GAMOW and by YČAS in this volume that the code which translates information from the four-symbol DNA code via RNA to the twenty-symbol protein code is based on triads of nucleotide pairs. Indeed it can be seen that it must be at least the triads since a twenty-symbol alphabet carries 4.32 bits per symbol whereas the pairs in a four-symbol alphabet carry exactly four bits per symbol, assuming no intersymbol constraints. The triads carry six bits per symbol and so this represents some inherent redundance. It would be desirable to express this formalism in terms of the DNA triads of nucleotide pairs; however, this requires a knowledge of the DNA to protein code. These data are missing. Our objective is to develop the mathematical formalism in as simple a way as possible so it appears more appropriate to consider the communication of specificity from DNA to RNA. Here we are dealing with a coding between two four-symbol alphabets.

Suppose we are considering an ensemble of organisms which is isogenic, and further that this means that each organism is characterized by exactly the same order of nucleotides in the DNA of its nucleus. We shall now show that this situation is unstable and that therefore a real ensemble of organisms will be represented by an ensemble of messages recorded in its DNA. From this it will follow that there is a distribution in the message entropy, characteristic of any ensemble of organisms, even one which is isogenic.

The message entropy is

$$H = H_o - H_n \tag{1}$$

where H_o is the message entropy of the genetical information and H_n is the loss of information due to noise. That is, H_n is the loss of information from

some fault either in the duplication process in the germ line or the somatic line or from incorrect read-off of any kind. H_n may be expressed in terms of the read-off or transition probabilities (3) of a letter of kind i to a letter of kind j, $p_i(j)$. The probability of letter i is $p(i)$.

$$H = H_o + \sum_{i,j} p(i)\, p_i(j) \log_2 p_i(j) \qquad (2)$$

Consider the case where these probabilities are a function of some variable λ. In the application of these considerations λ is the measure of some deleterious influence such as dose of ionizing radiation. Form the derivative $dH/d\lambda$:

$$dH/d\lambda = \log_2 e \sum_{i,j} \{p(i)\,(d/d\lambda)\, p_i(j) + p(i) \log_e p_i(j)\,(d/d\lambda)\, p_i(j) + p_i(j) \log_e p_i(j)\,(d/d\lambda)\, p(i)\} \qquad (3)$$

The absolute value of $dH/d\lambda$ will become indefinitely large because of the second term in equation (3) as any $p_i(j)$ approaches zero if $p(i) \neq 0$ and $(d/d\lambda)\, p_i(j) \neq 0$. This may happen, in particular, if any $p_i(j)$ approaches one for then all $p_i(k)$, $(j \neq k)$ approach zero. This situation $(p_i(j) = 1)$ corresponds to the assumption that there is always a correct reproduction in the DNA duplication or in the RNA read-off. Under these circumstances the first term is finite and the third term is zero.

WATSON and CRICK regard a mutation as being reflected by a change in order of the nucleotide bases in DNA. This is apparently always possible; they have suggested a biochemical scheme by which this can be affected. This means that in a real biological system $p(i) \neq 0$ and $(d/d\lambda)\, p_i(j) \neq 0$. A real ensemble of organisms will be represented by an ensemble of genetic messages. This will be true even if the ensemble is isogenic. Some noise must exist in the genetical information; if the noise is less than equilibrium it is quickly introduced.

There is some experimental evidence in support of this conclusion. BURDETTE (4) prepared populations of isogenic *Drosophila*. One strain had the same low incidence of tumors in both sexes (about 4 per cent) and the other had a high incidence (about 60 to 80 per cent) even greater in males than in females. The tumor incidence of the isogenic strains was initially much lower in each case than the stock from which it originated. But in each case, by the twelfth generation, the tumor incidence of the isogenic strain had returned to about the same rate as that of the original stock. Tumor incidence is a morphological malfunction and, as shown in this and other experiments, is under genetic control.

The fact that all flies were not tumor bearing and the gradual return of the isogenic strains to the tumor incidence of the strains from which they were selected, reflects the accumulation of errors in the genome. The results of the experiment are in accord with the proposition proved above.

Representation of the Ensemble of Organisms by a Probability Distribution in H: $\rho(H, \lambda)$

If we grant that perfect systems do not exist, the other side of the coin is, how imperfect may they be? This question was first discussed by DANCOFF and QUASTLER (5) and their conclusion, which is known as Dancoff's principle, states that the amount of redundance is just that required to reduce the error

rate to a tolerable level. According to this principle, we may expect that errors will continue to accumulate in the genome of a given organism until at some point serious difficulty including death will occur. This will be reflected by some value of H, which we call H_d, limited by viability. An argument for a lower limit H_d has been given previously (6).

Errors will accumulate in the genome but at the same time there is a favorable selection for those members of the ensemble which have low equivocation. This represents a certain reserve capacity to withstand the insults of existence. It may therefore be expected in general that the message entropy of the ensemble of organisms will be described by a probability distribution. This distribution can, perhaps, be calculated from first principles, at least for simple cases, when more is known about the storage and transfer of genetical information.

Death of an organism is defined in different ways in various fields of biology. Permanent loss of reproductive power is the definition of death usually expressed or implied in bacteriology (7). This is the definition chosen in spite of the fact that there are many intermediate stages between the active living cell and the dead cell. It is known that yeast cells which have lost the power to multiply may still be able to ferment (8). ZELLE and HOLLAENDER (7) have recently pointed out that attempts to explain the bactericidal effects of irradiation on the basis of one mechanism are unrealistic. In the case of animals the cessation of metabolism, not the loss of fertility, is the criterion of death. These criteria of death are not really different or antagonistic. Since loss of function is implied by loss of information content any experimentally convenient definition of lethality may be used to suit the problem at hand. The lower end of the distribution in message entropy will therefore be determined by the specificity required by the environment.

A communications analogy may clarify the notion further. Suppose we have a message, with redundance, which is sent through a communication channel with a small but finite noise level. The message contains instructions to perform some necessary task. A recording is made and the message is sent through again, and so forth. Eventually, depending on the noise level of the channel and the redundance in the message, it will be just barely intelligible. No further recordings can be made without loss of part of the required information content. The ensemble of recordings is analogous to the ensemble of organisms. It will be seen in either case that there is a distribution of information content among the elements of the ensemble.

Individuality finds a place in the theory developed here in a very natural way. This feature corresponds more to reality (9) than theories which must explain non-uniform response as fluctuations. Besides the experiments of Burdette mentioned above it will suffice to note one other example of biological individuality.

Consider the experiments of SCHOTT (10, 11), HETZER (12), LAMBERT (13), GOWEN (14), discussed by GOWEN (15), on *Salmonella typhimurium* in mice and *Salmonella gallinarum* in fowl. The host population is exposed to the pathogen and the survivors are chosen for further breeding. The case for mice is typical. The survival ratio improved from 18 per cent to 93 per cent in six generations, but remained nearly constant after that. One hundred per cent survival was not achieved. The survival ratio is characteristic of the ensemble not of the

individual. GOWEN (15) also prepared six strains of mice by sibling matings for twenty or more generations. When survival was tested the survival ratios were 1, 14, 34, 63, 64, 83 and 88 per cent. These results again stress the importance of individuality as Gowen pointed out.*

Point Mutations and Chromosome Aberrations

We have now arrived, via our discussion, at territory familiar to the radiation biologist. This is the controversy over the role played by point mutations and chromosomal aberrations induced by deleterious agents such as x-rays. This subject has been ably discussed recently by MULLER, KAUFMANN, GILES, CARLSON, SWANSON and STADLER, and by KIMBALL (16). The point of view of these authors varies. Kimball takes the stand with LEA (17) that the death of cells is due to chromosome aberrations which become effective at cell division. SWANSON and STADLER point out that the two effects occur together and that a clear cut separation has not yet been accomplished. MULLER points out some difficulties with the mutation by breakage interpretation. RUSSELL (18) states that gross chromosomal aberrations, although they cause early death of embryos, are probably not an important radiation hazard to man.

From the point of view of this article each of these effects is a way of introducing disorganization in the genome. The point mutation mechanism is the biological analogue of the 'white noise' of the communications engineer. The other extreme is not found in communication engineering but involves a strong correlation between errors and is reflected as a loss of whole paragraphs or other gross mutilation of the message. Each of these extreme cases will be important in applications of information theory in biology. Unfortunately, the second case has not been studied mathematically and so it is not known how to calculate the equivocation it introduces.

It is therefore necessary to proceed with the calculation of only the part of the equivocation which corresponds to point mutations. Since one of our objectives is to develop a fundamental theoretical treatment of radiation hazard to man, RUSSELL's comment encourages one to think that this procedure is worthwhile. It should be remembered that equivocation from these two extreme conditions may have the same dependence on the deleterious influence. This is a point which requires further mathematical study.

The Interaction of the Deleterious Agent with DNA and the Decay of H

According to the Watson and Crick model of DNA there seems to be no biochemical reason why there should be an interaction between nucleotide pairs. The biological requirements for protein specificity do not seem to demand an intersymbol influence (19). The matter is not closed, but the evidence favors regarding the interaction of a deleterious agent with a nucleotide pair to be of the first order.

We have previously suggested that the action of ionizing radiation or other deleterious agent may be such that the nucleotide pair is altered in such a way that it mimes another symbol as far as protein synthesis is concerned (6). It

* Individuality as an integral feature in biology has been emphasized recently by ROGER J. WILLIAMS: in *Biochemical Individuality*, J. Wiley and Sons, New York, Chapman & Hall, London (1956).

may be thrown into an excited tautomeric form from which it recovers by relaxation. Possibly one can account for biological recovery by such a mechanism. The consideration of recovery is omitted from this paper for simplicity and we shall need only the notion expressed in the first sentence of this paragraph.

In view of the above remarks we may write the following equation for the rate of change of $p_i(j)$ with λ:

$$(d/d\lambda)\, p_i(j) = -J_{ij}(\lambda)\, p_i(j) + c_{ij}(\lambda) \tag{4}$$

The first term represents the loss in nucleotides responsible for the (i, j) transition. The second term is due to the gain in nucleotides engaging in the (i, j) transition coming from other nucleotides altered by the deleterious agent. This can be brought into sharper focus by thinking of the binary case. Suppose q is the correct and p is the incorrect read-off probability. We are calculating the equivocation, or damage to the message, resulting from point errors. This means that, accordingly, a letter is not deleted but is read off either correctly or incorrectly. This letter switching process may continue until half the letters are correct and half are incorrect; at that point $p = 1/2$ and $q = 1/2$. The information content vanishes. In the case of a four letter alphabet a letter which is acted upon and which may therefore change or may retain its original read-off character has an *a priori* probability of 1/4 to remain or to become a correct letter. Thus the second term is required by the normalization condition.

Equation (4) describes the effect of the interaction of the deleterious agent, say the x-ray dose, with the information bearing molecules in the cell. It corresponds to current views of reaction kinetics. Should it be discovered that some effect, for example, inter-symbol influence, should be taken into account then equation (4) may be altered suitably. The following argument would then still be cogent except that the new form of equation (4) would be used. Present experimental evidence substantiates equation (4) and we have no present justification for greater complication. In fact the $J_{ij}(\lambda)$ and $c_{ij}(\lambda)$ represent more detail than is available. Sum equation (4) over all j:

$$\sum (d/d\lambda)\, p_i(j) = -\sum_j J_{ij}(\lambda)\, p_i(j) + \sum_j c_{ij}(\lambda) \tag{5}$$

Since

$$\sum_j p_i(j) = 1; \qquad \sum_j (d/d\lambda)\, p_i(j) = 0 \tag{6}$$

$$0 = -\sum_j J_{ij}(\lambda)\, p_i(j) + \sum_j c_{ij}(\lambda) \tag{7}$$

If the $J_{ij}(\lambda)$ and the $c_{ij}(\lambda)$ may be replaced by an average value $J(\lambda)$ and $c(\lambda)$, equation (7) becomes, for a four-letter alphabet:

$$0 = -J(\lambda) + 4\, c(\lambda) \tag{8}$$

$$c(\lambda) = +\tfrac{1}{4}J(\lambda) \tag{9}$$

Equation (4) may be written as follows:

$$(d/d\lambda)\, p_i(j) = -J(\lambda)\, p_i(j) + \tfrac{1}{4}J(\lambda) \tag{10}$$

Given $(d/d\lambda)\, p_i(j)$ as some function of λ, equation (3) may be regarded as a differential equation for $H(\lambda)$. This equation has a simple form if the $J_{ij}(\lambda)$ and the $c_{ij}(\lambda)$ may be replaced by their averages $J(\lambda)$ and $\frac{1}{4}J(\lambda)$.

$$(dH/d\lambda) = \log_2 e \sum_{i,j} \{p(i)\, J(\lambda)[-p_i(j) + \tfrac{1}{4}] + p(i)\, J(\lambda)\, [-p_i(j) + \tfrac{1}{4}] \log_e p_i(j) + p_i(j) \log_e p_i(j)\, (d/d\lambda)\, p(i)\} \quad (11)$$

$$\begin{aligned}(dH/d\lambda) = {} & -J(\lambda) \log_2 e \sum_{i,j} p(i)\, p_i(j) \log_e p_i(j) \\ & + \tfrac{1}{4} J(\lambda) \log_2 e \sum_{i,j} p(i) \log_e p_i(j) \\ & + \log_2 e \sum_{i,j} p_i(j) \log_e p_i(j)\, (d/d\lambda)\, p(i) \end{aligned} \quad (12)$$

Substituting equation (2) in equation (12) and rearranging we have

$$(dH/d\lambda) + J(\lambda) H = J(\lambda) H_o + \tfrac{1}{4} J(\lambda) \sum_{i,j} p(i) \log_2 p_i(j) + \sum_{i,j} p_i(j) \log_2 p_i(j)\, (d/d\lambda)\, p(i) \quad (13)$$

The third term on the right of equation (13) is negligible for biological systems. To show this we must discuss first the method of calculating the $(d/d\lambda)\, p(i)$. By definition (3) the following relation holds:

$$p(i) = \sum_j p(j)\, p_j(i). \quad (14)$$

Form the derivative with respect to λ and substitute equation (4):

$$(d/d\lambda)\, p(i) = \sum_j [p(j)\, (d/d\lambda)\, p_j(i) + p_j(i)\, (d/d\lambda)\, p(j)] \quad (15)$$

$$(d/d\lambda)\, p(i) = -\sum_j J_{ji}\, p_j(i)\, p(j) + \sum_j c_{ji}\, p(j) + \sum_j p_j(i)\, (d/d\lambda)\, p(j) \quad (16)$$

The equations (16) are a set of differential equations for the $p(i)$. They may be rearranged in the usual form:

$$(d/d\lambda)\, p(i) - \sum_j p_j(i)\, (d/d\lambda)\, p(j) = -\sum_j J_{ji}\, p_j(i)\, p(j) + \sum_j c_{ji}\, p(j) \quad (17)$$

We are interested in the conditions when the $(d/d\lambda)\, p(i)$ vanish. The condition is of course that the terms on the right of the equations (17) are all equal and that the determinant of the coefficients of the $(d/d\lambda)\, p(i)$ be different from zero. Among the circumstances in which this will occur are those where all $p_i(i) = q$ and all $p_i(k) = p$ $(i \neq k)$. That is, all letters are equally probable and one kind of error is as likely as the other. In my paper in Part V the behavior of $dH/d\lambda$ under the much stronger conditions that the J_{ij} and c_{ij} vanish at $\lambda = 0$ will be needed. Then, of course, providing that the determinant of the coefficients of the $(d/d\lambda)\, p(i)$ be different from zero, all $(d/d\lambda)\, p(i) = 0$. It may therefore be expected that except under most exceptional and special conditions the $(d/d\lambda)\, p(i)$ will be very small or will vanish.

It can be further shown that for a nearly perfect system the coefficients of the $(d/d\lambda)\, p(i)$ in equation (13) are small compared to one. Dancoff and

QUASTLER (5) have estimated the error rate per cell per generation to be some 10^{-1} to 10^{-2} times the spontaneous mutation rate per cell generation (10^{-4} to 10^{-13}). Taking this to mean that

$$p_i(i) = q \approx (1 - p) \quad \text{and} \quad p_i(j) = p \approx 10^{-6} \, (i \neq j)$$

we see that

$$p_i(i) \log_2 p_i(i) = +\log_2 (1 - p) \approx -p = -10^{-6}$$

$$p_i(j) \log_2 p_i(j) = -6 \times 10^{-6} \log_2 10 \approx -10^{-5} \tag{18}$$

Because of the discussion given above this term in equation (13) may be neglected.

Equation (13) gives the value of $(dH/d\lambda)$ at the values of $p_i(j)$ corresponding to H_d. Let these values be $p_i'(j)$.

$$\left.\frac{dH}{d\lambda}\right|_{H_d} = J(\lambda)[H_o - H_d + \tfrac{1}{4} \sum_{i,j} p(i) \log_2 p'_i(j)] \tag{19}$$

The coefficient of $J(\lambda)$ will be a constant so that $\left.\frac{dH}{d\lambda}\right|_{H_d}$ will behave as a function of λ like $J(\lambda)$. This result will be needed in my article in Part V.

REFERENCES

1. J. D. WATSON and F. H. C. CRICK: Genetical implications of the structure of deoxyribose nucleic acid. *Nature, Lond.* **171**, 964–967 (1953).
J. D. Watson and F. H. C. Crick: The structure of DNA. *Cold Spr. Harb. Symp. Quant. Biol.* **18**, 123–131 (1953).
J. D. Watson and F. H. C. Crick: Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid. *Nature, Lond.* **171**, 737–738 (1953).
2. G. GAMOW: Possible mathematical relation between deoxyribonucleic acids and proteins. *Biol. Medd., Kbh.* **22**, (3), 1–13 (1954).
3. C. SHANNON and W. WEAVER: *The Mathematical Theory of Communication*, University of Illinois Press, Urbana (1949).
4. W. J. BURDETTE: Incidence of tumors in isogenic strains. *J. Nat. Cancer Inst.* **12**, 709–714 (1952).
5. S. M. DANCOFF and H. QUASTLER: The information content and error rate of living things. In: *Information Theory in Biology*, ed. by Henry Quastler, 263–373, University of Illinois Press, Urbana (1953).
6. H. P. YOCKEY: An application of information theory to the physics of tissue damage. *Radiat. Res.* **5**, 146–155 (1956).
7. M. R. ZELLE and A. HOLLAENDER: Effects of radiation on bacteria. In: *Radiation Biology*, ed. by A. Hollaender, Vol. 2, chap. 10, McGraw-Hill, New York (1955).
8. O. RAHN: The order of death of organisms larger than bacteria. *J. Gen. Physiol.* **14**, 315–337 (1930).
9. R. PEARL: On the distribution of differences in vitality among individuals. *Amer. Nat.* **61**, 113–131 (1927).
10. R. G. SCHOTT: The inheritance of resistance to *Salmonella aertrycke* in various strains of mice. Thesis, Iowa State College Library, 1–59 (1931).
11. R. G. SCHOTT: The inheritance of resistance to *Salmonella aertrycke* in various strains of mice. *Genetics* **17**, 203–229 (1932).
12. H. O. HETZER: The genetic basis for resistance and susceptibility to *Salmonella aertrycke* in mice. *Genetics* **22**, 264–283 (1937).

13. W. V. LAMBERT: Genetic investigations of resistance and susceptibility to disease in laboratory animals. *Rep. Agric. Res., Iowa Agric. Exp. Sta.*, 89–90 (1931); 91–92 (1932); 115 (1933); 142–143 (1934); 158–159 (1935); 147–148 (1936).
14. J. W. GOWEN: Genetic investigations of resistance and susceptibility to disease in laboratory animals. *Rep. Agric. Res., Iowa St. Coll. Agric. Exp. Sta.*, 158–159 (1937); 151–153 (1938); 156–160 (1939); 192–194 (1940); 171–172 (1941); 189–190 (1942); 178–182 (1943); 204–210 (1944); 278–283 (1945); 257–260 (1946); 230–232 (1947).
15. J. W. GOWEN: Significance and utilization of animal individuality in disease research. *J. Nat. Cancer Inst.* **15,** 555–570 (1954).
16. A. HOLLAENDER (ed.): *Radiation Biology*, McGraw-Hill, New York (1955).
17. D. E. LEA: *Actions of Radiations on Living Cells*, Cambridge University Press, Cambridge, England (1955).
18. W. L. RUSSELL: Genetic effects in mammals. In: *Radiation Biology*, ed. by A. Hollaender, Chap. 12, McGraw-Hill, New York (1955).
19. M. YČAS: The protein text. This volume.

PART II

STORAGE AND TRANSFER OF INFORMATION

A CENTRAL issue in modern biology, which touches in some degree all branches of that science, is the problem of species specificity and its relation to protein-specificity and synthesis. The subject can be approached from many points of view but the one adopted by the authors of the papers in Part II is to seek the solution in terms of the properties of a communication system

The justification for considering, from this point of view, a phenomenon which looks, at first sight, to be purely biochemical lies in the recent discovery that protein specificity is expressed as an exact order of amino acid residues. If this is even substantially the case then it is germane to discuss such problems in these terms. In fact, a number of current papers on protein synthesis and specificity have recourse, at one point or another, to the language of information theory. Since the specificity of proteins is thought to be coded in the exact order of pairs of nucleotide bases in DNA, the relationship of DNA, RNA, and proteins can be considered from aspects which are mathematical rather than purely biochemical.

GAMOW was the first to notice these mathematical aspects. He and YČAS pursue in this part some of the issues which they reveal. The influence one hopes these considerations will have on the experimentalist is clear. Additional data on the amino-acid residue sequences and other structural data for a large number of proteins can be put to immediate practical use in solving for the protein code, and therefore in understanding more about protein synthesis. Unfortunately, mainly due to the lack of sufficient protein text, few definite answers can be given. But it is possible to eliminate some past errors and to phrase the question in a sharper fashion than before.

The notion that an abstract quantity such as information is stored in the genetic material and is transferred to proteins during their synthesis raises immediate questions as to how this is done, how much is transferred, and how this quantity is affected by changing experimental conditions. These questions are attacked from different analytical and experimental points of view by the papers by AUGENSTINE, by MAHLER, WALTER, BULBENKO and ALLMANN, and by KOCH and by GLINOS.

The information theoretic properties of communication systems of particular concern to the papers in this part are the coding problem, the representation theorem, and redundance. Each paper deals with issues of its own but in terms of these ideas to a greater or lesser degree. It is in this way, among others, that information theory may grow to be as useful to the biologist as thermodynamics is to the chemist, whether his subject is clearly one in communication as is that of FRISHKOPF and ROSENBLITH or somewhat less clearly that of protein specificity.

H. P. Y.

THE CRYPTOGRAPHIC APPROACH TO THE PROBLEM OF PROTEIN SYNTHESIS

GEORGE GAMOW and MARTYNAS YČAS
University of Colorado and University of New York

Abstract—The WATSON and CRICK suggestion concerning the role of DNA in replication, mutation, and protein synthesis requires a coding between the four-letter DNA alphabet and the twenty-letter protein alphabet. An attempt has been made to discover this code by cryptographic methods. Various schemes have been worked out but no success obtained at this writing. There is hope that as the number of protein sequences increases this problem will be solved.

SPEAKING about information storage and transfer in a living cell, one always likes to compare the cell with a large factory. The cell nucleus is the manager's office, directing the work of the factory, and the chromosomes are the file cabinets in which all blue prints and production plans are stored. The cytoplasm is the plant itself with the workers and machinery carrying out the actual production; those are, of course, the enzymes catalyzing various biochemical reactions. If something goes wrong with the information stored in the chromosome, the corresponding enzyme will also do a wrong thing. Consider, for example, an enzyme which produces the pigment necessary for color vision. If the particular section of chromosome carrying the directions for producing that pigment is defective, the enzyme will not get the correct instructions, and will not produce the right type of pigment. As a result, the individual will be color blind.

The materials of chromosomes and of enzymes are chemically different, except that in both cases we deal with long molecular chains formed by the repetition of a comparatively small number of different units. DNA (deoxyribonucleic acid), forming the chromosomes, is a sequence of *four* different units or 'bases': namely, adenine, thymine, guanine, and cytosine. For sake of picturesque presentation, we may associate them with four suits of cards: spades, clubs, diamonds and hearts. Each DNA molecule is equivalent to a sequence of cards many thousand units long, and the way in which different suits follow each other contains, in code form, the instructions to the original cell (fertilized ovum) and its descendants to develop into a rosebush, a skunk, or a man.

The first question is this. How is information which is carried by DNA molecules of the chromosomes duplicated when the cell goes through the process of division? An answer can be given on the basis of the model of DNA proposed about three years ago by J. WATSON and F. CRICK (1). They started with the fact, first noticed by E. CHARGAFF (2), that the number of adenines in any given DNA molecule is always equal to the number of thymines, while the number of guanines is always equal to the number of cytosines (3). In the playing card analogy there are as many spades as there are clubs, and as many diamonds as hearts. This suggests that we deal here with a double-stranded

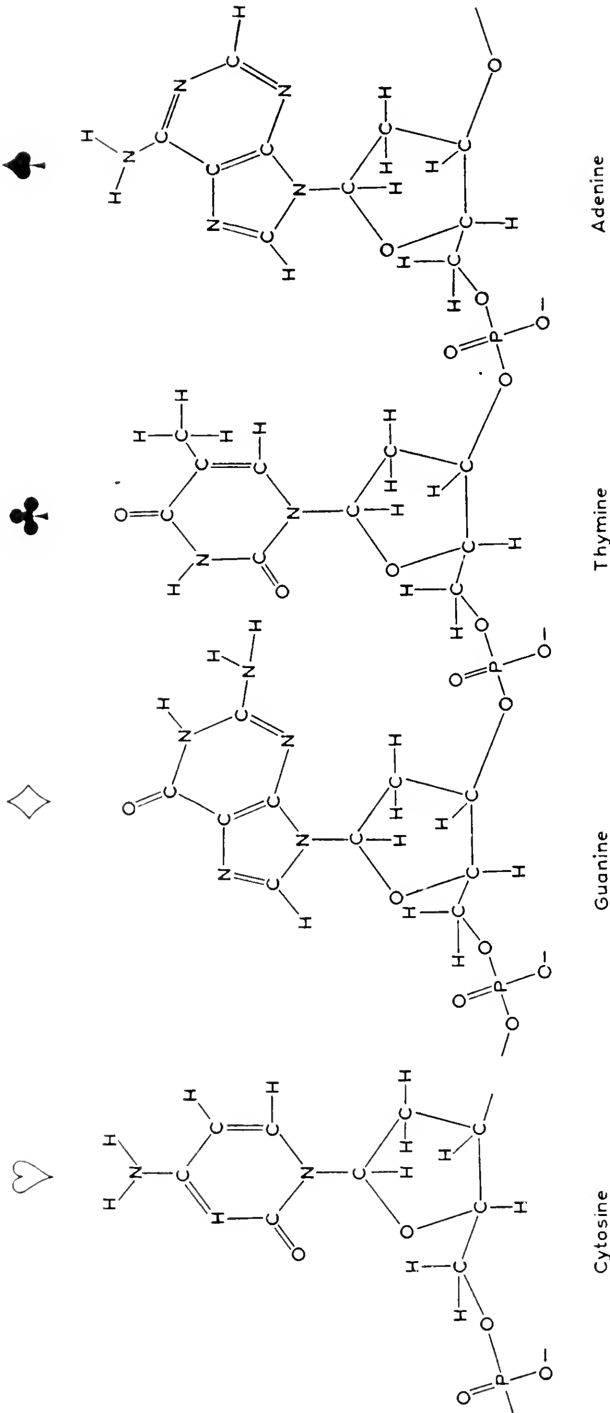


FIG. 1.

sequence in which red and the black cards are paired together. A heart is always paired with a diamond (and vice versa), while a spade is always paired with a club (and vice versa). The fact that DNA molecules also contain one sugar (ribose) and one phosphate for each 'base' suggests a molecular model similar to a rope ladder. The vertical ropes on both sides are formed by 'sugar-phosphate- sugar-phosphate-' sequences, while the paired bases form rigid horizontal steps attached to sugars on both sides. The reason why the above-mentioned pairing of bases takes place is two-fold. Cytosine and thymine (hearts and clubs) are 'pyrimidines', being formed by a single C—N— ring with different atomic groups attached to them. Adenine and guanine (spades and diamonds) are 'purines', and contain in their structure two connected rings, one with six atoms, and the other with five.

The chain shown in Fig. 1 is a sequence of sugars and phosphates. To each sugar is attached a 'base', and in this section of the molecule you see four different bases. Two of them (hearts and clubs) are short, and two others (spades and diamonds) are long. Now, in order to run the second strand beside it in the parallel way, we should attach short bases to long ones, and long bases to short ones. Of course, in the playing card analogy again, one could also join a heart to a spade and a club to a diamond. But this is excluded because in these cases hydrogen atoms will be in the wrong places to form proper hydrogen bonds between these two bases.

The evidence supplied by an x-ray diffraction pattern indicates in addition that the DNA molecule has a helical shape, being twisted around its central axis by 36° each step. Thus, it makes a complete turn each 10 steps.

The WATSON and CRICK (4) theory of duplication of DNA molecules proceeds as follows. When the cell is ready to divide, there appears a large number of free nucleotides in the nucleoplasm surrounding the chromosomes. A nucleotide is defined as one of the four bases with a sugar and a phosphate attached to it. At that time the double stranded DNA molecule splits into two single strands along its main axis, and each strand is regenerated by catching the corresponding free nucleotides from the surrounding medium. Thus, each heart separated by splitting from its diamond gets another diamond from the solution, and each diamond gets another heart. As the results, we get two new double stranded DNA molecules, each identical with the original one. Once in a while a mistake may be made in this duplication process, and we call it a mutation. So much for the structure and functioning of DNA molecules.

Now we come to the problem of information transfer from the chromosomes to the enzymes. How does the sequence of bases (card suits) in DNA determine the structure of the enzyme? Enzymes are proteins, and are formed by long sequences of twenty different chemical groups known as *amino acids*. It is well known that there are as many as twenty-four or twenty-five amino acids, but, as Dr Yčas tells us in more detail in the next paper, one can show that the extra ones in the original protein synthesis are modifications of the original twenty which take place *after* the protein molecule is synthesized. Thus, for example, 'proline' is an original amino acid used in protein synthesis, whereas 'hydroxyproline' is its postsynthetic modification. Since we symbolized four bases of nucleic acid molecule by four playing card suits, it is reasonable to symbolize the twenty basic amino acids, which have complicated chemical

names, by twenty letters of a (reduced) English alphabet. Thus, one protein molecule may look like:

...arreducesugarreducesug...

and another like:

...akeacoloruisionpigmentma...

Just to give an example of how the sequence of amino acids in protein molecules may affect their biochemical activity, we will give the example of two closely related hormones: oxytocine and vasopressin. Both are formed by a sequence of only nine amino acids:

Oxytocine—Cys-Tyr-*Ileu*-Glun-Aspn-Cys-Pro-*Leu*-Gly
Vasopressin—Cys-Tyr-*Phe*-Glun-Aspn-Cys-Pro-*Arg*-Gly

The two sequences are identical except for the substitutions in the third and eighth place. However, their functions are rather different. Oxytocine has the property of causing the contraction of the uterus in the process of childbirth. If you inject it into the blood of a cow, even if the cow is not pregnant it will go through all motions it would go through if a calf were to be born. Vasopressin, on the other hand, has rather different properties: it contracts the blood vessels and causes increased blood pressure. Thus, simply by changing two amino acids out of nine, the action of the hormone is completely changed.

Whereas replacement of some amino acids in a protein may completely change its biological function, there also exist replacements which distinguish the same protein taken from different species of animals. Thus, for example, insulin A, which is formed by a sequence of amino acids with twenty-one members, differs for cattle and swine in the eighth and tenth place. Human insulin, which has not yet been analyzed, possibly differs slightly from that extracted from cattle and swine. Nevertheless, the latter are successfully used on human patients.

Since there must exist a definite relation between the sequence of bases in nucleic acid and the sequence of amino acids in proteins, we can ask ourselves what this relation is. Here we have to return to our analogy of a factory. The workers from the factory do not walk into the manager's office to find out what to do, and the manager also does not go to the plant to instruct workers personally. There are people, called foremen, who get the information from the manager's office and tell the workers. In the cell the role of foreman is carried out by RNA molecules (ribonucleic acid) which are, presumably, very similar to the molecules of DNA. They are different only in that one oxygen atom is missing in each sugar of DNA, and there is a slight change in one of the four bases, which in RNA is called urosil instead of thymine. RNA is presumably synthesized by DNA inside the nucleus and receives the set of instructions carried by DNA. Then it passes out into the cytoplasm, and is incorporated into the so-called microsomes, i.e. foremen's offices, where the synthesis of proteins takes place.

We do not yet have a model of the RNA molecule. It seems, however, that in this case the pairing rules of adenine to thymine (urosil), and guanine to cytosine do not hold, which suggests that RNA molecules are single-stranded.

Since RNA serves as an intermediary between DNA and proteins, we have here two problems. First, how is RNA formed by DNA? Second, how are proteins synthesized by RNA? The first problem may turn out not to be very difficult because of the close similarity between the two molecules. For example, RNA may be a non-regenerated half of DNA with small changes in sugars and in one of the bases. It may be that the absence of the oxygen atom in RNA's sugar is responsible for the failure to form a double-stranded configuration. However, we still do not know the answer to this question.

The second problem concerning the synthesis of proteins by RNA molecules presents more challenge to the imagination. How can a sequence formed by four different units (four bases) be translated in a unique way into a sequence formed by twenty units (twenty amino acids)? Here is a possibility which seems to us to be very likely. Suppose one plays a game of poker in which only three cards are dealt, and pays attention only to the suit of the card. How many different hands will one have? Well, one can have a 'flush', i.e. three cards of the same suit. There are four different flushes: three hearts, three spades, etc. Then one can have a 'pair', i.e. two cards of the same kind, and one different. How many of those are there? One has four choices for the suit of the pair, and three choices for the third card. Thus, there are altogether twelve possibilities. The poorest hand will be a 'bust', i.e. three different suits. There are four different busts: no hearts, no diamonds, etc. We have altogether twenty different possibilities. This 'magic number' 20 is just the number of amino acids participating in the primary process of protein synthesis. We may imagine that each amino acid in the synthesized protein is determined by *a triplet* of bases in the RNA template.

Since the distances between neighboring amino acids in the extended polypeptide chain are equal to the distances of neighboring bases in the polynucleotide chain (both being equal to 37 Å), it was at first natural to suppose that the correlation between the two chains looks in a way shown in Fig. 2,

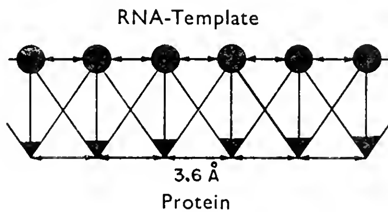


FIG. 2.

where individual bases are shown by circles and the amino acids by triangles. This represents the so-called *over-lapping code* in which the neighboring amino acids have in common two bases in the RNA template. If the transfer of information from nucleic acid to protein is carried out according to such an overlapping code, there must exist a definite inter-symbol correlation between the amino acids constituting protein molecules. Thus, for example, if a certain amino acid is determined by two adenines and some other base, its neighbors will be preferably amino acids which also contain adenine in their template transcript. In order to see whether or not such a correlation between the

neighbors really exists in the known protein sequences, it is necessary to test all possible assignments between the twenty amino acids and the twenty possible base triplets. The number of all possible assignments of that type is $20! = 3.10^{17}$. Since 3.10^{17} represents the age of our universe (5 billion years) expressed in seconds, the straightforward test of that kind would require quite a considerable time even if we could test one assignment each second! However, as it often happens in cryptographic problems, one can sometimes find parts of the message which reduce quite considerably the amount of necessary work. Thus the code messages sent by German spies during the war were likely to contain the combinations of letters corresponding to various possible ports of embarkation of American expeditionary troops. The same happens in protein sequence. For example, the adrenocorticotropin molecule contains the sequence:

—Lys—Lys—Arg—Arg—Pro—Val—Lys—Val—

In this sequence there are two identical amino acids in succession followed by another pair of identical ones. In the English language there are not many words having such a property. (Tennessee is one of the rare examples!) Then lys repeats again three steps later, and has identical neighbors (val) on both sides. These facts simplify the problem to such an extent that, instead of spending five billion years, it was possible to find a single assignment between the amino acids in the above sequence, and the base triplets in the course of an afternoon. At first it was thought the problem had been solved, but, when one tried to extend these assignments to the other parts of the ACTH molecule and to the other known protein sequences, one was led to direct contradictions. In the course of subsequent decoding work, other examples leading to similar contradictions were found, and it became clear that the thing just will not work. In fact, as Dr Yčas discusses in the following article, it seems that there is no correlation between the neighboring amino acids whatsoever.

This negative result can only mean that the original hypothesis represented in Fig. 2 was incorrect, and that in the process of protein synthesis the nucleic acid molecule is not present in its extended form. If, as seems to be true, we deal here with a "non-overlapping code" in which each amino acid is determined by an individual base triplet of its own (Fig. 3), we are forced to assume that

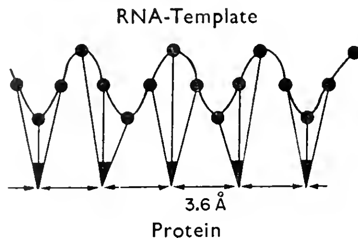


Fig. 3.

the RNA molecule is shrunk by a factor of three. We can imagine, for example, that during the process of protein synthesis the RNA molecule has the shape of a spiral as shown in Fig. 4.

Closely connected with the problem of a non-overlapping code is the problem

of "punctuation". Indeed, a sequence of bases can be broken into a set of non-overlapping triplets in three different ways depending upon the base with which we start. The three different readings of the same template can be described mathematically as 3n, 3n/1, and 3n/2 (3n/3 being the same as 3n).

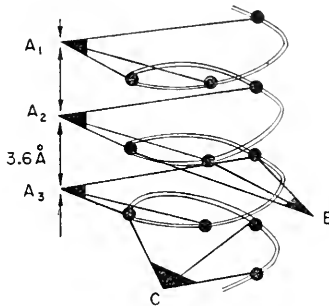


Fig. 4.

As was suggested by Dr Barbara Law, three possible readings of the same RNA template may explain an interesting regularity first noticed by Dr Martynas Yčas. He observed about two years ago that, in a case of seven proteins for which the sequences of amino acids were known, the total number of amino acids in the protein molecule was a multiple of three: nine amino acids in oxytocine and vasopressin, twenty-one in insulin A, thirty in insulin B, thirty-nine in ACTH, 126 in ribonuclease, etc. This could be explained if one assumes that each RNA template synthesizes the proteins in all three possible ways, and that these three different readings are afterwards united in one linear sequence. If this were true, there must exist a cryptographic correlation between the first, second, and third "thirds" of each protein molecule. One thinks of how such a correlation could be checked, but it seems to be very difficult indeed. Recently, though, the existence of such a correlation became rather doubtful, since two protein sequences published recently contain 29 and 124 amino acids.

In summing up, we should say that the problem of finding the nature of the correlation between polynucleotide chains of nucleic acids, and the polypeptide chains of the proteins is still unsolved, although various methods for establishing such a correlation have been worked out. We may hope, however, that with the increased number of known protein sequences, this problem will be solved in one way or another.

REFERENCES

1. J. D. Watson and F. H. C. Crick: Molecular structure of nucleic acids. *Nature, Lond.* **171**, 737–738 (1953).
2. E. Chargaff: for reference see S. Zamenhof, G. Brawerman, and E. Chargaff: On the deoxypentose nucleic acids from several micro-organisms. *Biochim. Biophys. Acta.* **9**, 402–405 (1952).
3. G. R. Wyatt: Nucleic acids of some insect viruses. *J. Gen. Physiol.* **36**, 201–205 (1952).
4. J. D. Watson and F. H. C. Crick: Genetical implications of the structure of deoxyribose nucleic acid. *Nature, Lond.* **171**, 964–967 (1953).

THE PROTEIN TEXT

MARTYNAS YČAS
Department of Microbiology, State University of New York
Upstate Medical Center, Syracuse, New York

And strange to tell, among that Earthen Lot
Some could articulate, while others not:
And suddenly one more impatient cried—
'Who is the Potter, pray, and who the Pot?'
The Book of Pots

Abstract—The sequence of residues in proteins, regarded as a text written in a twenty symbol alphabet, is examined. The following tentative conclusions are drawn:

1. Twenty amino acids are distinguished by the protein-forming mechanism. Supernumerary amino acids arise from the regular twenty by secondary modification of protein-bound residues.
2. Each residue in the protein has a separate genetic representation.
3. There is no intersymbol correlation between adjacent residues.
4. Natural selection is not the only factor determining the frequency of occurrence of the various kinds of residues. It is suggested that the method of encoding protein sequence information in nucleic acid imposes differences in frequency of occurrence on the different kinds of residues.
5. Peptide chains are not multiples of some fixed number of residues.

The encoding and transfer of genetic (DNA) information to RNA and protein is discussed, as well as the problem of the independent reproduction of RNA viruses. While the data set certain limits on the possible ways of encoding and transferring information, they are not sufficient for a unique solution of these problems.

RIBONUCLEIC acid of Tobacco Mosaic Virus (TMV) has been shown to determine the sequence of amino acid residues in the protein of the virus (1, 2, 3). It seems logical therefore to believe that the sequence of other proteins is also determined by RNA.*

Since RNA is essentially a linear sequence of four kinds of nucleotides, while proteins are linear sequences of about twenty kinds of amino acid residues, the RNA molecule can be regarded as a text, written in a four-symbol alphabet, which encodes another text, the protein, written with about twenty symbols.

* The following abbreviations will be employed. RNA—ribonucleic acid; DNA—deoxyribonucleic acid; Ad—adenylic acid; Gu—guanylic acid; Cy—cytidylic acid; Ur—uridylic acid; ala—alanine; arg—arginine; asp—aspartic acid; aspn—asparagine; asx—aspartic acid or asparagine; cys—cysteine; glu—glutamic acid; glun—glutamine; glx—glutamic acid or glutamine; gly—glycine; his—histidine; ileu—isoleucine; leu—leucine; lys—lysine; met—methionine; phe—phenylalanine; pro—proline; ser—serine; thr—threonine; try—tryptophan; tyr—tyrosine; val—valine; Hlys—hydroxylysine; Hpro—hydroxyproline; serP—phosphoserine. Peptides are written with the amino group to the left, the symbols being connected by a dash (—). The sign (*) signifies a terminal residue. Sequences considered uncertain are in parentheses ( ). Symbols in parentheses, with commas between (ala, gly) mean that the sequence is not known.

Several attempts, none completely convincing, have been made to determine the coding system employed (4, 5, 6, 7). Cryptography must be based on a study of texts, and I shall therefore attempt an examination of protein molecules from this point of view. The following aspects of protein structure will be examined:

1. The number of kinds of amino acids which occur in proteins.
2. The effect of mutations on amino acid sequence.
3. Whether intersymbol correlations exist between adjacent residues.
4. The frequency of occurrence of the various amino acid residues.
5. Whether any restrictions exist on the length of peptide chains.

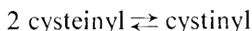
After considering the empirical evidence, I shall indicate its bearing on the problem of encoding protein sequence information into the RNA molecule.

1. THE NUMBER OF AMINO ACIDS OCCURRING IN PROTEINS

In previous studies (6, 7) it has been assumed that proteins are composed of exactly twenty different kinds of residues. Since in fact more than twenty kinds of residues occur in proteins, the assumption requires some justification.

All organisms, from viruses to mammals, use the same building blocks for their proteins. With minor qualifications this is also true of the nucleic acids, but not true of the third major class of biologically-occurring high polymers, the polysaccharides. The amino acids which invariably occur in all organisms and virtually all proteins are the following: ala, arg, asp, aspn, cys, glu, glun, gly, his, ileu, leu, lys, met, phe, pro, ser, thr, try, tyr, val. The number in this list is exactly twenty.

It will be noted that I omit cystine from this list. Because of its structure, cystine corresponds to two residues. The structure of insulin (8) shows that one cystinyl residue can occupy non-adjacent positions in a peptide chain or even participate in two different chains. Cystine is best regarded as an oxidation product of cysteine, formed after incorporation of the cysteinyl residue into the peptide chain. This view is supported by the recent discovery of an enzyme which reversibly catalyzes the reaction



$$2 \text{ cysteinyl} \rightleftarrows \text{cystinyl}$$

when these residues are protein bound (9). Another example of such a reaction may be the cyclic oxidation and reduction of protein SH groups during the various stages of cell division (10).

In addition to the above twenty, other alpha amino acids occur in nature. Some of these, such as homocysteine, citruline and ornithine are well known biochemical intermediates but do not occur in proteins. It is clear that the number of amino acids which occur in proteins is limited by an inability to incorporate, rather than make, amino acids. Hydroxyglutamic acid and norleucine, previously believed to be protein constituents, have been shown not to exist as natural products (11). Alpha amino-adipic acid has been isolated from an impure protein hydrolyzate, but it has not been demonstrated that it is a protein constituent in the same way as other amino acids (12). Diamino pimelic acid, commonly occurring in bacteria, appears to be associated with the polysaccharide material of the cell wall (13, 14).

Nevertheless, there are amino acids, other than the twenty enumerated,

which certainly occur in proteins. These include hydroxylysine and hydroxyproline (in collagen), phosphoserine (in a number of different proteins (15)), thyroxine (in thyroglobulin) and tyrosine —O— sulphate (in fibrinogen) (16). The distribution of these amino acids is different from the regular twenty. Whereas the twenty amino acids occur in virtually all proteins, the supernumerary ones have an erratic distribution, being confined to one or to a few. The suggestion was first made by CRICK, that the supernumerary amino acids are the result of modifications of some of the regularly occurring amino acids after these have been incorporated into a peptide chain. The biochemical evidence for this is as follows.

When one of the twenty regularly occurring amino acids is presented labeled to an organism, it is rapidly incorporated into protein and most of the label is found in the corresponding residue. It should be noted that glutamine and glutamic acid are separately incorporated and do not arise one from another by addition or subtraction of amide groups after incorporation (17). (A similar demonstration for the analogous case of asparagine and aspartic acid is still lacking.) Clearly, therefore, these amino acids are the precursors of the corresponding protein-bound residues.

The supernumerary amino acids behave differently. Thus lysine is the precursor of hydroxylysine (18), but C^{14} or tritium-labeled hydroxylysine is not incorporated into collagen (19). Similarly, proline is the precursor of hydroxyproline, but proline is a much better precursor of the hydroxyprolyl of collagen than is hydroxyproline itself (20, 21). These amino acids, then, are not incorporated as such, but presumably are formed by oxidation of protein-bound proline and lysine. Phosphoserine likewise is formed by phosphorylation of protein-bound serine (22). Thyroxine is apparently formed from the tyrosine residues of thyroglobulin (23). There is no information at present on the metabolism of tyrosine —O— sulfate.

Since not all appropriate residues are secondarily modified, this interpretation implies that the enzymes catalyzing such conversions show specificity for sequence in the protein. At least one enzyme is known which shows such specificity. Prostatic phosphatase dephosphorylates phosphoserine in the sequence asx-serP-glx-ileu-ala, but not in glx-serP-ala (24). It is therefore suggestive of some enzyme specificity that hydroxyproline in collagen occurs mainly, if not exclusively, before glycine (25) (Table IV). Other amino acids, as shown later, show no such neighbor preferences. The region determining whether proline is to be oxidized or not probably includes more than three residues, as indicated by the isolation from collagen of the tripeptides ala-pro-gly; ala-Hpro-gly and ser-pro-gly; ser-Hpro-gly (Table IV).

The biochemical evidence thus appears to indicate that the protein-forming mechanism selects exactly twenty different kinds of amino acids, and that the supernumerary ones arise by secondary modification of protein-bound residues. A possible cause for error in this conclusion should be noted. It is virtually certain that amino acids are not incorporated as such, but in the form of some sort of activated derivative. If the same amino acid were to form more than one derivative, the number of items to be selected would of course exceed twenty. There is no evidence for this at present, and only further advances in biochemistry can decide whether this is the case.

II. GENETIC EFFECTS ON PROTEINS

There is an increasing body of evidence indicating that the details of protein structure are genetically determined. A study of the effect of mutations on proteins should therefore tell us something both about the nature of mutations and the protein forming mechanism. Known cases of genetic effects on proteins are listed below.

1. In man hemoglobin occurs in several electrophoretically distinguishable forms, the presence of each being apparently controlled by alleles of a single gene (26). Hemoglobin C differs significantly in amino acid composition from hemoglobin A (27). Hemoglobin A and S have been degraded in a controlled fashion with trypsin and the resulting peptides separated. The difference between these hemoglobins is apparently confined to a short section of the molecule (28).

2. Two electrophoretically different hemoglobins occur in sheep. Their presence is determined by alleles of a single gene (29).

3. Two forms of lactoglobulin occur in cow's milk, and like the hemoglobins are determined by different alleles of one gene. Crystallographic investigations indicate unit cells of the same size, but there are very slight differences in the diffraction pattern, which the investigators attribute, possibly, to the substitution of a few amino acid residues by others (30).

4. Mutants of *Neurospora* and *Escherichia coli* produce abnormally heat-labile forms of tyrosinase (31) and a panthothenic acid synthesizing enzyme (32), respectively. It is clear that a change in the proteins has occurred, but unfortunately there is no further information on its physico-chemical nature.

The genetic evidence indicates that there is no interaction between alleles controlling the synthesis of different variants of one protein. If both alleles are present, both types of protein are formed. A possible exception should be noted. The N-terminal groups of wheat gliadin are reported to be phe, of rye gliadin phe and glx, but unexpectedly the gliadin of wheat × rye hybrids was found to have no amino or carboxyl terminal ends, indicating, possibly, a cyclic protein (33). This case obviously needs further study*.

The evidence cited above shows that the properties of proteins are gene-determined, but it does not indicate clearly what these properties are. More detailed information is available on this point from a comparison of homologous proteins of related species, if it is assumed, as is usually done, that species differences are the result of gene mutations.

Available evidence on amino acid sequence of homologous proteins is

* There is considerable confusion as to the N-terminal residues of wheat gliadin. Fraenkel-Conrat (51) misquotes Deich and Soreni (33) as stating that the N-terminal residues are phenylalanine and histidine, apparently because of a misunderstanding in *Chemical Abstracts* (138). Koros, whose paper I was able to consult only in abstract (139), reports histidine as N-terminal. Ramachandran and McConnell (140), working with wheat gliadin but failing to specify the species, also find histidine. Deutsch (the same as Deich quoted above, the difference in spelling being due to transliteration from the Cyrillic) reports that gliadin from *Triticum durum* and *Triticum vulgare* has N-terminal phenylalanine (141). This is misquoted as tyrosine, and tyrosine and glutamic acid, respectively, by Ramachandran and McConnell (140). The original paper of Deutsch (141) was also unavailable to me.

collected in Table I. Mutations (as inferred from differences between homologous proteins) do not produce a general scrambling of protein sequence, but a replacement of one or more residues, leaving the rest of the sequence unchanged. Since homologous proteins can differ by a one residue replacement, it is clear that individual residues, rather than groups of residues, are represented in the genetic material.

Table I. Sequences in Homologous Proteins from Different Species

Protein	Species
Insulin (34)	
. . . cys-thr-ser-ileu-cys . . .	Pig
. . . cys-ala-ser-val-cys . . .	Cattle
. . . cys-ala-gly-val-cys . . .	Sheep
. . . cys-thr-gly-ileu-cys . . .	Horse
. . . cys-thr-ser-ileu-cys . . .	Whale
Myoglobin (35)	
*val . . .	Finback whale
*val . . .	Sperm whale
*gly . . .	Horse
*gly . . .	Seal (*Phoca vitulina*)
Protamine (36) (Composition, not sequence)	
$gly_2ser_2ala_2val_4ileu_1$	*Salmo irideus*
$gly_2ser_2ala_2val_5ileu_0$	*Salmo trutta*
Serum albumin (37, 38)	
asp-ala . . . . leu	Man
asp-thr . . . . ala	Cattle
Cytochrome c (39)	
. . . cys-ala-glun . . .	Horse, Cattle, Pig, Salmon
. . . cys-ser-glun . . .	Chicken
Vasopressin (40)	
. . . pro-arg-gly-NH_2*	Cattle
. . . pro-lys-gly-NH_2*	Pig

Protein	Species
Hemoglobin (41)	
*val-leu . . . *val-gly . . . *val-glun . . .	Horse, Pig
*val-leu . . . *val-gly . . . *val-asx . . .	Dog
*val-leu . . . *met-gly . . .	Cattle, Goat, Sheep
*val-leu . . . *val-ser . . . *val-asx . . .	Guinea pig
*val-leu . . . *val-gly . . .	Rabbit, Snake
*val-leu . . .	Chicken
Gliadin (33)	
*phe . . . *phe . . .	Wheat
*phe . . . *glx . . .	Rye
Fibrinogen (42)	
*tyr . . . *ala . . .	Man
*tyr . . . *glx . . .	Cattle
ACTH (43, 44, 45)	
. . . pro-ala-gly-glu . . . . . . pro-gly-ala-glu . . .	Sheep Pig
. . . glu-ala-ser-glu . . . . . . glu-leu-ala-glu . . .	Sheep Pig
Hypertensive p·ptide (46, 47)	
. . . val . . . . . . ileu . . .	Cattle Horse

Protein	Species
Virus (48)	
. . . thr-ser-gly-pro-ala-thr*	TMV (M, YA strains)
. . . thr(thr,ala)pro-ala-thr*	TMV (HR strains)

It is possible that a mutation may suppress an amino acid determining site altogether. This is indicated by the tentative finding of AKABORI (quoted in (41)), that the 'B' chain of fish insulin has the sequence . . . pro-lys*, as compared with the sequence . . . pro-lys-ala* in cattle.

In some cases (ACTH, TMV), two adjacent replacements differentiate one homologous protein from another. It is not probable that this is due to two independent but adjacent mutations, but rather that a single mutational event has affected two residue-determining sites. Such a view is made plausible by the work of BENZER (49). He has shown that mutations in bacteriophage involve small sections of DNA, of molecular dimensions, but that these sections can be of different lengths. Presumably the length of the mutated section determines the number of residues changed in the protein. It is perhaps not too sanguine to hope that eventually it may become possible to measure crossover values in terms of distance in residues along a protein chain, and thus obtain an estimate of the number of bases in DNA determining a single residue selecting site. The present difficulties of such an approach are of course obvious (50).

It would be of interest to determine if there are any restrictions on the replacement process. Restrictions might be expected on the following grounds. More than one nucleotide must determine an amino acid site. If the process of mutation were predominantly to change some, but not all nucleotides determining a site, then obviously not all sites would be interconvertible in one step. A study of any such restrictions would be of great value, since their nature would depend on the coding principle and could be used to infer the latter.

Table II. Replacements Inferred from Table I and their Frequency of Occurrence

Occurrence	Replacement
3	val ↔ ileu
2	ala ↔ thr
2	ala ↔ ser
2	ala ↔ gly
2	ala ↔ leu
2	ser ↔ gly
1	ala ↔ glx
1	val ↔ gly
1	val ↔ met
1	phe ↔ glx
1	glun ↔ asx
1	arg ↔ lys

Known replacements in homologous proteins are collected in Table II. In the small sample we have (nineteen replacements), half recur twice or more, suggesting strongly that the process, as observed, is not a random one. Unfortunately, the sample is not unbiased. Certain replacements are lethal or semi-lethal (hemoglobin S, for example), and are, without doubt, selected against. What we actually observe has therefore passed through the sieve of selection. The direct genetic approach to this problem is tedious, because of the difficulty of determining the phenotype (the amino acid sequence), and rapid progress is scarcely to be expected. A much larger body of data on homologous proteins may, however, enable us to reach a decision on whether the replacement process is intrinsically restricted or not.

An additional point emerges from a consideration of such protein molecules as consist of more than one chain (Table III). It will be noted that there

Table III. Terminal Residues of Proteins having more than one Peptide Chain

(The exact number of chains is not indicated.)

Protein	N-terminal	C-terminal	Reference
Cytochrome c	his his		(51)
Growth hormone	phe ala	phe phe	(51)
Triosephosphate-dehydrogenase	val val	met met	(51)
Collagen		gly ala	(51)
Gliadin (wheat)	phe phe		(33)
Gliadin (rye)	phe glx		(33)
β lactoglobulin	leu leu	ileu ileu	(51)
Fibrinogen (man)	tyr ala		(51)
Fibrinogen (cattle)	tyr glx		(51)
Hemoglobin (horse)	val val		(41)
Hemoglobin (cattle)	val met		(41)

is a strong tendency for the terminal residues of such proteins to be identical. This is certainly not due to the chains being identical in all cases, since the hemoglobins, for example, do differ in the penultimate positions (Table I). Rather it appears to indicate that multi-chain proteins arise by reduplication of genetic material, so that the several chains start out by being identical, but gradually diverge in the course of evolution in the same way as homologous proteins of different species. This hypothesis, as applied to the hemoglobins and insulin, has been previously discussed (6). Determinations of the residue sequence along different chains of one protein may therefore throw additional light on the replacement process.

Table I shows that the process by which replacements become established is very slow. Elucidation of the sequence of homologous proteins may therefore make it possible to determine phylogenetic relations between large groups such as phyla, which cannot now be certainly determined from morphological and embryological evidence.

III. CORRELATIONS BETWEEN ADJACENT RESIDUES

Are there any forbidden combinations of adjacent residues? An examination of the sequence of residues in proteins (Table IV) could provide an answer to this question.

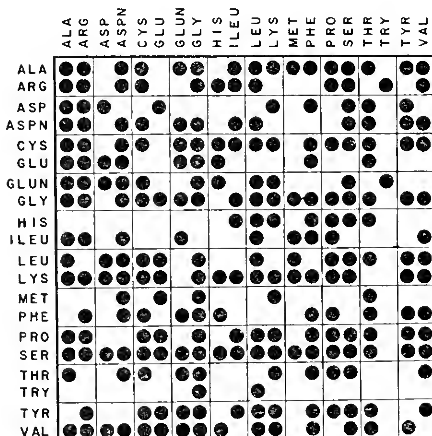


FIG. 1. Dipeptide sequences now known to occur in proteins, compiled from Table IV. The N-terminal amino acids are plotted in the rows, the C-terminal in the columns.

There are of course 400 possible pairs of the twenty amino acids. The known protein sequences in Table IV have been broken down in the following way. A sequence, say, of ala-arg-gly is broken down into the dipeptides ala-arg, arg-gly, and the appropriate cells in Fig. 1 are then filled, the N-terminal residues being represented by the rows, the C-terminal by the columns. Using all the data available in Table IV, Fig. 1 shows that somewhat more than half of all possible dipeptide combinations are known to occur. The question

Table IV. List of Known Sequences in Proteins

Actin (52)
. . . his-ileu-phe*

Adrenocorticotropin (45)
ser-tyr-ser-met-glu-his-phe-arg-try-gly-lys-pro-val-gly-lys-lys-arg-arg-pro-val-lys-val-tyr-pro-ala-gly-glu-asp-asp-glu-ala-ser-glu-ala-phe-pro-leu-glu-phe

Carboxypeptidase (53)
*aspn-ser; ser-thr

α Casein (54, 55)
serP-glx; *lys-leu-val-ala-glx-asx

Chymotrypsinogen (56, 57)
leu-ser-arg-ileu-val; aspn-ser-gly-(glun-ala)

Clupein (58, 59)
pro-ser-arg; ser-ala-arg-arg; arg-arg-arg-arg;

Collagen (60, 61)
ala-Hpro-gly; ala-pro-gly; glx-arg; glx-Hpro-gly; gly-asx-gly; gly-glx; gly-pro-ala; gly-pro-glx; gly-pro-gly; gly-pro-Hpro; ser-Hpro-gly; ser-pro-gly; ala-gly-ala; gly-gly; ser-gly; thr-gly; ala-asx; asx-asx; asx-glx; asx-gly; glx-ala; glx-glx; glx-gly; glx-gly-gly; glx-met; glx-phe; ser-asx; val-glx; ala-arg; arg-gly-gly; arg-val-gly; ser-arg; val-arg; ala-lys; asx-arg; lys-gly; pro-ser; pro-thr; ser-ala; thr-ala; lys-pro-gly; leu-ala; ala-ala-gly;

Cytochrome c (39)
. . . val-glun-lys-cys-ala-glun-cys-his-thr-val-glu . . .

γ Globulin (rabbit) (62)
*ala-leu-val-asx . . .

Glucagon (63)
his-ser-glun-gly-thr-phe-thr-ser-asp-tyr-ser-lys-tyr-leu-asp-ser-arg-arg-ala-glun-asp-phe-val-glun-try-leu-met-aspn-thr

Hemoglobin (41)
*val-glun-leu; *val-leu; (horse). *val-gly; *met-gly; *val-ser; *val-glx; *val-asx; (various species, see Table I.)

Hypertensive peptide (46)
asp-arg-val-tyr-val-his-pro-phe-his-leu

Insulin (cattle) (8)
'A' chain: *gly-ileu-val-glu-glun-cys-cys-ala-ser-val-cys-ser-leu-tyr-glun-leu-glu-aspn-tyr-cys-aspn*
'B' chain: *phe-val-aspn-glun-his-leu-cys-gly-ser-his-leu-val-glu-ala-leu-tyr-leu-val-cys-gly-glu-arg-gly-phe-phe-tyr-thr-pro-lys-ala*

β Lactoglobulin (64)
his-ileu*

Lysozyme (65, 66, 67, 68)
thr-asx-val-glx-ala; ileu-glx-leu-ala-leu; asx-glx-ala; leu-thr-ala; glx-asx-ileu; thr-glx-ala-gly; ser-asx-gly-met-asx; asx-ala-met-lys-cys-arg; val-thr-pro-gly-ala; ser-asx-arg; lys-phe-glx-gly; arg-cys-glx-ala; ser-phe-asx-glx; thr-asx-arg-arg; thr-gly-asx-val; ser-val-cys-ala-lys-gly; gly-cys-asx; leu-gly-ala-val; asx-ileu-pro-cys; arg-cys-lys-gly; ser-val-asx-cys-ala; asx-leu-cys-asx; arg-asx-cys-ileu; ser-arg-leu; ser-asx-cys-arg-leu; arg-asx; arg-gly; asx-asx; gly-leu; ileu-arg; ileu-asx; ileu-val; leu-leu; ser-ala; ser-leu; val-ala; *lys-val-phe-gly-arg; arg-his-lys; asx-gly-ala-asx-leu*; glx-ser-phe-asx; ala-lys-phe-glx; asx-tyr-arg-gly; arg-gly-tyr-ileu-leu; asx-ala-tyr-gly-ser-leu-asx; leu-pro; ala-ala-met;

Melanophore expanding hormone (69, 70)
asp-glu-gly-pro-tyr-lys-met-glu-his-phe-arg-try-gly-ser-pro-pro-lys-asp

Myoglobin (71)
*gly-leu

Ovalbumin (72, 15, 73, 74, 64)
val-ser-pro*; asx-serP-glx-ileu-ala; glx-serP-ala; ala-gly-val-asx-ala-ala; cys-ala; cys-val; cys-gly; cys-phe; thr-cys; ser-cys; cys-glx; glx-cys; phe-cys; asx-cys; val-cys;

Oxytocin (75)
cys-tyr-ileu-glun-aspn-cys-pro-leu-gly-NH_2

Papain (76)
*ileu-pro-glu

Pepsin (77, 15)
*leu-gly-asx-asx-his-glx; thr-serP-glx;

Prolactin (78)
*thr-pro-val

Ribonuclease (79)
lys-glu-thr-ala-ala-ala-lys-phe-glun-arg; lys-ser-arg-aspn-leu-thr-lys-asp-arg; lys-aspn; tyr-glun-ser-tyr; tyr-lys; lys-his; asp-ala-ser-val

Salmine (80, 81)
*pro-arg-arg; arg-pro-val-arg-arg; pro-ileu-arg; val-gly; arg-val-ser-arg; arg-ileu-arg; arg-ala-ser-arg; arg-gly-gly-arg; arg-ser-ser-arg; val-gly;

Serum albumin (37)
*asp-ala (man); *asp-thr (cattle);

Silk fibroin (*Bombyx*) (82, 83, 84)
gly-ala-gly-ala-gly-[ser-gly-(ala-gly)$_n$]$_8$-ser-gly-ala-ala-gly-tyr
n usually 2, mean value always 2.
gly-val-gly; tyr-gly; phe-gly; gly-ser-pro-tyr-pro; tyr-pro-ser-tyr;

Tobacco mosaic virus (48)
thr-ser-gly-pro-ala-thr*

Tropomyosin (52)
ala-ileu-met-thr-ser-ileu*

Trypsinogen (85)
*val-asp-asp-asp-asp-lys-ileu

Vasopressin (40)
cys-tyr-phe-glun-aspn-cys-pro-arg-gly-NH_2

Wool (86)
ser-cys; gly-cys; thr-cys; ala-cys; leu-cys; cys-gly; cys-thr; cys-ala; cys-val; cys-leu; cys-phe;

remains whether any of the blank cells represent forbidden combinations, or whether they are merely the result of accidents of sampling.

To answer this question statistically, the frequencies of occurrence of various combinations have been plotted in Fig. 2. There are more blank cells here than in Fig. 1, as a portion of the data has been discarded to avoid obvious sources of bias. Thus the sequences of silk, collagen, wool and protamine have been omitted, since these proteins have an obviously aberrant structure. Likewise, sequences of less than three residues have not been used, since the ease of

isolation of various dipeptides varies, making it possible that the frequencies of some peptides have been systematically over- or underestimated.

Figure 2 can now be treated as a contingency table with 761 degrees of

	ALA	ARG	ASP	ASPN	CYS	GLU	GLUN	GLY	HIS	ILEU	LEU	LYS	MET	PHE	PRO	SER	THR	TRY	TYR	VAL	a
ALA	4			1			2	3		1	3	3	2	1		3	1		1	1	26
ARG	1	3		2	2			4		1	2				1			1		1	18
ASP	1	2	4			2						1		1		1			1		13
ASPN	2	2			4		3	2		2	3					1	1		2	2	24
CYS	4	2		3	1		1	2	1	1		1			1	1			1		19
GLU	4	1	1	2	1		1	1	1					1			1				16
GLUN	4	1	1	1	1			2	1		3	1				2		1			18
GLY	3	1		1	1	2	1			1		2	1	1	2	2	1		1	1	21
HIS										1	3			1	1	1	1				8
ILEU	1						2				1		1	1	2					2	10
LEU	1		1	1	2	2		3					1			1	2		2	4	20
LYS	1	1	2		2	1		2		1	1	1		3	1	1			1	2	20
MET				2		1						1					1				5
PHE		1		1			4	1	1					1	1		1		1	2	14
PRO	2	1			1	1		1			2	2		1	1				1	3	16
SER	1	4	1	3		1	1	2	1	1	2	1	1	1	2				2	4	28
THR	2			2			1	1				1		1	3	4				1	16
TRY								1			1										2
TYR		1			1		2	1		2	2	1		1	1	2	1			1	16
VAL	1		1	2	3	3	4	1	1			1		1		1	1		2		22
b	32	20	11	21	19	13	22	27	6	11	23	16	6	15	16	20	11	2	15	24	330
c	58	38	24	45	38	27	40	48	14	21	43	36	11	29	32	48	27	4	31	46	

FIG. 2. Frequencies of occurrence of dipeptide sequences in proteins, plotted as in Fig. 1. The sequences of clupein, collagen, salmine, silk fibroin and wool have not been used. Sequences of less than three residues, as well as those where the acid and amide forms of glx and asx are not differentiated, were also not used. On the basis of the study of OHNO (68), glx and asx in lysozyme are assigned to glun and aspn, respectively. The seven-residue sequence common to ACTH and MEH was counted only once. a—marginal totals of rows; b—marginal totals of columns c—marginal totals of rows and columns.

freedom, and the null hypothesis, that there is no correlation between adjacent residues, tested. The deviation λ from the expected distribution in Fig. 2 is calculated as:

$$\lambda = \sum_{ij} \frac{\left[a_{ij} - \frac{(a_{i.} + a_{.i})(a_{j.} + a_{.j})}{4n}\right]^2}{\frac{(a_{i.} + a_{.i})(a_{j.} + a_{.j})}{4n}} \tag{1}$$

where n is the sum of the marginal totals (330), a_{ij} the value of a cell in column i and row j, $a_{.i}$ and $a_{i.}$ the marginal totals in column and row respectively of the residue defining the column, $a_{.j}$ and $a_{j.}$ the analogous values for the residue defining the row. For computational purposes (1) reduces to:

$$\lambda = n\left(\sum_{ij} 4\frac{a_{ij}^2}{(a_{i.} + a_{.i})(a_{j.} + a_{j.})} - 1\right) \tag{2}$$

From Fig. 2, $\lambda = 392$. The value of t, which is calculated from

$$t = \sqrt{2\lambda} - \sqrt{761} \tag{3}$$

is 0.414, which is less than 1.645, the 5 per cent confidence limit.

It may therefore be concluded that there is no evidence for any intersymbol correlation between nearest neighbors. Inspection of sequences reveals likewise no obvious correlations of residues more than one removed from each other, but to decide this question definitely will require more knowledge of longer sequences than is now available.

GAMOW, RICH and YČAS (6) have previously studied this question of intersymbol correlation. They examined a grid diagram, similar to Fig. 2 but embodying fewer data, to see whether the frequencies of entries follow the POISSON distribution. This method is invalid, since it does not take into account the fact that different amino acids occur with very different frequencies. I am glad to avail myself of this opportunity to correct these authors.

IV. FREQUENCY OF OCCURRENCE OF DIFFERENT AMINO ACIDS

Amino acids occur with different frequencies in proteins. Some, like leucine, are consistently abundant, others, like methionine, consistently rare. The frequency of occurrence of the various amino acids in the bulk protein of a whole organism, *Escherichia coli*, is shown in Fig. 3.

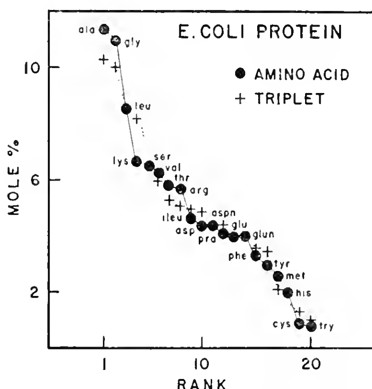


FIG. 3. Composition of bulk protein of *Escherichia coli* (87), amino acids arranged in order of abundance. The values for glu, glun and asp, aspn arbitrarily taken as half of glx and asx, respectively. The value of cysteine taken from ROBERTS and COWIE (88). 'Triplets' refers to the frequencies of triplets of nucleotides, calculated according to the hypothesis of GAMOW and YČAS (7) from the composition of *E. coli* RNA (89).

Data on the composition of twenty-three proteins are summarized in Table V. This table shows that the composition of individual proteins is not too different from that of bulk protein. The most abundant amino acid usually has a frequency of about 0.10 to 0.12, the least 0.005 to 0.01.

Table V suggests the possibility that the differences in composition of various proteins may be merely the result of chance fluctuations from a mean, and not importantly related to biological function. This notion may not be as far-fetched as might appear at first sight. The most important function of proteins is catalysis, and the enzymatically active site probably involves only a few amino acids. In addition, proteins of a given organism appear to have

an important mutually complementary relation to each other which enables them to be retained by the cells. This is shown by experiments with injected catalase. Homologous catalase injected into guinea pigs is absorbed by the tissues, but heterologous catalase is rejected (108). Similarly, homologous antibodies readily pass the fetal barriers in rabbits, heterologous pass much less readily (109). This phenomenon is probably connected with the antigenicity of proteins. The antigenically active sites of proteins are probably also small, and therefore the exact sequence and composition of the major part of the protein may be irrelevant to function. It might be expected, then, that the exact structure of small parts of a protein molecule would be rigidly determined, and any mutation affecting this portion would be eliminated by selection. Mutations affecting the 'irrelevant' portions may not affect the viability of the organism, and the same protein in different species may therefore diverge by a process of 'evolutionary drift.' That this process is real is strongly suggested by the facts known about cytochrome *c*. This enzyme serves the same function and has the same prosthetic group in both yeast and mammalian tissues, but the two cytochromes have very different elution volumes from ion exchange resin columns (110), almost certainly indicating a large difference in amino acid composition.

If for each kind of residue there is a characteristic rate of replacement by mutation, the proteins should approach a definite equilibrium composition, if selection is a minor factor. More definitely, each protein will constitute a 'random grab' from a universe of amino acids, the frequencies of the amino acids in this universe being determined by the equilibrium condition.

Qualitative considerations suggest that there is something other than selection which tends to make a given amino acid occur with a certain frequency. Certain amino acids, alanine, leucine, isoleucine and valine have aliphatic side chains lacking any obvious reactive functional group. The data on replacements (Table II) indicate, apparently, that one is as good as another, as far as their function in a protein is concerned. Yet leucine is systematically more abundant than isoleucine. These two amino acids are so similar that it is difficult to separate them by paper chromatography. Each of the other aliphatic amino acids has its own characteristic frequency. likewise.

Quantitatively, if a sample of n items is drawn at random from a population where an item of type A occurs with frequency p, the distribution of A in a large series of samples is given by the binomial $(p + q)^n$, where $q = 1 - p$. In particular, the variance σ^2 of the distribution of A is given by

$$\sigma^2 = npq \tag{4}$$

If the hypothesis of a 'random grab' is correct, then in a collection of proteins the variances of amino acids should be related to the mean value of their frequencies and to the size of the proteins, expressed as the number of residues per molecule.

An immediate difficulty is that the sizes of the proteins listed in Table V are not known, and these certainly differ one from another, It should be particularly noted that the relevant size is *not* necessarily that obtained from physical measurements of diffusion, osmotic pressure and sedimentation. This is because there is ample evidence that physical molecules can be the result of

aggregation of smaller, chemically identical units. Furthermore, from the evidence presented in Table III, the several peptide chains constituting some proteins may not be identical, but are nevertheless quite similar. The statistically relevant size of hemoglobin would then be somewhere between 600, the approximate number of residues in the whole molecule, and 150, the average size of the four subunits.

Disregarding this difficulty, I have plotted the variance of each amino acid, calculated from Table V, against pq (Fig. 4). All points (except glx) fall within

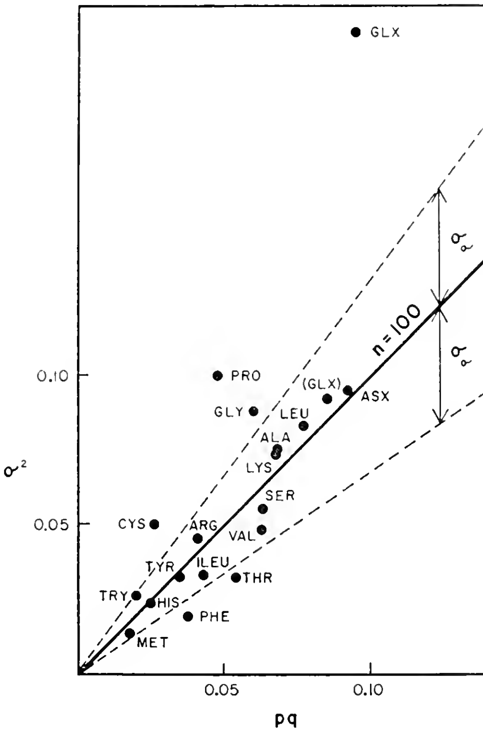


FIG. 4. Plot of variances of amino acids against pq, where p = mean frequency of occurrence of amino acid, $q = 1 - p$. Line $n = 100$ calculated variance for sample size (protein) of 100 residues. (glx) is plot with the values from tropomysin and γ casein omitted.

(or very close to) one standard error of the line for $n = 100$. The fact that the sizes of the proteins are not identical tends to scatter the points, making agreement with the hypothesis somewhat more significant. The large deviation of glx is due to its abundance in two proteins, γ casein and tropomyosin. If these are omitted the agreement is good.

The evidence therefore permits (but of course does not prove) the hypothesis that the composition of proteins is mainly determined not by selection, but rather approximates to a 'random grab' from a single universe of amino acids.

There is of course no question that selection *can* produce proteins of very

Table V. Composition of Nonstructural Proteins
in moles per cent, p-mean frequency, as fraction of 1. σ^2—variance

	Transphosphorylase (90)	Prothrombin (91)	Actin (92)	Tropomyosin (92)	Hemoglobin (93)	Myoglobin (93)	Lysozyme (94)	α Lactalbumin (95)	Alcohol Dehydrogenase (96)	Salivary amylase (97)	γ Casein (98)	Papain (99)	γ Globulin (100)	Papaya lysozyme (101)	Carboxypeptidase (102)	Aldolase (103)	Glyceraldehyde dehydrogenase (103)	Ribonuclease (79)	Nerve protein (104)	Myeloma globulin (105)	β Lactoglobulin (106)	Serum albumin (106)	Phosphorylase (107)	p	σ^2
ala	2.6	6.2	8.2	12.9	12.2	7.3	7.8	3.1	9.2	3.4	3.0	7.1	4.4	9.2	6.4	10.6	8.4	9.7	5.9	6.0	9.2	8.1	6.7	0.073	0.076
arg	7.1	6.4	4.4	4.9	2.3	1.3	8.6	0.8	3.5	7.6	1.3	4.9	4.4	5.7	3.2	4.0	3.3	3.2	4.6	3.5	1.9	3.9	8.3	0.043	0.045
asx	11.4	10.2	9.5	10.5	8.3	6.7	15.6	16.9	8.0	16.2	3.5	9.3	8.1	9.6	9.6	8.1	10.5	12.9	13.7	8.1	10.0	9.5	8.7	0.102	0.095
cys	1.7	3.6	1.3	0.8	1.6	0.0	7.8	6.2	2.3	3.6	0.0	4.4	2.6	3.5	1.3	1.0	1.0	6.5	0.0	2.0	3.3	6.3	0.3	0.027	0.050
glx	12.0	11.8	11.7	24.8	5.5	12.7	3.1	10.8	5.6	8.2	18.5	9.3	10.1	4.8	8.0	8.6	5.1	9.7	12.3	9.0	15.0	13.0	11.5	0.105	0.216
gly	4.2	8.1	7.8	1.5	6.3	8.7	8.6	5.4	9.7	5.0	2.3	12.6	4.2	11.4	7.4	8.3	8.9	2.4	6.9	7.3	2.1	2.8	6.3	0.064	0.088
his	4.9	1.6	2.2	0.6	6.0	5.3	0.8	2.3	2.0	2.8	2.9	0.5	1.3	1.3	2.6	3.0	3.6	3.2	5.1	1.7	1.2	3.0	2.7	0.026	0.023
ileu	3.0	3.2	6.6	3.4	0.0	4.0	4.7	6.1	6.6	4.9	4.0	4.9	3.7	4.8	6.4	6.6	7.7	2.4	5.4	2.2	5.2	2.3	6.2	0.045	0.033
leu	9.9	7.8	7.3	11.2	12.5	12.0	6.2	10.8	6.7	4.8	10.8	4.9	6.6	5.3	8.0	9.7	5.7	1.6	8.3	8.0	13.7	10.9	10.0	0.084	0.083
lys	10.4	5.3	6.0	12.9	8.1	14.7	4.7	9.2	6.6	5.4	4.9	4.4	5.5	4.4	5.8	7.2	7.2	8.1	4.2	6.4	9.9	10.2	6.2	0.073	0.075
met	2.7	1.8	3.5	1.9	0.9	2.0	1.6	0.8	4.2	2.0	3.6	0.0	1.2	1.8	0.3	0.9	2.0	3.2	0.0	0.8	2.5	0.6	2.2	0.018	0.013
phe	3.8	3.9	3.4	0.4	6.3	5.3	2.3	3.1	4.3	6.2	4.2	2.2	4.7	5.3	4.8	2.0	3.7	2.4	5.3	3.6	2.6	4.6	4.7	0.039	0.019
pro	5.3	6.5	5.1	0.0	4.9	5.3	1.6	1.5	5.2	2.9	17.4	4.9	7.0	7.9	3.5	5.5	3.5	4.0	4.8	6.7	5.2	4.8	5.1	0.051	0.100
ser	4.2	7.3	6.5	4.7	5.4	4.7	7.8	5.4	6.8	6.2	6.2	6.0	8.8	7.0	10.6	6.9	7.1	12.1	6.1	12.9	4.4	4.7	3.6	0.068	0.055
thr	4.4	5.3	6.8	3.3	5.5	2.7	5.5	5.4	6.7	3.7	4.4	3.8	10.9	5.7	8.7	6.0	6.4	8.1	5.6	7.0	4.8	5.7	4.5	0.057	0.032
try	2.8	2.2	1.2	0.0	1.8	2.0	6.2	3.8	1.0	6.4	0.8	2.7	2.4	3.1	1.9	1.3	1.1	0.0	1.8	0.9	1.1	0.3	1.2	0.020	0.026
tyr	3.7	3.3	3.7	1.8	1.8	1.3	2.3	3.8	4.1	4.8	1.8	9.3	5.8	5.7	6.4	3.2	2.8	4.8	2.6	4.5	2.3	3.2	4.1	0.038	0.032
val	6.0	5.6	4.9	4.5	10.4	4.0	4.6	4.6	7.3	5.7	10.7	8.2	8.4	3.5	5.1	7.0	11.8	7.3	7.2	9.4	5.5	5.9	7.8	0.068	0.068
amide	1.3	9.2	7.6	7.5	7.8	8.7	14.1	11.5	—	—	13.5	10.4	—	9.2	6.1	7.2	7.5	13.7	12.3	1.26	8.9	6.5	10.7		

unusual composition. This occurs mainly in cases where the mechanical properties of protein fibers are important, as in keratin, collagen and silk. These have been omitted from Table V. The most extreme case known to me is the silk of the Congolese moth *Anaphe maloneyi*, where glycine and alanine together constitute 94 per cent of the entire protein (111).

Fox and Homeyer (112) have also noted the general similarity of composition of various proteins, but have interpreted it in a quite novel manner. Their suggestion is that proteins are similar because the time that has elapsed since the origin of life has been too short to allow more differences to develop between the various proteins, all of which are presumed to be descendants of a single molecule. I believe the composition of silk tends to indicate that there has been ample time for any conceivable differentiation.

V. LENGTH OF PEPTIDE CHAINS

I have previously called attention to the apparent fact that the number of residues in naturally occurring peptide chains is an exact multiple of three (113). Since then, a more exact determination of the composition of ribonuclease (79) and the elucidation of the structure of glucagon (63) have shown that this statement is incorrect (Table VI). In view of the predominance of chain lengths

Table VI. Length of Protein and Peptide Chains in Number of Residues

(Note: Cystine counted as two cysteine residues.)

Protein or peptide	Number of residues	Reference
Oxytocin	9	(75)
Vasopressin	9	(40)
Melanophore expanding hormone I (hog)	18	(69, 70)
Insulin 'A' chain	21	(8)
Glucagon	29	(63)
Insulin 'B' chain	30	(8)
Melanophore expanding hormone II (hog)	30	(114)
Melanophore expanding hormone (ox)	48	(114)
Ribonuclease	124	(79)

that are multiples of three, it might perhaps be suspected that the exceptions are due to secondary removal of residues, as occurs, for example, in the activation of pepsinogen, trypsinogen, chymotrypsinogen and fibrinogen. The tentative finding of Akabori (quoted in (41)), that the B chain of fish insulin has twenty-nine residues, rather than the thirty found in cattle insulin, makes it doubtful that secondary removal of residues is the explanation. Since twenty-nine (the number of residues in glucagon) is a prime number, and not a factor in the chain lengths of other peptides, it seems reasonable to conclude that peptide chains are not multiples of some fixed number of residues.

VI. THE CODING PROBLEM

Having examined the protein text, we can now discuss what conclusions we may draw as to the storage, transfer and replication of the information contained in the protein molecule.

The gene, and by inference DNA, is thought to contain the information which eventually appears as a sequence of amino acid residues in the corresponding protein. As shown by a study both of the replacement process and of the amino acid sequences, each residue has an independent genetic representation. These representations are presumably aligned in linear order on the DNA molecule. There is in fact no evidence at present that the gene is anything other than a linear sequence of amino acid determining sites, although the possibility that it may also determine the structure of immunopolysaccharides in an analogous fashion cannot yet be dismissed.

Recent biochemical evidence (which I shall not discuss here) indicates that it is RNA, not DNA, which is directly involved in the process of protein formation. Transfer of information therefore involves at least two steps: DNA to RNA, and RNA to protein.

The straightforward inference would thus be that DNA serves as a template for the formation of RNA. Absence of cytoplasmic inheritance supports the view that RNA is not a self-replicating structure. This is also supported by four lines of biochemical evidence:

1. The initial rate of incorporation of labeled precursors into nuclear RNA is much greater than into cytoplasmic RNA (115).
2. In *Amoeba* depleted of RNA, RNA only regenerates if a nucleus is present (116).
3. A one-way flow of RNA from nucleus to cytoplasm can be demonstrated (117).
4. The rate of RNA formation is minimal at the time DNA is replicating (118).

Unfortunately, this conclusion may be an oversimplification. There is no lack of biochemical evidence pointing in the opposite direction:

1. The composition of nuclear and cytoplasmic RNA is not identical (119).
2. The time curves of precursor incorporation into RNA do not indicate that the nuclear fraction is the precursor of the cytoplasmic (115).
3. Radioactive precursor is incorporated into the RNA of enucleated *Acetabularia* plants (120).
4. Different strains of RNA viruses are self-replicating. This is difficult to explain if RNA is the product of a DNA template.

The problem is to reconcile these apparently discordant facts. Consider first the determination of RNA structure by DNA. Since both DNA and RNA are texts written in a four symbol alphabet, it is natural to suppose that the coding problem is very simple. It is sufficient to assume that one nucleotide of DNA determines one nucleotide of RNA (121). Recent evidence indicates, however, that this is incorrect.

It is possible to suppress protein synthesis in susceptible bacteria with chloramphenicol. When this is done using amino acid-requiring strains, it can be demonstrated that amino acids are required for RNA synthesis, even though no protein synthesis is taking place (122, 123, 124). The natural inference, supported by several converging lines of evidence, is that it is not the nucleotides themselves which are the precursors of RNA, but rather compounds containing both a nucleotide and an amino acid. This leads to a unitary picture of the synthesis of RNA and of protein. When such precursors are lined up on a

protein-synthesizing template (RNA), the amino acids polymerize to form protein; when lined up on DNA, the nucleotide portions polymerize to form RNA (Fig. 5).

If this is correct, an obvious conclusion follows. Since omission of a single amino acid stops RNA synthesis, the RNA-forming mechanism must distinguish not four, but a minimum of twenty different kinds of items. But since the product contains only four, the RNA in general must contain less information than the template that made it. Several nucleotides in DNA must be involved in selecting a single nucleotide of RNA. Since the template must contain more information than the product, RNA cannot be the template for itself; i.e. it cannot be self-replicating. There is an important exception to this statement.

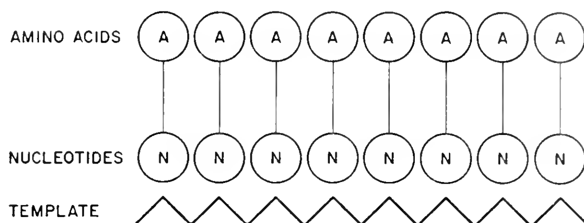


FIG. 5. Schematic representation of the synthesis of RNA and protein from common precursors (see text). The nature of the template is presumed to determine whether the aligned precursors polymerize to produce protein or RNA.

If the information in the template is reduced below a certain level, it is possible to obtain a product identical to the template itself. The formalization is as follows.

While in process of formation, the RNA molecule can be visualized as a sequence of nucleotides to which amino acids are attached (Fig. 5). Before removal of amino acids on polymerization the informational content of the 'proto-RNA,' of length n, is $n \log_2 20$. After removal of the amino acids the information content is reduced to $n \log_2 4$. If restrictions of some kind exist on the number of combinations allowed, the number possible for 'proto-RNA' will be reduced to $b[n \log_2 20]$; $(b < 1)$. Such restrictions on 'proto-RNA' will result in less severe restrictions on the RNA itself, since in general one configuration of RNA can correspond to numerous different configurations of 'proto-RNA'. Therefore, if there are 20^{bn} possible configurations of 'proto-RNA', RNA itself has 4^{cn} possible configurations available $(1 > c > b)$.

The information content of RNA will equal that of 'proto-RNA'

$$bn \log_2 20 = cn \log_2 4 \quad (5$$

when $1 > c \geqslant 2.16b$. Since the information content of 'proto-RNA' is now the same as that of RNA, an RNA template could, formally, be self-replicating.

It is now possible to reconcile the genetic and biochemical facts outlined above. Assume that the synthesis of RNA proceeds in two steps. At the first step, a strand of RNA is synthesized using a DNA template. Information is thus transferred from DNA to RNA. The next step is supposed to occur in the cytoplasm. RNA material is added to the nuclear-synthesized RNA, but in a manner which does not add to the informational content. A model for

this process could be the building up of a complementary strand of DNA, as in the WATSON and CRICK scheme for DNA reproduction (125).*

Normally, the process stops at this stage, since the RNA molecule has insufficient information to act as a template for itself. In the case of viruses, however, the cytoplasmic process of adding new material to the original RNA

Table VII. The Composition of the Protein and RNA of Viruses

Composition of protein in moles per cent, of RNA as fractions of 1. † value assumed. It should be noted the influenza virus contains lipid, and the protein analysed may in part be of host provenance.

Protein	Tobacco Mosaic (126)	Tomato Bushy Stunt (127)	Turnip Yellows (128)	Southern Bean Mosaic (129)	Influenza A (130)
ala	9.6	8.5	6.6	7.5	5.9
arg	6.4	5.3	1.6	6.4	6.0
asx	10.3	11.1	4.2	7.3	11.7
cys	0.7	0.8	2.3	0.9	—
glx	8.7	5.7	7.1	6.8	7.0
gly	3.9	8.6	4.2	8.9	7.0
his	0.0	1.2	1.5	1.3	1.9
ileu	5.2	3.3	9.0	6.2	8.3
leu	7.1	10.9	8.6	8.3	8.5
lys	1.1	3.4	8.0	3.0	5.2
met	0.0	0.8	2.1	2.6	3.2
phe	5.7	3.6	2.5	3.5	4.7
pro	5.5	3.9	10.2	5.3	4.7
ser	10.0	8.6	8.4	8.7	4.4
thr	11.6	11.0	13.9	11.5	6.5
try	1.1	0.5	0.6	1.0†	1.1
tyr	2.4	2.8	1.5	4.1	3.6
val	10.8	10.0	7.9	6.7	6.1
amide	12.7	11.4	8.0	—	—
RNA	(131)	(127)	(132)	(131)	(133)
Ad	0.30	0.26	0.22	0.26	0.23
Gu	0.25	0.29	0.18	0.26	0.20
Cy	0.19	0.21	0.38	0.23	0.24
Ur	0.27	0.26	0.22	0.25	0.33

results in the production of material identical to the template itself. From this point of view, an RNA virus can be regarded as a specialized RNA molecule, which because of restrictions on the sequence of 'proto-RNA' can act as its own template, utilizing the normal RNA-synthesizing mechanism of its host.

The composition of the RNA of viruses lends some support to these ideas.

* It is obvious that until more is known about RNA structure the question of its replication can be discussed only in general terms. If RNA is a double-stranded structure, the nucleotide composition shows that bases in the two chains cannot be uniquely paired as in DNA, but each base must pair with one of two others, as shown by the equality of 6-keto and 6-amino groups (89). In attempting to elucidate the details of RNA reproduction information on the number of strands, whether each strand contains all the information of the whole structure, and where the complementary strand is synthesized, is of crucial importance.

Normally the number of 6-keto (Gu + Ur) and 6-amino (Ad + Cy) groups in RNA is equal (89). Virus RNA does not necessarily obey this rule, indicating that it differs in this respect, at least, from all the others (Table VII).

This hypothetical scheme is presented to show that the apparent contradictions of the genetic and biochemical evidence do not make it logically necessary to abandon a unitary view of RNA reproduction.

The coding of protein information into RNA has attracted considerable attention, but cannot as yet be considered as solved. Study of the protein text indicates that any solution will have to meet several requirements.

Firstly, since exactly twenty amino acids are incorporated into protein, it is clear that at least three nucleotides are needed to determine an amino acid. GAMOW (134) has proposed that 20 is a 'magic' number, which is the result of the existence of twenty possible sites of three nucleotides each. Four kinds of items, taken three at a time, give twenty different combinations, if order is disregarded.

CRICK, GRIFFITH and ORGEL (135) point out, however, that there is at least one other way of deriving a 'magic' 20 number. They start by considering the problem of what it is that delimits one amino acid-determining site from another, the 'punctuation mark problem'. Assuming that three bases determine a site, it is a problem why the $3n + 1$, $3n + 2$, $3n + 3$ bases represent a site, while $3n + 2$, $3n + 3$, $3n + 4$ do not. They solve this problem by assuming that only certain triplets of nucleotides correspond to an amino acid (sense sites), while others do not (non-sense sites). The criterion separating these two types of sites is the following. The set of sense sites are all triplets which, when placed next to each other in any possible combination, give sense sites only at positions $3n + 1$, $3n + 2$, $3n + 3$, but not otherwise. For example, the triplet AAA is a non-sense site, since when placed next to itself it gives the sequence AAAAAA. The site is not unambiguously defined, as AAA occurs both at the 1–3 position and at the 2–4 position. They find that there are exactly twenty triplets (out of sixty-four) which satisfy the criterion of sense sites, as follows:

ABA	BCA	ADC	BDD
ABB	BCB	ADD	CDA
ACA	BCC	BDA	CDB
ACB	ADA	BDB	CDC
ACC	ADB	BDC	CDD

Other ways of selecting twenty sense sites are also possible. The sense sites, these authors suggest, may correspond to amino acid-selecting sites of RNA.

The 'punctuation mark problem' could, of course, also be solved if amino acids were selected in a sequential manner starting from one end of the template.

Secondly, besides the requirement that at least three nucleotides are required to determine an amino acid site, the study of proteins indicates that these amino acid determining sites are independent and share no nucleotides with their neighbors. This conclusion follows from the absence of any intersymbol correlations in the protein text, and also from the fact that a mutation (as inferred from a study of homologous proteins) can result in a change at one site only, leaving the rest of the sequence unchanged. The number of nucleotides

in the template must therefore exceed the number of residues in the corresponding protein by a factor of at least three.

Absence of intersymbol correlation shows that the 'overlapping' codes discussed by GAMOW, RICH and YČAS (6) do not correspond to reality.

The third requirement is somewhat more hypothetical. From the evidence presented above, it would appear that selection is not the sole factor determining the frequency of occurrence of the various amino acids. This is strongly suggested by the different frequencies of amino acids with aliphatic side chains, and particularly by the characteristic preponderance of leucine over isoleucine. It is therefore reasonable to believe that the coding principle itself imposes certain differences in frequency on the various amino acids.

If only one configuration of nucleotides corresponds to each amino acid, the coding *per se* cannot make some amino acids frequent and others rare. This can be done, however, if some amino acids have more than one configuration of nucleotides to which they correspond. For this reason I am inclined to believe that the type of coding proposed by CRICK, GRIFFITH and ORGEL (135) does not correspond to reality.

GAMOW and YČAS (7) have proposed a code that formally meets these three requirements. An amino acid is presumed to be determined by three nucleotides, taken without regard to order. In addition, the number of nucleotides in the RNA is assumed to be three times the number of amino acid residues in the corresponding protein. This has the following consequences:

1. There are twenty such triplets, the same as the number of amino acids.
2. Neighboring triplets share no nucleotides between them. Any sequence of amino acids is thus permitted.
3. The frequencies of various amino acids, calculated on the assumption that the sequence in RNA is random, are unequal. This is because the expected frequency of any triplet is given by the product of the frequencies of the component nucleotides *and* the number of configurations for the given composition. Thus there are six triplets (all presumed to determine the same amino acid) of the type ABC, three of AAB and one of AAA.

The pattern of frequency distribution of the various triplets, calculated in this manner, corresponds very closely to the amino acid distribution, as shown, for example, in Fig. 3 for the case of *E. coli*.

I believe that this type of coding, even if not itself the one which actually occurs, is similar to the one that corresponds to reality. The most striking defect is that it provides no explanation, in fact contradicts, the requirement that in RNA the number of 6-keto groups should equal the number of 6-amino groups. H. A. SIMON (136) has proposed a modification to take care of this difficulty. If RNA is a paired structure, somewhat similar to DNA, and 6-keto bases pair with 6-amino ones, then the following four pairs of nucleotides exist (again disregarding order):

Ad–Gu; Ad–Ur; Cy–Gu; Cy–Ur.

If one takes these pairs, rather than the individual nucleotides, as units, one can maintain an hypothesis of determination by sextuplets, analogous to determination by triplets. The frequency distribution of sextuplets, calculated for a random RNA sequence, is very similar to that obtained for the triplet

distribution. This suggests that a whole series of codes of this type may exist, all having similar general properties.

At present the major difficulty is not to produce a coding principle that explains the known facts, but rather to make a choice between the many that are possible.

The correctness of a coding principle can, in general, be ascertained from a consistency of correspondence of the RNA and protein texts. Unfortunately, such a direct approach is not at present possible. Except perhaps in the case of RNA viruses, it is not possible to isolate a pure RNA corresponding to a pure protein, and were this possible, the sequence of nucleotides could not be determined by any method currently available.

If the composition only of a series of RNA's and the corresponding proteins is known, it is theoretically possible to check some coding schemes as follows: If the coding scheme is correct, the various configurations of nucleotides can be assigned to the amino acids in such a manner as to give, when summed over the protein, the experimentally determined RNA composition, and this consistently for all RNA-protein pairs. No assumption need be made that the RNA sequence is random. Actual application of this method requires a large number of RNA protein pairs of accurately determined composition, obviously differing as much as possible from each other, and the facilities of an electronic computer.

The electronic computer is much the easier of the two to provide. At present the data are hopelessly inadequate, although analyses of the proteins and RNA's of viruses may eventually make such an approach possible. However, in attempting a correlation of viral RNA and protein (Table VII), it should be remembered that some viral RNA's do not show the equality Ad + Cy = Gu + Ur characteristic of non-viral RNA (89). This suggests that normal RNA may be multi-stranded, while viral may not be. It is therefore not impossible that viral RNA may contain all the information, but not all the material of a protein determining structure, and hence differ in composition from it. An additional difficulty is that it is not certain that all viral RNA is concerned in the determination of the protein which eventually appears in the virus particle.

In lieu of anything better, I have attempted to make consistent assignments of triplets to amino acids on the assumption that the sequence in RNA is random. The random frequencies of triplets were calculated for liver (Fig. 5), Tobacco Mosaic and Turnip Yellow virus. I then tried to assign each triplet to an amino acid in such a manner that each member of the pair would have approximately the same frequency in the three cases. No satisfactorily consistent assignments could be obtained by this method. Assuming that the RNA's and proteins actually correspond, failure indicates one or more of the following:

1. The coding principle used is false.

2. The RNA is not a random sequence.

3. The proteins of viruses are so small that relatively large deviations from expected frequencies may be found. The molecular weight of TMV protein is about 17000 (48, 137), that of Southern Bean mosaic about 26000 (129). Several of the amino acids occur as only a few residues per molecule, so that a

difference of one or two residues from the statistically expected value produces very large relative deviations.

Since the frequency of occurrence of an individual amino acid is small, even a larger protein such as hemoglobin may be too small to be a statistically valid sample for the purpose of calculating frequencies on the basis of a random RNA sequence. The following case is of interest. The RNA's of liver and of reticulocytes are virtually identical in composition, and therefore the proteins (bulk liver protein and hemoglobin) would be expected to have a very similar composition. Actually, this is not the case (Fig. 6). Considerable differences

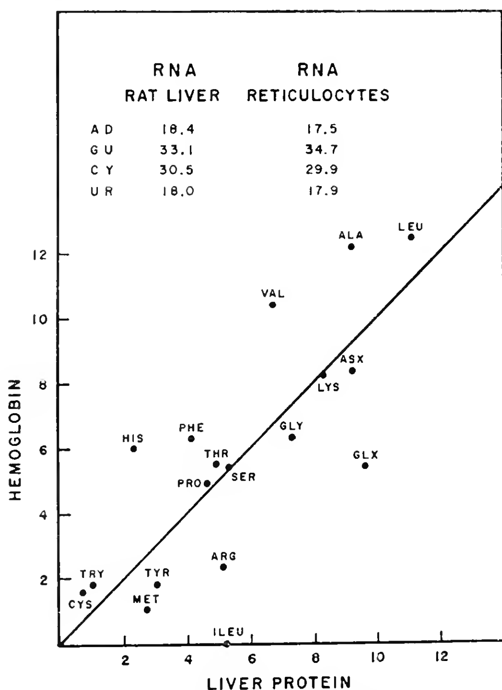


FIG. 6. The composition of bulk liver protein (142) and hemoglobin (93). The RNA composition of liver from (89) of reticulocytes (143). All in moles per cent.

exist, as can be seen from the deviations of the points from the line of slope 1.

It would be better to use for this purpose the bulk RNA's and proteins of whole organisms and organs, were it not for the fact that bulk protein and RNA from various sources is so similar that no strong check on the coding principle is possible.

The method of assignments from the assumption of a random RNA sequence fails, then, either strongly to confirm or to deny any proposed coding principle.

It is possible that as more information becomes available some light may be thrown on the coding problem from a study of replacements of residues in homologous proteins, if replacements prove to be nonrandom.

The reader will not fail to notice that the inadequacy of the data render most of my conclusions tentative. More information of the type considered

here will, of course, become available in the future and will not fail to clarify matters. I have attempted to organize and analyse such data as exist, in the hope that the value of this sort of information might become clearer, and in order to facilitate their examination as more become available.

Obviously, data on composition and sequence are not the only possible sources of information bearing on coding. Strong hints will eventually be obtained from a study of RNA structure and sequence, as well as from other, more conventional, biochemical approaches. The solution of these problems will surely not be long delayed.

Acknowledgment—It is a pleasure to acknowledge the collaboration of GEORGE GAMOW. I have profited from discussions of various aspects of these problems with Drs F. H. C. CRICK, BEATRICE S. MAGDOFF and HERBERT A. SIMON. Dr LOUIS J. COTE has given valuable assistance with statistical problems. The errors, of course, remain my own.

REFERENCES

1. H. FRAENKEL-CONRAT: The role of the nucleic acid in the reconstitution of active tobacco mosaic virus. *J. Amer. Chem. Soc.* **78**, 882–883 (1956).
2. A. GIERER and G. SCHRAMM: Infectivity of ribonucleic acid from tobacco mosaic virus. *Nature, Lond.* **177**, 702–703 (1956).
3. A. GIERER and G. SCHRAMM: Die Infektiosität der nucleinsäure aus tabakmosaikvirus. *Z. Naturf.* **11b**, 138–142 (1956).
4. A. L. DOUNCE: Duplicating mechanism for peptide chain and nucleic acid synthesis. *Enzymologia* **15**, 251–258 (1952).
5. D. SCHWARTZ: Speculations on gene action and protein specificity. *Proc. Nat. Acad. Sci., Wash.* **41**, 300–307 (1955).
6. G. GAMOW, A. RICH, and M. YČAS: The problem of information transfer from the nucleic acids to proteins. *Advances in Biological and Medical Physics* **4**, 23–68, Academic Press, New York (1956).
7. G. GAMOW and M. YČAS: Statistical correlation of protein and ribonucleic acid composition. *Proc. Nat. Acad. Sci., Wash.* **41**, 1011–1019 (1955).
8. A. P. RYLE, F. SANGER, L. F. SMITH, and R. KITAI: The disulfide bonds of insulin. *Biochem. J.* **60**, 541–556 (1955).
9. W. J. NICKERSON and G. FALCONE: Identification of protein disulfide reductase as a cellular division enzyme in yeasts. *Science* **124**, 722–723 (1956).
10. D. MAZIA: Materials for the biophysical and biochemical study of cell division. *Advances in Biological and Medical Physics* **4**, 69–118, Academic Press, New York (1956).
11. E. E. HOWE: Properties of amino acids. In *Amino Acids and Proteins*, ed. by D. M. Greenberg, pp. 13–55, Charles C. Thomas, Springfield, Ill. (1951).
12. E. WINDSOR: α-Amino adipic acid as a constituent of a corn protein. *J. Biol. Chem.* **192**, 595–606 (1951).
13. E. WORK and D. L. DEWEY: The distribution of α, ε diaminopimelic acid among various micro-organisms. *J. Gen. Microbiol.* **9**, 394–409 (1953).
14. H. SMITH, R. E. STRANGE, and H. T. ZWARTOUW: α, ε-Diaminopimelic acid in the peptide moiety of the cell wall polysaccharide of bacillus anthracis. *Nature, Lond.* **178**, 856–866 (1956).
15. M. FLAVIN: The linkage of phosphate to protein in pepsin and ovalbumin. *J. Biol. Chem.* **210**, 771–784 (1954).

16. F. R. BETTELHEIM: Tyrosine-*o*-sulfate in a peptide from fibrinogen. *J. Amer. Chem. Soc.* **76**, 2838–2839 (1954).
17. J. M. BARRY: Use of glutamine by the mammary gland for the synthesis of casein. *Nature, Lond.* **174**, 315–316 (1954).
18. F. M. SINEX and D. D. VAN SLYKE: The source and state of the hydroxylysine of collagen. *J. Biol. Chem.* **216**, 245–250 (1955).
19. F. M. SINEX: personal communication.
20. M. R. STETTEN: Some aspects of the metabolism of hydroxyproline, studied with the aid of isotopic nitrogen. *J. Biol. Chem.* **181**, 31–37 (1949).
21. G. WOLF, W. W. HECK, and J. C. LEAK: The metabolism of hydroxyproline-α-C^{14} in the intact rat. Radioactivity in amino acids from proteins. *J. Biol. Chem.* **223**, 95–105 (1956).
22. G. BURNETT and E. P. KENNEDY: The enzymatic phosphorylation of proteins. *J. Biol. Chem.* **211**, 969–979 (1954).
23. J. ROCHE and R. MICHEL: Thyroid hormones and iodine metabolism. *Annu. Rev. Biochem.* **23**, 481–500 (1954).
24. M. FLAVIN: The linkage of phosphate to protein in pepsin and ovalbumin. *J. Biol. Chem.* **210**, 771–784 (1954).
25. A. RICH and F. H. C. CRICK: The structure of collagen. *Nature, Lond.* **176**, 915–916 (1955).
26. H. A. ITANO: The hemoglobins. *Annu. Rev. Biochem.* **25**, 331–348 (1956).
27. T. H. J. HUISMAN, J. H. P. JONIX, and P. C. VAN DER SCHAAF: Amino acid composition of four different kinds of human haemoglobin. *Nature, Lond.* **175**, 902–903 (1955).
28. V. M. INGRAM: A specific chemical difference between the globins of normal human and sickle-cell anaemia haemoglobin. *Nature, Lond.* **178**, 792–794 (1956).
29. J. V. EVANS, J. W. B. KING, B. L. COHEN, H. HARRIS, and F. L. WARREN: Genetics of haemoglobin and blood potassium differences in sheep. *Nature, Lond.* **178**, 849–850 (1956).
30. D. W. GREEN, A. C. T. NORTH, and R. ASCHAFFENBURG: Crystallography of the β-lactoglobulins of cow's milk. *Biochim. Biophys. Acta.* **21**, 583–585 (1956).
31. N. H. HOROWITZ and M. FLING: The role of the genes in the synthesis of enzymes. In *Enzymes: Units of Biological Structure and Function*, ed. by O. H. Gaebler, pp. 139–145 Academic Press, New York (1956).
32. W. K. MAAS and B. D. DAVIS: Production of an altered pantothenate-synthesizing enzyme by a temperature-sensitive mutant of *Eschericia coli*. *Proc. Nat. Acad. Sci., Wash.* **38**, 785–797 (1952).
33. T. L. DEICH and E. T. SORENI: Aminokontsevie gruppi gliadinov i ich izmenenie pod vlianiem mezhrodovoi gibridizatsii. *C. R. Acad. Sci. U.R.S.S.* **98**, 623–626 (1954).
34. J. I. HARRIS, F. SANGER, and M. A. NAUGHTON: Species differences in insulin. *Arch. Biochem. Biophys.* **65**, 427–438 (1956).
35. J. C. KENDREW, R. G. PARRISH, J. C. MARACK, and E. S. ORLANS: The species specificity of myoglobin. *Nature, Lond.* **174**, 946–949 (1954).
36. K. FELIX: Zur chemie des zelkerns. *Experientia* **8**, 312–317 (1952).
37. E. O. P. THOMPSON: The N-terminal sequence of serum albumins; observations on the thiohydantoin method. *J. Biol. Chem.* **208**, 565–572 (1954).
38. W. F. WHITE, J. SHIELDS, and K. C. ROBBINS: C-terminal sequence of crystalline bovine and human serum albumins: relationship of C-terminus to antigenic determinants of bovine serum albumin. *J. Amer. Chem. Soc.* **77**, 1267–1269 (1955).
39. H. TUPPY and S. PALEUS: Study of a peptic degradation product of cytochrome c. I. Purification and chemical composition. *Acta Chem. Scand.* **9**, 353–364 (1955).
40. V. DU VIGNEAUD, H. C. LAWLER, and E. A. POPENOE: Enzymatic cleavage of glycinamide from vasopressin and a proposed structure for this pressorantidiuretic hormone of the posterior pituitary. *J. Amer. Chem. Soc.* **75**, 1880–1881 (1953).
41. H. OZAWA and K. SATAKE: On the species difference of N-terminal amino acid sequence in hemoglobin I. *J. Biochem.* **42**, 641–648 (1955).

42. L. Lorand and W. R. Middlebrook: Species specificity of fibrinogen as revealed by end-group studies. *Science* **118**, 515–516 (1953).
43. P. H. Bell: Purification and structure of β-corticotropin. *J. Amer. Chem. Soc.* **76**, 5565–5567 (1954).
44. W. F. White and W. A. Landmann: Studies on adrenocorticotropin. XI. A preliminary comparison of corticotropin-A with β-corticotropin. *J. Amer. Chem. Soc.* **77**, 1711–1712 (1955).
45. C. H. Li, I. I. Geschwind, I. D. Raacke, J. I. Harris, and J. J. Dixon: Amino acid sequence of alpha-corticotropin. *Nature, Lond.* **176**, 687–689 (1955).
47. W. S. Peart: Composition of a hypertensin peptide. *Nature, Lond.* **177**, 132 (1956).
46. L. T. Skeggs, W. H. Marsh, J. R. Kahn, and N. P. Shumway: Amino acid composition and electrophoretic properties of hypertensin I. *J. Exp. Med.* **102**, 435–440 (1955).
48. C. I. Niu and H. Fraenkel-Conrat: C-terminal amino acid sequences of four strains of tobacco mosaic virus. *Arch. Biochem. Biophys.* **59**, 538–540 (1955).
49. S. Benzer: Fine structure of a genetic region in bacteriophage. *Proc. Nat. Acad. Sci., Wash.* **41**, 344–354 (1955).
50. G. Pontecorvo and J. H. Roper: Resolving power of genetic analysis. *Nature, Lond.* **178**, 83–84 (1956).
51. H. Fraenkel-Conrat: The chemistry of proteins and peptides. *Annu. Rev. Biochem.* **25**, 291–330 (1956).
52. R. H. Locker: C-terminal groups in myosin, tropomyosin and actin. *Biochim. Biophys. Acta.* **14**, 533–542 (1954).
53. E. O. P. Thompson: The N-terminal sequence of carboxypeptidase. *Biochim. Biophys. Acta.* **10**, 633–634 (1953).
54. T. Posternak and H. Pollaczek: De la protection contre l'hydrolyse enzymatique exercée par les groupes phosphoryles. Étude de la dégradation enzymatique d'un peptide et d'un polyose phosphoryles. *Helv. Chim. Acta.* **24**, 921–930 (1941).
55. N. Seno, K. Murai, and K. Shimura: Studies on the N-terminal lysylpeptides of α-casein. *J. Biochem.* **42**, 699–704 (1955).
56. H. Neurath and W. J. Dreyer: The activation of chymotrypsinogen. Isolation and identification of a peptide liberated during activation. *J. Biol. Chem.* **217**, 527–539 (1955).
57. F. Turba and G. Gundalch: Aminosäure sequenz in der umgebung des reaktiven serinrestes in chymotrypsin-molekül. *Biochem. Z.* **327**, 186–188 (1955).
58. (This article consists of paragraphs by different authors) Erweitertes makromolekulares kolloquium. *Angew. Chem.* **65**, 349–352 (1953).
59. K. Felix, R. Hirohata, and K. Dirr: Über clupein. *Hoppe-Seyl. Z.* **218**, 269–279 (1933).
60. W. A. Schroeder, L. M. Kay, J. LeGette, L. Honnen, and F. C. Green: The constitution of gelatin. Separation and estimation of peptides in partial hydrolysates. *J. Amer. Chem. Soc.* **76**, 3556–3564 (1954).
61. T. D. Kroner, W. Tabroff, and J. J. McGarr: Peptides isolated from a partial hydrolysate of steer hide collagen. II. Evidence for the prolylhydroxyproline linkage of collagen. *J. Amer. Chem. Soc.* **77**, 3356–3359 (1955).
62. R. R. Porter: A chemical study of rabbit antiovalbumin. *Biochem. J.* **46**, 473–478 (1950).
63. W. W. Bromer, L. G. Sinn, A. Staub, and O. K. Behrens: The amino acid sequence of glucagon. *J. Amer. Chem. Soc.* **78**, 3858–3860 (1956).
64. C. I. Niu and H. Fraenkel-Conrat: Determination of C-terminal amino acids and peptides by hydrazinolysis. *J. Amer. Chem. Soc.* **77**, 5882–5885 (1955).
65. A. R. Thompson: Amino acid sequence in lysozyme. 2. Elution chromatography of peptides on ion-exchange resins. *Biochem. J.* **61**, 253–263 (1955).
66. R. Acher, U. R. Laurila, and C. Fromageot: Contribution à l'étude de la structure du lysozyme d'oeuf de poule. Peptide aromatique obtenue par hydrolyse enzymatique. *Biochim. Biophys. Acta* **19**, 97–109 (1956).

67. K. OHNO: On the structure of lysozyme. III. On the carboxyl-terminal peptide. *J. Biochem.* **42,** 615–625 (1955).
68. K. OHNO: On the structure of lysozyme II. Characterization of aspartyl, asparaginyl, and glutaminyl residues in lysozyme. *J. Biochem.* **41,** 345–350 (1954).
69. I. I. GESCHWIND, C. H. LI, and L. BARNAFI: Isolation and structure of melanocyte-stimulating hormone from porcine pituitary glands. *J. Amer. Chem. Soc.* **78,** 4494–4495 (1956).
70. J. I. HARRIS and P. ROOS: Amino-acid sequence of a melanophore stimulating peptide. *Nature, Lond.* **178,** 90 (1956).
71. V. M. INGRAM: The application of Edman's peptide degradation method to horse myoglobin and haemoglobin. *Biochim. Biophys. Acta* **16,** 599–600 (1955).
72. M. FLAVIN and C. B. ANFINSEN: The isolation and characterization of cysteic acid peptides in studies on ovalbumin synthesis. *J. Biol. Chem.* **211,** 375–390 (1954).
73. M. OTESSEN and A. WOLLENBERG: Stepwise degradation of the peptides liberated in the transformation of ovalbumin to plakelbumin. *C. R. Lab. Carlsberg (Chim.)* **28,** 463–475 (1953).
74. M. FLAVIN: Cysteine and phosphoserine containing peptide sequences of ovalbumin. *Nature, Lond.* **173,** 214 (1954).
75. V. DU VIGNEAUD, C. RESSLER, and S. TRIPPETT: The sequence of amino acids in oxytocin, with a proposal for the structure of oxytocin. *J. Biol. Chem.* **205,** 949–957 (1953).
76. E. O. P. THOMPSON: Crystalline papain. IV. Free amino groups and N-terminal sequence. *J. Biol. Chem.* **207,** 563–574 (1954).
77. M. B. WILLIAMSON and J. M. PASSMAN: The amino acid sequence at the N terminus of pepsin. *J. Biol. Chem.* **222,** 151–157 (1956).
78. D. COLE and C. H. LI: N-terminal sequence of prolactin. *Fed. Proc.* **14,** 195 (1955).
79. C. H. W. HINS, W. H. STEIN, and S. MOORE: Peptides obtained by chymotryptic hydrolysis of performic acid-oxidized ribonuclease. A partial structural formula for the oxidized protein. *J. Biol. Chem.* **221,** 151–169 (1956).
80. R. MONIER and M. JUTISZ: Contribution a l'étude de la structure de la salmine d'Oncorhyncus. *Biochim. Biophys. Acta* **14,** 551-558 (1954).
81. R. MONIER and M. JUTISZ: Contribution a l'étude de la structure de la salmine d'Onchorhynchus. II. Étude de quelques peptides résultant de l'hydrolyse trypsique. *Biochim. Biophys. Acta* **15,** 62–68 (1955).
82. F. LUCAS, J. T. B. SHAW, and S. G. SMITH: Amino-acid sequence in a fraction of Bombyx silk fibroin. *Nature, Lond.* **178,** 861 (1956).
83. L. M. KAY and W. A. SCHROEDER: The chromatographic separation and identification of some peptides in partial hydrolysates of silk fibroin. *J. Amer. Chem. Soc.* **76,** 3564-3568 (1954).
84. E. ABDERHALDEN and A. BAHN: Isolierung von tyrosyl-seryl-prolyl-tyrosin beim stufenweisen abbau von seidenfibroin (Bombyx Mori). *Hoppe-Seyl. Z.* **219,** 72–81 (1933).
85. E. W. DAVIE and H. NEURATH: Identification of a peptide released during autocatalytic activation of trypsinogen. *J. Biol. Chem.* **212,** 515–529 (1955).
86. R. CONSDEN and A. H. GORDON: A study of the peptides of cystine in partial hydrolysates of wool. *Biochem. J.* **46,** 8–20 (1950).
87. A. POLSON: Quantitative partition chromatography and the composition of *E. coli. Biochim. Biophys. Acta* **2,** 575–581 (1948).
88. R. B. ROBERTS, D. B. COWIE, P. H. ABELSON, E. T. BOLTON, and R. J. BRITTEN: Studies of biosynthesis in *Escherichia coli.* Carnegie Institution of Washington Publication **607,** p. 28, Washington, D.C. (1955).
89. D. ELSON and E. CHARGAFF: Evidence of common regularities in the composition of pentose nucleic acids. *Biochim. Biophys. Acta* **17,** 367–376 (1955).
90. F. FRIEDBERG: The amino acid composition of adenosine triphosphate-creatine transphosphorylase. *Arch. Biochem. Biophys.* **61,** 263–266 (1956).

91. K. LAKI, D. R. KOMINTZ, P. SYMONDS, L. LORAND, and W. H. SEEGERS: The amino acid composition of bovine prothrombin. *Arch. Biochem. Biophys.* **49,** 276–282 (1954).
92. D. R. KOMINTZ, A. HOUGH, P. SYMONDS, and K. LAKI: The amino acid composition of actin, myosin, tropomyosin and the meromyosins. *Arch. Biochem. Biophys.* **50,** 148–159 (1954).
93. A. ROSSI-FANELLI, D. CAVALLINI, and L. DE MARKO: Amino acid composition of human crystallized myoglobin and hemoglobin. *Biochim. Biophys. Acta* **17,** 377–381 (1955).
94. G. JOLLES and C. FROMAGEOT: La proteine lysante II de la rate du lapin. II. Composition en acides amines. *Biochim. Biophys. Acta* **14,** 219–227 (1954).
95. W. G. GORDON and J. ZIEGLER: Amino acid composition of crystalline α-lactalbumin. *Arch. Biochem. Biophys.* **57,** 80–86 (1955).
96. K. LANGE: Aminosäurezusammensetzung kristallisierte alkohol-dehydrogenase aus bäckerhefe. *Hoppe-Seyl. Z.* **303,** 272–275 (1956).
97. J. MUUS: The amino acid composition of human salivary amylase. *J. Amer. Chem. Soc.* **76,** 5163–5165 (1954).
98. W. G. GORDEN, W. F. SEMMETT, and M. BENDER: Amino acid composition of γ-casein. *J. Amer. Chem. Soc.* **75,** 1678–1679 (1953).
99. E. L. SMITH, A. STOCKELL, and J. R. KIMMEL: Crystalline papain. III. Amino acid composition. *J. Biol. Chem.* **207,** 551–561 (1954).
100. E. L. SMITH, M. L. MCFADDEN, A. STOCKELL, and V. BUETTNER-JANUSCH: Amino acid composition of four rabbit antibodies. *J. Biol. Chem.* **214,** 197–207 (1955).
101. E. L. SMITH, J. R. KIMMEL, D. M. BROWN, and E. O. P. THOMPSON: Isolation and properties of a crystalline mercury derivative of a lysozyme from papaya latex. *J. Biol. Chem.* **215,** 67–89 (1955).
102. E. L. SMITH and A. STOCKELL: Amino acid composition of crystalline carboxypeptidase. *J. Biol. Chem.* **207,** 501–514 (1954).
103. S. F. VELICK and E. RONZONI: The amino acid composition of aldolase and d-glyceraldehyde phosphate dehydrogenase. *J. Biol. Chem.* **173,** 627–639 (1948).
104. B. A. KOECHLIN and H. D. PARISH: The amino acid composition of a protein isolated from lobster nerve. *J. Biol. Chem.* **205,** 597–604 (1953).
105. E. L. SMITH, D. M. BROWN, M. L. MCFADDEN, V. BUETTNER-JANUSCH, and B. U. JAGER: Physical, chemical and immunological studies on globulin from multiple myeloma. *J. Biol. Chem.* **216,** 601–620 (1955).
106. W. H. STEIN and S. MOORE: Amino acid composition of lactoglobulin and bovine serum albumin. *J. Biol. Chem.* **178,** 79–91 (1949).
107. J. F. VELICK and L. F. WICKS: The amino acid composition of phosphorylase. *J. Biol. Chem.* **190,** 741–751 (1951).
108. R. N. FEINSTEIN, M. HAMPTON, and G. J. COTTER: Species specificity of catalase. *Enzymologia* **16,** 219–225 (1953).
109. I. BATTY, F. W. R. BRAMBELL, W. A. HEMMINGS, and C. L. OAKLEY: Selection of antitoxins by the foetal membranes of rabbits. *Proc. Roy. Soc. B* **142,** 452–471 (1954).
110. B. HAGIHARA, T. HORIO, M. NOZAKI, I. SEKUZU, J. YAMASHITA, and K. OKUNUKI: Comparison of properties of crystalline cytochrome c from yeast, beef heart and pig heart. *Nature, Lond.* **178,** 631–632 (1956).
111. F. LUCAS, J. T. B. SHAW, and S. G. SMITH: The chemical constitution of some silk fibroins and its bearing on their physical properties. *Shirley Institute Memoirs* **28,** 77–89 (1955).
112. S. W. FOX and P. G. HOMEYER: A statistical evaluation of the kinship of protein molecules. *Amer. Nat.* **89,** 163–168 (1955).
113. M. YČAS: Numerology of peptide chains. *Naturwissenschaften* **43,** 197–198 (1956).
114. B. J. BENFREY and J. L. PURVIS: Purification and amino acid analysis of melanophor-expanding hormone from hog and ox pituitary glands. *Biochem. J.* **62,** 588–593 (1956).

115. R. M. S. SMELLIE: The metabolism of the nucleic acids. *The Nucleic Acids*, ed. by E. Chargaff and J. N. Davison, vol. 2, pp. 393–434, Academic Press, New York (1955).
116. J. BRACHET: Action of ribonuclease and ribonucleic acid on living amoebae. *Nature, Lond.* **175,** 851–853 (1955).
117. L. GOLDSTEIN and W. PLAUT: Direct evidence for nuclear synthesis of cytoplasmic ribose nucleic acid. *Proc. Nat. Acad. Sci., Wash.* **41,** 874–880 (1955).
118. K. LARK and O. MAALOE: Nucleic acid synthesis and the division cycle of *Salmonella typhimurium. Biochim. Biophys. Acta* **21,** 448–458 (1956).
119. B. MAGASANIK: Isolation and composition of the pentose nucleic acids and the corresponding nucleo-proteins. *The Nucleic Acids*, ed. by E. Chargaff and J. N. Davidson, vol. 1, pp. 373–407. Academic Press, New York (1955).
120. J. BRACHET and D. SZAFARZ: L'incorporation d'acide orotique radioactif dans des fragments nucléés et anucléés *d'Acetabularia mediterranea. Biochim Biophys. Acta* **12,** 588–589 (1953).
121. L. S. LOCKINGEN and A. G. DEBUSK: A model for intracellular transfer of DNA (gene) specificity. *Proc. Nat. Acad. Sci., Wash.* **41,** 925–934 (1955).
122. A. B. PARDEE and L. S. PRESTIDGE: The dependence of nucleic acid synthesis on the presence of amino acids in *Escherichia coli. J. Bact.* **71,** 677–683 (1956).
123. F. GROS and F. GROS: Rôle des aminoacides dans la synthese des acides nucléiques chez *Escherichia coli. Biochim. Biophys. Acta* **22,** 200–201 (1956).
124. M. YČAS and G. BRAWERMAN: Interrelations between nucleic acid and protein biosynthesis in micro organisms. *Arch. Biochem. Biophys.* **68,** 118–129 (1957)
125. J. D. WATSON and F. H. C. CRICK: Genetical implications of the structure of deoxyribose nucleic acid. *Nature, Lond.* **171,** 964–967 (1953).
126. F. L. BLACK and C. A. KNIGHT: A comparison of some mutants of tobacco mosaic virus. *J. Biol. Chem.* **202,** 51–57 (1953).
127. D. DE FREMERY and C. A. KNIGHT: A chemical comparison of three strains of tomato bushy stunt virus. *J. Biol. Chem.* **214,** 559–566 (1955).
128. E. ROBERTS and G. B. RAMASARMA: Amino acids of turnip yellow virus. *Proc. Soc. Exp. Biol. Med.* **80,** 101–103 (1952).
129. B. S. MAGDOFF R. J. BLOCK, and D. B. MONTIE: Amino acid composition of southern bean mosaic virus. *Contr. Boyce Thompson Inst.* **18,** 371–375 (1956).
130. C. A. KNIGHT: Amino acid composition of highly purified viral particles of influenza A and B. *J. Exp. Med.* **86,** 125–129 (1947).
131. R. W. DORNER and C. A. KNIGHT: The preparation and properties of some plant virus nucleic acids. *J. Biol. Chem.* **205,** 959–967 (1953).
132. R. MARKHAM and J. D. SMITH: Chromatographic studies of nucleic acids. 4. The nucleic acid of the turnip yellow mosaic virus, including a note on the nucleic acid of the tomato bushy stunt virus. *Biochem. J.* **49,** 401–406 (1951).
133. G. L. ADA and B. T. PERRY: Specific differences in the nucleic acids from A and B strains of influenza virus. *Nature, Lond.* **175,** 854 (1955).
134. G. GAMOW: Possible mathematical relation between deoxyribonucleic acid and protein. *Dansk. Biol. Medd.* **22,** No. 3 (1954).
135. F. H. C. CRICK, J. S. GRIFFITH, and L. E. ORGEL: Codes without commas. *Proc. Nat. Acad. Sci., Wash.* **43,** 416–421 (1957).
136. H. A. SIMON: personal communication.
137. I. HARRIS and C. A. KNIGHT: Studies on the action of carboxypeptidase on tobacco mosaic virus. *J. Biol. Chem.* **214,** 215–230 (1955).
138. T. L. DEǏCH and E. T. SZORÉNYI: *Chem. Abstr.* **49,** 1882 (1955).
139. ZOTÁN KOROS: Free amino groups of gliadin. *Magyar Kim. Folyoirat* **56,** 131–136 (1950).
140. L. K. RAMACHANDRAN and W. B. MCCONNELL: The terminal amino acids of wheat gliadin. *Canad. J. Chem.* **33,** 1463–1466 (1955).

141. T. DEUTSCH: N-terminal amino acids of gliadins from wheat and rye. *Acta Physiol. Acad. Sci. Hung.* **6**, 209–224 (1954).
142. B. S. SCHWEIGERT, B. T. GUTHNECK, J. M. PRICE, J. A. MILLER, and E. C. MILLER: Amino acid composition of morphological fractions of rat livers and induced liver tumors. *Proc. Soc. Exp. Biol. Med.* **72**, 495–501 (1949).
143. G. ROST: Zusammensetzung der ribonucleinsäure der reticulozyten. *Naturwissenschaften* **43**, 499 (1956).

DISCUSSION

KOCH: I should like to comment on the result of some recent tracer experiments that have been conducted in Dr SWICK's laboratory at the Argonne National Laboratory (1, 2, 3). What we have tried to do is to ask ourselves something about the total balance of the turnover of RNA, DNA, and protein in the tissue which is most often studied by the biochemist; namely, rat liver. The interesting thing that comes out of this is that when suitable tracer experiments are done, you can make the definite statement that in a single cell DNA is synthesized when it is produced and DNA stays as a cell compound until the death of the cell, whereas on the other hand it is very easy to show that all of the RNA in the cell is turned over, and it is turned over essentially with about the same half-life that all of the proteins are turned over in the cell; that is, there are no special classes of proteins that are not turned over, especially classes of RNA that are not turned over in this tissue.

The immediate conclusion from this is that, inasmuch as the amount of protein is many times more than the amount of RNA, on a molar or other basis, there can be no one-to-one hand-off of this kind. In other words, you cannot take the DNA and make the RNA from it without using it over and over again in a different way than has been suggested here.

YČAS: While it may be true that there is turnover of RNA in rat liver, I believe, on the basis of work with micro-organisms, that there is no obligatory turnover of RNA associated with protein synthesis. The RNA, which is part of the protein forming mechanism, is a passive template, and apparent coupling or dissociation of protein and RNA turnover is adequately explained, I think, by the assumption that both have common precursors.

KOCH: I would just like to add that in the case of micro-organisms it is fairly clear that the protein turnover does not occur (4). It is also pretty well established that DNA and RNA turnover do not occur in an actively growing culture. So the concept of turnover in the micro-organism is not a relevant one. But what it does mean is that you cannot accept some of the proposals that have been described that inherently require the obligatory breakdown of something (RNA), concomitant to the synthesis of another type of molecule (protein).

MOROWITZ: I would like to introduce some evidence for an alternative approach to the problem of intersymbol influence. In some work recently published by SIDNEY FOX (5) analyses are reported on the total protein of soybean, corn, wheat, and rye. These analyses indicate that a very high proportion of the protein molecules have lysine in an N-terminal position and arginine in the next position. This approach to statistical constraints involves an experimental analysis of a population of proteins from a single source as contrasted to Dr YČAS' theoretical analysis of a population of unrelated proteins.

We have attempted to determine if any constraints are to be found in *E. coli* protein. The preliminary results indicate that methionine is found in N-terminal positions in a proportion consistent with a chance distribution. Cystine and cysteine in N-terminal positions may show a considerably greater constraint.

YČAS: I think that the method used by FOX and yourself introduces an obvious source of bias, if what you are trying to do is look for intersymbol correlations. The abundances of different species of protein in a cell are not equal, and more abundant proteins contribute more end groups. You have to examine the proteins one by one, giving the same statistical weight to each.

A similarity in end groups of proteins from related species indicates not an effect of intersymbol correlation, but rather descent from a common ancestor. As can be seen from the data I summarized, proteins change only slowly in evolution.

BRANSON: There is one question which has been opened up by Dr GAMOW's and Dr YČAS' comments; namely, the whole problem of redundancy in protein molecules. The evidence is fairly conclusive, I believe, that so far as the antigenic action of a protein is concerned, the

active region is approximately 15 Å on a side. If the same is true of other biological functions, a great deal of surface area in a protein is passive. At least it is passive for a given specific function. Thus it is reasonable to inquire how much of a protein molecule you can whittle away and keep a given biological property.

There is a fairly convincing teleological explanation for this redundancy. In the early history of living systems, the membranes containing the living material might have been rather leaky. Thus to retain the small biologically-active components within the cell, they had to be associated with a large but inactive structure which would not pass out through the large spaces. In the evolutionary scheme, then, there remain many large units where really the functional part is relatively small. So that when one amino acid is taken out and another put in, the substitution does not make much difference so long as it is not in the essential small functioning unit of the protein molecule.

YČAS: I am also of the opinion that mere size of an enzyme may be quite important for the totality of its biological functions, even if it seems to make no difference to the catalytic function as measured in a test tube. Which part of a protein is significant and which is not is a matter of what function we are measuring. I doubt that at present we know all the functions of a protein from the point of view of the organism itself.

REFERENCES

1. R. W. SWICK and D. T. HANDA: The distribution of fixed carbon in amino acids. *J. Biol. Chem.* **218**, 557 (1956).
2. R. W. SWICK, A. L. KOCH, and D. T. HANDA: The measurement of nucleic acid turnover in rat liver. *Arch. Biochem. Biophys.* **63**, 226–242 (1956).
3. R. W. SWICK and A. L. KOCH: The measurement of nucleic acid phosphorus turnover in rat liver by the constant exposure technique. *Arch. Biochem. Biophys.* **67**, 59–73 (1957).
4. A. L. KOCH and H. R. LEVY: Protein turnover in growing cultures of *Escherichia coli. J. Biol. Chem.* **217**, 947–951 (1955).
5. S. FOX: Evolution of protein molecules and thermal synthesis of biochemical substances. *Amer. Scient.* **44**, 347–359 (1956).

PROTEIN STRUCTURE AND INFORMATION CONTENT*

L. G. Augenstine
Brookhaven National Laboratory, Upton, New York

I. INTRODUCTION

In stating that a given system has an information content of a certain number of bits, care must be taken to specify not only the context within which this number has been derived but also an attempt must be made to give meaning and utility to this measure. Specifying the context is particularly important since for most systems there are many levels at which the information content can be derived. For example, the information content for a cell is very low, if one is concerned only whether it is living or dead, but it is very large if one is interested in specifying the parameters of each of its individual elementary particles. In this article, estimates will be made of the information content of given proteins by taking into account that they are a sequence of amino acids which can assume only a discrete number of configurations. An attempt will be made to study some of the factors which affect the information content and the types of constraints which must operate in the elaboration of proteins. Some idea of the magnitude and types of the constraints pertinent to proteins can be obtained from parallel studies on proteins and printed English (for which the constraints are known). Finally, the information content based upon structure will be compared with estimates of information content obtained within the context of protein function.

Although the fact has not always been fully appreciated, information measures are usually more effective in selecting among alternative hypotheses than in suggesting new ones. This particular trait arises from the fact that information estimates, which depend only upon the probabilities associated with a class of experimental outcomes, will often describe the degree to which a number of variables interact but indicate little or nothing about the behavior of the individual variables. As a result no novel synthetic procedures or selection principles are advanced here to explain the manner in which polypeptide sequences and/or configurations are determined. Rather, in this paper information theory considerations have been used to evaluate alternative explanations of some aspects of protein construction.

II. ESTIMATION OF STRUCTURAL INFORMATION CONTENT AND CONSTRAINTS

At the structural level the total information content (I_t) of a protein will be treated as the sum of two terms; one (I_s) depends upon the amino acid sequence

* Research carried out at Brookhaven National Laboratory under the auspices of the U.S. Atomic Energy Commission.

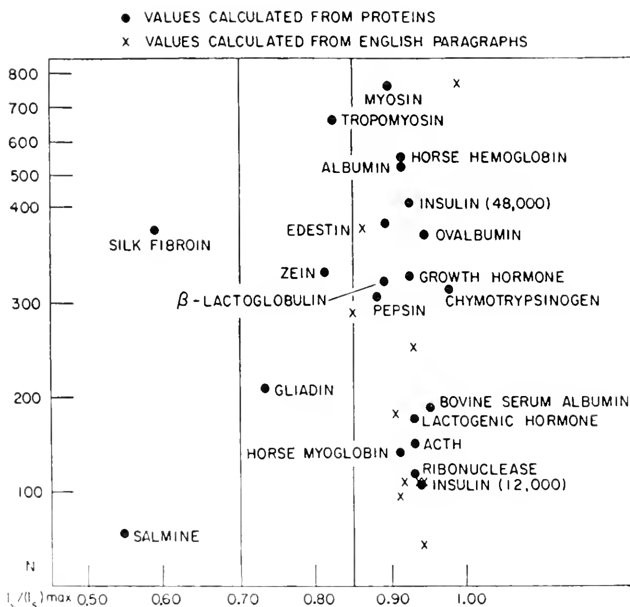


FIG. 1. Values of $I_s/(I_s)_{\max}$ as a function of the number of symbols, N in proteins and paragraphs.

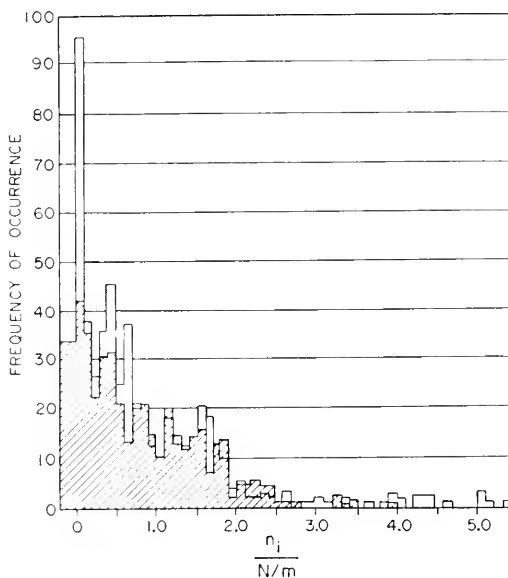


FIG. 2. Distribution of the normalized frequency, $\frac{n_i}{N/m}$ of letters and amino acids in the language and protein samples. See the text for further discussion.

and the other (I_c) upon the configurations of the polypeptide chain in the native molecule. Treating sequence and configuration independently should lead to overestimates of I_t, since the permissible configurations will depend upon the sequence. However, care has been taken to reduce the interaction of the two terms as much as possible, so that for the purposes of this paper no significant discrepancies should occur.

$I_{sequence}$: There are twenty amino acids which are most commonly incorporated into proteins. Therefore the maximum value of I_s is 4.32 bits ($\log_2 20$) per amino acid residue.* It would occur when the twenty amino acids occur equiprobably. Values less than the maximum would occur due to any constraints upon the amino acid sequence. Branson (1) calculated I_s of twenty-six proteins for which the frequency of occurrence of the twenty amino acids had been determined (disregarding possible sequential dependencies). He found that those which formed part of a living structure of an organism had an I_s which was greater than 0.70 of the maximum value. His analysis is shown by the dots in Fig. 1. The X's show the result of a similar analysis on language samples. The language study was based on ten paragraphs chosen from diverse sources such as want ads, newspaper articles, textbooks, and magazines and differs from that usually used in analysis of language in that it is based on the paragraph rather than on large continuous samples.† In this case, letters have been treated like amino acids and paragraphs like proteins. Except for the single value of 0.99 the values from proteins and paragraphs agree quite well.

Similarities between the distribution of amino acid frequencies and letters can be seen further in Fig. 2. There the ordinate indicates the number of times that a particular normalized frequency occurs; the normalized frequency is the number of times, n_i, that the ith symbol (either amino acid or letter) occurs, divided by N/m, the expected number of times that each type of symbol should occur if all m different kinds of symbols had equiprobable occurrence in the sample of N symbols. As can be seen in Fig. 2 the distribution of the normalized frequencies $\frac{n_i}{N/m}$ for the letters (solid line) and the amino acids (shaded area) are almost identical except for the higher incidence of rarely-used letters in language. This small difference might not have occurred if some of the rarer amino acids, for which assays are difficult, had been included in the data.

Constraints—The fact that the distribution of amino acids in non-structural proteins deviates from equiprobability about the same as (or possibly a little less than) the letters in written English, indicates that the constraints producing such unequal frequencies should be of the same order of magnitude as (or slightly less than) those governing English texts. However, this tells nothing about the

* This value disregards any influence of residue 'complexions'. However, it is difficult to see how factors other than the identity of the residues can be very important, when one considers the freedom of rotation of the R-groups with respect to the polypeptide chain.

† It was felt that such a small-sample statistics study was preferable to one based upon large samples (such as a determination of confidence intervals for I_s as a function of the paragraph size), since by essentially duplicating the analyses applied to proteins, insightas to the limitations of that procedure could be observed.

nature of the constraints or the manner in which they arise. The obvious question arises—is the unequal distribution due to unequal availability of the amino acids or is it due to constraints imposed in the processes of synthesis, i.e. by 'intersymbol influence'?* Is the make-up of the pool of amino acids available to the protein-synthesizing centers indicative of the nature of the processes involved in amino acid synthesis or have these processes become adapted to the peculiar demands of the proteins being synthesized? This is essentially the same as looking at a collection of printer's type and asking the question, did the printer select his supply of type because this particular distribution of letters was all that was available to him or did he purposely purchase his particular assortment because he had found that it satisfied his needs?

The possibility that the unequal availability of amino acids in the cellular pool may produce the unequal distribution does not seem likely. The experiments of ROBERTS, COWIE *et al.* (2, 3) at the Carnegie Institution indicate that it requires a five to thirty-fold excess of exogenous amino acids, such as valine, leucine and isoleucine, before the incorporation of these amino acids into protein is seriously affected in *E. coli.* In fact, once a substance has been incorporated into the amino acid pool of yeast, 1000 times the normal concentration of exogenous amino acid does not affect its incorporation into protein (COWIE). Although these are excellent experiments they do suffer from problems of cell membrane permeability, intracellular diffusion, etc.; however, they, along with numerous experiments involving amino acid deficient mutants, suggest that as long as the minimum required amount of each amino acid is present the frequency distribution of the amino acids in the pool has a relatively small influence on the distribution of amino acids incorporated into protein.

Two methods have been utilized in searching for intersymbol influence in proteins. In the first (reported previously (4)), the behavior of the normalized amino acid frequencies $\frac{n_i}{N/m}$ were studied in individual proteins. The average normalized frequency of the individual amino acids for the twenty-six proteins was tabulated. Comparing the normalized frequency for the individual amino acids in *particular* proteins with the corresponding average value from the 26 proteins indicated large deviations in many cases. The gross deviations were examined for correlations between pairs of amino acids, both for positive and negative effects. Examination of the 26 proteins indicated that although there are some correlations between the frequencies of individual amino acids combined in single proteins, none was strong enough to be measurable with any degree of confidence for a sample as small as 26 proteins.

Similar examinations of the normalized letter frequencies in paragraphs were investigated for significant deviations of pairs or groups of letters. Although strong intersymbol influences are known to exist between letters (e.g. between

* 'Intersymbol influence' is a term commonly used to designate sequential dependencies, i.e. influences upon the identity of a particular element by neighbouring elements, which are not the only types of constraints which might be imposed by a synthesizing center. It is easy to imagine the possibility of unequal 'acceptability' for different symbols at individual sites on a template in which the factors affecting the specifications of each location are independent of the neighbors.

q and u) no significant results were detected. Thus it can be concluded that such analyses do not exclude intersymbol influences of the same type or order of magnitude as those in language.*

Gamow, Rich, and Yčas (5) have made a more exacting study of possible inter-symbol influences affecting amino acids. They treated the known amino acids as a series of dipeptides which they tallied into a 20 × 20 matrix similar to the 26 × 26 digram matrices common in language analyses. The distribution for nonstructural proteins in such a 20 × 20 matrix followed quite closely a Poisson distribution. This they state is compatible with the assumption that the occurrence of a given amino acid does not affect the identity of its nearest neighbor. Their comparable analysis for English language gave a distribution which deviated from a Poisson.

The Poisson distribution associated with the amino acid dipeptide analysis is not too significant since the sample of experimentally determined sequences is not necessarily a reliable representation of the bulk of amino acid sequences in nature. As Gamow, Rich, and Yčas point out, their available sample is strongly affected by the composition of ACTH, lysozyme and insulin for which the complete sequences have been determined and the shorter sequences from other proteins are biased due to differential bond labilities within the protein which give rise preferentially to certain amino acids occurring as terminal peptides in the sequences isolated.

It was felt that a possible explanation of the difference noted between digram analysis of letters and amino acids was that amino acids were also grouped into word-like structures but that the average number of symbols per 'word' was different than that found in English. Therefore, separate digram analyses were performed on English words having two to five letters, six to nine letters and those having ten or more. All the samples were selected so that the average cell density in the 26 × 26 matrix was 0.44, the same as that of Gamow, Rich, and Yčas, and these also all showed significant deviations from a Poisson distribution.

Morowitz (6) and some of the Biophysics group at Yale have been investigating the possibility that a polypeptide chain is a segment selected from either a single or a small number of repeating sequences which are invariant for a given chromosomal complement. The particular segments chosen and the unique fashion in which they are combined and folded would then account for the highly specific properties of the individual proteins. The possibility also exists that there was an initial long, or at least restricted, set of sequences from which present day polypeptide sequences have evolved in a manner similar to that by which organisms have evolved. Gamow, Rich, and Yčas (5) have pointed out the most striking evidence for a "phylogenetically common ancestral sequence" in their comparison of the A and B chains of insulin, where the same amino acids occur in equivalent positions in both chains four times.

The known sequences containing five amino acids or more (from Table I, ref. 5) were examined for repeating or matching sequences. (This was done by superposing the sequences in all possible permutations.) These data indicate that for proteins from a given species any single repeating sequence must

* See the discussion by Dr Platt at the end of this paper.

be at least forty amino acid residues or longer. Comparing the sequences of different types of proteins indicated that (a) there is not a master sequence operating among species, or (b) evolution, i.e. amino acid substitution, has been so extensive as to make it undetectable, or (c) the master sequence is 200 residues or longer. The additional sequences (for hormones of sub-protein size) cited by YČAS (7) show that short polypeptide sequences with only minor amino acid differences do occur in cells of different species. Thus, the occurrence of repeating or a restricted number of amino acid sequences may be an explanation of the unequal amino acid frequencies observed.

This possible restriction provides a basis for estimating the minimum value of I_s. A single, long, completely-determined sequence would provide a situation of minimum information content for polypeptides selected from it. To select N residues from a sequence of S amino acids would require $\leq \log_2 S$ bits to find N and $\leq \log_2 (S - N)$ bits to determine the starting point; or by another selection procedure, $\leq \log_2 (S - 1)$ to find the starting point and roughly $\log_2 S/2$ to determine the end point. Either of these methods of selection gives an estimate of the minimum of I_s which is of the order of $2 \log_2 S$ bits. This is a very low minimum since according to the best present estimate (which is obviously too low) $S \approx 200$ and thus $2 \log_2 S \approx 15$. Therefore, the minimum of I_s is of the order of 0.1 bit/residue since $N \geq 100$ for proteins. Even if S is found to be 10^6, $(2 \log_2 S)/N$ will still only be ~ 0.4 bits/residue. Thus, the search (6) for long master sequences of amino acids is of considerable interest with respect to information content considerations.

Summarizing for I_s, we can say that for nonstructural proteins the potential information due to the amino acid sequence should be of the order of 0.85–0.95 of the possible maximum value. Although the constraints necessary to produce such an effect should be of the same order of magnitude as those in printed English, tests comparing language and the available proteins for which amino acid composition or sequences are known indicate that the constraints operating in the elaboration of proteins are probably different from those associated with language. Further, it seems unlikely that the unequal frequency of amino acids in proteins is due to unequal availability of the amino acids in the cellular pool. The possibility that polypeptide chains are segments selected from a single or restricted number of repeating sequences may be an explanation of the unequal frequencies, in which case I_s/residue would be close to zero.

$I_{configuration}$: With the present state of knowledge the factors affecting I_c are much more difficult to assess. The number of states available to a polypeptide chain whose bonds retained all of the lability they had as uncombined amino acids would be essentially innumerable. In fact, about the only configurations ruled out would be those resulting in closure of the chain upon itself. However the D- and L- forms do not both exist in nature and as has been pointed out by PAULING, COREY and BRANSON (8), the α-C, N and O group in the backbone of the polypeptide chain is essentially the planar, resonance

structure —C---N—. Other than these primary restrictions the polypeptide chain, in the absence of intramolecular or secondary bonding structures is essentially a random structure.

KAUZMANN (9) has given an excellent discussion of the known types of intramolecular bonds which are responsible for protein folding and which should therefore affect I_c. The most common type is the H—bond, especially those formed between the carboxyl O and the amide H. These are essentially non-specific bonds which can form between any pair of amino acid residues in which the C—O and N—H bonds are oriented at the proper angle. A stronger, more specific, but less common H—bond can form between the phenolic OH groups of tyrosine and the carboxyl group of glutamic or aspartic acid (9, 10). Another common type of bond stems from the van der Waals forces, which can exist between the atoms in different portions of the same or neighboring chains. The third type discussed by KAUZMANN is the so-called hydrophobic bond, which is distinct from the more commonly discussed van der Waals bonds. This results from the tendency of the more hydrophobic amino acid residues to avoid the aqueous phase and adhere together to form a sort of intramolecular micelle. These bonds, although they possess a low order of specificity, may contribute a good deal of stability since they arise as a result of the fact that the more hydrophobic amino acids cannot participate in the strong H-bonding with the solvent water molecules. Salt bridges, which are the ionic bonds formed between the negatively charged (glutamic and aspartic) and positively charged (lysine and argenine) residues, are another type. However, JACOBSEN and LINDERSTROM-LANG (11) have presented evidence which indicates that these bonds are of negligible importance as intramolecular protein bonds. One of the most important types of intramolecular bond (at least according to current theories (12)) is the highly specific S—S bond formed between cysteine residues in different portions of the same or neighboring chains. The formation of disulfide bonds as well as the 'strong' H-bonds greatly reduces the number of physical states available to the molecule since they can only be formed at a very few sites in the molecule. Since these two types of bond are the most specific of the intramolecular bonds, they are undoubtedly the most effective in determining variations in structure between different kinds of proteins.

Repetitious Structures: Intramolecular bonds formed in such a fashion as to produce repetitious structures reduce I_c tremendously. In the helical or pleated sheet structures proposed by PAULING, COREY and BRANSON (8) (and illustrated in (13)) the number of free parameters necessary to describe the configuration completely is extremely low and therefore the information content, I_c, is also very low. In the helices it is only necessary to specify the length (that is, the total number of residues R), the pitch (3.7 or 5.1 residues per turn) and the exact orientation of the helix with respect to a reference point in the protein.

An estimate of the lower bound of I_c can be obtained from these factors as follows: 1) To find the exact number of residues, R, in a helix requires about $2 \log_2 R$ bits.* 2) The pitch requires 1 bit (3.7 or 5.1 residues/turn of the helix).

* It is rather interesting that the determination of the value of any integer, either + or − (other than zero), requires exactly $2n$ bits, where $2^{n-1} < R \leq 2^n$ (which is close to $2 \log_2 R$): n bits are necessary to find that $|R|$ is in the range indicated, $n - 1$ bits to find $|R|$ and 1 bit to determine R, i.e. the sign. For example, let $R = -48$: six questions which can be answered by yes or no will show that $|R|$ is 33–64; five more questions will determine that of the 32 possible values $|R| = 48$ and one yes or no question determines $R = -48$. Thus, $2^5 < R \leq 2^6$ and $2n = 12$ bits.

3) A reasonable value for the number associated with specifying the interhelical bonds would seem to be $R/2$ bits. This arises by assuming $R/4$ interhelical bonds, i.e. one bond per turn of the helix, and the previous discussion of intramolecular bonding indicates that the identity of each interhelical bond requires about 2 bits of information. Another reasonable value for this factor is $R/4$; this would occur for 1 one-bit interhelical bond or 1 two-bit bond every other turn, which attempts to take into account that disulfide and "strong" H-bonds are probably the most important interhelical bonds. Actually this factor could be zero since it may *not* be possible to specify interhelical bonds independent of the sequence. 4) The information necessary to specify the orientation of each helix with respect to some reference point in the protein is the most difficult factor to estimate. It may be almost zero, since the interhelical bonds may unequivocally determine the orientation of the helix. On the other hand, it should not be larger than ($\log_2 R + 30$) bits, where $\log_2 R$ bits is sufficient to determine a specific residue and 30 bits to specify its orientation. The 30 bits would be assigned to the six parameters associated with the two vectors necessary to specify orientation. An average 'grain' of 1:32 is undoubtedly too coarse for specifying the orientation of a *single* isolated helix, but is probably adequate for specifying a helix which is oriented in relation to others in the same molecule.

The $R/2$ and 30 and the zero terms have been combined to give 'high' and 'low' values for the estimation of the *minimum* of I_c. These are calculated as I_c/residue (in bits) by

$$I_c/\text{residue} = \frac{3 \log_2 R}{R} \qquad LOW \qquad (1)$$

and

$$= 0.50 + \frac{30 + 3 \log_2 R}{R} \qquad HIGH \qquad (2)$$

The results as a function of R are shown in Fig. 3. PAULING, COREY and BRANSON (8) cite examples of helical polypeptides for which R is 11, 18 and 36. The corresponding region of Fig. 3 has been shaded. From these considerations it would appear that the minimum value of I_c should be about 1 to 4 bits/residue depending upon R.

Although many proteins appear to be helical in nature, there are others, such as ribonuclease (RNase), which from the available evidence would seem not to be. In RNase the structural specificity appears to be determined predominantly by the S—S bonds with the other intramolecular bonds adding stability to the structure. A further discussion of the relative importance of the specific and non-specific intramolecular bonds in maintaining structure will be presented later.

It is obvious that an upper limit cannot be assigned to I_c as readily as to I_s. However, since the structures proposed by PAULING, COREY and BRANSON probably represent polypeptide configurations for which I_c is near minimum, it would appear that one bit/residue is a reasonable lower limit for I_c. From the estimates of I_s and I_c presented here, it appears that for the proteins of general interest I_t should have a value in excess of 4.5 bits/residue although

if it is found that polypeptides are chosen from a single, long master sequence the value could be as low as 1.0 bits/residue.

Estimates of 4.5 bits per residue or greater at the structural level give a total information content, I_t, for the non-structural proteins in excess of 500 bits (or in excess of 100 bits if the minimum estimate turns out to be the true one). Such an estimate is in sharp contrast to the estimates of 10 bits or less

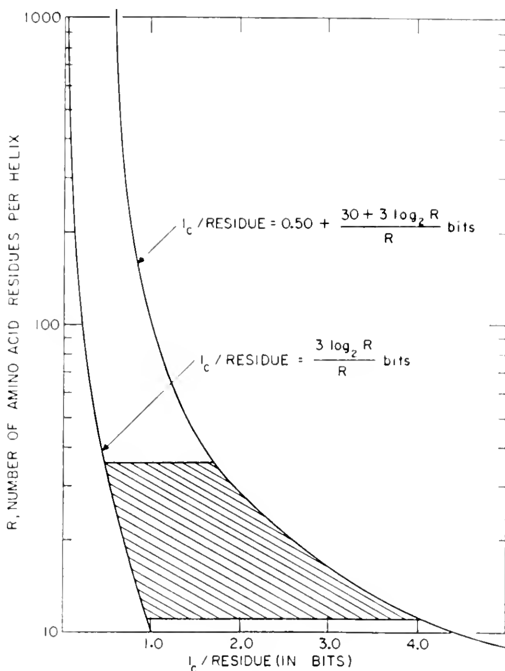


FIG. 3. Limits for estimates of the minimum of I_c as a function of the number of residues per helix. The shaded area indicates helical polypeptide sizes reported by PAULING, COREY and BRANSON (8).

obtained by QUASTLER and his co-workers (14) as the amount of information which must be transmitted for the proper functioning of most protein-controlled systems (e.g. enzymes, immune bodies).

III. ESTIMATION OF STRUCTURAL INFORMATION CONTENT NECESSARY FOR FUNCTION

A disparity of at least one order of magnitude or more in passing from one context or level of organization to another is of considerable interest. The ten-fold difference indicates that only a small part of the information potential is actually utilized in information transmission.

Does this indicate that information transmission in such systems is very noisy and therefore organisms obtain good transmission by utilizing a very high degree of redundancy? DANCOFF (15) proposed a principle of maximum

error in which he postulated that an organism (or for instance a protein-controlled system) will commit as many errors as are consistent with normal function, but that the inherent error rate, which is probably quite high for such reactions, is maintained at a tolerable level by the use of redundancy. Resorting again to the language analogy—a protein corresponds to a paragraph in complexity and its function may correspond to the thought which is conveyed by a paragraph.

Does the difference in information content between the two contexts mean that in the process of evolution the organisms found that particular polypeptide configurations contained structures which could perform useful functions, but that these polypeptide permutations contained a large amount of excess and useless information which has been perpetuated along with the small amount of information associated with the necessary structure?

Does it indicate that much of the protein structure is involved in secondary features of information transmission (e.g. the acquisition, concentration, and transport of energy) and only a small part of the total information content of the protein is intimately engaged in the process of information transmission?

Or does it indicate that each enzyme or protein is capable of mediating many reactions and our experimental ingenuity has not been able to determine more than just a few of them? (This is analogous to attempting to measure the information transmitted by a source which is transmitting through many channels, by monitoring only a single channel.)

The discussion which follows will attempt to throw some light on these questions. However, two important considerations must always be borne in mind when one is dealing with proteins. They are first and foremost colloidal in nature and therefore much of their activity falls in the realm of surface reactions. In the globular proteins it is quite likely that much of the total structural information content is in the interior of the molecules and therefore is unavailable to participate in information transfer occurring at their surface and can only participate in secondary operations similar to those mentioned above. The second consideration involves the question, just what is required for the transmission of one bit of information by a protein system? It seems very likely that one bit of potential structural information will not always transmit the same amount of information; rather, the efficiency of transmission will depend upon the context within which the performance is measured. For example, it is probably much simpler to attach either a *hydroxyl* or *methyl* group to a benzene molecule (which would involve one bit of determination) than it is to construct either a 3.7 or 5.1 helix (which also involves one bit of determination). This is somewhat analogous to the relative difficulties of determining whether a symbol is 0 or 1, or to determining whether one should get married or not!

I_s necessary: It appears in some cases that a fairly large fraction of the potential surface information due to the amino acids present is superfluous. For instance, it has been found in insulin that a large fraction of the residues cannot be critical for function. Iodination, sulfonation and chelation, each of which can mask surface *R*-groups, have been found not to affect insulin activity. Those residues which are species-specific can also be ruled out as being critical for function. Unfortunately, it is difficult to determine the exact

degree to which a particular type of residue is masked by a given treatment, so that it is impossible to state exactly the fraction of surface residues which are not critical. In a similar manner, it is possible to mask the lysine and arginine residues on the surface of trypsin without destroying its activity (16). In fact, acetyltrypsin is available commercially (17) and has the ideal feature that with its lysine and arginine *R*-groups masked, its ability to act as a substrate for other molecules of trypsin is decreased. HAUROWITZ (18) has also pointed out that some of the antigenic properties of proteins are in many cases not affected by iodination or sulfonation of receptive surface groups.

The work of RAACKE (19) has shown that a certain amount of surface heterogeneity (as demonstrated by electrophoretic behavior) is still compatible with a fully active protein. Her results plus the uncertainty found in the analyses of amino acid compositions indicate that an uncertainty of the order of 3 to 10 per cent can occur in the amino acid complement without loss in characteristic function. The results of ROBERTS and COWIE (mentioned previously) involving competition in the amino acid pool also indicate that about 3 to 20 per cent variability in amino acid incorporation can occur. However, it should be borne in mind that each position in the polypeptide sequence may not have a 3 to 10 per cent tolerance associated with it; rather, those residues which participate in active sites likely have a zero tolerance.

I_t *necessary:* KALNITSKY and ROGERS (20) have reported that approximately 15 per cent of the ribonuclease molecule can be digested off with carboxypeptidase before activity is lost. Work reported by ANFINSEN (10, 21) indicates that this estimate may be a little high. Rather, he reports that the carboxyterminal three amino acids (valine, serine, alanine) can be removed with no loss in activity; but, that digestion with pepsin which splits off these three plus their neighbor, aspartic acid, and also ruptures a "strong" hydrogen bond in the vicinity produces loss in activity. Partial digestion by subtilisin (10, 22), which apparently digests central portions of the polypeptide chain, leaves the activity of the RNase intact as long as the digested portion is not oxidized. It is also known that fragments obtained either by hydrolysis or partial enzymatic degradation from myosin (23–25), trypsin (26), chymotrypsin (27, 28), lysozyme (29), papain (30) and pepsin (31, 32) retain their activity in certain situations. The results with pepsin and papain are particularly striking. HILL and SMITH report no loss in the molar activity of papain (toward a synthetic substrate) after an average of 120 of its 180 residues had been removed by leucine-aminopeptidase (an N-terminal type enzyme). PERLMANN has reported that some of the dialyzable fragments (which represent 20 per cent of the total original protein) resulting from pepsin auto-digestion retained 1 to 5 per cent of the original activity toward hemoglobin, but about 75 per cent of the activity of the intact pepsin when tested against the synthetic substrate acetyl l-phenylalanyl diiodotyrosine. These latter results indicate strongly that pepsin, at least, has more than one active site and the site specific for peptide linkages adjacent to an aromatic amino acid depends upon the integrity of only a small portion of the molecule.

I_c *necessary:* Of parallel interest to the above considerations is the question of how much configurational information, I_c, is necessary for function? The work of ANFINSEN and others (10, 33) indicates that the configuration of RNase

can be considerably disrupted without loss in activity. They found that reversible denaturation in 8 M urea did not cause permanent loss in activity; in fact the RNase was still active in 8 M urea in which its specific viscosity was 8.9 as compared with 3.3 in aqueous solution. This large increase in specific viscosity indicates that the so-called native configuration can be opened considerably without destruction of activity. However, ANFINSEN reports that oxidation with performic acid, which disrupts the disulfide bonds, causes irreversible inactivation and an increase in specific viscosity to 11.6.

The phenomenon of complete loss in activity upon the appearance of the full sulfhydryl titer has been observed in most proteins. It has also been known for a number of years that different degrees of loss in characteristic activity can occur. A number of workers (34, 35) have studied reversible inactivation of enzymes in which it has been observed that a partial unfolding of the molecule can occur with a rise in specific viscosity, change in the optical rotation of the protein solutions, changes in solubility, etc., which upon the proper treatment can be reversed. The thermodynamics for reversible denaturation shown in Fig. 4 indicate that quite likely the first step is common from protein to protein since ΔF^* is remarkably constant for all proteins. Reversible denaturation invariably shows an increase in entropy. However, ΔS^* is not constant from protein to protein but varies by a large amount as shown by the unhatched areas to the right in Fig. 4.

The author has proposed (12, 36) and discussed elsewhere in this volume (37) a hypothesis involving three steps, which attempts to explain this phenomenon by ascribing the constant ΔF^* to the initial opening of a disulfide bond. This first step is followed by the rupture of a number of neighboring intramolecular bonds (step 2) with a resulting opening of the molecule indicated by the increase in entropy. According to the proposal, this opening of the molecule is sufficient to disrupt the spatial arrangement of critical amino acids causing loss in activity, but enough stability and configuration is retained so that under the proper conditions the original native structure, or at least a structure compatible with activity, can restitute. In this hypothesis the rupture of a second disulfide bond (step 3) allows irreversible inactivation to proceed with essentially complete destruction of the characteristic protein structure.

A conversion (using an equivalence derived in reference (38)) has been made in Fig. 4 from ΔS^* to ΔI_c. By assuming an average amino acid residue

Table I

Protein	M.W. $\times 10^{-3}$	$N = \frac{\text{M.W.}}{120}$	ΔI_c (bits)	$\Delta I_c/N$ (bits/residue)
Pepsin	36	300	78	0.26
Trypsin	20	167	30	0.18
Emulsin	38	317	48	0.15
Amylase	59.5	496	36	0.07
Hemoglobin	67	558	110	0.20
Egg albumin	40	333	226	0.68
Lacto peroxidase	93	775	340	0.44
Insulin	12	103	18	0.18

weight of 120, ΔI_c/residue is given in Table I for those proteins in Fig. 4 for which the molecular weights are available.

Thus ΔI_c for the loss in specific activity is of the order of 0.25 bits/residue (the 0.68 value for egg albumin does not correspond to a loss in specific activity). This indicates that destruction of the right 5 to 25 per cent of I_c (assuming I_c is close to our minimum estimate of 1 to 4 bits/residue) causes loss of function, which may be reversible or irreversible depending upon which intramolecular bonds are disrupted.

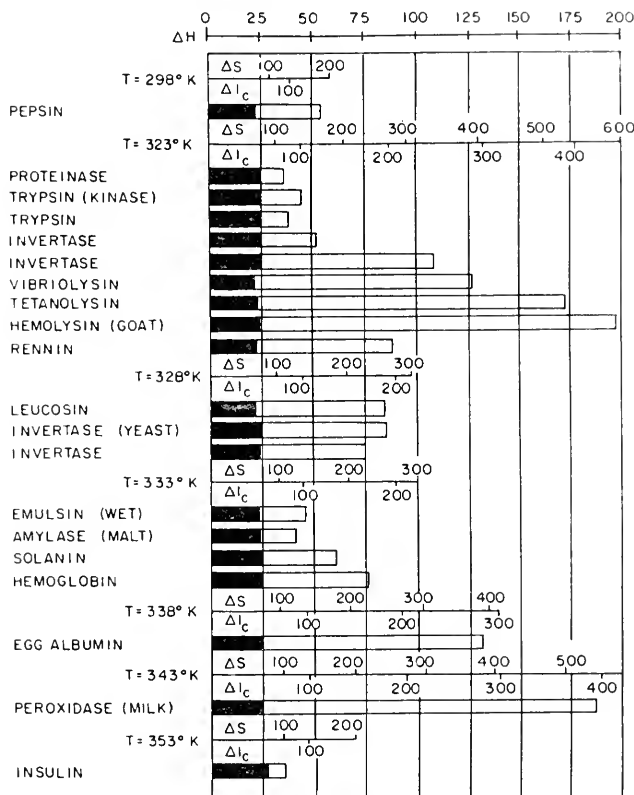


FIG. 4. The equivalence between ΔS^* and ΔI_c for thermal inactivation. The shaded areas to the left represent ΔF^* and the clear areas to the right ΔS^*. (Adapted from Fig. 1, ref 12, by courtesy of University of Illinois Press.)

Summary

The above discussions indicate that redundancy considerations are not the explanation of the large excess of structural information content; rather, that only a small fraction of the potential information on the surface of the molecule is actively utilized in information transfer. HAUROWITZ (18), for instance, has pointed out that experiments with substituted antigens indicate that the antigenic specificity resides in an area on the surface of the protein which is approximately 10 to 15 Å in diameter. Results cited here suggest

that the four or so amino acid residues which would occupy such a surface area (13) may occur as neighbors on the same chain (30–32). Other results mentioned previously (20, 21, 33) suggest that the critical amino acids do not occur in sequence in a single polypeptide chain. This follows from the consideration that digestion should be able to consume an average of about 50 per cent of the protein molecule before an active site composed of four or five adjacent amino acids would be encountered; whereas one of four or five amino acids making up an active site should be encountered, on the average, after about a 20 to 25 per cent digestion of the molecule if the amino acids are distributed roughly at random. In addition, KENNEDY and KOSHLAND (39) has found that phospho-glucomutase when placed in 6 M urea *loses* its activity but recovers it upon dilution, which also indicates separated locations for the critical amino acids. Therefore it may not be possible to state a general rule concerning the relationship between the loci of critical amino acids within polypeptide chains.

It seems that the role of intramolecular bonds is to insure that the amino acids which are critical for function are maintained in the proper spatial relationship to each other so that function can occur. Here again it is impossible to state a general rule as to how many of these intramolecular bonds can be disrupted before loss of function occurs, since apparently all of the hydrogen bonds can be broken in RNase without loss in function but not so in phosphoglucomutase. However, the integrity of the more specific secondary bonds (such as S—S) seems to be much more critical for the maintenance of function. The digestion experiments with pepsin and papain indicate further that it is important where in the molecule the bonds are destroyed.

Other than ruling out redundancy as a possible reason for the discrepancy between the large potential information and the measured performance, it is difficult to choose among the other possibilities mentioned. The results with pepsin and papain, which have been mentioned, suggest strongly that much of the information content may be unnecessary for function, but has been perpetuated along with the critical content. However, the results with pepsin indicating that multiple sites do exist makes it impossible to assign a certain fraction of the information content as 'garbage'. How much of the polypeptide chain is involved in secondary features of information transmission and the structural complexity necessary for transmitting one bit of information are factors which are now being actively investigated by a number of workers.

The various estimates of I_t, I_s and I_c are tallied in Table II.

Table II

	I_{total} =	$I_{sequence}$ +	$I_{configuration}$
Maximum	—	4.32	—
Plausible	>4.5	3.5	>1.0
Minimum	1.0	15/N	1.0
Necessary for performing a single specific function	10–90%	25%	35–90%

IV. CONJECTURES

Some of the results considered in preparing this paper lead to rather interesting speculation. The repetitious minimum entropy polypeptide structures proposed by PAULING, COREY and BRANSON (8) have already been mentioned. Such configurations may be generally applicable to macromolecules, since helical structures have also been proposed for desoxyribonucleic acid (DNA) polymers (40) and some viruses (41). CRANE (42) states that helical configurations occur in linear (uni-dimensional) crystals, i.e. structures where progression from each sub-unit to its essentially identical neighbor is by a repeated process of translation and rotation. LUMRY and EYRING (43) predict that once hydrogen-bonded secondary structures are formed the characteristic protein 'conformation' is determined by tertiary folding such that the free energy is minimized. However, this does not explain why crystallization should initially occur and be maintained in solution; and to the author's knowledge no one has advanced arguments which provide a complete basis to account for the apparent prevalence of minimum entropy biostructures, although there have been discussions of how living organisms produce 'order from disorder' or 'order as a result of order' (44). Considering the innumerable configurations available to biological polymers, the question arises 'Are there criteria which determine that the seemingly improbable, highly ordered structures occur spontaneously?' or 'Are these structures imposed at some specific stage in biosynthesis?'

Studies on the reversible denaturation of proteins (34, 35) suggest that the latter possibility is more probable: that is, mild mistreatment can be reversed; whereas, once a certain molecular disarray or instability occurs, an unfolded state results from which the characteristic, native structure does not reconstitute. NEURATH *et al.* (35) make the interesting point, that even if denaturation is complete enough so that physical properties such as solubility, crystallizing ability, or diffusion constants are seriously affected, some of the molecules may subsequently revert to a biologically active form; whereas, others will tend to reverse the molecular disarray by forming a more condensed state but without successfully restoring the native biological properties. This suggests that, although polypeptide chains have an inherent tendency to form semi-condensed configurations, the highly ordered, biologically-active structures are probably not only imposed during biosynthesis, but represent quasi-stable structures with built-in constraints which tend to cause small fluctuations to revert, i.e. a limited amount of disorder can be restrained without the inexorable Second Law prevailing. NEURATH (35) has also reported that the amount of disarray compatible with reversibility depends upon the type of denaturation. Further, denaturation is not reversible under all conditions but may await a change in pH or temperature. However, it is interesting that although an entropy increase is invariably associated with denaturation, removal of the denaturing agents can cause a decrease, which appears to contradict the Second Law; we will later resolve this apparent contradiction.

The quasi-stability of native configurations is suggestive of the situation in diatomic molecules where stability conditions are readily depicted as a local 'well' (relative to the surroundings) or null area in a two-dimensional energy-configuration plot. However, since two dimensions would allow only

a very gross specification of the myriad degrees of freedom of macromolecules, some form of multi-dimensional space will be necessary to represent their stability conditions. The biologically significant portion of such a macromolecular space will also be a 'well', but in a multi-dimensional surface rather than a line plot and will be centered near the locus of native structures in configuration space. A fraction of the well will represent conditions consistent with an active macromolecule and the remainder, conditions characteristic of reversible inactivation. Anything outside the well will correspond to states inconsistent with the restitution of a native configuration.

The multi-dimensional space can be of sufficient dimensionality so that all configurations differing by a 'single step' are neighbors. In such a 'fine-grain' specification each microstate and its probability density (as a function of energy, for example) can be represented. However, such a scheme has drawbacks: first, it has little novelty since *any* situation can be completely described by a sufficient number of parameters; second, a model dealing only with microstates would be extremely difficult to test experimentally; and third, the excessive dimensionality makes it useless as an aid in envisioning possible mechanisms of macromolecular rearrangements.

Thus, a 'coarse-grain' specification, which requires reducing the dimensionality by transforming the microstates into a more useful set of macrostates, is desirable. This general operation can be schematized by the use of the following contingency table:

Table III

←Microstate→ ←Macrostate→		←Molecular Energy→ E_1 E_2 .... E_k .... E_n
010..........111	110.............001	α_{111} α_{112}..........α_{11k}..........α_{11n}
010..........111	111.............001	α_{121} α_{122}..........α_{12k}..........α_{12n}
.		
.		
.		
.		
010..........111	001.............110	α_{1j1} α_{1j2}..........α_{1jk}..........α_{1jn}
110..........111	110.............001	α_{211} α_{212}..........α_{21k}..........α_{21n}
.		
.		
.		
...M_i..........	...m_{ij}.............	α_{ij1} α_{ij2}..........α_{ijk}..........α_{ijn}
.		
.		
.		
.		
.		

A plausible specification for a multi-dimensional space is given in Table III, where a sufficient number of binary digits is used so that each microstate can be unequivocally identified, e.g. the two atoms involved in each bond as well as the bond length and angle could be identified. Each α_{ijk} represents

the probability density of a given microstate for molecular energy state E_k, where the ranges of *i*, *j* and *k* can be essentially infinite.

A transformation to a 'coarse-grain' scheme which seems worth consideration is as follows. Each macrostate, M_i (depicted by the leftmost column of digits in Table III) designates *only* which bonds *exist* in the macromolecule, e.g. sulfur atom no. 7 is hooked to carbon no. 179 and sulfur no. 11, C–563 to C–564 and N–201, etc. Mechanistically all microstates, m_{ij}, contained in a given macrostate, M_i, are grouped together by ordering the digits (or analogously ordering the axes in space). To complete the transformation other bond properties, e.g. length and orientation (the other column of digits in Table III), and their associated probabilities (the right hand portion of Table III) are lumped into two gross categories to provide an intuitively manageable representation. This 'lumped fine structure' for each macrostate, M_i can be represented on an 'energy-deviation' (ED_i) plane at the locus (in transformed configuration space) corresponding to M_i: 'deviation' is a measure of instability, i.e. the extent to which individual microstates, m_{ij}, deviate from the configuration m_{is} corresponding to maximum stability for macrostate M_i. An example of a method for constructing such values is: (a) find the set of digits m_{is} in the middle column of Table III which represents maximum stability for macrostate M_i and (b) determine how many of the corresponding digits of m_{is} and m_{ij} differ. This number provides an excellent measure of 'deviation' because each microstate has a unique *D*-value and 'neighboring' microstates have adjacent *D*-values. Assigning probabilities to pairs of 'energy' and 'deviation' values completes the "fine" to 'coarse-grain' transformation. This requires summing the probabilities, α_{ijk}, of those microstates associated with a particular *D*-value. The probability densities for *E* and *D* values can be arranged into contours of equal probability to avoid further complications of adding a third coordinate to the *ED* plane. These contours will possibly be quite irregular in shape and may well be discontinuous, since the only obvious restriction on their form is that they be non-intersecting.

It should be noted that 'lumping' on to 'energy-entropy' planes would have provided a simpler transformation than that to the 'energy-deviation' planes. The microstates corresponding to a given 'deviation' can be equated to an entropy value by the usual $-\Sigma p_i \log p_i$ procedure, where the p_i's are the probabilities (properly normalized) associated with the microstates. Such a scheme was considered, but was found to be intuitively less useful than the *ED* transformation.

The 'energy-deviation' scheme is of considerable interest when one considers possible mechanisms of both protein inactivation and enzymatic activity. Suppose, for instance, that the energy of a molecule in a native configuration is slowly raised, e.g. by external heat: the point representing 'molecular state' will be driven to new loci in multi-dimensional space. Undoubtedly a trajectory is followed such that the locus resides, 'statistically', on the contour which has the maximum probability permissible or consistent with its energy content and macrostate at any instant. This means that the locus first progresses over the ED_i plane of the particular native configuration, M_i. Eventually a locus will be reached where the probability contour occupied is lower than the corresponding contour on an adjacent *ED* plane. The molecular state will then

jump to that adjacent macrostate by some form of bond rearrangement.* Even without an immediate change in molecular energy due to external heat, the jump will likely be followed by an instantaneous migration of the molecular state locus on the new *ED* plane. This would be anticipated since the new locus might not be the position of maximum probability for that instantaneous molecular energy. A sufficient increase in temperature would eventually drive the trajectory out of the *fraction* of the null region corresponding to an active molecule: with sufficient mistreatment the locus would be driven completely out of the null region into the portion of configuration space representing irreversibly inactivated molecules.

Molecular energy will decrease when external heat is removed, and the molecular rearrangements will be reversed or not depending upon the symmetry of the multi-dimensional surface of the well. Where denaturation is reversed merely by reversing the denaturing conditions, apparently the inactivation trajectory is retraced or else the null region is a smooth "well" with no intervening metastable positions in the reversal trajectory. Thus, for reactivation the two trajectories would not have to be identical but need only form a closed loop.

Asymmetry in the probability contours of even one of the *ED* plots traversed, could cause the inactivation and reversal trajectories to diverge sufficiently so that metastable, non-active configurations would result. Such situations have been observed experimentally; for instance, thermal denaturation at alkaline pH is not reversed upon cooling until the pH is adjusted to acidic conditions (35). Since a change in pH should alter the *ED* contours it is easy to envision how it could make the reversal of denaturation more likely by changing the transition probabilities between macrostates and thus alter the reversal trajectory. Such an alteration would resolve the apparent contradiction of the Second Law: a changed pH would act as a 'Maxwell Demon guiding the footsteps of the reversal trajectory'.

Considering its likely statistical nature, it is probable that much of the trajectory of the locus of molecular states proceeds along essentially negligible probability gradients, not only with respect to transitions from one macrostate to another but more particularly with respect to instantaneous displacement from the locus of arrival on a new *ED* plane. Such transitions should be readily reversible and in general of limited consequence except as they lead to regions of larger gradients. However, a 'low-gradient' region would allow considerable leeway in trajectories. This would permit multiple pathways which would account for the spectrum of effects often observed following physical denaturation. In those transitions involving bonds which latch large segments of the molecule together (12) (e.g. interhelical bonds) gross molecular rearrangements could occur so that the trajectory would pass through regions of large probability gradients. Such transitions would not be instantaneously reversible and would therefore be relatively *important* in driving the trajectory away from the "active" portion of or even out of the 'well'.

My proposed inactivation hypothesis discussed later (37) attempts to

* Somewhat more rigorous discussions of factors affecting the trajectory of the locus of molecular state in similar multi-dimensional plots have been given by TELLER (45) and LUMRY and EYRING (46).

specify the identity and sequence of high-gradient transitions. On this basis energy from an absorbed quantum, ionization or thermal process would migrate through the molecule in a fashion represented mainly by a 'low gradient' trajectory. However, once the energy or charge becomes localized in a bond of low ionization potential involved in latching large segments of the molecule together, a 'high-gradient' transition, not readily reversible, would occur. The inactivation efficiency of absorbed energy will thus be a function both of the locus of the molecular state at the time energy is absorbed as well as its resulting trajectory; where the trajectory depends upon the amount of energy introduced, the point of absorption and any external factors which affect the contours on the *ED* planes. For instance, the quantum efficiency of UV varies considerably with pH for a number of enzymes (47).

The interdependence of energy, configuration and probability proposed here provides a formalism for depicting enzyme action. It is fairly typical of enzyme, as well as other types of catalysis, that reactions proceed which are normally not feasible because of steric or energetic hindrances. It is entirely possible that because of their large size, enzymes act as large energy reservoirs whose function is to "deliver" a quantity of energy to a particular site or complex in an irreversible fashion. Another possibility is that energy may not be delivered *per se* but as a change in configuration of the enzyme with a corresponding alteration in the spatial relationship between reactants complexed to the enzyme. Within these proposals the formation of the enzyme-substrate complex could have an important function. It could act as an external agent affecting the *ED* contours so as to cause a directed alteration in trajectory, leading finally to a completed enzyme catalysis. Effective, i.e. rapid and essentially irreversible, enzyme catalysis will likely depend upon (1) an *E—S* complex formation which involves a high-gradient transition, so as to enhance a drastic alteration in the trajectory of molecular state, and (2) the directed trajectory passing through a high-gradient region, preferably just before completion of catalysis, in order to make reversibility unlikely.

REFERENCES

1. H. Branson: Information theory and the structure of proteins. In: *Information Theory in Biology*, ed. by H. Quastler, 84–104, University of Illinois Press, Urbana (1953).
2. R. Roberts: *Carnegie Institution Yearbook*, No. 55, 110–148 (1956).
3. D. Cowie: *Carnegie Institution Yearbook*, No. 55, 110–148 (1956).
4. L. Augenstine, H. Branson, and E. Carver: A search for intersymbol influences in protein structure. In: *Information Theory in Biology*, ed. by H. Quastler, 105–118, University of Illinois Press, Urbana (1953).
5. G. Gamow, A. Rich, and M. Yčas: The problem of information transfer from the nucleic acids to proteins. In: *Advances in Biological and Medical Physics* **4**, ed. by J. H. Lawrence and C. Tobias, 23–68 (1956).
6. H. Morowitz, *et al.*: personal communication.
7. M. Yčas: (preceding paper in this volume).
8. L. Pauling, R. Corey, and H. Branson: The structure of proteins. *Proc. Nat. Acad. Sci., Wash.* **37**, 205–211, 235–285 (1951).
9. W. Kauzmann: In: *The Mechanism of Enzyme Action*, ed. by W. McElroy and B. Glass, Johns Hopkins Press, Baltimore (1954).

10. C. Anfinsen: Informal lecture at the Fourth Buena Vista Conference on Protein Synthesis (1956).
C. Anfinsen: *Advances in Protein Chemistry* **11,** 1–100 (1956).
D. Steinberg, M. Vaughan, and C. Anfinsen: Kinetic aspects of assembly and degradation of proteins. *Science* **124,** 389–395 (1956).
11. C. F. Jacobsen and K. Linderstrøm-Lang: Salt linkages in proteins. *Nature, Lond.* **164,** 411–412 (1949).
12. L. Augenstine: Structural interpretations of denaturation data. In: *Information Theory in Biology*, ed. by H. Quastler, 119–124, University of Illinois Press, Urbana (1953). "Trypsin monolayers at the air-water interface," PhD thesis at the University of Illinois (1956).
13. L. Augenstine: Remarks on Pauling's protein models. In: *Information Theory in Biology*, ed. by H. Quastler, 75–83, University of Illinois Press, Urbana (1953).
14. H. Quastler: The specificity of elementary biological functions. In: *Information Theory in Biology*, ed. by H. Quastler, 170–190, University of Illinois Press, Urbana (1953).
15. S. Dancoff and H. Quastler: The information content and error rate of living things. In: *Information Theory in Biology*, ed. by H. Quastler, 263–273, University of Illinois Press, Urbana (1953).
16. J. Sri Ram, L. Terminiello, M. Bier, and F. Nord: On the mechanism of enzyme action. LVIII. Acetyltrypsin, a stable trypsin derivative. *Arch. Biochem. Biophys.* **52,** 464–477 (1954).
17. Worthington Biochemical Corp.: Descriptive manual 8, Freehold, N.J.
18. F. Haurowitz: Protein synthesis and immunochemistry. In: *Information Theory in Biology*, ed. by H. Quastler, 125–146, University of Illinois Press, Urbana (1953).
19. I. Raacke: Heterogeneity studies on several proteins by means of zone electrophoresis on starch. *Arch. Biochem. Biophys.* **62,** 184–195 (1956).
20. G. Kalnitsky and W. Rogers: The activity of ribonuclease after digestion with carboxypeptidase. *Biochim. Biophys. Acta* **20,** 378–386 (1956).
21. C. Anfinsen: The inactivation of ribonuclease by restricted pepsin digestion. *Biochim. Biophys. Acta* **17,** 593–594 (1955).
22. S. Kalman, K. Linderstrøm-Lang, M. Ottesen, and F. Richards: Degradation of ribonuclease by subtilisin. *Biochim. Biophys. Acta* **16,** 297 (1955).
23. J. Gergeley: Studies on myosin-adenosinetriphosphatase. *J. Biol. Chem.* **200,** 543–550 (1953).
24. A. Szent-Györgyi: Meromyosins, the subunits of myosin. *Arch. Biochem. Biophys.* **42,** 305–320 (1953).
25. E. Milhalyi: Trypsin digestion of muscle proteins. II. The kinetics of the digestion. *J. Biol. Chem.* **201,** 197–209 (1953).
E. Milhalyi and A. Szent-Györgyi: Trypsin digestion of muscle proteins. I. Ultracentrifugal analysis of the process. *J. Biol. Chem.* **201,** 189–196 (1953).
26. F. Nord, M. Bier, and L. Terminiello: On the mechanism of enzyme action. LXI. The self digestion of trypsin, Ca-trypsin and acetyltrypsin. *Arch. Biochem. Biophys.* **65,** 120–131 (1956).
27. M. Kunitz: Formation of new crystalline enzymes from chymotrypsin. *J. Gen. Physiol.* **22,** 207–237 (1938).
28. J. Gladner and H. Neurath: Carboxyl terminal groups of proteolytic enzymes. II. Chymotrypsins. *J. Biol. Chem.* **206,** 911–929 (1954).
29. J. Harris, C. Li, P. Condliffe, and N. Pon: Action of carboxypeptidase on hypophyseal growth hormone. *J. Biol. Chem.* **209,** 133–143 (1954).
30. R. Hill and E. Smith: Crystalline papain. *Biochim. Biophys. Acta* **19,** 376–377 (1956).
31. G. Perlmann: Formation of enzymatically active dialysable fragments during autodigestion of pepsin. *Nature, Lond.* **173,** 406 (1954).

32. G. Perlmann: Discussion of paper by D. Koshland in *J. Cell. Comp. Physiology*, **47,** Supplement 1, 217–234 (1956).
33. C. Anfinsen, W. F. Harrington, Aa. Hvidt, K. Linderstrom-Lang, M. Otteson, J. Schellman: Studies on the structural basis of ribonuclease activity. *Biochim. Biophys. Acta* **17,** 141–142 (1955).
34. A. Stearn: Kinetics of biological reactions with special reference to enzymic processes. *Advances in Enzymology* **9,** 25–75 (1949).
35. H. Neurath, G. Cooper, and J. Erickson: The denaturation of proteins and its apparent reversal. *J. Phys. Chem.* **46,** 203–211 (1942).
H. Neurath, J. Greenstein, F. Putnam, and J. Erickson: The chemistry of protein denaturation. *Chem. Rev.* **34,** 157–265 (1944).
36. L. Augenstine and R. Ray: Trypsin monolayers at the air-water interface. III. Structural postulates on inactivation. *J. Phys. Chem.* **61,** 1385–1388 (1957).
37. L. Augenstine: Discussion of a proposed mechanism of protein inactivation. (Part IV of this volume.)
38. L. Augenstine: Information and thermodynamic entropy. In: *Information Theory in Biology*, ed. by H. Quastler, 16–20, University of Illinois Press, Urbana (1953).
39. E. Kennedy and D. Koshland: Properties of the phosphorylated active site of phosphoglucomutase. *J. Biol. Chem.* **228,** 419–433 (1957).
40. J. D. Watson and F. H. C. Crick: Molecular structure of nucleic acids. *Nature, Lond.* **171,** 737–738 (1953).
41. H. Fraenkel-Conrat: Rebuilding a virus. *Sci. Amer.* **194** (6), 42–47 (1956).
42. H. Crane: Principles and problems of biological growth. *Sci. Monthly* **70,** 376–389 (1950).
43. R. Lumry and H. Eyring: Conformation changes of proteins. *J. Phys. Chem.* **58,** 110–120 (1954).
44. E. Schrödinger: *What is Life?* Cambridge University Press, Cambridge, England (1946).
45. E. Teller: The crossing of potential surfaces. *J. Phys. Chem.* **41,** 109–116 (1937).
46. R. Lumry and H. Eyring: Energy exchange in photoreactions. In: *Radiation Biology III,* ed. by A. Hollaender, 1–70, McGraw-Hill Book Co. (1956).
47. P. Finkelstein and A. McLaren: Photochemistry of proteins VI. pH dependence of quantum yield and UV absorption spectrum of chymotrypsin. *J. Poly. Sci.* **4,** 573–582 (1953).
48. H. Simon: On a class of skew distribution functions. *Biometrika* **42,** 425–440 (1955).

DISCUSSION

Platt: Simon (48) has shown that skewed distributions (Yule distributions), such as those in Fig. 2, can be obtained from models based on probability assumptions much weaker than those we were looking for. Thus our inability to determine constraints from a study of the distribution of amino acid and letter frequencies in proteins and words is not surprising. However (in agreement with our summarizing statement for that section), Simon points out that the occurrence of a Yule distribution does not obviate more stringent constraints as the underlying probability mechanism.

SPECIFIC MECHANISMS OF PROTEIN SYNTHESIS AND INFORMATION TRANSFER IN THE DEVELOPING CHICK EMBRYO*

H. R. Mahler, H. Walter, A. Bulbenko and D. W. Allmann
Department of Chemistry, Indiana University, Bloomington, Indiana

Abstract—Some preliminary data on precursors and pathways of protein biosynthesis in chick embryos have been presented. The tentative conclusions stated are:

1. Egg white proteins are not utilized for the synthesis of embryonic proteins up to and including the ninth day. Soluble proteins added to the yolk are incorporated effectively, and preferentially to some of the yolk proteins proper.
2. Proteins, peptides and amino acids injected into the yolk sac are incorporated at approximately equal rates. Considering the relative available pool sizes of the various precursors present in the egg, added proteins have to be regarded as the preferred amino acid source of embryonic proteins.
3. A common precursor formed efficiently from proteins and relatively slowly from added amino acids and peptides is considered a likely intermediate in the process.
4. Homogenates of adult organs injected into embryos can be used to elicit a response previously reported for organ transplants, i.e. the apparently specific transfer of labeled material from donor organs to the corresponding organ in the embryonic host. The supernatant fraction of the cytoplasm appears to be, at least in part, responsible for the results observed.

I. INTRODUCTION

It is the purpose of this contribution to describe, in brief, some preliminary experiments on a controlled biosynthetic activity, namely, the precursors and pathways of protein formation. It differs from most of the papers in this symposium in dealing with phenomena rather than with concepts and in the absence of any attempt to establish a functional correlation between these biological phenomena and information-theoretical abstractions. It shares with other papers in this volume the properties of being highly tentative, and in presenting data and comments on a subject to which it is felt information theory should eventually make significant contributions. With the hope that arrival of that time might be hastened and that thought and discussion might be stimulated, our data are presented for consideration. Some of the results are derived from single experiments only and thus lack further confirmation. All of the approaches and conclusions reported are still under active investigation and thus subject to revision and modification.

Embryos were chosen for the experiments since their cells exhibit two fundamental and related properties, both apparently controlled by the nuclear

* The investigations reported have been supported by grants-in-aid of the National Heart Institute, National Institutes of Health, U.S. Public Health Service (Grant No. H 2177) and of the National Science Foundation. This article is contribution No. 746 from the Department of Chemistry, Indiana University.

machinery, which set them apart from other cells of higher organisms. These are: the capacity for replication, that is, rapid yet controlled growth; the capacity for differentiation, that is, continuous yet controlled change and evolution (1). Therefore, one might consider this the system of choice for attempts at discovering how the information content of the hereditary material, the genetic potentialities, are translated into progressive biochemical capabilities and thus into physiological and morphological realities (2). The experiments were done with chick embryos *in ovo* because of the ease of handling and the essentially closed and self-contained nature of the experimental system. Furthermore, there is a relative paucity of reliable, modern information available about their metabolism and that of embryos of higher vertebrates in general, as contrasted to the large body of knowledge derived from experimental embryology.

Our eventual aim is to study the initiation, the mode, and the control of synthesis of highly specific, respiratory enzymes as an indicator of controlled biosynthetic events; however, our initial investigations deal with the more modest one of a definition of parameters for embryonic protein synthesis (3). For any protein formed *de novo*, as has been pointed out by SPIEGELMAN (4) essentially three different mechanisms may be envisaged:

1. The rearrangement of pre-existing protein molecules; namely, the urprotein hypothesis of NORTHROP (5), with suitable modifications.
2. The accretion of amino acids on to pre-existing proteins or peptides.
3. *De novo* synthesis from amino acids.

In the special case of the formation of induced enzymes in rapidly dividing bacterial cells and cell-free systems derived therefrom, the evidence is overwhelmingly in favor of the third alternative (4, 6). The situation is not nearly as straightforward in the vertebrate systems studied. On the one hand, for example, WORK and collaborators investigated the synthesis of milk proteins (7), VELICK, SIMPSON and co-workers the synthesis of several specific enzyme proteins for muscle (8, 9), and LOFTFIELD and HARRIS the synthesis of liver ferritin (10). All this work was *in vivo* and by different experimental techniques, but all these authors presented strong evidence for the last alternative and against the first two. On the other hand ANFINSEN and his co-workers, working with hen's oviduct *in vitro*, have demonstrated that in short term incubations incorporation of amino acids into freshly formed ovalbumin is non-uniform, which is suggestive of the second alternative, but that after longer periods there is a redistribution towards uniformity (11). Similar results have also been obtained for ribonuclease and insulin synthesis by pancreas slices.

In the case of the proteins of the chick embryo proper, FRANCIS and WINNICK have presented data on the incorporation of labeled amino acids in free and protein-bound form as possible precursors of cardiac muscle protein grown in tissue culture (12). The amino acids of the proteins did not exchange with large pools of the corresponding unlabeled acid in the medium, and from this and from experiments with doubly-labeled proteins it was concluded that proteins could be transferred from a nutrient embryo extract medium to heart muscle protein without release of free amino acids. Tracer experiments of this sort, as will be discussed later, do not, however, prove the direct transfer

of protein, but solely suggest that there may not be free equilibration between the free added amino acid pool and amino acids formed and utilized metabolically during precursor protein breakdown and product protein formation respectively.

Another potentially very fruitful line of investigation is provided by some experiments of EBERT's, the results of which tentatively suggest the incorporation of organ specific adult proteins into those of embryos subsequent to chorio-allantoic grafts of the donor organs (3, 13). These researches were the outgrowth of findings by MURPHY (14) and by DANCHAKOFF (15), made some forty years ago, that such transplants of adult chicken spleen lead to a specific enlargement of the host organs. A systematic re-investigation of the phenomenon by WEISS led to the conclusion that transplants of kidney and liver, as well as injections of organ breis of six-day old chick embryos into four-day old hosts, could lead to similar effects (2). WEISS correctly pointed out that experiments of this sort did not permit a choice between a 'template' or a 'specific precursor' type of mechanism. EBERT's investigations are designed to shed some light on this question as well as on the more general ones of protein synthesis and organ specific growth control in embryonic development.

In our own investigations we have made use of S^{35}-labeled organ homogenates, isolated proteins, peptides, and amino acids to gain some insight into the pattern of embryonic protein biosynthesis. In this work we have been interested not only in the immediate but also in the original precursors, which in this case must consist of all or part of the egg white and yolk proteins. Preliminary accounts of some aspects of this work have appeared (16).

II. METHODS AND RESULTS

1. *Preparation of Labeled Precursors*

In the experiments to be reported in this and subsequent sections S^{35}-labeled proteins, peptides, and a mixture of amino acids were prepared biosynthetically as follows: *Torulopsis utilis* was grown on S^{35}-sulfate (obtained from Oak Ridge National Laboratory), according to WOOD and PERKINSON. (17) After extraction with organic solvents (18) the yeast protein was hydrolysed with a 1:1 mixture of 6N HCl and 90 per cent formic acid. Humin was removed by centrifugation and a portion of the neutralized hydrolysate, which also served as source of amino acids in the experiments to be reported, corresponding to 50 mc of the original S^{35}, was injected intraperitoneally into a laying White Leghorn hen in two doses, about five hours apart. Eight hours after the second injection the blood was withdrawn by heart puncture, allowed to clot, and serum albumin and serum globulin prepared (19). The oviduct was removed from the hen, and ovalbumin prepared essentially as described by STEINBERG and ANFINSEN (11). All proteins were treated with cysteine at a pH of 8.0 to 8.5 to assure removal of exchangeable S^{35}, and then dialysed. Peptides were prepared by peptic hydrolysis of the proteins. Aliquots of the radioactive amino acids, peptides, and proteins were prepared by standard methods and counted. In the tracer experiments, 0.05 to 0.1 ml aliquots of the radioactive precursor solutions, containing 0.3 to 1.8 mg and 6000 to 25,000 counts per minute each, were injected into the yolk or the albuminous portion of some two

to three dozen unincubated, embryonated White Rock eggs. The punctures were sealed with paraffin wax and the eggs then incubated at 38° C under conditions of controlled humidity. Starting with the fifth and ending with the ninth day after the injection, embryos were harvested and a number pooled. The mixture was homogenized for about three minutes in a Potter-Elvehjem homogenizer in Ringer's isotonic saline solution, made up to 10 ml (fifth and sixth days) or 20 ml (seventh through ninth days), and precipitated with trichloracetic acid (final concentration, 8 per cent). Dry protein powders were then prepared and counted (20).

2. *Is There Evidence for Selective Utilization of Egg-white or Yolk Proteins?*

In the first set of experiments, chicken serum albumin injected into yolk or egg-white was used as a protein tracer. Table I shows the results of two

Table I. Injection of Chicken Serum Albumin into Embryonated Eggs

Day after injection	Egg white		Egg yolk	
	% of injected activity found per embryo	Protein wt of embryo in mg	% of injected activity found per embryo	Protein wt of embryo in mg
5	.006	5.5	0.79	5
	.008	7	1.12	6.5
6	.012	13	1.34	11
	.100	16	0.31	19
7	.015	28	2.84	17
	.029	29	1.58	27
8	.016	45	4.04	43
	—	—	3.35	48
9	.088	72	2.86	53
	.133	79	7.28	87

series of experiments. The spread of the data is indicative of the precision, reliability, and reproducibility usually obtained in experiments of this sort.

Let us now make the following assumptions: (a) that the injected protein is a true tracer for egg-white and yolk protein respectively, i.e. that no permeability or other pool barriers exist for its equilibration with the corresponding unlabeled egg proteins; and (b) that there is no selectivity in the uptake mechanism of the embryo either for or against a serum albumin tracer as a typical precursor protein. Now we can calculate data shown in Table II and compare the observed mean of the amount of protein actually formed, with that expected on the basis of the above assumptions. The latter value is calculated by multiplying the weight of total yolk or egg-white protein, about 3000 mg each, by the per cent of the injected activity incorporated per embryo (from Table I).

There are profound discrepancies between the calculated and the observed

values. Those for the egg white are only a small fraction of those expected, while those for the yolk are uniformly about two-fold greater. It is thus apparent that at least one of the assumptions cited cannot be valid. The simplest modification would be to postulate that assumption (b) is not true, and that over the time-period studied egg white proteins are not precursors

Table II. Amounts of Embryonic Protein Formed Compared to that Calculated from Tracer Data

Day after injection	Protein (mg/embryo)		
	Observed	Calculated*	
		Egg-white	Yolk
5	6	0.21	28.8
6	15	1.68	24.9
7	29	0.66	66.3
8	45	0.48	111.0
9	76	3.30	152.0

* From injected albumin tracer.

of embryonic proteins. Soluble proteins injected into the yolk can be utilized for this purpose, and may be more efficient than some of the yolk proteins proper.

3. *Is There Evidence for Selective Utilization of Amino Acids, Peptides or Proteins?*

In the next series of experiments we compared serum albumin, albumin peptides and amino acids all injected into the yolk, with the same precursors injected into egg white. The design of the experiment was the same as before and the results of one run are summarized in Table III.

*Table III. Incorporation of Protein Precursors into Chick Embryos**

Day after injection	Precursors injected into YOLK			Precursors injected into EGG-WHITE		
	albumin	albumin peptides	amino acids	albumin	albumin peptides	amino acids
5	0.75	0.44	0.34	0.0063	0.35	1.19
6	1.30	0.90	1.53	0.013	0.56	3.03
7	2.80	1.70	3.86	0.015	1.59	3.48
8	4.05	4.72	5.15	0.016	2.32	4.94
9	2.85	8.52	9.18	0.088	5.94	5.65

* Expressed as per cent of injected activity recovered per embryo.

We see that except for albumin injected into egg-white, which has already been discussed, all the precursors tested appear to be utilized with approximately equal efficiency regardless of whether they are injected into the yolk or the egg white. This is not limited to serum albumin, but holds true equally well for serum globulin and ovalbumin and their peptides as is shown in Table IV.

*Table IV. Incorporation into Embryos of Proteins and Peptides Injected into the Yolk**

Day after injection	S. albumin	S. globulin	Ovalbumin	S. albumin peptides	S. globulin peptides	Ovalbumin peptides
5	0.75	1.10	0.45	0.44	0.20	0.95
6	1.30	1.75	0.80	0.90	0.55	1.65
7	2.80	2.35	0.40	1.70	1.15	2.20
8	4.05	2.55	1.45	4.72	2.20	—
9	2.85	4.50	2.95	8.52	4.50	6.60

* Expressed as per cent of injected activity recovered per embryo.

4. *Is There Evidence for Organ-specific Transfer?*

In order to test the hypothesis of organ-specific transfer advanced by EBERT we have attempted to extend investigations of this sort to the use of S^{35}-labeled adult chicken liver and heart homogenates. These were prepared from deep-frozen organs of a White Leghorn hen injected with a mixture of S^{35}-amino acids, and treated as described above.

After several months the tissues were thawed and homogenized in a tris-(hydroxymethyl)-aminomethane buffer solution at pH 7.4 containing 0.9 per cent KCl, first in a Waring blender and then in a Potter-Elvehjem homogenizer. The liver and heart homogenates, made up to 10 per cent (weight/volume) with the same buffer solution, were then treated with cysteine at a pH of 8.0 to 8.5 to assure removal of all exchangeable S^{35}. After dialysis, some undissolved material was removed by low-speed centrifugation, and the relatively clear supernatant fluid was used for intravenous injection into 9-day-old chick embryos. Embryonated White Rock eggs were incubated at 38° C under controlled humidity conditions for a period of 9 days. They were then candled, and the location of the blood vessels was marked on the shell of each egg. An area of about 1 cm^2 of the shell above the vessel was carefully cut out by means of a dental drill and burr without injuring the membrane, and the small square was removed with a razor blade. A drop of mineral oil was placed on the membrane to render it transparent, and 0.1 ml of the liver or heart homogenate was intravenously injected in the direction of blood flow. The eggs were reincubated for 24 hours and the embryos were excised. Hearts and livers were removed, the organs were pooled, and homogenized; dry protein powders were prepared for counting as described before. Similarly aliquots of the homogenates used for injection were prepared and counted.

The results of these experiments are given in Table V. In all, two series of experiments make up the Table. In the first series, twenty-four embryos each were injected with heart and liver homogenates; of these, twenty-two and eleven respectively survived.

In the second series, forty-four out of forty-seven embryos injected with the heart preparation survived, while the number of survivors was twenty-two out of twenty-eight for the liver homogenate. Thus the table summarizes data obtained on 99 survivors out of 123 embryos that were injected: 66/71 for heart; 33/52 for liver.

It can be seen that the relative specific activity of hearts is higher than that of livers when chicken heart homogenate is injected, whereas the relative specific activity of the livers is higher than that of hearts when chicken-liver homogenate is injected.

Table V. Incorporation of Activity from Adult-Tissue Homogenates into Nine-Day Embryos after Twenty-four-hour Incubation

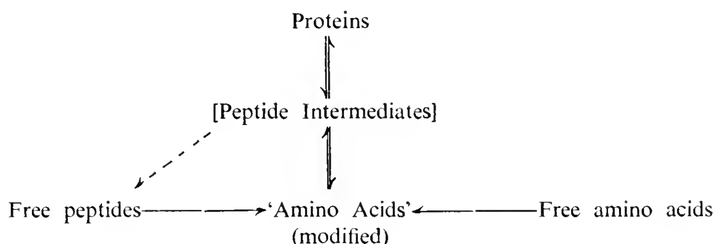
Item	Injection							
	Chicken heart homogenat				Chicken liver homogenate			
Count/min per embryo injected	398			398	2780			2780
mg injected per egg	0.1			0.1	0.1			0.1
Organs investigated	Hearts	Livers	Hearts	Livers	Livers	Hearts	Livers	Hearts
No. of organs cut out	22	11	22	11	11	11	11	11
Dry protein wt of organs obtained (mg)	38.2	72.0	38.8	70.0	84.7	20.9	77.6	22.4
Wt counted (mg)	18.3	29.8	23.4	30.0	30.1	11.6	30.2	12.6
Count/min observed*	21	24	22	19	366	173	389	214
Corrected count/min per 30 mg	28	24	25	19	365	286	386	340
Relative specific activity	1.00	0.86	1.00	0.76	1.00	0.78	1.00	0.87

* Counts per minute are within 5 per cent standard deviation.

III. CONCLUSIONS

The experiments on soluble protein tracers added to yolk and egg-white demonstrate quite clearly that proteins added to the egg-white or, probably, egg-white proteins themselves are incorporated with such low efficiency as to rule out any important contribution from this source to the protein of the developing embryo, at least up to and including the ninth day. Incorporation of protein from the yolk is rapid, and soluble proteins injected into this source may be utilized preferentially to some of the yolk proteins themselves. This utilization of yolk rather than egg-white proteins as a source of embryonic protein during this period is in accord with other investigations, notably the quantitative protein depletion studies of RUPE and FARMER (21). For the intervals studied, amino acids, peptides and proteins, even those of relatively 'foreign' origin such as the serum proteins, all apparently provide an equally acceptable source of S^{35} for embryonic protein synthesis (within an order of magnitude or so), provided they are injected into the yolk. Now the protein tracer must be diluted by at least a portion of the 3.0 g or so of yolk protein—an estimate of approximately 50 per cent would appear reasonable in view of the results reported above. On the other hand, amino acids or peptides cannot be diluted to any appreciable extent since the pools of these substances in the egg are vanishingly small (22). From this one might conclude that proteins themselves or substances easily formed from them must be the preferred precursors of embryonic proteins. Since the egg protein ovalbumin is used no more efficiently than the more "foreign" serum proteins, the pathways of assimilation for these precursors, available to the embryo, must have at least some intermediates in common. The data on peptides may find a similar interpretation.

These intermediates are not free amino acids, as evidenced by their relatively low incorporation rates. They may be small peptides or activated forms of amino acids, formed readily and reversibly from protein precursors, but not identical and not in equilibrium with the pool of added low-molecular weight precursors. This view would be in accord with the findings of FRANCIS and WINNICK (12), although not with their interpretation. The occurrence of pools of modified amino acids, incapable of equilibrating with those in the medium, has been demonstrated in micro-organisms. Thus GALE, working with *Staphylococcus aureus*, found that added glutamic acid could be so transformed, and the modified form used for protein synthesis (24). Similarly COWIE and WALTON (25) have presented evidence that the pools of amino acids formed metabolically in *Torulopsis utilis* and utilized as effective precursors in protein synthesis, are present in some modified form, possibly as complexes adsorbed onto macromolecules, and do not equilibrate freely with added amino acids in the medium. In all the cases presented, this metabolically active form of the amino acids may be formed by a variety of pathways as indicated below.



Recent investigations, especially by ZAMECNIK and his collaborators, (26) have disclosed that free amino acids are first 'activated' by enzymes in the soluble portion of the cytoplasm (27), probably through mixed anhydride formation with adenylic acid (27, 29, 30) prior to their incorporation into a protein-bound form (30, 31), which takes place in RNA-rich granules associated with the microsomal fraction of homogenates (32, 33, 34). Whether or not the metabolically active form of amino acids alluded to above can be equated with these aminoacyl adenylates has not yet been established.

An alternative explanation, which has been invoked to account for apparent preferential utilization of proteins over amino acid precursors in the formation of specific proteins, postulates proteolysis and protein synthesis sites in such close spatial juxtaposition as to permit ready transfer of intermediates from breakdown to synthesis site at the expense of penetration of the latter by added amino acids. This has been suggested by LOFTFIELD and HARRIS (10) as the mechanism operative in ferritin synthesis, and by WALTER *et al.* (20) in the transformation of serum into organ proteins. Purely spatial factors of this sort are probably not the determining ones in the present instance, since it can be demonstrated that the bulk of the proteolytic activity is centred in the yolk (23), and thus remote from the synthetic activity which is, presumably, occurring in the embryo itself. It is hoped that critical experiments now in progress will permit a choice to be made between the various alternatives suggested.

We have shown that the organ-specific localization phenomenon, previously observed with chorio-allantoic transplants, can be duplicated by the injection of homogenates of adult tissue. Similarly TUMANISHVILI *et al.* (35) found almost simultaneously that host organ enlargement could also be elicited by the same technique. This demonstration of the essential similarity of two approaches clears the way for an investigation of the problem by means of relatively straightforward biochemical and enzymological techniques rather than the more demanding ones of experimental embryology. Obviously only a bare beginning has been made. The findings will have to be confirmed and extended and several relatively trivial explanations excluded. Among such explanations are, for instance, the transfer of whole cells on the one hand, and differential composition and/or incorporation rates with respect to cystine and methionine in the two tissues studied, on the other. EBERT claims to have eliminated both these alternatives in his transplantation experiments; in the light of the available information, they are not very likely in the present case. Nevertheless they will have to be rigorously excluded. Our tentative interpretation of the preliminary results described is identical with that advanced by EBERT: that we are dealing with a specific transfer of rather large units from the donor preparation to the embryonic organ.

Preliminary experiments indicate that the injection of either heart or liver (donor) homogenates leads to an increase in specific activity in the liver as compared to the heart. The effect in this case is therefore non-specific and possibly related to the higher mitotic and synthetic activity of liver relative to heart, i.e. to fuller differentiation. Another line of approach which promises to be of some interest is to determine the cell fraction or fractions, if any, responsible for eliciting the effect both with respect to the donor and the acceptor organ. Impetus is added to this approach by the recent experiments which have focussed attention on the soluble and microsomal fractions as being involved in the initial phases of protein synthesis. In preliminary experiments with fractionated, dialysed heart homogenates the data of Table VI were

Table VI. Transfer of Label from Donor Heart Fractions into Organs of Recipient Embryos

Fraction	Relative specific activity of embryonic organs (heart/liver)
Homogenate	1.17, 1.32, 1.23
Nuclei	0.65, 0.74
Mitochondria	0.22 (?)
Microsomes	2.56
Soluble	1.85, 2.50, 1.49

obtained. The number of data in each row corresponds to the number of experiments actually performed. Thus the results for the microsomal and mitochondrial fractions must be regarded as exceedingly tentative. With this proviso, components of the soluble fraction of the cytoplasm might be regarded as responsible for the phenomenon observed with whole heart homogenates.

A similar observation has been reported by KUTSKY who found the supernatant fraction of embryo extract to be most active in stimulating the growth of heart fibroblasts *in vitro* (36).

REFERENCES

1. For a very stimulating and up to date review the reader is referred to: B. EPHRUSSI: Enzymes in Cellular Differentiation, *in:* O. H. Gaebler: *Enzymes: Units of Biological Structure and Function*, Academic Press, New York City, 29–40 (1956).
2. For an intriguing hypothesis involving control of organ development by diffusible components see the following articles: P. WEISS, *in* A. K. Parpart: *The Chemistry and Physiology of Growth*, Princeton University Press, 135–186 (1949).
   P. WEISS: Self regulation of organ growth by its own products. *Science* **115**, 487–488 (1952).
   P. WEISS: Some introductory remarks on the cellular basis of differentiation. *J. Embryol. Exp. Morph.* **1**, 181–211 (1953).
3. The present state of knowledge, with special emphasis on protein biosynthesis, is admirably reviewed in the following: J. D. EBERT: Some aspects of protein biosynthesis in development, *in:* D. Rudnick: *Aspects of Synthesis and Order in Growth*, Princeton University Press, 69–112 (1956).
4. S. SPIEGELMAN: On the nature of the enzyme-forming system, *in:* O. H. Gaebler: *Enzymes: Units of Biological Structure and Function*, Academic Press, New York City, 67–89 (1956).
5. R. S. ALCOCK: The synthesis of proteins *in vivo*. *Physiol. Rev.* **16**, 1–18 (1936).
   J. H. NORTHROP, M. KUNITZ, and R. M. HERRIOTT: *Crystalline Enzymes*, 2nd ed., Columbia University Press, New York (1948).
   S. B. KORIK and H. CHANTRENNE: The relationship of ribonucleic acid to the *in vitro* incorporation of radioactive glycine into the proteins of reticulocytes. *Biochim. Biophys. Acta* **13**, 209–215 (1954).
6. D. S. HOGNESS, M. COHN, and J. MONOD: Studies on the induced synthesis of β-D-galactosidase in *Escherichia coli*. The kinetics and mechanism of sulfur incorporation. *Biochim. Biophys. Acta* **16**, 99–116 (1955).
7. P. N. CAMPBELL and T. S. WORK: The biosynthesis of protein. Uptake of glycine, serine, valine and lysine by the mammary gland of the rabbit. *Biochem. J.* **52**, 217–227 (1952).
   B. A. ASKONAS, P. N. CAMPBELL, and T. S. WORK: The biosynthesis of proteins. Synthesis of milk proteins by the goat. *Biochem. J.* **58**, 326–331 (1954).
   B. A. ASKONAS, P. N. CAMPBELL, C. GODIN, and T. S. WORK: Biosynthesis of Proteins. Precursors in the synthesis of casein and β-lactoglobulin. *Biochem. J.* **61**, 105–115 (1955).
   C. GODIN, and T. S. WORK: Biosynthesis of proteins. The effect of intravenous peptides on casein synthesis in a lactating goat. *Biochem J.* **63**, 69–71 (1956).
8. M. V. SIMPSON and S. F. VELICK: The synthesis of aldolase and glyceraldehyde-3-phosphate dehydrogenase in the rabbit. *J. Biol. Chem.* **208**, 61–71 (1954).
   M. V. SIMPSON: Further studies on the biosynthesis of aldolase and glyceraldehyde-3-phosphate dehydrogenase. *J. Biol. Chem.* **216**, 179–183 (1955).
9. M. HEIMBERG and S. F. VELICK: The synthesis of aldolase and phosphorylase in rabbits. *J. Biol. Chem.* **208**, 725–730 (1954).
   S. F. VELICK: The metabolism of myosin, the meromyosins, actin and tropomyosin in the rabbit. *Biochim. Biophys. Acta* **20**, 228–236 (1956).
10. R. B. LOFTFIELD and A. HARRIS: Participation of free amino acids in protein synthesis. *J. Biol. Chem.* **219**, 151–159 (1956).
11. D. STEINBERG and C. B. ANFINSEN: Evidence for intermediates inovalb umin synthesis. *J. Biol. Chem.* **199**, 25–42 (1952).
    C. B. ANFINSEN and D. STEINBERG: Studies on the biosynthesis of ovalbumin. *J. Biol. Chem.* **189**, 739–744 (1951).

M. Vaughan and C. B. Anfinsen: Nonuniform labeling of insulin and ribonuclease synthesized *in vitro*. *J. Biol. Chem.* **211**, 367–374 (1954).
M. Flavin and C. B. Anfinsen: The isolation and characterization of cysteic acid peptides in studies of ovalbumin synthesis. *J. Biol. Chem.* **211**, 375–390 (1954).
D. Steinberg, M. Vaughan, and C. B. Anfinsen: Kinetic aspects of assembly and degradation of proteins. *Science* **124**, 389–395 (1956).
12. M. D. Francis and T. Winnick: Studies on the pathway of protein synthesis in tissue culture. *J. Biol. Chem.* **202**, 273–289 (1953).
13. J. D. Ebert: The effects of chorioallanteic transplants of adult chicken tissues on homologous tissues of the host chick embryo. *Proc. Nat. Acad. Sci., Wash.* **40**, 337–347 (1954).
14. J. B. Murphy: The effect of adult chicken organ grafts on the chick embryo. *J. Exp. Med.* **24**, 1–6 (1916).
15. V. Danchakoff: Equivalence of different hematopoietic anlagen (by method of stimulation of their stem cells). II. Grafts of spleen on the allantois and response of allantoic tissues. *Amer. J. Anat.* **24**, 127–189 (1918).
16. H. Walter, A. Bulbenko, and H. R. Mahler: Precursors of embryonic chick proteins. *Nature, Lond.* **178**, 1176–1177 (1956).
H. Walter, D. W. Allmann, and H. R. Mahler: Influence of adult tissue homogenates on formation of similar embryonic proteins. *Science* **124**, 1251–1252 (1956).
17. J. L. Wood and J. D. Perkinson Jr.: Yeast biosynthesis of radioactive sulfur compounds. *J. Amer. Chem. Soc.* **74**, 2444–2445 (1952).
18. R. B. Williams and R. M. C. Dawson: The biosynthesis of L-cystin and L-methionine labeled with radioactive sulphur S^{35}. *Biochem. J.* **52**, 314–317 (1952).
19. W. Friedberg, H. Walter, and F. Haurowitz: The fate in rats of internally and externally labeled heterologous proteins. *J. Immunol.* **75**, 315–320 (1955).
20. H. Walter, F. Haurowitz, S. Fleischer, A. Lietze, H. F. Cheng, J. E. Turner, and W. Friedberg: Metabolic fate of injected homologous serum proteins in rabbits. *J. Biol. Chem.* **224**, 107–119 (1957).
21. C. O. Rupe and C. J. Farmer: Amino acid studies in the transformation of proteins of the hen's egg to tissue proteins during incubation. *J. Biol. Chem.* **213**, 899–906 (1955).
22. A. L. Romanoff and A. J. Romanoff: *The Avian Egg*. J. Wiley and Sons, New York City. (1949).
23. A. L. Romanoff: Membrane growth and function. *Ann. N.Y. Acad. Sci.* **55**, 288–301 (1952).
24. E. F. Gale: Assimilation of amino acids by gram-positive bacteria and some actions of antibiotics thereon. *Advances in Protein Chemistry* **8**, 285–391 (1953).
25. D. B. Cowie and B. P. Walton: Kinetics of formation and utilization of metabolic pools in the biosynthesis of protein and nucleic acid. *Biochim. Biophys. Acta* **21**, 211–226 (1956).
26. M. D. Hoagland, E. B. Keller, and P. C. Zamecnik: Enzymatic carboxyl activation of amino acids. *J. Biol. Chem.* **218**, 345–358 (1956).
27. P. Siekevitz: Uptake of radioactive alanine *in vitro* into the proteins of rat liver fractions. *J. Biol. Chem.* **195**, 549–565 (1952).
28. J. A. DeMoss and G. D. Novelli: An amino acid dependent exchange between inorganic pyrophosphate and ATP in microbial extracts. *Biochim. Biophys. Acta* **18**, 592–593 (1955).
29. P. Berg and G. Newton: Acyl adenylates: the interaction of adenosinetriphosphate and L-methionine. *J. Biol. Chem.* **222**, 1025–1034 (1956).
30. E. B. Keller and P. C. Zamecnik: The effect of guanosine diphosphate and triphosphate on the incorporation of labeled amino acids into proteins. *J. Biol. Chem.* **221**, 45–59 (1956).
31. H. Sachs and H. Waelsch: The effect of pyrophosphate on amino acid incorporation into rat liver microsomes. *Biochim. Biophys. Acta* **21**, 188–189 (1956).

32. J. W. Littlefield, E. B. Keller, J. Gross, and P. C. Zamecnik: Studies on cytoplasmic ribonucleoprotein particles from the liver of the rat. *J. Biol. Chem.* **217,** 111–123 (1955).
33. M. V. Simpson and J. R. McLean: The incorporation of labeled amino acids into the cytoplasmic particles of rat muscle. *Biochim. Biophys. Acta* **18,** 573–575 (1955).
34. G. C. Webster and M. P. Johnson: Effect of ribonucleic acid on amino acid incorporation by a particulate preparation from pea seedlings. *J. Biol. Chem.* **217,** 641–649 (1955).
35. S. Tumanishvili, K. M. Djandier, and I. K. Skanidze: Specific stimulation of the growth of the chicken embryo by the effects of tissue extracts. *C.R. Acad. Sci., U.R.S.S.* **106,** 1107–1109 (1956).
36. R. J. Kutsky: Growth stimulating effects by nucleoprotein and cell fractions on chick heart fibroblasts *in vitro. U.S. Atom. Energ. Comm., U.C.R.L.* No. 2270 (1953).

DISCUSSION

Quastler: It is useful to compare the informational requirements of various alternative methods of protein synthesis.

If the whole protein is synthesized directly from amino acids, then each locus on the template must carry sufficient information to specify a single amino acid, or approximately four bits; this is well within the informational capacities of chemical reactions. If the incorporation occurs in two steps, as has been suggested, then each step might have to specify no more than two bits.

If the protein is synthesized from peptide chains, then the informational requirements are much more stringent. Consider the linking of two peptide chains of, say, five amino acids each. If each of the ten amino acids can be any one of the whole set of amino acids, then the linking operation must, in some way, identify ten amino acids, for a total of about forty bits—which is a very large amount of information to be processed in a single act. The requirements are greater—in fact, almost certainly too great—if two chains of ten amino acids are to be linked. The following possibilities exist which allow the use of large fragments without imposing high informational requirements: (a) the terminal amino acid in a chain identifies automatically the other members—this would imply very strong sequential dependencies within peptide chains, and consequently a low informational capacity of the whole amino acid sequence; (b) linkages are formed without reference to the nature of residues remote from the locus of linkage, and the resulting proteins are torn down again if not functional—in this case, the probability of producing functional sequences by chance is small, and the efficiency of protein synthesis is low; or (c) the protein studied is such that the exact sequence of residues is irrelevant.

THE MECHANISM OF ACTION OF METHYL XANTHINES IN MUTAGENESIS

ARTHUR L. KOCH
Department of Biochemistry, College of Medicine, J. Hillis Miller Health Center, University of Florida, Gainesville, Florida

Abstract—The biochemical findings relating to the action of methyl xanthines on bacteria and bacterial extracts have been reviewed. These observations, together with those of NOVICK and SZILARD on the mutagenic activity of these substances, have suggested that the biological action results from an inhibition of enzymes of nucleic acid biosynthesis. Consequences of this hypothesis have been discussed relative to the regulation of growth of cell constituents. Alternative hypotheses are enumerated.

I. INTRODUCTION

A NUMBER of agents, both chemical substances and radiations, cause mutations. One particular class appears to be potentially most fruitful in an attempt to understand the genetic replication process. This class includes purines and related compounds. Particularly important are the plant alkaloids responsible for the pharmacological effects of coffee, tea and cocoa. If these substances are added to a continuously growing culture of bacteria, the mutation rate is caused to increase markedly (1, 2).

If we compare the structures (Fig. 1) of the alkaloids caffeine, theobromine and theophylline, with the purine bases normally present in nucleic acids of

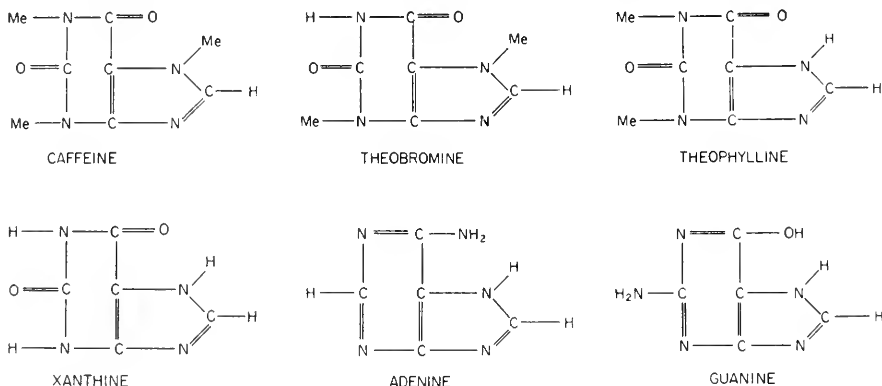


FIG. 1. The structure of purine derivatives

all species, adenine and guanine, the similarity is readily apparent. The former are methyl derivatives of xanthine, the latter amino and deoxy derivatives of xanthine. It is tacitly assumed that these agents are mutagens because of this similarity.

II. TRACER STUDIES

The first possibility to test was that these compounds, or products derived from them, are utilized for the synthesis of the nucleic acid of the host (3). To do this we prepared these substances as well as some others, labeled with carbon 14 in the 8-position of the heterocyclic nucleus. These were then added to growing cultures of *Escherichia coli* under conditions similar to those employed by NOVICK and SZILARD (1, 2) in their studies.

In Table I the data so obtained are presented. Adenine and guanine as well as the deaminated derivatives are very well incorporated into the nucleic

Table I. Incorporation and Mutagenicity of Various Purines

	RSA* of DNA purines	Mutagenicity
Adenine	0.3	+
Guanine	0.20	±
Hypoxanthine	0.30	−
Xanthine	0.20	−
Theobromine	0.00002	+++
Caffeine	0.00001	+++
Theophylline	0.00001	+++

*RSA = relative specific activity = ratio of the specific activity of the purine isolated from the bacteria to that of the growth medium.

acids of both the RNA and DNA type, whereas all methylated substances are incorporated only to a very small extent, if at all. On the other hand, the correlation of mutagenesis is the reverse.

A mutation is a very rare event, and though these agents, when present in quite high concentration, may raise the mutation rate by a factor of fifteen or so, this still only corresponds to one event in 10^7 duplications.

The small amount of radioactivity that is found associated with the DNA from cells grown in the presence of radioactive mutagens is probably experimental contamination. However, although these experiments are technically excellent, they cannot begin to exclude the possibility that a methylxanthine molecule is incorporated into the DNA molecule in the process of the rare mutational event itself, since the resultant incorporation for one locus would be many orders of magnitude below the trace amount observed here. Consideration of the structures of these substances, however, makes this possibility rather unlikely.

In the formation of the normal 9-*N*-riboside or 9-*N*-deoxyriboside linkages, the single replaceable hydrogen which may be in either the 7- or 9-position is replaced by the glycosyl residue. In the case of caffeine or theobromine, which are 7-methyl derivatives, this is not possible because of the prior replacement of the hydrogen by the methyl group. Thus even though the methyl group is

attached to the 7-position it prevents bond formation at the 9-position. Consequently, the methyl group must be removed if the molecule is to be incorporated into the nucleic acids.

The isotopic data, as well as other information, are adequate to demonstrate that there is not a single molecule of enzyme present in these bacteria capable of removing this methyl group (3). Therefore it would appear that certain of the mutagenic materials are not and cannot be converted into a form in which they can be linked covalently to cell materials, not at least by the 9-*N*-glycosyl bond which has been universally found in biological materials.

III. PURINE METABOLISM IN *ESCHERICHIA COLI*

The next possibility we investigated was that the mutagens act by interfering with nucleic acid biosynthesis. First, however, it is necessary to discuss the metabolism of the organism under study. Fig. 2 summarizes, from the

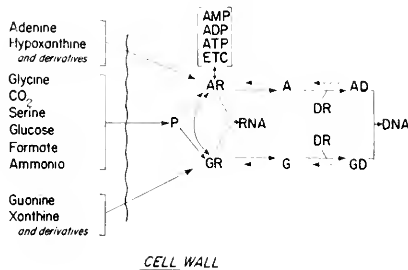


FIG. 2. The purine metabolism of *Escherichia coli*

available tracer data, the pathways of purine synthesis in growing cultures of the test organism (4, 5, 6, 7, 8). C^{14}-labeled CO_2 (4), glycine (8), and serine or formate (unpublished) lead to the formation of RNA adenine, DNA adenine, RNA guanine, and DNA guanine, all of equal specific activity. The activity in the purines derived from CO_2 and glycine is such as to indicate that the well-accepted scheme for purine biosynthesis is the major pathway in this organism (4). C^{14}-labeled adenine and hypoxanthine and their derivatives yield adenine samples of equal, but lower, specific activity in both RNA and DNA. From these facts it is inferred that there are three pools at which purine metabolism branches, namely, a 'purine' pool which is common to all cellular purines, and an 'adenine' and a 'guanine' pool which are precursors of the corresponding purine in both types of nucleic acid. So far, attempts to find a precursor which enters purine metabolism at some point beyond the 'adenine' or 'guanine' pool have failed. Even when the intracellular adenine-C^{14} ribonucleotides were specifically labelled (5), the incorporation into the purines of the ribose nucleic acid was equal to that in the deoxyribose nucleic acids.

It should be mentioned that in organisms under conditions of rapid growth, the soluble intermediate pool concentrations relevant to this scheme are small (5). It was impossible to demonstrate guanosine, adenine deoxyriboside, guanine deoxyriboside or phosphorylated derivatives.

Although the tracer data delineate the pathways, they do not define the intermediates. It is, however, possible to conclude from available enzyme data that 'adenine' and 'guanine' pools are made up at least in part of the free bases themselves. This follows from the fact that the known enzymes of purine metabolism which might be involved in the conversion of the hypothetical 'purine' precursor to the two types of nucleic acids catalyze reactions involving the free purine base. The purine nucleoside hydrolases, purine nucleoside phosphorylases, purine *N*-trans-glycosidases, and purine nucleotide pyrophosphorylases yield the free purine base. These enzymes and the postulated pathway of direct reduction of the riboside to the deoxyriboside constitute the only pathways of interconversion of ribose and deoxyribose purine compounds that can be imagined at present. Since the reductive pathway is known not to occur in *E. coli* (9) (although the interesting work from VOLKIN's laboratory may be relevant (10)), it appears quite likely that the free purine base is involved in the 'adenine' and 'guanine' pools.

In addition to these general considerations, the specific observation of LAMPEN and MANSON (11) that purine deoxyriboside phosphorylase is inhibited by adenine led us to investigate the inhibition of phosphorylases of *E. coli* by methyl xanthines.

IV. ENZYMATIC INHIBITION STUDIES

The main conclusion from these studies (12, 13) was that the organism possesses enzymes, particularly nucleoside phosphorylases of both types (ribose and deoxyribose), that are inhibited by purines generally but specifically by the mutagenic substances. It was also found that even in the presence of

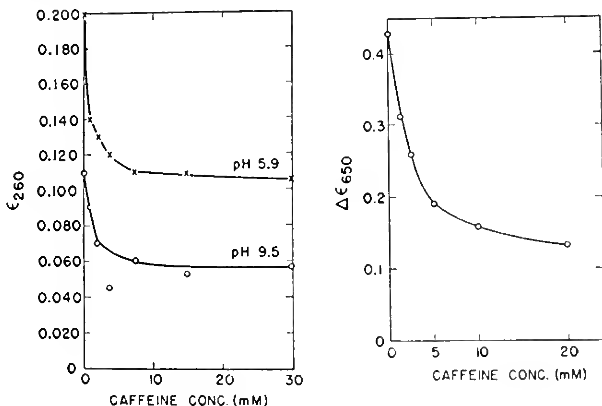


FIG. 3. The inhibition of purine nucleoside phosphorylase.
The effect of caffeine concentration on the arsenolysis of adenine riboside is shown at the left, and on adenine deoxyriboside on the right. The systems contain arsenate to prevent the complication of back reaction.

large amounts of inhibitors enzyme action was not completely repressed (Fig. 3). In all cases this suggested the presence of more than one enzyme catalyzing the reaction under study. Studies of the effect of pH and the separation of the bacteria into several chemical fractions supported this notion.

The activity in various fractions was differently affected by caffeine and this effect was different in acid and at neutrality and at alkaline reaction (see Table II). This finding explains the relatively low toxicity and bacteriostatic power of the plant alkaloids.

Table II. Inhibition of Inosine Arsenolysis by Caffeine

Enzyme preparation* No.	Inhibition produced by 10 μ moles caffeine per ml			Distribution of activity (measured at pH 7)
	pH 5.0	pH 7.0	pH 9.0	
	per cent	per cent	per cent	per cent
6–1 (soluble)	29	59	35	67
6–2 (particulate)	64	97	46	17
6–3 (phosphate extract)	78	78	6	16
			TOTAL	100

* Enzyme Preparation 6–1 was most active at pH 5, preparation 6–2 at pH 9, and preparation 6–3 at pH 7.

In more recent work (13) three new enzymes have been demonstrated in extracts of this organism: an inosine hydrolase, a purine-pyrimidine transribosidase, and a purine-purine transribosidase. All are inhibited to some degree by various purines. The results of the enzymatic studies are summarized in Table III.

Table III. Enzymes of Nucleic Acid Metabolism

Type	Specificity	Inhibition by methyl purines
Adenosine deaminase	Ribose	0
	Deoxyribose	0
Cytidine deaminase	Ribose	0
	Deoxyribose	0
Purine phosphorylases	Ribose	some
	Deoxyribose	some
Pyrimidine phosphorylase	Ribose	0
	Deoxyribose	0
Inosine hydrolase	Ribose	+
Purine-pyrimidine *trans*-glycosidase	Ribose	++
Purine-purine *trans*-glycosidase	Ribose	+

V. WORKING HYPOTHESIS

The mutagenic agents do inhibit enzymes that appear to be directly linked to the path of nucleic acid synthesis, but how can such an interference affect the

mutation probability? We have proposed (12) that this may result from a change in the steady state concentrations of the intermediates that are to be assembled together to form the macromolecular DNA. This must happen without any change in the flow of intermediates, in accord with the experimental fact that the growth rate of the bacteria is not affected significantly by the mutagens when present at concentrations that give rise to large changes in the mutation rate (1).

Let us first consider the consequences of lowering of the concentration of whatever adenine deoxyriboside or guanine deoxyriboside derivative is involved in the polymerization reaction leading to macromolecular DNA. The WATSON-CRICK model for DNA assumes that the specificity lies in the formation of two or three hydrogen bonds between specific pairs of nucleotides: adenine and thymine, and guanine and cytosine. It has been suggested by WATSON and CRICK (14) that the mutational event is the entry of a heterocylic base which is not complementary. This would yield a double helix which is energetically less stable. Upon subsequent duplication this yields two stable molecules, one of the parental type and one of a new mutant type.

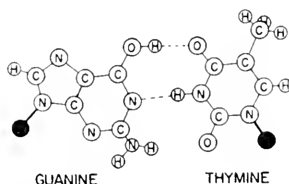


FIG. 4.

It is to be recognized that the mutational event is an improbable one, and therefore quite improbable structures may be involved. Two options for the unfavorable pairing are available. First, two pyrimidines or two purines may become situated opposite each other. This gives structures that should be capable of forming hydrogen bonds, but are either too long or too short. Alternatively, a purine and a pyrimidine may pair, but the purine may occur in the uncommon tautomeric form and consequently pairing will occur abnormally. WATSON and CRICK (14) suggested adenine in the lactim form binding with cytosine, more probable is the pairing of guanine with thymine (Fig. 4). This pair has the proper dimensions; there are no steric difficulties. In this structure guanine is written with the oxygen in the 6- position in an enol form. X-ray-diffraction workers have concluded that guanine is ordinarily found in the keto form, but the evidence is not strong that the keto form is even dominant (15), and considerations of the resonance possibilities indicate a considerable stabilization of the enol form because the latter allows aromaticity of the heterocyclic ring.

Thus, guanine-thymine pairing might well be of likely occurrence. With this in mind, we have attempted in our enzyme studies to find differences of the effects of mutagens on the inhibition of reactions of the adenine compounds, as opposed to the guanine ones, that would be implied if this structure were to

account for the mutational activity of these methylated purines. So far we have been unable to detect any such differences. We may have been examining the wrong systems.

For the present we shall tentatively suggest the pair thymine-cytosine (Fig. 5) as the culprit. This pair is shorter than the conventional structures. In the very interesting paper by DONOHUE (16) a large number of possible pairings are suggested. For our purposes most of these are unsatisfactory because they give rise to helices possessing a two-fold axis parallel to the helical axis, whereas

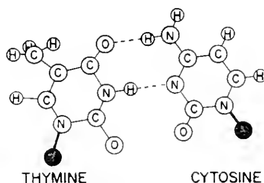


FIG. 5.

in the WATSON-CRICK structure this two-fold axis is perpendicular to the helical axis, and thus consistent helices formed by substitution between the two types can not occur. One structure (DONOHUE'S no. 22) would fit into the symmetry of the WATSON-CRICK model and it is the pairing suggested in Fig. 5.

VI. STEADY-STATE CONSIDERATIONS

Whatever may be the critical or quantitatively most significant substitution in this type of mutational change, the hypothesis we have proposed requires that the concentration of terminal pools be altered. The experimental data that we have obtained have been primarily with purine ribonucleoside phosphorylase which catalyzes a step which is clearly non-terminal in DNA synthesis, and very likely the reaction catalyzed by purine deoxyriboside phosphorylase is also not the transformation of the last small-molecular-weight intermediate into DNA.

Although it may be that the terminal processes are inhibited, let us examine some possible situations that might lead to an alteration of the steady-state concentration of the penultimate substance without influencing the steady-state flux of DNA synthesis. To do this, the question of bacterial growth itself must be raised. Bacteria grow autocatalytically. HINSHELWOOD (17) as well as others have pointed out that this results from an interaction of catalytic units. Thus, if the amount of one component, P (protein), controls the rate of synthesis of another component, N (nucleic acid), then

$$\left.\begin{aligned} \frac{dP}{dt} &= k_1 N \\ \frac{dN}{dt} &= k_2 P \end{aligned}\right\} \qquad (1)$$

where k_1 and k_2 are characteristic constants. The steady-state solution of this pair of equations is

$$\left.\begin{aligned} P &= P_0\, e^{+\sqrt{k_1 k_2}\, t} \\ N &= N_0\, e^{+\sqrt{k_1 k_2}\, t} \end{aligned}\right\} \qquad (2)$$

where P_0 and N_0 depend on the initial conditions and the constants k_1 and k_2. Thus both P and N increase exponentially at the same rate and each therefore appears to be 'autocatalytic'.

Clearly, processes of this kind are responsible for the maintenance of constant growth rates and constant composition of cells during the exponential growth of bacteria. However, the control of the system by this type of interaction cannot explain the regulation of synthesis of intermediates for the biosynthesis of either P or N. Additional regulatory processes must be considered. From equation (2) it is evident that for any constituent of the cell (intermediate or enzymatic catalyst) the steady concentration increases autocatalytically. If expressed as amount per unit number of bacteria or per unit bacterial mass, any cell constituent may be considered constant. Thus, if such a transformation is made, we can consider a system with time-invariant concentrations of intermediates and catalysts and also time-invariant fluxes. Thus, the steady-state treatment of reaction rates is immediately applicable to our problem. The most general formulation is that of CHRISTIANSEN and has been well described by HEARON (18, 19).

In essence the rate expression for each step of a concatenated reaction scheme, in which a substance is produced in one step and utilized in the next, is written down. Each of the terms in these expressions is the product of the intermediate with a rate constant and also with either unity or with the concentration(s) of the other chemical reactant(s). If the product of the two latter factors is set equal to a quantity W, bearing suitable subscripts to identify the term, and if the usual steady-state assumptions are made, then the solutions for both the flux of the system or the over-all reaction rate v and the concentration of each intermediate $[X_i]$ may be computed. If the very last reaction is irreversible, equations (3) and (4) are obtained.

$$v = \frac{W_1\, W_2\, W_3 \cdots W_n}{(W_2\, W_3 \cdots W_n) + (W_{-1}\, W_3 \cdots W_n) + \cdots + (W_{-1}\, W_{-2} \cdots W_{-(n-1)})} \qquad (3)$$

$$[X_i] = v \left[\frac{(W_{i+2} \cdots W_n) + (W_{-(i+1)}\, W_{i+3} \cdots W_n) + \cdots (W_{-(i+1)} \cdots W_{-(n-1)}}{W_{i+1} \cdots W_n}\right] \qquad (4)$$

The assumption of the irreversibility of the last step is made necessary by the well-known metabolic stability of DNA. Recent experiments (20) demonstrate the extreme irreversibility in the normal adult rat. The evidence for growing cultures of *E. coli* is less stringent (21, 22) but does permit this assumption in comparison with the tremendous synthetic rate in these organisms.

Now if in addition we assume that some step is either rapid in the direction of synthesis or irreversible, then it may easily be seen that the reaction velocity v, is completely independent of subsequent steps. Thus, the synthetic rate can be made to depend on the level of a few catalysts or other reactants involved earlier in the sequence. Consequently, increased protein synthesis would cause increased synthesis of a very few enzymes critical for nucleic acid biosynthesis, and this would lead smoothly to increased DNA synthesis without requiring exact synchronization in the increase of each enzyme on the biosynthetic pathway. The concentration of the last intermediate X_{i-1} can be seen from equation (4) to be v/W_n, and thus is completely independent of any step that has no effect on the reaction velocity, v.

This case does not therefore satisfy the requirements suggested above to explain the mutagenic effects of the plant alkaloids. The independence of growth rate in the presence of caffeine could be explained simply by assuming that the inhibition occurs *after* some fast or irreversible reaction; but the action of the inhibitor on any but the final step has no effect on the concentration of the immediate precursor of the macromolecule, and thus cannot affect the probability of mutation.

The scheme considered above has two desirable features: it permits a reciprocal control of nucleic acid by the level of protein synthesis, and it prevents the accumulation of large amounts of intermediates. Let us now turn to a possible mechanism that will do these two things but also will fulfill the conditions imposed by our ideas of the mutation event. Such a mechanism occurs in systems showing product inhibition. Here the rate of production of the final product will depend on the level of some enzyme catalyzing a step late in the reaction sequence, but at the same time, the inhibition prevents the unlimited synthesis of earlier intermediates.

Product inhibition is of common occurrence. It has been suggested as having metabolic significance in two cases (23, 24) in which the product of a reaction sequence inhibits some earlier reaction than its own formation. In the present case it has been shown that adenine deoxyriboside is an inhibitor of the phosphorylase (12) as well as purine bases. Let us assume that all of these agents are competitive inhibitors of enzyme action, although this remains to be demonstrated conclusively.

Under such conditions the reaction velocity is given by the well-known expression for competitive inhibition (see, for example, (25))

$$v = \frac{V(S)\, K_i}{K_s\, K_i + K_s(I) + K_i(S)} \tag{5}$$

where V is the maximal velocity obtainable, K_s is the Michaelis-Menten constant for the substrate S, and K_i is the constant for the binding of the enzyme with the inhibitor, I. If $K_s(I)$ is the dominant term in the denominator, this expression simplifies to give:

$$v = \frac{K_i V(S)}{K_s(I)} \tag{6}$$

In the present case, adenine deoxyriboside is the inhibitor which is formed from the substrate adenine and deoxyribose-1-PO_4. Now, if the net rate of

removal of adenine deoxyriboside is to be maintained constant and determined solely by the process of removal, then a steady-state will quickly ensue in which $(S) \propto (I)$, and in which the rate of formation of I is dependent only on the rate of utilization. The concentration of I will become adjusted to establish such a condition.

In the presence of the mutagen, the total inhibitor is effectively derived from three sources; deoxyribosides, free normal bases, and the mutagen. While maintaining constant synthesis of DNA, the effect of the mutagen will then be to decrease the level of the normal reaction product, adenine. Similar relations will hold for guanine deoxyriboside.

It should be noted that in this case, although not in the case considered above, any number of intermediates may occur between the step under consideration and the polymerization step, if these reactions are rapidly reversible. Then a change in adenine deoxyribose concentration will lead to a proportional change in the precursor immediately used for the formation of the macromolecule.

This model can then utilize the enzymatic finding, and the biological facts. There is, however, one additional fact that should be introduced, *viz.* certain specific substances, the purine ribosides (26), are anti-mutagens. That is, these substances will prevent the action of caffeine and related compounds in causing mutations. Moreover, they will decrease the so-called ‘spontaneous mutation’ rate.

This can be tentatively explained on the basis that these substances are substrates or immediate precursors of the substrates of the key step, and that their increase simply affects the system so as to cause an increase in the concentration of purine deoxyribosides and thus a decrease in the mutation rate.

VII. ALTERNATIVE HYPOTHESES

In concluding, I should like to list various hypotheses that one should consider in this type of chemical mutagenesis. They will be considered in order of the intimacy of the mutagen with the duplication process.

1. *The mutagen is incorporated into the nucleic acid.* This is tentatively rejected as indicated above, from the tracer evidence, and the argument that methylation in the imidazole ring prevents *N*-glycoside formation. It should be noted that production of a self-duplicating ‘methylated gene’ can be rejected because the mutants cannot metabolize methyl purines and certainly do not require them (3).

2. *The mutagens inhibit enzymes of nucleic acid biosynthesis, and this causes a change in the concentration of intermediates.* This latter effect changes the probability of mutation. This is the hypothesis we favor, but it is clear that a great deal of work will be required to establish it or some variant thereof. It is also clear from what has been said above that special circumstances must occur in order that the proposed mechanism can work.

3. *The mutagen causes some change in the general metabolism of the organism and this leads to a change in the mutation probability.* It is certainly true that the mutation probability is dependent on a great many factors. Kihlman (27,

28), working with plants, has suggested such a mechanism to explain chromosome breakage induced with caffeine derivatives. He proposes that ATP is necessary for the aberrations produced by the compound 8-ethoxy caffeine. However, there appear to be considerable differences between the two systems; with the bacteria one *thinks* the process involved is one of 'point mutation', but certain clearcut differences are evident in the two types of material with regard to the interaction of oxygen tension and ionizing radiations. (Compare (2) and (27).

4. *The mutagen causes the organism to 'adapt' to its presence, and thus causes widespread alterations in the amount of enzymes and intermediates.* This could lead to a change in mutation rate. This may be in fact the explanation of the effect of adenine (12). This substance inhibits the growth of bacteria which have previously been grown in its absence. Growth resumes when the organism has 'adaptively' produced an 'adenine deaminase' activity which is not demonstrable in bacteria grown in its absence. This shift in metabolism can then be envisioned to lead to changes in the mutation rate.

This list is probably sufficiently inclusive to include the right answer if there is only one, but at least the necessary research, both with test tubes and with pencil and paper, to test these possibilities is feasible.

REFERENCES

1. A. NOVICK and L. SZILARD: Experiments on spontaneous and chemically induced mutations of bacteria growing in the chemostat. *Cold Spr. Harb. Symp. Quant. Biol.* **16,** 337–343 (1951).
2. A. NOVICK: Mutagens and anti-mutagens. *Brookhaven Symp. Biol.* No. 8, 201–214 (1956).
3. A. L. KOCH: The metabolism of methyl purines by *Escherichia coli. I.* Tracer studies. *J. Biol. Chem.* **219,** 181–188 (1956).
4. A. L. KOCH, F. W. PUTNAM, and E. A. EVANS Jr.: The purine metabolism of *Escherichia coli. J. Biol. Chem.* **197,** 105–112 (1952).
5. A. L. KOCH: Biochemical studies of virus reproduction. XI. Acid soluble purine metabolism. *J. Biol. Chem.* **203,** 227–37 (1953).
6. M. E. BALIS, C. T. LARK, and D. LUZZATI: Nucleotide utilization by *Escherichia coli. J. Biol. Chem.* **212,** 641–645 (1955).
7. E. BOLTON: Biosynthesis of nucleic acid in *E. coli. Proc. Nat. Acad. Sci., Wash.* **40,** 764–772 (1954).
8. A. L. KOCH: The kinetics of glycine incorporation by *Escherichia coli. J. Biol. Chem.* **217,** 931–945 (1955).
9. I. A. ROSE and B. S. SCHWEIGERT: Incorporation of C^{14} totally labeled nucleosides into nucleic acids. *J. Biol. Chem.* **202,** 635–644 (1953).
10. E. VOLKIN and L. ASTRACHAN: Phosphorus incorporation in *Escherichia coli* ribonucleic acid after infection with bacteriophage T2. *Virology* **2,** 149–161 (1956).
11. J. O. LAMPEN: *Symposium on Phosphorus Metabolism,* vol. II, ed. by W. D. McElroy and B. Glass, Johns Hopkins Press, Baltimore, 363–380 (1952).
12. A. L. KOCH and W. A. LAMONT: The metabolism of methyl purines by *Escherichia coli.* II. Enzymatic studies. *J. Biol. Chem.* **219,** 189–201 (1956).
13. A. L. KOCH: Some enzymes of nucleoside metabolism of *Escherichia coli. J. Biol. Chem.* **223,** 535–549 (1956).
14. J. D. WATSON and F. H. C. CRICK: The structure of DNA. *Cold Spr. Harb. Symp. Quant. Biol.* **18,** 123–131 (1953).

15. D. O. JORDAN: The physical properties of nucleic acids. In: *The Nucleic Acids*, ed. by E. Chargaff and J. N. Davidson. Vol. 1, 447–492, Academic Press, New York (1955).
16. J. DONOHUE: Hydrogen-bonded helical configurations of polynucleotides. *Proc. Nat. Acad. Sci., Wash.*, **42,** 60–65 (1956).
17. C. N. HINSHELWOOD: *The Chemical Kinetics of the Bacterial Cell*, Clarendon Press, Oxford (1946).
18. J. Z. HEARON: The steady-state kinetics of some biological systems. I. *Bull. Math. Biophys.* **11,** 29–50 (1949).
19. J. Z. HEARON: Rate behavior of metabolic systems. *Physiol. Rev.* **32,** 499–523 (1952).
20. R. W. SWICK, A. L. KOCH, and D. T. HANDA: The measurement of nucleic acid turnover in rat liver. *Arch. Biochem. Biophys.* **63,** 226–242 (1956).
21. A. D. HERSHEY: Conservation of nucleic acids during bacterial growth. *J. Gen. Physiol.* **38,** 145–148 (1954).
22. A. L. KOCH and H. R. LEVY: Protein turnover in growing cultures of *Escherichia coli*. *J. Biol. Chem.* **217,** 947–957 (1955).
23. R. A. YATES and A. B. PARDEE: Control of pyrimidine biosynthesis in *Escherichia coli* by a feedback mechanism. *J. Biol. Chem.* **221,** 757–770 (1956).
24. H. F. UMBARGER: Evidence for a negative-feedback mechanism in the biosynthesis of isoleucine. *Science* **123,** 848 (1956).
25. P. W. WILSON: Kinetics and mechanism of enzyme reactions, in *Respiratory Enzymes*, ed. by H. A. Lardy, 22–67, Burgess Publishing Co., Minn. (1949).
26. A. NOVICK and L. SZILARD: Anti-mutagens. *Nature, Lond.* **170,** 926–927 (1952).
27. B. KIHLMAN: Chromosome breakage in *Allium* by 8-ethoxy-caffeine and X-rays. *Exp. Cell Res.* **8,** 345–368 (1955).
28. B. A. KIHLMAN: Oxygen and the production of chromosome aberrations by chemicals and X-rays. *Hereditas* **41,** 384–404 (1955).

EVIDENCE FOR A NEGATIVE FEEDBACK SYSTEM CONTROLLING LIVER REGENERATION

André D. Glinos
Growth Physiology Laboratory,
Walter Reed Army Institute of Research, Washington, D.C.

Abstract—Cell division was induced in the resting liver of the rat by lowering the concentration of serum constituents through plasmapheresis, and was inhibited in the regenerating liver by increasing the concentration of the serum by fluid intake restriction.

Electrophoretic analysis of serum proteins and histochemical investigation of the organization of cytoplasmic ribonucleoprotein of the liver cells during regeneration suggest that plasma proteins may participate as information-carrying agents in a negative feedback system controlling the growth of liver cells.

Liver is an excellent tissue for investigating mechanisms of growth control because it regenerates very rapidly. In the rat, removal of up to two-thirds of the total mass of the liver is followed by active cell division leading to complete restoration of the organ within two weeks.

As early as 1923 Akamatsu (1) reported that tissue cultures of rabbit liver grew better in plasma from partially hepatectomized animals than in normal control plasma, and more recently it was shown that cell division can be induced in the resting liver of a parabiotic rat by a partial hepatectomy performed on its partner (2, 3, 4). These findings were considered to indicate the presence or the increase of growth-stimulating factors in the plasma of partially hepatectomized animals.

In our own studies on the possible participation of the humoral system of communication in the control of this growth, blood serum from animals undergoing liver regeneration was assayed in tissue culture (5). These cultures showed a comparable outgrowth in a high concentration of serum of partially hepatectomized rats and in a low concentration of normal serum. A high concentration of normal serum showed inhibitory effects. Based on these findings a hypothesis was formulated with regard to the induction of the regenerative process in the liver which follows partial hepatectomy.

According to this hypothesis, certain constituents of normal blood serum exert a growth-inhibitory action at their normal concentration. Partial hepatectomy would be expected to result in a decrease of the serum concentration of these constituents. Thus in turn regenerative growth is initiated. During regeneration, as the number of liver cells increases, the concentration of these constituents will also increase. When the original equilibrium between a given number of liver cells and a given concentration of the serum constituents is restored, further growth is expected to cease. The evidence for a negative feedback system of this type should satisfy the following two conditions:

1. Induction of growth in the resting tissue by plasma dilution.
2. Inhibition of growth in the regenerating tissue by plasma concentration.

Figure 1 illustrates the application of the classical method for plasma dilution, plasmapheresis, and the results obtained. Normal adult male rats were used. Blood was withdrawn every twelve hours corresponding to 31 to 38 per cent of the initial total blood volume of the animals in the first group

RATE	TIME IN HOURS			
	36-48	60	72-96	MEAN
0%	<.002 .003	.002 <.002	.004 .002	.002
31-38%	<.002 .022	.002 .048	.009 .002	.014
39-46%	.168 .002	.019 .054	.441 .197	.147
MEAN	.048	.031	.163	

FIG. 1. Induction of cell division in the resting liver by plasmapheresis. A total of eighteen adult male rats was used.
The rate of plasmapheresis is expressed as the percentage of the initial total blood volume of the animal replaced by saline per 12 hours. In the control group 0 rate refers to the fact that blood was merely withdrawn and re-injected, with the animals submitted to the same stressful conditions of restriction, anesthesia, venipuncture etc. as the experimental groups. The mitotic activity was obtained by counting 50,000 cells, and expressed as the per cent mitotic index. When no mitosis was found, the mitotic index was recorded as <0.002.

and 39 to 46 per cent in the second. The bleedings were followed by re-injections of the blood cells suspended in an equal volume of saline. Under such conditions cell division was induced in the resting liver of adult rats and was intensified with increasing dilution of the plasma. In this experiment, then, the evidence obtained satisfies the first condition for a negative feedback system.

With respect to the second condition, the method used to achieve plasma concentration was restriction of fluid intake as illustrated in Fig. 2. Two experimental groups were used, differing with regard to the weight of the animals and the extent of the partial hepatectomy. All animals were partially hepatectomized and tube-fed an identical isocaloric fluid diet containing 3 per cent water. The controls were given drinking water *ad libitum* but the experimental animals were deprived of water for the duration of the experiment, which was sixty-four hours, starting sixteen hours prior to the operation and continuing for forty-eight hours postoperatively, at which time the animals were sacrificed. A measure of total body-water loss obtained by this regimen is given by the difference in weight change between experimental and control animals in each group. A measure of the plasma concentration achieved is given by the difference in total protein change. In both the experimental groups an effective inhibition of cell division in the liver was obtained; this inhibition became greater with increasing concentration of the serum. On the other hand mitosis in the intestinal epithelium was not affected. The evidence obtained in this experiment, then, satisfies the second condition for a negative feedback system.

The smaller extent of total body-water loss and plasma concentration in the first group can be ascribed to the greater initial weight of the animals in

this group. It is well known that when dehydration proceeds slowly the maintenance of plasma volume at the expense of extravascular fluid may be quite successful. This is significant since the extravascular fluid of the liver must participate in the transmission of information to the liver cells. The serum albumin fraction in this experiment was found to be low when liver cell division was present and normal or slightly increased when liver cell division was absent. In the framework of the present discussion this feature is somewhat suggestive

DEGREE OF HEPATECTOMY	TREATMENT	NO. OF RATS AND WEIGHT	WEIGHT % CHANGE	SERUM PROTEIN % CHANGE	SERUM ALBUMIN % CHANGE	MITOSES /50000 CELLS
30 % (MEDIAN LOBE)	CONTROLS	10 457	- 7.9	- 4.7	- 16.0	81.4
	FLUID RESTRICTION	9 458	- 12.9	+ 5.0	- 14.0	18.0
10 % (CAUDATE LOBE)	CONTROLS	8 331	- 9.1	+ 1.8	- 26.2	16.5
	FLUID RESTRICTION	8 313	- 17.6	+ 28.9	+ 8.1	0.1

FIG. 2. Inhibition of cell division in the regenerating liver by fluid restriction. The experimental variables are defined in the text.
The percentage changes refer to differences in weight, total serum protein, and serum albumin between the values obtained before treatment and those obtained at sacrifice.

when it is considered that albumin is synthesized in the liver. In view of these facts it was thought that an investigation of changes in protein metabolism of the liver cells, early after partial hepatectomy, could help in elucidating the possible role of the serum proteins and the extravascular fluid of the liver in the transmission of information to the liver cells.

In this, we took advantage of the many observations showing histochemically detectable changes in the organization of cytoplasmic ribonucleoprotein with increasing demands on the protein synthetic mechanism of the cells (6, 7). Briefly stated, these changes consist of the disappearance from the cytoplasm of discrete basophilic bodies which are associated with ribonucleoprotein; the cytoplasm then stains uniformly with basic dyes. Rats were sacrificed at frequent intervals after partial hepatectomy and their livers fixed and stained with gallocyanin chrome alum. Within thirty minutes after partial hepatectomy the ribonucleoprotein-associated basophilic bodies started disappearing from the cells in the periportal area. This change proceeded gradually toward the center, so that eight hours after the operation all cells, even those in the centrolobular area, were affected. After this time reconstruction of the basophilic bodies proceeded in the opposite direction from the center of the lobule towards the periphery. At 24 hours cells in the centrolobular area had completed the cycle

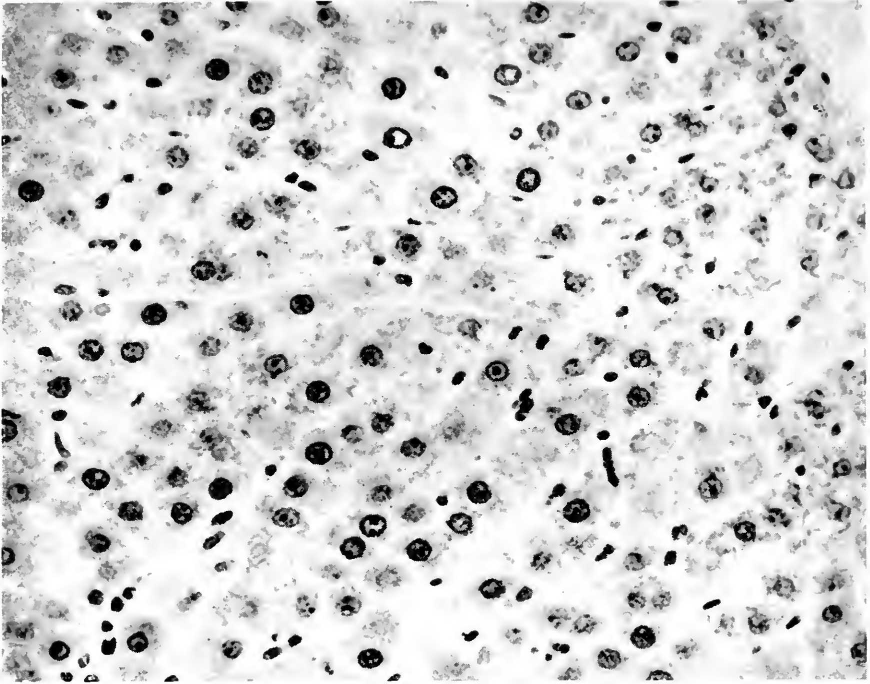


FIG. 3. Regenerating liver twenty-four hours after partial hepatectomy. Central vein at lower left corner. Adjacent centrolobular zone with cells containing ribonucleoprotein-associated basophilic bodies in their cytoplasm. Middle and periportal zones with mostly altered cells having a uniformly basophilic cytoplasm. Two mitotic figures in the middle zone among altered cells.

to face p. 150

showing well organized basophilic bodies, whereas cells in the middle and the periphery of the lobules remained altered (Fig. 3).

Confirming the earlier data of HARKNESS (8), we found that cell division begins between 16 and 24 hours postoperatively in the periportal area. This is significant because cells in this area remained altered for the longest time. The changes in cytoplasmic ribonucleoprotein organization indicate an activation of the protein synthesizing mechanism of the liver cells after partial hepatectomy, proceeding in a topographical pattern related to the direction of the intralobular blood flow. According to the Law of Mass Action these changes would be expected to appear with decreased protein concentration in the immediate environment of these protein-secreting cells. The cells in the periphery of the lobules would be expected to react faster and longer since the ones more centrally located are in an environment richer in protein produced by the more peripheral cells. This interpretation was, in part, verified experimentally by

TREATMENT	FLUID	SERUM PROTEIN CHANGE	LIVER RIBONUCLEOPROTEIN CHANGE
ADDITION	SALINE	− 11.8	0
	DEXTRAN	− 31.2	+
	SERUM	+ 7.9	0
REPLACEMENT	SALINE	− 19.2	+
	DEXTRAN	− 37.8	+
	SERUM	− 8.7	0

FIG. 4. Induction of cytoplasmic ribonucleoprotein changes in the liver by plasma dilution.
A total of six male adult rats was used.
Addition refers to a single intravenous injection of 5.5 ml of fluid. Replacement refers to a 5.5 ml single plasmapheresis treatment. All animals were sacrificed two hours after treatment.
Serum protein change refers to the percentage difference between the values obtained before treatment and those obtained at sacrifice.
Liver ribonucleoprotein change refers to the disappearance of the basophilic bodies from the cytoplasm of the cells in the periportal area.

showing that changes in the cytoplasmic ribonucleoprotein of the cells in the periportal area appear rapidly after a sudden decrease of the serum protein concentration (Fig. 4). After partial hepatectomy, however, these histochemical changes occur as we have seen within thirty minutes before any appreciable changes in the plasma proteins.

The relationships between increased pressure in the portal system following

partial hepatectomy and regeneration have been demonstrated by GRINDLAY and BOLLMAN (9). It is conceivable that, under conditions of increased pressure immediately following partial hepatectomy, the transfer of protein and water from the intravascular to the extravascular space is altered and results in a rapid lowering of the protein concentration of the interstitial fluid of the liver. This leads within a short period to increased protein production in the liver cells and sometime later to cell division.

REFERENCES

1. N. AKAMATSU: Über Gewebskulturen von Lebergewebe. *Virch. Arch.* **240,** 308-311 (1923).
2. B. G. CHRISTENSEN and E. JACOBSEN: Studies on liver regeneration. *Acta Med. Scand.* **234,** *Suppl.*, 103–108 (1949).
3. N. L. R. BUCHER, J. F. SCOTT, and J. C. AUB: Regeneration of the liver in parabiotic rats. *Cancer Res.* **11,** 457–465 (1951).
4. A. S. WENNEKER and N. SUSMAN: Regeneration of liver tissue following partial hepatectomy in parabiotic rats. *Proc. Soc. Exp. Biol. Med.* **76,** 683–686 (1951).
5. A. D. GLINOS and G. O. GEY: Humoral factors involved in the induction of liver regeneration in the rat. *Proc. Soc. Exp. Biol. Med.* **80,** 421–425 (1952).
6. S. LAGERSTEDT: Cytological studies on the protein metabolism of the liver in the rat. *Acta Anat., VII, suppl.* **9,** 1–116 (1949).
7. A. F. HOWATSON and A. W. HAM: Electron microscope study of sections of two rat liver tumors. *Cancer Res.* **15,** 62–69 (1955).
8. R. D. HARKNESS: The spatial distribution of dividing cells in the liver of the rat after partial hepatectomy. *J. Physiol.* **116,** 373–379 (1952).
9. J. H. GRINDLAY and J. L. BOLLMAN: Regeneration of the liver in the dog after partial hepatectomy. Role of the venous circulation. *Surg. Gynec. Obst.* **94,** 491–496 (1952).

FLUCTUATIONS IN NEURAL THRESHOLDS*

LAWRENCE S. FRISHKOPF and WALTER A. ROSENBLITH†
Research Laboratory of Electronics,
Massachusetts Institute of Technology, Cambridge, Massachusetts

Abstract—Over the past twenty-five years several independent investigations of the responsivity of nerve tissue have led to the conclusion that the threshold of a resting neuron fluctuates in time. The conclusion is based on the study of sensory and motor fibers, of monosynaptic arcs and neuromuscular junctions. A number of these studies have been reviewed and compared. The degree of threshold correlation among neurons of a given 'pool' or population has been considered for several systems. A number of possible sources of threshold fluctuation, giving rise to correlated and uncorrelated threshold variations, have been distinguished.

A mathematical model based on the concept of fluctuating thresholds has been described and applied to the problem of ensemble response from the peripheral auditory nervous system. The results of three experiments have been described and compared with the predictions of the model.

I. THE CONCEPT OF A FLUCTUATING THRESHOLD

THE threshold of a nerve fiber is defined as the minimum stimulus intensity that will cause an action potential to propagate. If the threshold of a nerve fiber were a fixed parameter—not changing in time—its value could be determined by presenting stimuli of increasing intensity. The fiber would fail to respond to all stimuli less than some value S_T, and would respond to all stimuli greater than S_T; S_T would then be the threshold of the fiber. However, careful experiments on a number of specific neural systems—sensory and motor, peripheral and central—have shown that such a unique value S_T does not exist; instead, there is a range of stimulus values, S_1 to S_2, such that a stimulus S lying within that range, when repeatedly presented at a rate well below that which would involve the refractory period of the fiber, sometimes evokes and sometimes fails to evoke a response. We find that the fiber responds in a fraction p of all trials and that $p(S)$ is a monotonic function that rises from zero to one as the stimulus increases from S_1 to S_2. Stimuli less than S_1 never evoke a response; stimuli greater than S_2 always evoke a response. We conclude that the threshold of a neuron which exhibits this behavior is a time-varying parameter. The value p approximates the fraction of the time that the threshold is somewhere below the stimulus value S. An equivalent statement is that p approximates (and for large sample size, approaches) the probability of finding the threshold of a fiber below the value S.

* This work was supported in part by the U.S. Army (Signal Corps), the U.S. Air Force (Office of Scientific Research, Air Research and Development Command) and the U.S. Navy (Office of Naval Research).

† Also in the Department of Electrical Engineering, M.I.T.

II. SUMMARY OF STUDIES OF OTHER WORKERS

The class of phenomena that we have been discussing was first observed by BLAIR and ERLANGER (1). They reported that an electric stimulus, repeatedly presented to a single sciatic nerve fiber of the frog, will for most stimulus values either always produce or always fail to produce a response. The transition between these two situations, however, is not sharp. Upon raising the shock intensity, a value is reached at which the fiber sometimes responds and sometimes fails to respond to repeated stimulation. In order to obtain a response every time it is necessary to raise the shock intensity an additional two per cent, far in excess of the uncontrollable variation in the stimulus. Moreover, BLAIR and ERLANGER were able, on occasion, to record simultaneously from two fibers whose potentials could be distinguished by their difference in latencies. On repeated testing with a near-threshold stimulus, sometimes both would respond, sometimes one, sometimes the other, and sometimes neither. Such a result cannot be accounted for on the basis of stimulus instability alone.

The most complete study of this kind that has been published to date was made by CHARLES PECHER (2) in 1939. Using a technique similar to that of

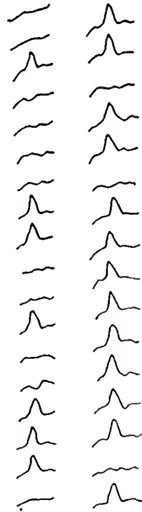


FIG. 1. Left: ink tracings of recordings from single units of frog sciatic nerve, showing occurrence and failure of response to repeated presentations of identical shock stimuli. Right: same, with amplitude of pulse producing the shock raised 4 per cent. Each series shown is part of a longer sequence of 100 presentations. Thirty-five responses were obtained with the weaker stimulus (left); 85 responses were obtained with the stronger stimulus (right). After PECHER (2).

BLAIR and ERLANGER, he also found a stimulus range in which a fiber sometimes responded and sometimes failed to respond to a constant stimulus. Some of his data appear in Fig. 1. In the column on the left we see the responses to successive identical stimuli, of which some produce a response and some fail to do so. In the second column the intensity was raised four per cent. In

Fig. 2 the percentage of responses of a fiber is plotted as a function of stimulus intensity. Again each point is based on 100 stimulus presentations. The total range of threshold variation is, on the basis of these data, about seven per cent. The function shown in Fig. 2 approximates the threshold probability function $p(S)$ that was discussed earlier.

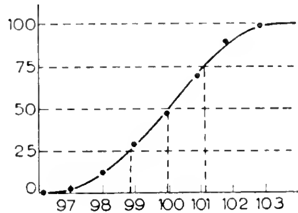


FIG. 2. Relation between stimulus intensity (abscissa) and the number of responses obtained in 100 presentations at a fixed intensity from a single unit of frog sciatic nerve (see Fig. 1). The interpolated solid line approximates the threshold probability function of a unit. From PECHER (2).

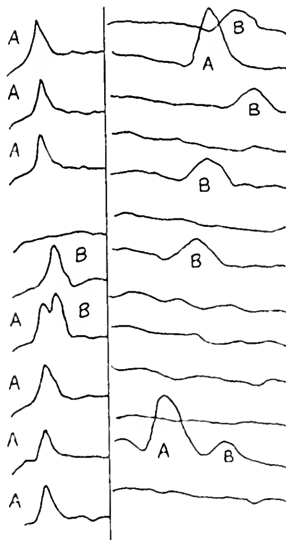


FIG. 3. Left: ink tracings of simultaneous recordings from two units of frog sciatic nerve to repeated presentations of identical shock stimuli. Units A and B are identified by their latencies. Right: same, but recording from two other units, identified by their amplitudes. After PECHER (2).

In the left column of Fig. 3 the responses of two different fibers were simultaneously recorded from a single electrode; the responses are distinguishable by their latencies. At a fixed level of stimulation all possible combinations of response occur: fiber A responds alone, fiber B responds alone, both respond, neither responds. On the right we see the responses from two other fibers; here the responses are distinguished by their amplitudes. Again, all possible

combinations occur. Such a result can only be explained as a result of spontaneous variation in fiber threshold. If threshold were fixed and the stimulus unstable, then only three of the four combinations could occur. That combination would be excluded in which the fiber with higher threshold fires alone.

When responses from two fibers can be distinguished, an opportunity is offered to test the degree of correlation of threshold fluctuation among different fibers. If fluctuations occur independently in two fibers, the probability of both firing to a single stimulus would be the product of their probabilities of firing separately. Any correlation in threshold variations would alter the probability of joint firing. These probabilities can be approximated by counting the number of times that fiber A fires, that fiber B fires, and that both fire, and dividing each by N. In the table below the results of such measurements by PECHER

Table I

Number of stimuli	Number of responses of fiber A	Number of responses of fiber B	Calculated number of simultaneous responses (independence assumed)	Observed number of simultaneous responses
100	78	25	19.5	19
188	129	26	17.8	18
285	205	33	23.7	18
222	150	79	53.4	56
370	214	93	53.8	50
194	113	34	19.8	19
155	110	62	44.0	40
218	168	87	67.0	59
236	152	24	15.5	17

are given for nine different fiber-pairs. In all of these instances the computed and observed frequencies of joint occurrence are in good agreement. The hypothesis of independent fluctuations is thus supported by this experiment.

PECHER tried to determine whether or not for a single fiber the 'response no-response' pattern to a sequence of periodic stimuli can be accounted for by the hypothesis that successive responses occur with equal and independent probability p. He chose a criterion of independence that relates the variables r and n_r, where n_r is the number of times that a sequence of r successive responses (bounded at each end by the absence of a response) occurs in a sample of length N ($r \ll N$):

$$r + k \ln n_r = K$$

where k and K depend on p and N (3). On the basis of samples of 1000 to 2000, PECHER concluded that within statistical limits a linear relation exists between r and $\ln n_r$ for rates of stimulation less than one per second. At higher rates the criterion was no longer satisfied. This is not a sufficient test for independence, since one could construct sequences which satisfy this relation and yet contain strong internal regularities.

PECHER also reported that the latencies of responses to identical shock stimuli exhibit variability of a sort that cannot be attributed to stimulus instability. The variability is greatest for stimuli near threshold.

ROSENBLITH (4) has obtained a threshold probability function for single units in the auditory system of the anesthetized cat. These results are shown in Fig. 4. Responses to repeated clicks were recorded by means of a microelectrode from a unit in the cochlear nucleus. The ratio of the number of responses to the number of stimuli is plotted for various stimulus values; the range of threshold variability is about 15 dB.

LLOYD and MCINTYRE (5) have investigated the variability in the responses of single ventral root motoneurons (*triceps surae*) to identical shock stimuli

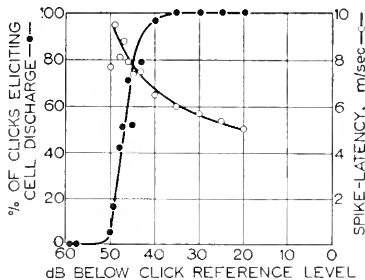


FIG. 4. Percentage of clicks eliciting a response from a single element in the cochlear nucleus of the cat as a function of click intensity. Each point is based on 10 to 40 click presentations. The interpolated solid line approximates the threshold probability function of a unit.

delivered to the gastrocnemius nerve in decapitated cats. Here, the impulse traverses a single synapse in the spinal cord. It was found that at every stimulus level there were neurons that sometimes responded and sometimes failed to respond. Some neurons always responded; others never responded. By raising the stimulus level, the latter could be brought into the range of partial response and, in some cases, of certain response. Different motoneurons receive different amounts of transsynaptic stimulation when a shock is applied to the sensory bundle. The strength of the effective stimulus is said, in this terminology, to depend on the 'transmitter potential' of the synapse. The 'firing index' of a motoneuron is defined as the percentage of trials in which it responds. LLOYD and MCINTYRE measured firing indices for 110 motoneurons under a variety of stimulus conditions. A histogram showing, for a constant stimulus, the number of motoneurons in each firing index interval is seen in Fig. 5. For the purpose of this histogram, units with firing indices of zero and 100 were not counted.

An appreciable change in stimulus strength changes the firing index of a particular motoneuron but affects the histogram very little. From this we can conclude that the distribution of motoneurons with respect to the effective stimulus level is approximately uniform. The situation may be visualized with the help of Fig. 6. Each vertical line represents the effective stimulus or 'synaptic drive' to one motoneuron; all motoneurons in this idealization are assumed to be identical, but subject to different effective stimuli. The curve represents

the threshold probability distribution common to all of the neurons. Units with synaptic drive to the left of the distribution have firing indices of zero; those to the right have firing indices of 100; and units with synaptic drive in the range of the distribution have intermediate firing indices. As the stimulus level is raised, the synaptic drive to every unit is shifted to the right by the

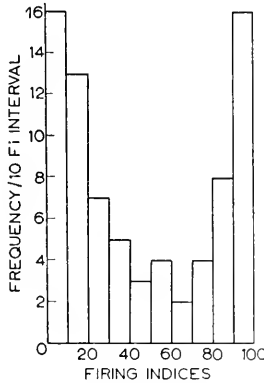


FIG. 5. Histogram showing the number of spinal motoneurons (*triceps surae*) within each firing index interval; responses were obtained by delivering repeated shocks to the gastrocnemius nerve. The firing index of a unit is the percentage of total stimulus presentations to which the unit responds. Units with firing indices of zero and 100 are not included in this diagram. From LLOYD and MCINTYRE (5).

same amount; thus some units with a firing index of zero will be shifted into the intermediate range; some with intermediate firing indices will be shifted into the range of firing index 100. But because the units are uniformly distributed the same number will move into the intermediate range as move out of it, and the distribution of intermediate firing indices will remain unchanged.

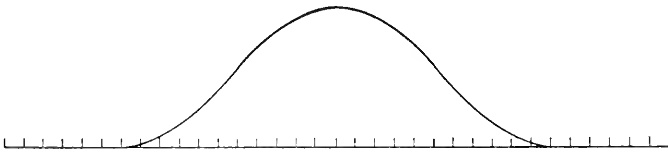


FIG. 6. Idealized relation between the threshold probability distribution of a motoneuron and the levels of synaptic drive to different motoneurons of a population (see text).

The particular choice of a bell-shaped probability distribution will lead to the U-shaped histogram of Fig. 5. For it is clear that if we divide the abscissa in such a way that equal areas under the distribution are subtended, those intervals will be largest near the tails of the distribution (firing indices near 0 and 100) and smallest at the center of the distribution (firing index near 50). Since the density of units along the abscissa is uniform, this means that many more motoneurons will have firing indices between 0 and 10 than between 45 and 55.

As in the study by PECHER, the degree of correlation of threshold variation for members of the same pool of motoneurons was investigated. The extent of correlated and uncorrelated fluctuations is a measure of the relative importance in producing fluctuations of events extrinsic and intrinsic to the fiber. In the spinal cord there is reason to believe that threshold fluctuation is, at least in part, the effect of background activity in other fibers. Such activity would presumably affect many fibers in a neighborhood; the threshold fluctuations of these fibers would therefore show definite correlations.

To determine the extent of correlated variation RALL and HUNT (6) recorded the response of a ventral root together with the response of a single motoneuron belonging to an adjacent root; an example of such a recording is shown in Fig. 7. Fig. 8 shows the results of an experiment based on a thousand

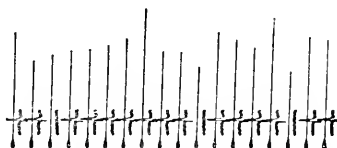


FIG. 7. Simultaneous recording of the responses of a single motoneuron (horizontal deflection) and of an adjacent ventral root (vertical deflection) upon repeated stimulation of the gastrocnemius nerve with identical shock stimuli. From RALL and HUNT (6).

such responses. The population response amplitudes were divided into class intervals, and the number of responses within each class interval was plotted. For each population response within a class interval, the occurrence or failure of a unit response was noted and the number of unit responses plotted (shaded area). The unit responded a total of 697 times out of 1000. If the population response and the unit response were not correlated, the firing index of the unit would be about the same in each class interval. This is clearly not the case. Instead, firing occurs infrequently when the population response is small, and more often as the population response grows. The probability of unit firing when the population response amplitude is in a given class interval —that is, the ratio of shaded to unshaded amplitude—is plotted in the lower part of the figure. If unit response and population amplitudes were uncorrelated this function would be a horizontal line at about 0.7. However, it is also clear that correlation of unit and population response is not complete. In other words, the thresholds of the units within the population vary with respect to one another, in addition to their collective (that is, correlated) fluctuation. If this were not so, a particular unit would respond only after all units of lower threshold had responded; therefore its probability of response would be zero if the population response were smaller than a certain value, and would be one if the population response were larger than that value. The lower curve would therefore be a step function.

III. POSSIBLE SOURCES OF THRESHOLD VARIATIONS

FATT and KATZ (7) have found that at motor endplates miniature end-plate potentials occur more or less randomly even though no stimulus is present.

They regard these potentials as being the result of spontaneous firings in the fine terminal branches of a motor nerve. The occurrence of an impulse in the nerve causes simultaneous firing in about a hundred such terminals, giving rise to the normal end-plate potential. Spontaneous firing implies the existence of a local source of varying excitation. FATT and KATZ compute that for fibers with a diameter of 0.1 μ thermal fluctuations in ionic concentrations

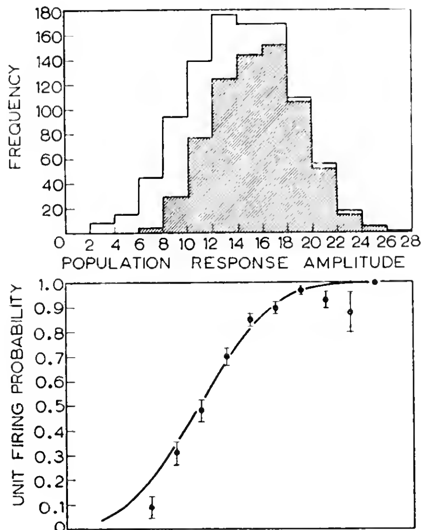


FIG. 8. Top: the upper curve is a histogram of population response amplitudes obtained as in Fig. 7 from *triceps surae* motoneurons by delivering repeated identical shock stimuli to the gastrocnemius nerve. The lower curve (shaded) was obtained from single-unit recordings like those shown in Fig. 7; the number of single-unit responses associated with population responses in each amplitude interval is plotted. Bottom: for a given population amplitude interval the number of single unit responses is divided by the total number of trials in that interval, and the ratio plotted as a function of population amplitude. The interpolated solid curve is a sigmoid fit to the data points and approximates the probability of unit response as a function of population amplitude. Note that when the population amplitude is large, the probability of unit response is large, and when the population response is small, the single unit probability is small, thus signifying a high degree of correlation among the thresholds of different units of the population. From RALL and HUNT (6).

could cause variations of resting potential of 1 mV to 2 mV. Though probably insufficient to produce excitation, such a variation would cause threshold fluctuations and contribute to spontaneous firing.

Both PECHER (2) and HUNT (8) have discussed possible sources of threshold fluctuation. PECHER considers in detail the apparent threshold variation that would result from statistical variations in the number of ions traversing the axon membrane when a constant potential is applied across it. Assuming that the excitatory current that he uses is uniformly distributed over a cross section of the nerve trunk, he concludes that at threshold about a million ions traverse a single nerve fiber. The statistical variation in this number of ions is

given by its square root, leading to a variation of about 0.1 per cent. This is several orders of magnitude below the range of threshold variation that he observed. However, he points out that the number of ions actually effecting excitation is probably considerably less than the value mentioned above and the resultant variability correspondingly greater. PECHER also considers as a possible source of threshold fluctuations local statistical variations of membrane potential, of the sort discussed by FATT and KATZ.

HUNT discusses two classes of possible sources of threshold fluctuation for spinal motoneurons: (a) sources with a local origin such as we have mentioned above, which give rise to an independent component of threshold variation and (b) sources whose effect is felt by many fibers and which therefore produce at least partially correlated variations in threshold. In the latter category are included the effects of activity of spinal interneurons. By using a drug (myanesin), in doses that block transmission through polysynaptic paths without reducing monosynaptic reflex responses, a considerable reduction in the range of variation of population response amplitudes was obtained. On the basis of this result it appears likely that internuncial activity is important in producing correlated threshold changes in spinal motoneurons.

IV. A MATHEMATICAL MODEL

Let us consider a mathematical model which is based on the concept of fluctuating thresholds, and which attempts to derive the ensemble behavior of large numbers of neural elements from assumed properties of neural units in a specific area of the nervous system (9, 10, 11).

This model is based on data obtained from the peripheral auditory system of the cat. When an electrode is placed near the round window of the cochlea, responses to clicks can be detected; such responses contain a component that represents the summated activity of peripheral auditory neurons. Fig. 9 shows such population responses at a number of intensities. In Fig. 10 the average peak-to-peak amplitude of such responses has been plotted as a function of stimulus intensity. The resultant 'intensity function' relates the number of units firing and the intensity of the stimulus.

The present version of the model (11) postulates the existence of several independent populations of neural units; within a population all units are identical. The threshold of a unit is a fluctuating parameter which can be described by a probability distribution; threshold variations in different units occur independently. At a rate of stimulation slower than one per second the 'response no-response' sequence obtained from a single unit is assumed to consist of a series of independent events. Thus we postulate units whose statistical properties resemble those found by PECHER in the frog's sciatic nerve.

The experiments used to test the model fall into three classes: two-click experiments (9, 10), measurements of variability of response amplitude (11), and studies of masking of click responses by noise.

When two clicks are delivered at an interval of less than approximately 100 msec the population response to the second click is smaller than it would be if the first click had not occurred. This effect is more pronounced the stronger the first click and the smaller the interclick interval, as illustrated

in Fig. 11. Consider the ratio of the response amplitude ${}_1R_2$ to a second click and the response amplitude R_2 to the same click presented alone. In Fig. 12 this ratio is plotted, for a fixed second-click intensity, as a function of the intensity of the first click. The parameter is the interval between clicks, $\Delta\tau$. If

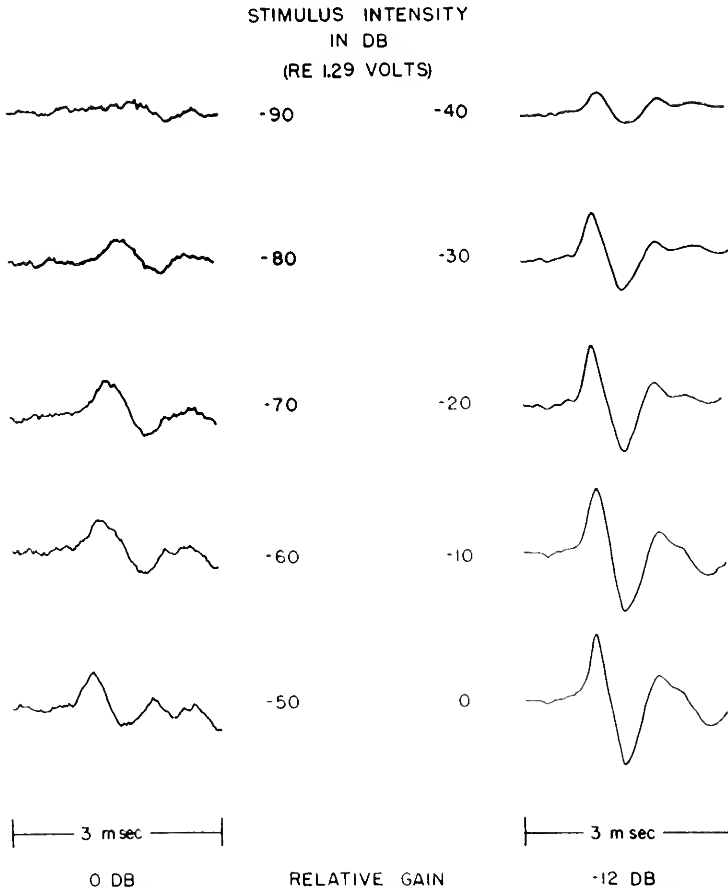


FIG. 9. Ink tracings of responses obtained from an anesthetized cat to clicks over a 90-dB range. The electrode was located near the round window. Note that the voltage gain of the recording equipment was reduced by 12 dB (factor of $\frac{1}{4}$) for stimulus intensities above −40 dB. The first peak represents the summated activity of first-order auditory neurons. With this calibration, click threshold for humans (verbal report) is about −95 dB.

we assume a one-population model, we obtain the result that the ratio ${}_1R_2/R_2$ is linearly related to the intensity function for the first click, provided that the second-click intensity (S_2) and $\Delta\tau$ are held constant. Specifically, we obtain

$$\frac{{}_1R_2}{R_2} = 1 - \frac{R(S_1)}{R_{\max}}\,[1 - g(S_2, \Delta\tau)] \tag{1}$$

Determination of a single intensity function therefore permits us to predict the dependence of this ratio on S_1 for any value of S_2 and of $\Delta\tau$. We may

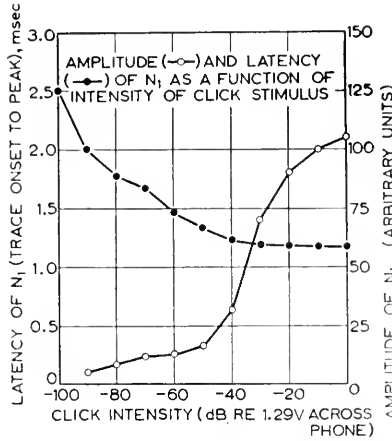


FIG. 10. Intensity function (open circles). N_1 is the first diphasic response component seen in the traces of Fig. 9. The amplitude measurement is made between the positive and negative peaks of N_1. Each plotted point is the median of about ten such measurements.

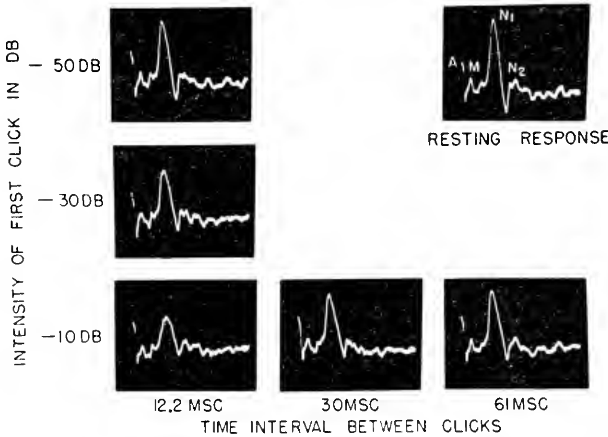


FIG. 11. Two-click paradigm: the responses shown are to a constant intensity (−45 dB) second click. The vertical set shows the effect of varying the intensity of the first click; the horizontal set shows the effect of varying the interval between clicks. Upper right: response to a −45 dB click presented alone. From McGILL (10).

in each case choose one constant, $g(S_2, \Delta\tau)$. Fig. 12 shows a number of fits to the data points which were obtained in this way; S_2 is constant and each curve corresponds to a different value of $\Delta\tau$.

In a second group of experiments the standard deviation of a hundred response amplitudes was computed at each stimulus intensity, and the result was plotted as a function of stimulus intensity. It is readily shown that N

independent units, each with a probability p of firing, will have a standard deviation of total response proportional to $\sqrt{Np(1-p)}$. As a function of p this quantity has minima at zero and one and has a maximum at $p = \frac{1}{2}$. The value of p at any stimulus intensity can be obtained from the intensity function.

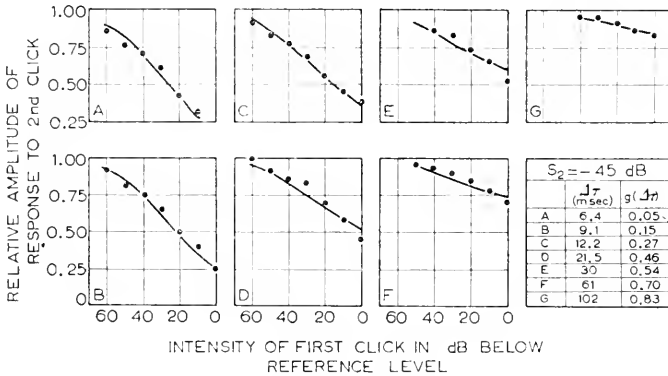


FIG. 12. ${}_1R_2/R_2$ (see text) as a function of first click intensity. In each block this ratio is plotted for a different interclick interval, as indicated at the lower right. The intensity of the second click was −45 dB throughout. The curves are obtained from the first click intensity function and eq. (1); the parameter $g(\Delta\tau)$, whose values are given at the lower right, is chosen in each case to give the best fit to the data. After MCGILL (10).

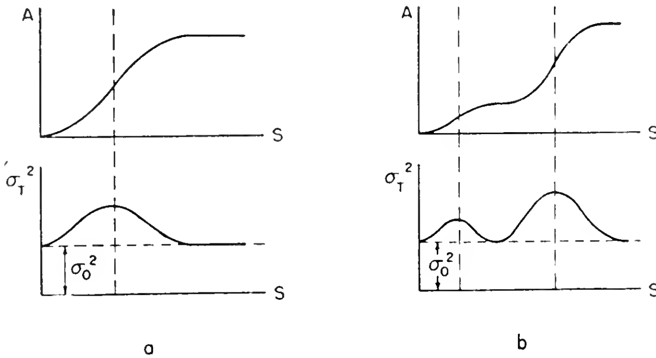


FIG. 13. Intensity function (upper) and the corresponding amplitude variance function predicted by the model: (a) for one population; (b) for two disjoint populations. σ_0 was chosen arbitrarily. Note that a peak of the variability function occurs at the stimulus value at which an intensity function component reaches half its maximum amplitude.

Fig. 13 shows the kind of variability function obtained by assuming one and two disjoint populations; σ_0 is the stimulus-independent component of variability arising from biological and non-biological sources. We have shown (11) that instability in stimulus intensity, which would also lead to a peaked variability function, can account for at most three per cent of the observed variability.

A detailed study of the shape of the intensity function led us to postulate

two populations of neural units, one consisting of 'sensitive' units and one of 'insensitive' units. In the three animals tested, variability measurements over the sensitive range are in good agreement with the theory stated above. One case is shown in Fig. 14. The intensity function and the probabilities obtained from it are shown with the derived standard deviation function. Here, σ_0 is determined from measurements of baseline variability in the absence of a stimulus; N is chosen to give the best fit to the data. Over the sensitive range

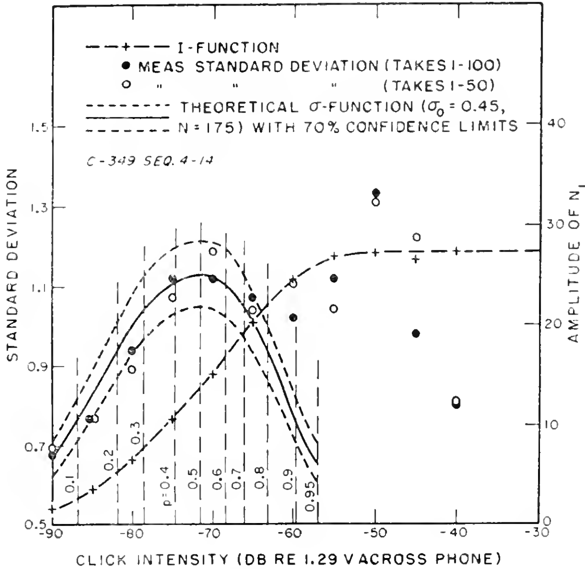


Fig. 14. Comparison of the theoretical variability function (with 70 per cent confidence limits) and the measured values of σ, over the range of initial growth of the intensity function. Each point represented by a solid circle is based on 100 responses; the open circles are based on the first fifty of these responses. The corresponding intensity function, and the probabilities obtained from it, are also shown.

(-100 dB to -60 dB) the data fall within the indicated confidence interval approximately seventy per cent of the time, as they should if the model is correct. Over the insensitive range of the intensity function (-60 dB to 0 dB), the standard deviation shows a complex behavior which cannot be simply reconciled with the idea of a single population over that interval.

The third aspect of this study concerns the masking of the neural responses to clicks by a background noise. Fig. 15 shows the effect of a constant noise level on response amplitude at several stimulus values. In Fig. 16 we have plotted these masked and unmasked intensity functions. The observation was made that a very weak level of continuous noise was sufficient to reduce almost to zero the N_1 response to a fairly intense click. A fixed threshold model would predict masking of only the units whose thresholds are below the noise level. If the threshold fluctuates, however, and does so rapidly, nearly all units of a given population will drop below the noise level and fire in a short

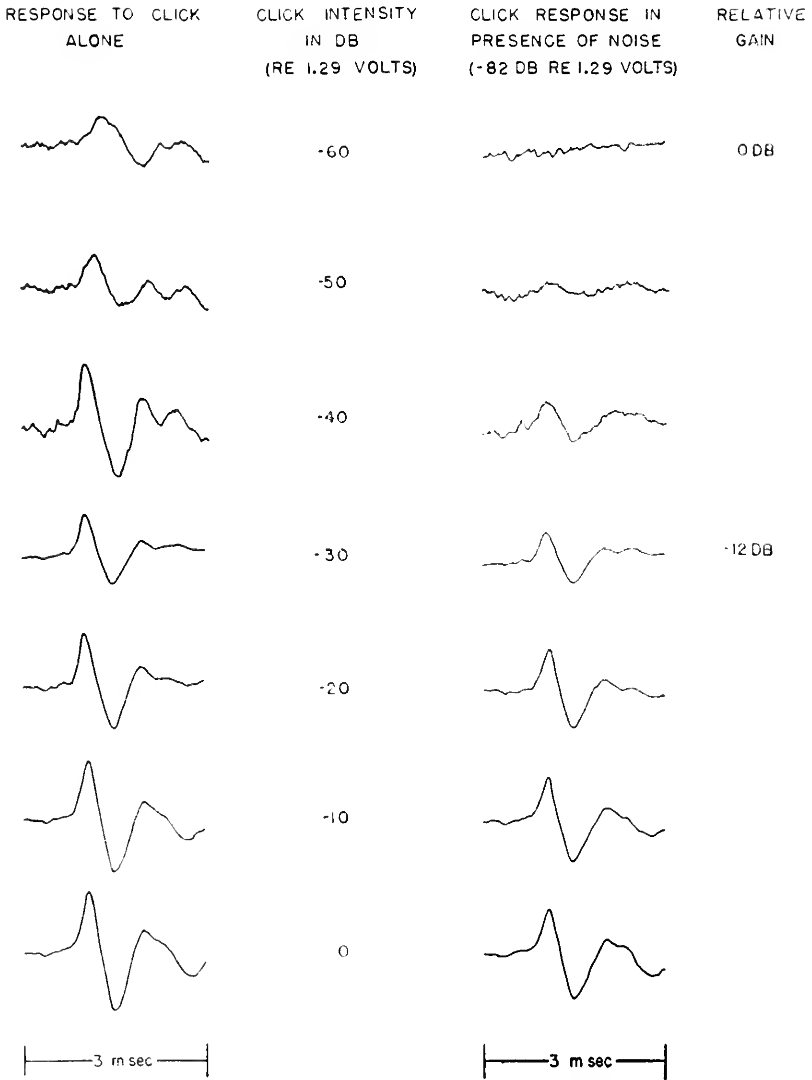


FIG. 15. Ink tracings of responses obtained from an anesthetized cat to clicks over a 60-dB range, with and without background noise; noise level, −82 dB. Note that the voltage gain of the recording equipment was reduced by 12 dB (factor of $\frac{1}{4}$) at a click intensity of −30 dB.

interval preceding the click; this assumes, of course, that the noise level lies within the range of threshold fluctuations of the unit.

* * * * * * * * * * * *

By a quantitative treatment based on these qualitative notions we have been able to show (a) that the hypothesis of a fixed threshold does not account for the observed data and (b) that over the sensitive range of the intensity

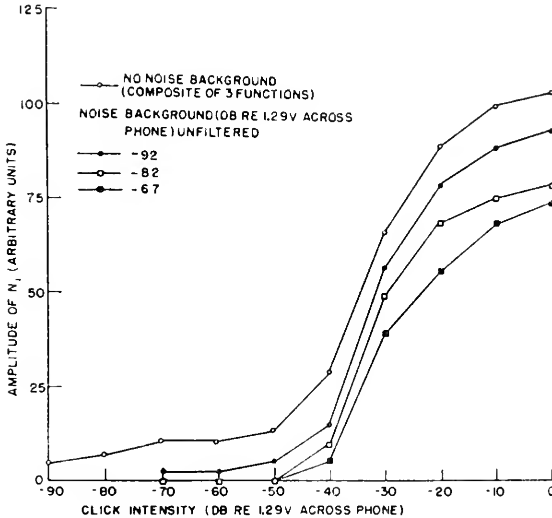


Fig. 16. Intensity functions for clicks, with and without noise background; noise levels −92, −82 and −67 dB. Each point of the masked functions represents the average N_1 amplitude of ten responses to identical stimuli. The upper curve was obtained by averaging the three unmasked functions which correspond to the masked functions shown; thus each point represents the average N_1 amplitude of thirty responses to identical stimuli. Typical data on which these curves are based are shown in Fig. 15.

function a single population of units making threshold 'jumps' at a rate of about 2000 times per second can account for the data. In addition, it is observed that low level noise has little effect on the intensity function over the insensitive range, except to reduce it by the constant contribution of the sensitive population. The need for a division of units into at least two populations is thus confirmed. When the noise level is raised into the insensitive range the observed effect is not nearly so marked, implying either that more than one population is involved in that interval or that the rate of threshold fluctuation is considerably slower than for the sensitive units.

It is noteworthy that population analyses based on two very different experiments, variability and masking, have a great deal in common.

REFERENCES

1. E. A. Blair and J. Erlanger: A comparison of the characteristics of axons through their individual electrical responses. *Amer. J. Physiol.* **106,** 524–564 (1933).

2. C. PECHER: La fluctuation d'excitabilité de la fibre nerveuse. *Arch. Int. Physiol.* **49,** 129–152 (1939).
3. A. HALD: *Statistical Theory with Engineering Applications,* p. 344 (eq. 13.3.6), J. Wiley and Sons, New York (1952).
4. W. A. ROSENBLITH: Some electrical responses from the auditory nervous system. *Proceedings of the Symposium on Information Networks,* Polytechnic Institute of Brooklyn, 223–247 (1954).
5. D. P. C. LLOYD and A. K. MCINTYRE: Monosynaptic reflex responses of individual motoneurons. *J. Gen. Physiol.* **38,** 771–787 (1955).
6. W. RALL and C. C. HUNT: Analysis of reflex variability in terms of partially correlated excitability fluctuation in a population of motoneurons. *J. Gen. Physiol.* **39,** 397–422 (1956).
7. P. FATT and B. KATZ: Spontaneous subthreshold activity at motor nerve endings. *J. Physiol.* **117,** 109–128 (1952).
8. C. C. HUNT: Temporal fluctuation in excitability of spinal motoneurons and its influence on monosynaptic reflex response. *J. Gen. Physiol.* **38,** 801–811 (1955).
9. W. J. MCGILL and W. A. ROSENBLITH: Electrical responses to two clicks: a simple statistical interpretation. *Bull. Math. Biophys.* **13,** 69–77 (1951).
10. W. J. MCGILL: A statistical description of neural responses to clicks recorded at the round window of the cat. Ph.D. Thesis, Harvard University (1952).
11. L. S. FRISHKOPF: A probability approach to certain neuroelectric phenomena. Research Laboratory of Electronics, Massachusetts Institute of Technology, Tech. Rep. No. 307 (1956).

PART III

DETERMINATION OF INFORMATION MEASURES

It is possible (as shown by several papers in this volume) to apply information theory to biology without introducing any actual information measures. Indeed, if one considers that it is very difficult to estimate information measures for living systems, and that the resulting measures are of an irreducibly relative nature, one might wonder whether it is worth-while to take such measures at all. However, it is difficult if not impossible to validate firmly the application of information theory without critical tests based on quantitative measurements; moreover, one hopes to discover lawful relations in the results of the measurements themselves. So, attempts are being made to estimate information contents associated with various biological structures and functions. All the papers in this part are chiefly concerned with such estimations; some from a general point of view, some with regard to particular systems, ranging in complexity all the way from simple molecules to whole men.

H. Q.

CHEMISTRY AND BIOCHEMISTRY AT LOW TEMPERATURES AND DISCRIMINATION OF STATES AND REACTIVITIES*

SIMON FREED

Chemistry Department, Brookhaven National Laboratory, Upton, New York

Abstract—In order to apply information theory to biochemistry and biology at the molecular level it is advantageous to reduce the number of classifications and specifications involved by reducing the temperature of the system. In this way the number of species and states with their reactivities is reduced. At the same time the chemical noise level falls and in consequence a resolution may be obtained between components whose properties are practically indistinguishable at ordinary temperatures. Weakly bonded systems and intermediates become more easily detectable not only because of an increase in their concentration, that is, an increase in their signal, but in addition because the noise level is weaker at the lower temperatuie.

Illustrations are given from chemistry where reactions in solutions proceed at the temperatures approaching that of liquid nitrogen. The information content of irreversible reactions at room temperature may be thought of as being stored in intermediates that participate in reversible reactions at the low temperatures.

Once the properties of the more stable states have been understood, the way is clear for investigating the system in its thermally active states since allowance can be made for the presence of the former. In this way, an ordering of experimentation according to temperature will bring into activity successive components of the system.

Examples have been selected mainly from work on the preservation of biological systems at low temperatures which indicate that biochemical and biological processes may likewise be investigated and that the finer discriminations and specificities associated with lower temperatures may be brought to light in these fields also.

IF we wish to measure a physical property, such as electrical conductivity or viscosity, with an instrument which we have no intention of modifying, there is little point in seeking the information content of the instrument. On the other hand, if we wish to employ chemical substances as probes for uncovering structures of enzymes by means of enzyme-substrate reactions, we are at once confronted by the need of the structural and functional information of our probes. In fact we are discussing properties at the molecular level. Pure substances at this level are mixtures composed of molecules in various energy states with their characteristic configurations, motions, and reactivities. The application of information theory to biology at the molecular level requires therefore a great expansion in the number of categories and specifications. It is to reduce this number in a systematic manner and make these categories more precise that I wish to draw upon the relation that has been recognized between information and entropy which asserts that the amount of information

* Research performed under the auspices of the U.S. Atomic Energy Commission.

required to specify the system will be less at lower temperatures. The system will redistribute itself from higher to lower energy levels so that only the more basic ones remain appreciably occupied. Fewer chemical species are now present and also active. There has been, in a sense, a reduction in chemical noise differing in its frequency spectrum from the continuum characteristic of an electrical conductor. Chemical noise reflects the structural properties of molecules and may consist of dominant discrete frequencies associated with virtual continua of modulations. Usually these represent coupling of the electronic system of the molecule in a given atomic configuration with its

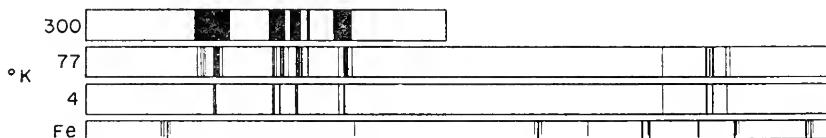


FIG. 1. The variation of absorption spectrum of praseodymium chloride with temperature. Line drawings of visible absorption spectra of crystals of anhydrous praseodymium chloride ($PrCl_3$) at room temperature, at that of liquid nitrogen, and that of liquid helium. Sharper spectra, improved resolution, and fewer lines are evident at lower temperatures. The fewer lines correspond to fewer energy states which are occupied by the praseodymium ions. At room temperature the blocks of diffuse spectra are actually not uniform in intensity but are more intense as a rule in those regions where the spectrum of the crystal at 77°K possessed its most intense line spectrum. The greater diffuseness of the lines and their increased numbers at the higher temperature may be regarded as chemical noise associated with the spectroscopic signals from the more stable states at the lowest temperature.

own vibrations, restricted rotations, etc. If the molecules are complex, fluctuations between different atomic configurations may contribute to the noise. In addition, coupling of the molecule in each of its states with the molecules of its environment in different configurations leads to more and more densely spaced energy levels which I referred to as the continua.

A reduction in temperature removes thermal energy required to activate some motions and effect changes in configurations, and reduces the number of perturbations of a given configuration. Not only are fewer species present but each species is more sharply defined; thus, less information is required for specifying the system than at higher temperature. Clearly, the system is now more specific in its reactions than at higher temperature and its specificity can be related to more sharply defined geometric configurations. The chemical system has become a more precise probe.

The following illustrations have been selected for the simplicity of their phenomena rather than for their direct relevance to biology.

The sharp absorption spectrum of a crystal of a rare earth salt (Fig. 1) shows very clearly that at the lower temperature fewer lines are present; they are sharper and more clearly resolved and the general diffuse background prominent at the higher temperature (not shown in the line drawing of the figure) becomes decidedly weaker. There are then fewer kinds of absorption centers at the lower temperature and, because the stable states are exposed to

more sharply defined environmental fields, there are fewer kinds of perturbations.

An especially vivid example of a solution showing somewhat similar phenomena is given by the fluorescence spectrum of solutions of europium chloride in ethanol at various temperatures (1). The spectra were taken to discover the discrete number of lines in the three separate sets which may furnish the point

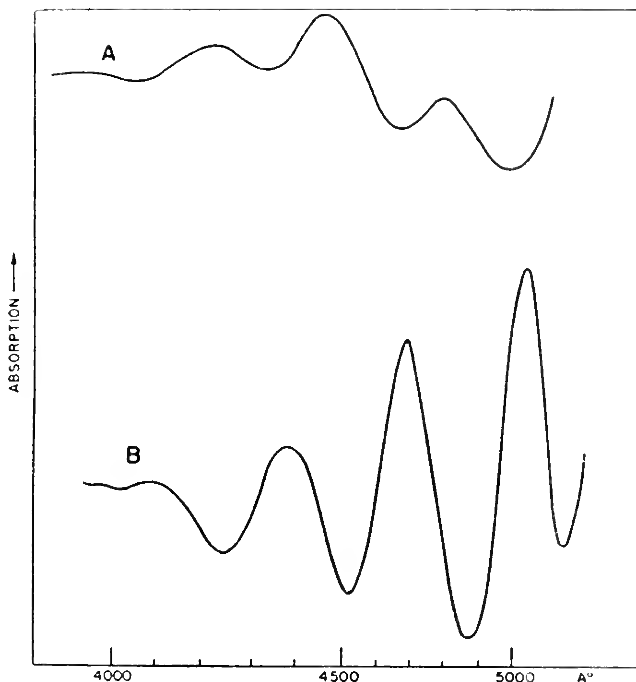


Fig. 2. Absorption spectra of carotene (90% alpha and 10% beta). A—In heptane at room temperature; B—In equal volumes of liquid propane and propene at 77°K.

group symmetry of the electrical fields about europium ion in the solution. It is clear that at room temperature the continuous noise is so great as to make enumeration impossible. As the temperature is lowered a few discrete lines can be resolved with such definiteness that they serve to eliminate some of the possible point group symmetries. At the temperature of liquid nitrogen and even at the temperature of dry ice adequate resolution is clearly achieved and the number of possible symmetries of the environmental fields is reduced to one only.

Figure 2 gives the absorption spectrum of a substance of some biological interest, β-carotene, and illustrates the increased contrast between absorption and transmission at the lower temperature, that is, the increased signal to noise ratio.

Figure 3 is presented to illustrate the resolution into components of what

is apparently a single species at room temperature. The figure reproduces the absorption spectra of chlorophyll *b* in ethyl ether and methanol (2). Our first inclination is to ascribe the differences in the spectra to the perturbations produced on the structure of the chlorophyll molecules by the two types of solvent molecules. Figure 3b is a magnification of the Soret band in the blue

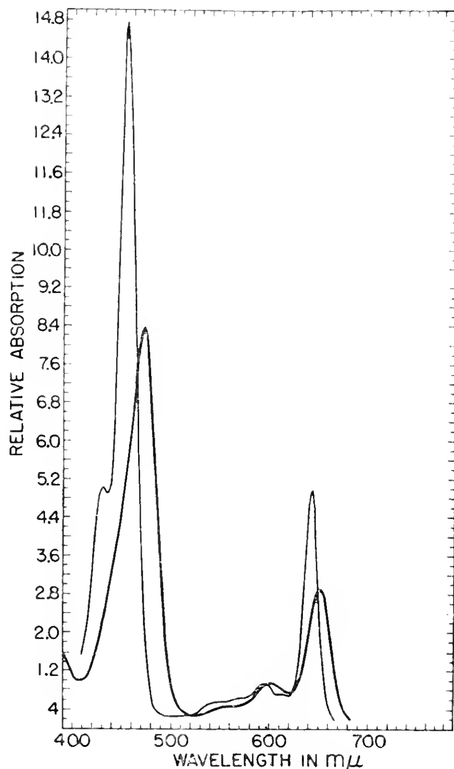


FIG. 3a. Absorption spectra of chlorophyll *b* at room temperature. The thin-lined curve with maxima at shorter wave-lengths represents a solution of chlorophyll in ethyl ether; the thick-lined curve gives the spectrum when the solvent is methanol.

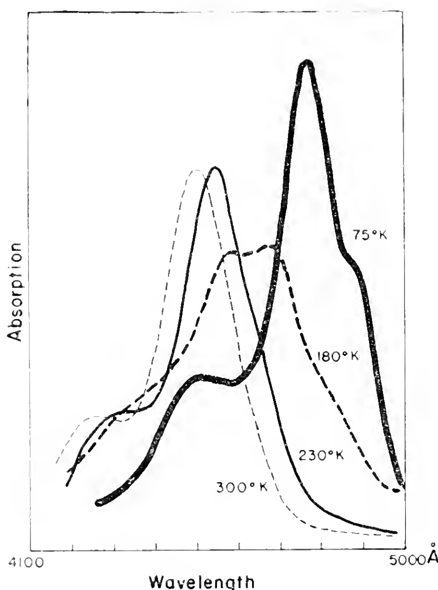


FIG. 3b. The dependence of the absorption spectra of chlorophyll *b* on temperature. Only the Soret band in the blue is shown. Enlarged scale of wave-lengths. At 300°K the solvent is 20% propyl ether, 80% hexane. At the lower temperature it is 20% propyl ether, 40% propane, and 40% propene. The hexane was substituted at 300°K for the hydrocarbons propane-propene since they are normally gases at room temperature.

region and shows that a solution of chlorophyll *b* in ether is really a mixture of two species (etherates) in equilibrium with each other in roughly equal amounts and clearly resolved at 180°K. A study of the dependence on temperature of the absorption spectrum of chlorophyll *b* in methanol reveals that in this solvent, chlorophyll *b* also exists as a mixture of solvates which are about equal in concentration at room temperature and together they yield the composite spectrum. However the spectrum of each alcoholate differs very little in shape from that of each etherate. Fig. 4 illustrates a form stable at a lower

temperature reacting to produce reversibly a stable intermediate but at still higher temperature ending in an irreversible reaction.

The following specific observations may prove worthwhile in illustrating what is probably a rather common phenomenon. Chlorophyll *b* dissolved in

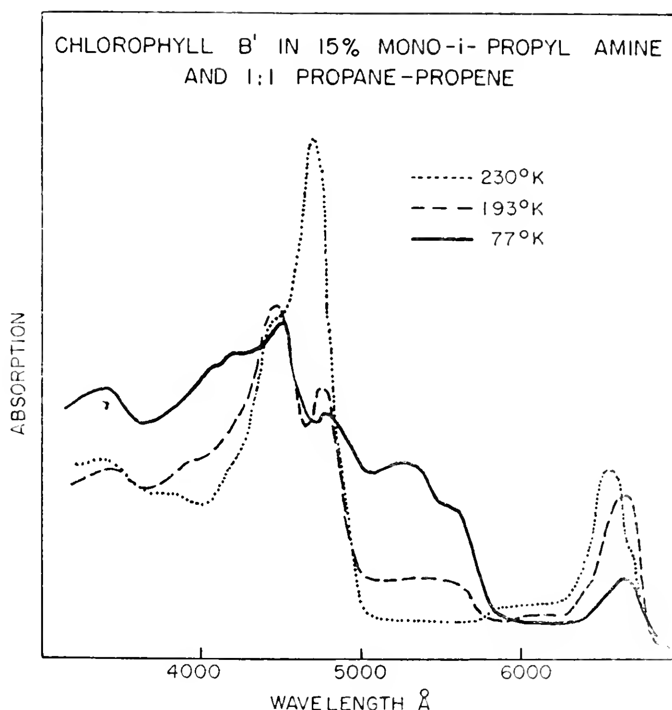


FIG. 4. Chlorophyll *b'* in 15% mono-*i*-propyl amine in 1 : 1 propane-propene. To show the presence of the red-brown intermediate stable at 193°K which is in equilibrium with the original chlorophyll. At temperatures higher than about 235°K, an irreversible reaction occurs.

ether is deposited as a green powder by pumping off the ether at room temperature. When the temperature of the powder is reduced to that of dry ice (about 193°K) and propylamine is condensed upon it at this temperature, it dissolves quickly, forming a red solution. Note in Fig. 4 the new absorption between 5000 Å and 6000 Å. A rise in temperature transforms the color into the green of chlorophyll with its characteristic spectrum which reverts back reversibly to the red substance when the temperature is reduced. However, if the temperature is kept any length of time at about 235°K or higher, an irreversible reaction sets in. For example, at room temperature the red color lasts only a fraction of a second. This evanescent red color is produced in the well known phase test for chlorophyll.

Figure 5 represents a chemical reaction which appears rapid even between 167°K and 75°K. Chlorophyll *b* dissolved in di-*iso*-propylamine is undergoing

transformation probably in an acid-base reaction. The quick readjustment to equilibrium is shown by the interchange in relative intensities of the bands in the red region. The band furthest towards the red grows in as the temperature is reduced, at the expense of the band near it toward shorter wavelengths.

That these reactions occur rapidly at such temperatures is not very surprising since little heat of activation is required for this type of reaction. Figure 6 depicts a type of oxidation-reduction at low temperatures. When iodine

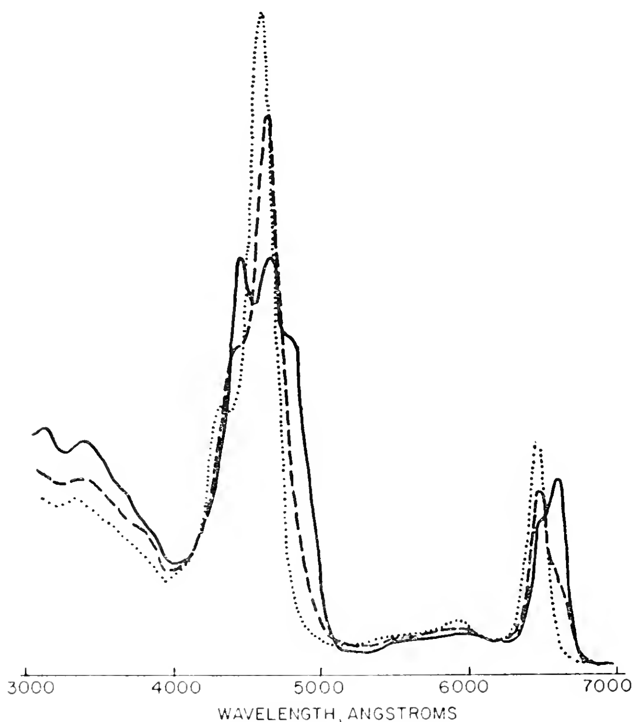


FIG. 5. Chlorophyll *b* in 15% dipropylamine diluted with equal proportions of propane and propene. A chemical readjustment toward equilibrium occurs between 170°K and 75°K.

is finely divided it rapidly dissolves in isoprene at the temperature of dry-ice, 193°K. A brown solution forms at the solid-liquid interface but it decolorizes very quickly, becoming colorless a little distance from the iodine surface. In the light of other investigations it was surmised that the solution is brown because of the presence of a 1:1 (molecular iodine–hydrocarbon molecule) addition compound which possesses a characteristic absorption band in the ultraviolet region. To build up any appreciable concentration of this compound it would evidently be necessary to make solutions of iodine in isoprene below 193°K. When a solution of isoprene in propane (to which propene had been added to increase the solubility of isoprene) at the temperature of liquid nitrogen (77°K) is mixed with a solution of iodine in propane and propene,

the new band anticipated in the ultraviolet does not appear within a day or two. Figure 6 indicates what happens when such a solution is warmed. At 146°K the absorption band shown is due to the iodine-propene molecular addition compound which has been identified in a previous experiment. At 150°K appears the anticipated new band arising from the compound iodine–isoprene. At 154°K, this band quickly disappears irreversibly and at the same time decoloration of the solution occurs. The molecular iodine has been removed, presumably by the halogenation of the double-bond system of isoprene, just

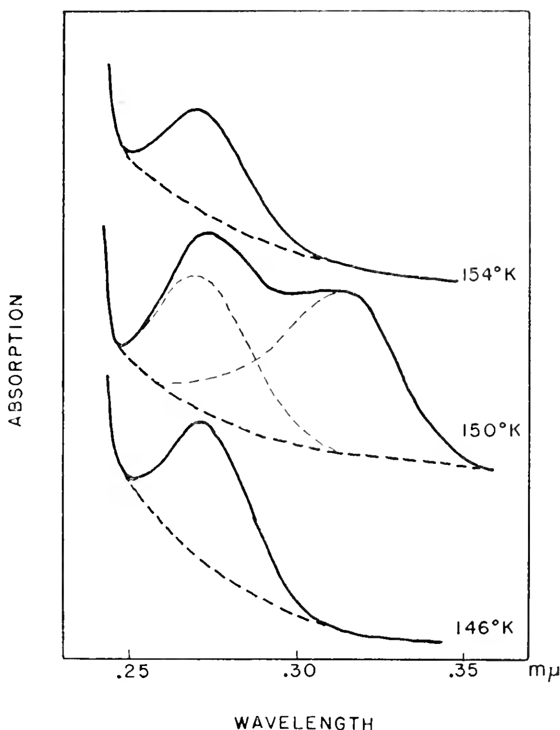


Fig. 6. Isoprene dissolved in 1 : 1 propane-propene to which iodine dissolved in 1 : 1 propane-propene has been added. The new absorption band which appears at 150°K is due to a 1 : 1 molecule addition compound of the iodine to isoprene. Its disappearance at 154°K is due to an irreversible reaction, probably halogenation across the double bond.

as had occurred when solid iodine reacted with isoprene at the temperature of dry ice. This oxidation appears to require the prior formation of the intermediate molecular addition compound stable at about 150°K at the concentrations employed.

By investigating the properties and reactions from the lowest practicable temperature upward we would observe the appearance of new thermally activated states and their subsequent reactions.

In analogy with the phenomena illustrated we would expect that a knowledge of biochemical and even biological processes of considerable value may be gained by investigations at low temperature. Support for these expectations comes mainly from recent investigations directed toward the preservation of cells, tissues, and entire organisms. Even more cogent for our purposes are the instances of partial preservation at low temperatures which becomes more effective at still lower temperatures. Unless explicit references are given, the following examples are drawn from the excellent review by AUDREY U. SMITH (3). For example, H. F. SMART found that twenty-one species of bacteria, yeasts, and molds continued to multiply in frozen media at 264.1°K. SIZER and JOSEPHSON found that lipase was active at 248.5°K, tryptic digestion proceeded at 258°K, and that invertase continued to hydrolyze sucrose at 255°K. At 203°K, however, they could detect no hydrolysis during several weeks. In the preservation of red blood cells, about ten per cent deterioration occurs per year at dry ice temperature, 193°K, but scarcely any loss is incurred when they are kept at the temperature of liquid air, 80°K. Ovarian tissue failed to survive nine days at 193°K but survived more than a year at 80°K under otherwise similar conditions (4). Revival of rats after cooling to 273.5°K was reported by ANDJUS (5, 6).

Irreversible reactions are then clearly progressing at low temperatures, in red blood cells and ovarian tissue at 193°K and at somewhat higher temperatures in the enzymatic reactions. If the simple reactions such as those of isoprene and iodine, chlorophyll and propylamine serve as models, the irreversible reactions are preceded in their first and intermediate stages by reversible reactions at still lower temperatures.*

BECQUEREL found that rotifers, spores of bacteria, non-sporing bacteria, algae lichens, mosses, and seeds of higher plants, after having been dried in a vacuum of 10^{-5} mm Hg over barium oxide, could be successfully kept at the temperature of liquid helium (4°K). PARKES showed that human spermatozoa survived exposure and storage at 80°K. Ovarian, testicular, pituitary, and adrenal tissue have given functional grafts after storage at 80°K, especially if glycerine was added. LUYET established that vinegar eels, spermatozoa muscle fibres of frogs, and hearts of embryonic chicks could be revived after sudden cooling to the temperature of liquid air (80°K). It is then not surprising that enzymes have been cooled to such temperatures without loss of subsequent potency. It would seem then that a number of biochemical and biological processes are available for study at low temperatures.

I shall consider both homogeneous and heterogeneous solutions. The first implies that solvents must maintain all the reactants in solutions fluid at low temperatures. It would seem well worthwhile to employ conventional solutions at as low temperatures as possible, and aqueous systems near zero degrees or under supercooled conditions. It has been shown (8) that proteins

* Lovelock (7) ascribes the deterioration of red cells to a physical mechanism rather than to a chemical process, namely, that the dissolution of lipoprotein and other components of the cell membrane proceeds more rapidly than the biochemical processes can repair them at the low temperature. Since the lipoprotein etc. is presumably bound as an integral part of molecules composing the membrane material, the physical process may also be initiated by reversible chemical transformations.

such as enzymes are soluble in some non-aqueous solvents and that a few enzymes can be recovered with virtually their full potency. Since some of the solvents have melting points below that of water they can be utilized for investigations of solutions of proteins at relatively low temperatures. It appears entirely possible that had the solution process been carried out at lower temperature a larger fraction of the enzymes would have been recovered without deterioration. Indeed it may prove fruitful to undertake studies at low temperatures of the first stages of reactions which are toxic at ordinary temperatures since the toxic substances may be removed at temperatures so low that little permanent injury is done to the enzyme or organism.

In analogy with the dissolution of finely divided chlorophyll and iodine by solvents at low temperatures it is to be expected that at low temperatures heterogeneous reactions are also possible between substances in solution and biological materials having high specific areas. Ready-made for such reactions with solutions seem sections of tissue with water removed by freeze-drying. Likewise BECQUEREL's procedure of removing water by pumping at room temperature would prepare material for reaction at low temperature. Some of the reactions with the surfaces constitute a generalized staining. Many staining processes are acid-base reactions and would be expected to be rather rapid at low temperatures. As has been remarked, molecular steric factors are as a rule more specific at the lower temperatures in general; hence finer discriminations between structures within the surfaces are to be anticipated.

REFERENCES

1. E. V. SAYRE, D. G. MILLER, and S. FREED: Symmetries of electric fields about ions in solutions. Absorption and fluorescence spectra of europic chloride in water, methanol, and ethanol. *J. Chem. Phys.* **26,** 109–113 (1957).
2. D. G. HARRIS and F. P. ZSCHEILE: Effects of solvent upon absorption spectra of chlorophylls A and B. *Bot. Gaz.* **104,** 515–527 (1943).
3. A. U. SMITH: Effects of low temperatures on living cells and tissues. In: *Biological Applications of Freezing and Drying*, ed. by R. J. Harris, 1–62, Academic Press, New York (1954).
4. A. S. PARKES and A. U. SMITH: Regeneration of rat ovarian tissue grafted after exposure to low temperatures. *Proc. Roy. Soc.*, (B) **140,** 455–470 London (1953).
5. R. K. ANDJUS: Sur la possibilité de ranimer le rat adulte réfroidi jusqu'a proximité du point de congélation. *C.R. Acad. Sci., Paris* **232,** 1591–1593 (1951).
6. R. K. ANDJUS and A. U. SMITH: Revival of hypothermic rats after arrest of circulation and respiration. *J. Physiol.* **123,** 66–67 (1954).
7. J. E. LOVELOCK: Physical instability and thermal shock in red cells. *Nature, Lond.* **173,** 659–666 (1954).
8. M. J. LOISELEUR: Sur quelques propriétés des protéides en solutions organique. *Bull. Soc. Chim. Biol.* **14,** 1088–1100 (1932).

J. J. KATZ: Anhydrous hydrogen fluoride as a solvent for proteins and some other biologically important substances. *Arch. Biochem. Biophys.* **51,** 293–305 (1954).

E. D. REES and S. J. SINGER: A preliminary study of the properties of proteins in some nonaqueous solvents. *Arch. Biochem. Biophys.* **63,** 144–159 (1956).

DISCUSSION

MAHLER: I can see where this might be useful in the study of the rate of formation of enzyme-substrate complexes. This is a reaction which proceeds much too rapidly to be measured by most ordinary techniques. It is only with very rare and very stable enzyme complexes and by using very interesting and very sensitive experimental devices that CHANCE*, for instance, has been able to study this at ordinary temperatures. But if one can find the right kind of solvent for both substrate and enzyme—there is no reason to assume that some of these solvents might not work—one might be able spectroscopically to study the rate of formation of enzyme-substrate complexes at low temperatures.

* B. CHANCE and G. R. WILLIAMS: The respiratory chain and oxidative phosphorylation. In: *Advances in Enzymology*, ed. by F. F. Nord **17**, 65–134. Interscience, New York. (1956).

INFORMATION CONTENT OF TRACER DATA WITH RESPECT TO STEADY-STATE SYSTEMS*

MONES BERMAN and ROBERT L. SCHOENFELD
Division of Biophysics, Sloan-Kettering Institute, New York

Abstract—A method for the quantification of information in data from tracer experiments on steady-state systems is presented. It is shown that if the system is represented by n compartments a point in an n^2 dimensional space can serve to represent a specific model. Furthermore, uncertainty about the system due to statistical fluctuations and incomplete data can be represented by regions in the n^2 dimensional hyperspace. A unit of information for such a system is defined and serves as a measure of the amount of information necessary to determine the system to within a desired accuracy.

In order to express the data in terms of the generalized n^2 dimensional space, a set of invariants is defined for the data. A concise matrix relation is shown to exist between the invariants of the data and the parameters that characterize the compartmental system. The matrix relation allows mappings between the data and the system.

The method presented is applicable to any compartmentalized system that shows linear kinetics.

I. INTRODUCTION

THIS paper is concerned with the quantification of information contained in data from tracer experiments performed on steady-state biological systems. In general, the same set of data may be analysed in terms of different systems of various degrees of complexity. To define the information content of the data, therefore, it is necessary to specify the system in terms of which the data are to be analysed.

It can be assumed for many tracer experiments that the system† consists of a discrete number of compartments (or pools) each representing a localization or chemical state of the labeled material, with exchange of molecules between compartments. The rate of exchange of the unlabeled molecules between compartments is in general a non-linear function of the amounts of material in the compartments. If, however, the system is in a steady state and the amount of the tracer is sufficiently small compared to its unlabeled isotope, the rate of exchange of the tracer may be treated as a linear function of the amounts of labeled material in the compartments (1).

The problems that arise in treating the data of tracer experiments are: first, to define the information content in the data, and second, to translate the information in the data into values of the system parameters (the turn-over rates of the compartments). In addition, it is desirable to have a measure of

* This work was supported in part by the U.S. Atomic Energy Commission Grant AT(30-1)-910.

† For this paper, the word 'system' will be used to mean a specific number of compartments independently of how they are interconnected. The word 'model' will refer to a specific configuration of the system.

uncertainty in the values determined for the system parameters. The uncertainty in these values arises from the fact that the collected data may not be sufficient to define the system completely and that the collected data have associated fluctuations.

A method for the quantification of the information in data and the systematic formulation of models consistent with it is presented here. The information content in the data is expressed by a set of invariants, and a concise matrix relation is shown to exist between the invariants of the data and the system parameters. Uncertainties in the data due to incompleteness or fluctuations are mapped into a generalized co-ordinate space which also represents the degrees of freedom of the system parameters and their uncertainty. The uncertainties in the data are expressed in terms of regions in the generalized co-ordinate space in such a way as to suggest a criterion for their quantification with respect to the system.

II. DATA INVARIANTS AND SYSTEM PARAMETERS

The response of the system to a tracer injected into any one compartment can be expressed in terms of the amounts of tracer in the various compartments as a function of time. If we define the probability per unit time for a transition from any compartment i to compartment j as λ_{ji}, then the kinetics of the tracer in the ith compartment of an n compartmental system can be represented by the following set of differential equations:

$$\frac{dq_i(t)}{dt} = -\lambda_{ii}q_i(t) + \sum_{\substack{j=1 \\ j\neq i}}^{n} \lambda_{ij}q_j(t) \qquad (i = 1, 2, \cdots, n) \tag{1}$$

where $q_i(t)$ is the amount of tracer material in the ith compartment at time t and

$$\lambda_{ii} \geqq \sum_{\substack{j=1 \\ \neq i}}^{n} \lambda_{ji} \tag{2}$$

is the probability per unit time that any molecule in compartment i will leave that compartment.

The inequality sign expresses the possibility that a molecule may leave the entire system from compartment i as in the case for open systems.

The solution of the set of differential equations (1) is:

$$q_k(t) = \sum_{j=i}^{n} A_{kj}\, e^{-\alpha_j t} \tag{3}$$

In a recent paper (2) we have pointed out that data expressed in the form of equation (3) have the following properties:

(a) There are at most n α_j in the data and these are invariants of the system and independent of the initial conditions or site of measurements.

(b) The A_{kj} represent n^2 independent variables in the data. Specification of the initial conditions reduces the A_{kj} to $(n^2 - n)$ independent variables which are a function of the system parameters only. The A_{kj} thus represent $(n^2 - n)$ invariants of the system parameters.

(c) The n α_j and n^2 A_{kj} comprise a necessary and sufficient set of data to define uniquely the parameters of the system.

(d) A simple matrix relation (3) exists between the A_{kj} and α_j of the data and the λ_{ij} of the system. This relation can be written:

$$|\lambda|\,|A| = |A|\,|\alpha| \tag{4}$$

or

$$|\lambda| = |A|\,|\alpha|\,|A^{-1}| \tag{5}$$

where

$$|\lambda| = \begin{vmatrix} \lambda_{11} & -\lambda_{12} & -\lambda_{13} & \cdots \\ -\lambda_{21} & \lambda_{22} & -\lambda_{23} & \cdots \\ -\lambda_{31} & -\lambda_{32} & \lambda_{33} & \cdots \\ \cdot & \cdot & \cdot & \cdots \end{vmatrix} \qquad |A| = \begin{vmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \cdot & \cdot & \cdot & \cdots \end{vmatrix}$$

$$|\alpha| = \begin{vmatrix} \alpha_1 & 0 & 0 & \cdots \\ 0 & \alpha_2 & 0 & \cdots \\ 0 & 0 & \alpha_3 & \cdots \\ \cdot & \cdot & \cdot & \cdots \end{vmatrix}$$

Equation (5) expresses the system parameters in terms of the invariants in the data. If these invariants are known, the fractional turnover rates, λ_{ij}, can all be determined. However, in most cases the experimental data are incomplete in that certain of the A_{kj} and α_j are not known. For these cases, an infinity of models mathematically consistent with the data can be obtained from equation (5) by inserting arbitrary values for the unknown A_{kj} and α_j, preserving the initial conditions and other constraints in the data. Most of these arbitrary models, however, will be physically meaningless because some of the fractional turnover rates will be negative. Consequently, it is necessary to investigate what range of values of the unknown A_{kj} and α_j correspond to physically meaningful models. This can be done by relating variations in A_{kj} and α_j to variations in the λ_{ij}.

One may define (2) a matrix $|P|$ in such a way that the product $|PA|$ will preserve the known A_{kj}. The number of variables in $|P|$ will be equal to the degrees of freedom in the A_{kj}. If both sides of equation (4) are premultiplied by the matrix $|P|$ this equation can be rewritten:

$$|P\lambda P^{-1}|\,|PA| = |PA|\,|\alpha| \tag{6}$$

which is of the form

$$|\lambda'|\,|A'| = |A'|\,|\alpha| \tag{7}$$

where

$$|A'| = |PA| \tag{8}$$

$$|\lambda'| = |P\lambda P^{-1}| \tag{9}$$

Equation (9) expresses a mapping of the matrix $|\lambda|$ corresponding to variations in the unknown A_{kj} only. It also represents a general solution of all models mathematically consistent with the data in terms of a minimum number of variables. This solution is expressed in terms of an arbitrary model represented by the matrix $|\lambda|$.

Similarly, we can define a matrix $|D|$ so that the product $|\alpha D|$ will preserve all the known α_j. Incorporating this into equation (4), we get

$$|\lambda A D A^{-1}|\,|A| = |A|\,|\alpha D| \tag{10}$$

which is of the form

$$|\lambda'|\,|A| = |A|\,|\alpha| \tag{11}$$

where

$$|\alpha'| = |\alpha D|$$

$$|\lambda'| = |\lambda A D A^{-1}| \tag{12}$$

Equation (12) represents a mapping of the matrix $|\lambda|$ in terms of the variations in the unknown α_j only.

By applying the restriction that every fractional turnover rate must be positive,

$$\lambda'_{ij} \geqq 0$$

$$\lambda'_{jj} \geqq \sum_{\substack{i=1 \\ i \neq j}}^{n} \lambda'_{ij} \tag{13}$$

equations (9) and (12) limit the range of values of the variables in the matrices $|P|$ and $|D|$. Since these variables are all independent, they represent a co-ordinate space of dimension equal to their number. Every point in this space specifies a set of values for the variables in the matrices $|P|$ and $|D|$ and, thus, defines a model through equations (9) and (12). The restrictions on the range of values of the variables as expressed by equation (13) correspond to a region in the co-ordinate space in which all physically meaningful models must lie.

The choice of the starting point for the transformations indicated above is completely arbitrary and does not affect the final result. Any mathematically consistent model leads to a region in the mapping space corresponding to proper physical models.

III. UNCERTAINTY MAPPINGS IN GENERALIZED SPACE

We now wish to examine the problem from a somewhat different point of view. The system is represented by n^2 λ_{ij}, generally independent of each other. We can, therefore, consider the λ_{ij} to represent an n^2 dimensional space, and any point in that space as a specific model of the system. It was also indicated earlier that the data could be represented by a set of invariants composed of n α_j and $(n^2 - n)$ A_{kj} or a total of n^2 invariants. Hence, the transformation from the data space to the λ_{ij} space is dimensionally consistent and unique.

This means that a complete set of A_{kj} and α_j corresponds to a point in the $|\lambda|$ space, and vice versa. By definition, however, the values of the λ_{ij} must all be positive. Consequently all the models must lie in a restricted region of the $|\lambda|$ hyperspace. This restriction carries over to the data space, limiting the region in which the A_{kj} and α_j may lie.

Any specified A_{kj} or α_j implies a one dimensional constraint in the data space. This carries over as a one dimensional constraint in the $|\lambda|$ space, and restricts all models to a surface in the hyperspace. If, however, the value of A_{kj} or α_j is known only within a certain range, the surface has correspondingly a certain thickness.

When several A_{kj} or α_j are known, the dimensions of the space in which all models must lie is reduced by a corresponding number. Statistical uncertainties

for any of the known values correspond to similar uncertainties along the appropriate co-ordinates in the hyperspace.

Thus, if all A_{kj} and α_j are known exactly, a point in the hyperspace of n^2 dimensions specifies the model. If all the data are known to within a certain statistical precision, the most likely model is estimated as a point in the n^2 dimensional space surrounded by a region that corresponds to the statistical uncertainty. If some A_{kj} or α_j are unknown, the corresponding dimensions in the n^2 dimensional hyperspace extend to the limits imposed by the relation that all λ_{ij} are positive.

IV. UNIT OF UNCERTAINTY

Based on the point of view presented, we can define a unit of uncertainty to be a certain volume of the hyperspace. The size of the volume so defined is arbitrary; it may correspond to a volume that is equivalent to the actual standard deviation in the data, or to some convenient standard deviation that may serve as a reference. The information necessary to define the system can then be expressed as the number of binary choices, or bits of information, necessary to reduce the total uncertainty space to the size of a defined unit.

V. CONCLUSION

The treatment presented provides a framework in which information in data from tracer experiments on steady-state systems can be quantified in terms of a compartmental system and its parameters. Before the information can be quantified, however, a number of compartments has to be chosen for the system. Unless this is known from independent sources, the method in choosing the number of compartments is based on the minimum number of exponential terms that 'reasonably' describe the data. This, at present, is by no means a unique procedure.

It was shown in this treatment that a model representing the system can be expressed as a point in a generalized co-ordinate space, and that any uncertainty in the system can be represented by a certain region in that space. The nature of the uncertainty (whether incomplete data or statistical fluctuations in the data) did not matter in the treatment.

There is, however, one difference in the regions of the hyperspace corresponding to these two sources of uncertainty. The difference is in the probability that any model in the region represents the true system. In the case of incomplete data, the probability density over the entire region is assumed constant; that is, every model in the region is considered equally probable. In the case of statistical fluctuations, however, a certain point or unit volume represents the most likely model, and the rest of the points or unit volumes decrease in probability in a manner governed by the statistics of the data.

The region in the $|\lambda|$ hyperspace can serve to define the information content in the data of the system as a whole or of each parameter of the system, namely the turn-over rates, separately. The latter can be obtained by investigating their values over the bounded region.

One need not necessarily deal with all the dimensions of the hyperspace. One can express the uncertainties in terms of a subspace whose dimensions are equal

to the degrees of freedom of the system, as implied by equations (9) and (12). In this case, however, the statistical variations of the collected data cannot be represented since their dimensions are omitted. Any new data to be collected, however, can be represented in this subspace. The significance of any new data can also be evaluated by the relative reduction in the size of the region in the subspace. A unit of uncertainty may be defined for this subspace as was done for the hyperspace.

In references (1) and (2) it was shown how information about the system from steady-state measurements and thermodynamic considerations can be combined with tracer data to form a unified methodology in reducing the uncertainty about the system. The treatment presented here can be extended to include such additional information.

Whereas the concepts presented here are relatively simple, the application to specific problems involves considerable work. One can handle two or three compartmental systems with few degrees of freedom fairly easily using a desk calculator. The handling of more complex systems becomes quite time consuming. It is hoped that a programming of this on digital computers can be worked out for routine applications.

REFERENCES

1. M. BERMAN: The formulation of biological models from tracer and steady-state data. Ph.D. Thesis, Polytechnic Institute of Brooklyn (unpublished) (1957).
2. M. BERMAN and R. SCHOENFELD: Invariants in experimental data on linear kinetics and the formulation of models. *J. Appl. Phys.*, **27**, 1361–1370 (1956).
3. H. MARGENAU and G. M. MURPHY: *The Mathematics of Physics and Chemistry*, chap. 10, Van Nostrand, New York (1943).

THE DOMAIN OF INFORMATION THEORY IN BIOLOGY*

HENRY QUASTLER
Brookhaven National Laboratory, Upton, New York

IN THE proper course of events, a theory is introduced to account for a specific body of facts; then nobody will presume to expatiate upon the domain of the theory. With information theory and biology, the situation is less simple. The modern development of the theory stems largely from C. E. SHANNON's concern with certain problems of communication engineering (1). I have heard Shannon say that he was somewhat dubious about the extension of his results to remote fields, and that he felt that people working in other disciplines might do better to develop their own theories. This is not what happened. Shannon's theory has been taken up with enthusiasm by psychologists, linguists, historians, planners, librarians, sociologists, and by biologists with a wide variety of interests. Motives for such generalizations were supplied by WIENER, who pointed out that all control (in the animal and in the machine) depended on communication, and that all communication involved measurable quantities of information (2); and by WEAVER, who emphasized the great generality of the information concepts in a searching study (1).

It appeared then that information theory was a tool made to order to deal with a vast variety of problems. This variety, however, is not limitless. Therefore, a discourse on the domain of information theory is indicated. One part of this discourse will deal with the negative domain, or with some of the limitations of the theory. The other part will be concerned with positive applications; it is largely an attempt to give clearer definition to the somewhat vague hopes most people have when proposing to apply information theory.

It is curious that applied information theory produces rather violent reactions, some of them negative. Certainly, it is entirely possible that every biologist who works with information theory, or any other systems theory, is wasting his time. But this, of course, applies to anybody who works with a new theory. It is difficult to see how applying information theory should irritate people—unless the cause should be the very pleasure of gently playing with the theory. Every scientist is aware that there is a 'difference between the labor of thought, and the sport of musing', and knows well the danger inherent in the latter. To go on with Dr Johnson: 'There is nothing more fatal to a man whose business is to think, than to have learned the art of regaling his mind with those airy gratifications .... This is a formidable and obstinate disease of the intellect, of which, when it has once become radicated in time, the remedy is one of the hardest tasks of reason and of virtue. Its slightest attacks, therefore, should be

* Research carried out at Brookhaven National Laboratory under the auspices of the U.S. Atomic Energy Commission.

watchfully opposed' (from *The Rambler*). Is this why so many scientists do not mind too much having collected a lot of useless data but dread to be found working with a useless theory?

I. APPLICATIONS

Every kind of structure and every kind of process has its informational aspect and can be associated with information functions. In this sense, the domain of information theory is universal—that is, information analysis *can* be applied to absolutely anything. The question is only what applications are useful.

1. *Use of Basic Concepts*

The basic concepts of information theory—measures of information, of noise, of constraint, of redundancy—establish the possibility of associating precise (although relative!) measures with things like form, specificity, lawfulness, structure, degree of organization. This alluring promise has introduced the information concepts into the thinking of many biologists. The results of conceptual applications range from harmless modernisms of language to very serious reasoning. In particular, the information concepts seem to lend themselves readily to dealing with the problems of emergence and destruction of order in complicated systems.

The problem of emergence of order is usually treated in terms of Darwinian machines, large more or less random assemblies of parts which can both function and, in some manner, register the results of their functioning. The resulting feedback loop produces some order amazingly fast (3, 4). The theory of random networks is a very active field, and some very competent men expect that the main contribution of information theory to biology (and to other fields concerned with very complicated systems) will come from this endeavour.

Closely related is the problem of destruction of orderliness. In biology, this is the problem of aging and decay; it is the topic of a major fraction of this conference (5, 6, 7).

2. *The Representation Theorem*

The use of the basic concepts of information theory becomes more powerful if one considers that the behavior of information measures follows certain rules; these rules are the theorems of information theory. There are two basic theorems which I like to call the 'representation theorem' and the 'noise-and-redundancy theorem'. The first has to do with the possibility of representing one kind of information by another kind of information. There are absolutely no qualitative limitations as to how information can be represented; but, there is a quantitative limitation: any physical entity can assume only a limited number of distinguishable states, and this limits the degree to which it can represent information. This degree is further modified by the rules of selecting successive states. The applicability of the representation theorem depends to a high degree on knowing the process by which states are selected.

The representation theorem applies every time information is transferred—because the transfer does involve representation of the information existing

in the transmitter, in the medium and, finally, in the receiver. It can thus be stated as follows: A source cannot transmit more information than it has, a receiver cannot register more information than it can display. This sounds trivial, but the point is that information contents can be precisely estimated in ways which are not trivial. The representation theorem implies that it is possible to establish an upper bound of the flow of information simply by investigating the terminals. It is, thus, a *one-sided conservation principle;* being one-sided, it is not as strong as the two-sided conservation principles which are so commonly used in physics. It becomes stronger in situations where one may assume that the inequality approaches an equality.

There are two conditions which are conducive to the establishment of full conservation of information: one, that information is a valuable and critical commodity, and two, that noise can be minimized. The concept that information is the most precious commodity for living things has been formulated strikingly by SCHROEDINGER in his assertion that 'living things feed on orderliness'—that they feed because they need fresh supplies of orderliness, not of energy or matter (8). The need for fresh supplies of orderliness presupposes that orderliness is somewhere lost, that is, that noise is present. This, however, does not mean that noise is present everywhere. Some processes may occur in 'clockwork fashion', without loss of information. That is the case which SCHROEDINGER classifies as 'generation of order from order'. He suspects that each individual act of transmission of genetic information from parent to offspring occurs without serious loss of information. This idea agrees with the current (Watson-Crick) model of DNA duplication; it recurs in GAMOW'S and YČAS' models of information transmission from genetic to somatic material (9).

3. *The Noise-and-Redundancy Theorem*

Information transfer from one body of information to another is not often with clockwork regularity. As a rule, interferences occur which will more or less affect the process of information interaction. Interference can be of many kinds: the worst kind of interference is one the results of which are not predictable in detail. In this case, some information will be irretrievably lost. However, in general some but not all order is lost. It is one of the most significant results of information theory to have shown that order and disorder can be measured by a common yardstick. Hence, it is possible to investigate the quantitative relations between total information, noise, and remaining orderliness. The second basic theorem of information theory states that the amount of information *effectively* transmitted is exactly the amount of information transmitted minus the amount of information lost because of noise. This implies that a source can transmit a certain amount of information reliably in the presence of noise provided it transmits more than the desired amount of information. This surplus must be distributed over the whole activity because it is never known which portions of the total activity will be interfered with by noise; necessarily, the surplus takes the form of *redundant information.* Thus, the second fundamental theorem states precisely the relation between amount of information to be transmitted, amount of information which will be lost through noise, and amount of redundant information needed to make up the loss. Like the first fundamental theorem, it is a one-sided conservation principle; it limits

the amount of order which can prevail in an 'order-from-disorder' situation. Again, the one-sided conservation principle will become more powerful if it can be assumed to approximate a two-sided conservation. However, very stringent conditions must be fulfilled if one expects to use the second theorem. There is some reason to believe that these conditions are at least approximated in some biological situations; this is stated in DANCOFF's principle (10).

Dancoff's principle deals with the economics of information. In 'noisy' situations, information is lost and errors will occur unless they are checked by redundant information. Now, errors may be costly, but so is redundant information; accordingly, the optimum amount of redundant information will be not that which makes all errors vanish, but that which minimizes the sum of the cost of errors plus the cost of redundant information, plus the cost—in information units—of error checking. Dancoff's principle asserts that any organism or organization which has gone through competitive evolution has approximated such an optimum; that is, it will commit as many errors as it can get away with, and use the minimum of redundant information needed to hold errors to this level. It follows from Dancoff's principle that the amount of redundant information in a system is bound to be limited, even if it is a system of enormous information content like a living thing. This is of great interest particularly in radiobiology, because what radiation does very effectively is to destroy information.

4. *The Estimation of Information Measures and the Search for Invariants*

It may well turn out that the qualitative and semi-qualitative applications of information concepts are going to be the most important contribution of information theory to biology. But, even successful qualitative applications have very little power in excluding the possibility that other sets of concepts could have been used just as successfully; besides, all scientists like to take measures. Thus, the problem arises of estimating information measures associated with biological structures and functions.

One fundamental difficulty appears immediately: information measures are relative and not absolute; hence, any information measure associated with a given set of biological objects will depend on the set itself and on the scientist who does the estimating. To be sure, one can establish objective bounds. Thus, if a certain genetic locus is known to be capable of having thirty-two distinct allelic states, which are transmitted to the offspring with equal probability given the proper conditions, then the information stored in this locus cannot be less than five bits. If it is also known that the region containing the locus under consideration comprises no more than, say, 20,000 atoms, then the total information stored cannot be more than about 60,000 bits (10). These brackets are safe, but they are too wide to be of interest. They can be very much reduced if one introduces specific assumptions. For instance, if the locus is known to contain no more than, say, 2×50 nucleic acid residues, and if one assumes that the genetic information is completely coded in the sequence of the residues on one strand of a double helix, with the information carried by each residue corresponding to unconstrained selection from four possibilities, then the upper bound is reduced to 100 bits—but its validity is less absolute.

Because of the relative nature of information measures, it will always be up to the ingenuity of the biologist to find ensembles which result in useful measures. In many cases, even the estimation of a limit is of interest: as in EHRET's demonstration that a few bits *could* be sufficient to specify the nature of cytoplasmic structures (11), or the result easily derived from D'ARCY THOMPSON's work (12) that apparently considerable differences in form *could* be coded in, say, a few nucleic acid residues.

The relativism of information measures is a basic difficulty in estimation; besides, the biologist will encounter a number of technical difficulties arising from the fact that 'message sets' and 'selection rules' are not perfectly known. A number of approximation methods for such situations have been worked out (13).

The relative nature of information measures and the technical difficulties of their estimation, cast some doubts on the usefulness of actual information measures in biology. Only experience will show whether these doubts are justified or not. Measures will be valuable if they lead to the discovery of invariants. In psychology, some invariants seem to be crystallizing out of a number of measurements: there seem to be invariant upper limits for the channel capacity for single activities; for the range of classes distinguishable in a single act, etc. (14). In biology, independent estimates of information transfer associated with three elementary biological functions (allelic, antigenic, enzymatic specificity) have yielded closely similar values (15). Much more material will be needed before we can draw definite conclusions.

The analysis which underlies the estimation of information measures presents certain novel features. Consider, for instance, the informational analysis of a hormonal control system. The traditional approach consists in isolating one hormonal function and one hormone after the other. In principle, this quest never ends—although physiologists might hope that some day they will run out of undiscovered hormones. The information theorist attacks the problem from the opposite end. He will argue that each hormone molecule constitutes a message from a control organ to a target organ, a message which is diffusely broadcast through the blood stream. In general each message must contain two parts, an address and an order. Actually, one or the other part can be omitted. We can imagine a hormonal control system in which only the addresses are specified—the 'order' may be completely determined in the target organ, and be executed automatically upon receipt of the only kind of hormone molecule with the proper address; or, the address may be unspecific, but the order such that only the right target organ can execute it. One would expect that the natural systems be somewhere between these two extremes. For the sake of simplicity we will consider a system in which only addresses are specified—the formal results have complete generality. Thus, each hormone will be represented only by the address of the target organ. In the interest of detailed and accurate control, it is desirable to have a maximum number of different addresses. Any duplication of addresses will lead to concomitant responses in other organs. On the other hand, the 'reading' of every single address involves distinguishing it from all other addresses; the greater the variety of addresses, the greater the labor in every single act of recognition. A compromise is indicated between the demand

for a great variety of addresses and the contradictory demand to keep each address simple. For any kind of system, there will be an optimum number of different hormones; the actual number will depend on the relative strength of the two competing needs. By Dancoff's principle, we expect that the actual number will not be too far from the optimum number.

We can add another line of considerations on the number of possible addresses. In order to fulfill its function, the hormone molecule has to enter into some kind of relation with the target organ; most likely, it has to form a complex. Now, the total surface area of any molecule that can enter into a specific complexing process is rather limited, and so is the number of molecular configurations available to living organisms; hence, a limited space accommodates only a limited number of significantly different configurations—and this limits the number of different hormones possible (and, incidentally, the number of distinct antigens and antibodies, enzymes and co-enzymes).

The example illustrates the concern with the whole system which is characteristic of many applications of information theory. It also illustrates a rather profound difference between the information theorist and many of his scientific colleagues. The information theorist will remain fairly cool at the news that another enzyme, or hormone, or vitamin has been isolated; his basic question is: 'How many more are there to be discovered?'

II. LIMITATIONS

Information theory could not possibly apply to a wide variety of situations if it were sensitive to every detail in every situation. Like thermodynamics (to which information theory is related) it has a vast domain of application, and like in thermodynamics, the vastness of the domain is paid for by a limited scope of every single application (16). The following four limitations deserve emphasis: (i) information measures refer to ensembles and not to single instances, (ii) they are relative and not absolute, (iii) informational capabilities are often not fully utilized, (iv) information measures are related to other aspects of systems such as utilities and mechanisms but the relations are not simple. None of these observations is particularly profound, but each one has been overlooked by competent investigators.

1. *Information Measures are Functions of Ensembles*

Information measures are not defined for particular historical occurrences or existing individual things; rather, they are defined for whole ensembles of events that could happen, or things that could be. The information measures are descriptive of the operations by which a particular item is selected from the set of possible items, and are associated with the whole set and not with any particular item that happened to be selected in a particular instance.

Ensembles are specified by their elements, by the classification to which these elements are subjected, and by the probability measures associated with the diverse classes. If these specifications are known, then the information functions can be derived—but not vice versa. For example: if it is known that a certain chemical system contains certain enzymes and certain substrates, if the probabilities of the various collisions and the probabilities of all possible

outcomes of such collisions are known, then it is possible to derive a number of information functions for this system; on the other hand, a given set of information functions is compatible with any number of chemical systems.

2. *Relativity of Information Measures*

In the early applications of information theory to problems of communication, the ensembles to be used were virtually defined by the situation. Thus, in dealing with Morse code, the element is clearly the single symbol, the classes are dot, dash, letter space and word space, the probabilities are the large sample frequencies. Similarly, in dealing with printed English as an objective phenomenon, one natural unit (not the only one, though!) is again the single symbol, and classes and probabilities can be determined from any large sample. The situation is immediately more complicated if we deal with a particular person's concepts of printed English; the 'subjective probabilities' are not the same as the objective relative frequencies. Much confusion has come to psychologists from disregarding the fact that the probability measures upon which a subject bases his operations are not necessarily those known to be correct—in one sense—to the experimenter (17).

There are situations where there is considerable leeway in defining the elements, classifications, and probability measures of an ensemble, and accordingly considerable variation in the information measures which can be associated with the situation. This is strikingly illustrated by the attempts to measure the information contents of molecules. Estimates have been based on considerations of structure (10, 18, 19, 20) or function (15, 22). Recently, RASHEVSKY and his associates (21) have shown that information measures can be associated with the topological representation of molecules. Each of these approaches yields some value of the information content of a molecule, and these values do not have to be identical. Yet, every one of them is a legitimate information measure. This may be disappointing, but like all abstractions, information measures are not 'right' or 'wrong'—they are only more or less useful. In the case under discussion, we may legitimately ask how the various ways of estimating information measures are related to the actual processes of information storage and transmission by molecules, to reaction rates, to the activity of antimetabolites, etc.

As a rule, the specifications of an ensemble do not result unequivocally from the given situation. Consequently, information measures are not properties but functions of a given situation—they are defined by the situation *and* the ensemble used in dealing with it. Information measures are irreducibly relative; they can be accurate and precise, but they cannot be absolute. The usefulness of a particular information measure in a particular context will depend on the way the defining ensemble is set up. Unfortunately, there exists no calculus, no set of hard and fast rules which tells one how to select the most appropriate elements, classifications, and probability measures. The choice must be made by guess, and its ultimate justification is only in the results it yields.

3. *Informational Capabilities and Performance*

An informational capability represents an upper bound to some class of informational performances—but a particular performance does not have to

approach this bound. One cannot transmit information through a channel at a rate higher than the channel capacity, but it is very easy to transmit at a lower rate. For instance: human capacity of transmitting information can be limited on the input side, on the output side, or centrally; if the limitation is central, then it can be due (a) to the limited channel capacity, but also to limitations of (b) the rate at which discrete acts of information-processing can be performed, of (c) the amount of information per single act, of (d) the number of information-carrying components considered in each act, of (e) the maximum amount of information per component, or finally, (f) to inefficient coding (14). Parallel situations are likely to exist in molecular biology. For instance, AUGENSTINE (18) discusses the fact that the amount of information which can be coded into an amino acid sequence is considerably greater than the amount of information needed to account for the functional specificity of a protein. This could mean that the channel capacity is only fractionally utilized, or that functional specificity is coded in an entirely different fashion.

4. *Information Measures and Other Aspects of Systems*

If the mechanism of a reaction is known, then the probabilities of all input-output associations can be computed, and the information measures derived from them. On the other hand, an information measure does not define a single mechanism—however, it imposes a condition with which input-output tables and, by implication, mechanisms have to comply. For instance, in the problem of the DNA-protein code studied by GAMOW and YČAS (9), the informational analysis furnishes conditions which the code must fulfill but does not yield the code itself. Accordingly, the informational analysis has served, repeatedly, to reject a proposed mechanism. It can, of course, never be used to prove a mechanism.

Amount of information is in general related to the utility of being informed —but the relation is not necessarily one of simple proportionality; in fact, the utility of information is not always a monotonically increasing function of its amount. Similarly, the information content of a structure is in general related to the difficulty of construction, but the relation is not one of simple proportionality.

The 'amount of information' in a statement is related to its capacity of carrying *semantic* information, but this capacity is rarely fully utilized (23).

III. CONCLUSION

I have tried to outline some of the applications and possible applications, and I hope to have shown that there is much promise in this field. I have tried to outline some of the limitations of applying information theory—and I hope to have shown that they are not serious, provided one is always aware of them. To make more progress, we need much more mathematical work, and we need very much more experimental work. In looking over the past of information theory in biology, a very strong emphasis on theory—more or less rigorous—is obvious; although more theory is needed, the most pressing need is now for a large body of good specific experiments. Also, it should be rewarding to examine closely other related possibilities in theoretical biology.

For some reason, our time has been (and still is) exceedingly fertile in producing theories dealing with systems, or in reviving and greatly expanding existing theories of systems. Information theory is one of several system theories—a very rewarding one, I believe, but one with very specific limitations; it should be useful to search specifically for theories with different limitations to supplement information theory.

REFERENCES

1. C. E. Shannon and W. Weaver: *The Mathematical Theory of Communication*, University of Illinois Press, Urbana (1949).
2. N. Wiener: *Cybernetics*, J. Wiley and Sons, New York (1948).
3. R. W. Ashby: *Design for a Brain*, J. Wiley and Sons, New York (1952).
4. M. Eden: A probabilistic model for morphogenesis. This volume.
5. H. P. Yockey: An application of information theory to the physics of tissue damage. *Radiat. Res.* **5**, 146–155 (1956); and, A study of aging, thermal killing, and radiation damage by information theory. This volume.
6. H. B. Jones: Some notes on aging. This volume.
7. H. A. Blair: A quantitative description of latent injury from ionizing radiation. This volume.
8. E. Schroedinger: *What is Life?* University Press, Cambridge, England (1944).
9. G. Gamow and M. Yčas: The cryptographic approach to problems of protein synthesis. This volume. M. Yčas: The Protein Text. This volume.
10. S. M. Dancoff and H. Quastler: The information content and error rate of living things. In: *Information Theory in Biology*, ed. by H. Quastler, 263–273, University of Illinois Press, Urbana (1953).
11. C. Ehret: Information content and biotopology of the cell in terms of cell organelles. This volume.
    G. Karreman: Topological information content and chemical reactions. *Bull. Math. Biophys.* **17**, 279-285 (1955).
    E. Trucco: A note on the information content of graphs. *Bull. Math. Biophys.* **18**, 129–135, (1956).
12. D'A. W. Thompson: *Growth and Form*, Cambridge University Press, Cambridge, England (1952).
13. H. Quastler: Approximate estimation of information measures. In: *Information Theory in Psychology*, ed. by H. Quastler, 124–139. The Free Press, Glencoe, Ill. (1955).
14. H. Quastler: Studies on human channel capacity. In: *Information Theory*, ed. by C. Cherry, Academic Press, New York (1956).
15. H. Quastler: The specificity of elementary biological functions and the measure of specificity. In: *Information Theory in Biology*, ed. by H. Quastler, 170–188; 41–71, University of Illinois Press, Urbana (1953).
16. L. J. Cronbach: On the non-rational application of information measures in psychology. In: *Information Theory in Psychology*, ed. by H. Quastler, 14-30. The Free Press, Glencoe, Ill. (1955).
17. H. Quastler (ed.): *Information Theory in Psychology*, The Free Press, Glencoe, Ill. (1955).
18. L. G. Augenstine: Protein structure and information content. This volume.
19. H. J. Morowitz: Some disorder-order considerations in living systems. *Bull. Math. Biophys.* **17**, 81–87 (1955).
20. H. Branson: Information theory and the structure of proteins. In: *Information Theory in Biology*, ed. by H. Quastler, 84–104, University of Illinois Press, Urbana (1953).
21. N. Rashevsky: Life, information theory and topology. *Bull. Math. Biophys.* **17**, 229–235 (1955).

22. P. D. KLEIN: Efficiency of information transmission by biochemical co-factors. This volume.
23. R. CARNAP and Y. BAR-HILLEL: *Tech. Rep. Res. Lab. Electr., Mass. Inst. Tech.*, no. 247 (1952).

DISCUSSION

A. RAPOPORT: It is admirable of biologists to look 'up to' physicists and mathematicians, but it is somewhat embarrassing to physicists and mathematicians to be looked upon with such confidence as a source of methodological gifts which can be immediately applied in biological investigations. It is noteworthy that the greatest discoveries of the physicists are stated in 'pessimistic' terms. They are statements about what *cannot* be done. For example the First Law of Thermodynamics is essentially a definitive demolition of an age-old dream. The law says in effect that the perpetual motion machine cannot be constructed. But it also holds out a hope of a machine that will keep on working provided only that a large supply of heat is available—the so-called perpetual motion machine of the second kind. The Second Law of Thermodynamics puts an end to that dream. It says that such a machine cannot be constructed either and prophesies the 'heat death' of the Universe. MAXWELL introduced his demon in the hopeful conjecture that the intervention of an intelligence might restore the order lost by the continual increase of entropy. SZILARD in his now classical paper showed that this too is an illusion, that the demon must pay for the restored order by becoming 'disordered' (a biologist would say 'denatured') himself.

Yet it would be a mistake to consider these discoveries as admissions of defeat only. Each has brought a broadened understanding; the First Law of Thermodynamics by revealing heat as a source of energy; the Second Law by revealing the role of entropy. SZILARD's investigation rests on quantum-theoretical principles and so provides an important juncture between thermodynamics, information theory, and quantum theory. It appears, therefore, that the grand discoveries of physics have a sobering effect. I think the principles of information theory are of a similar kind. Typically they are statements of limitations. Their constructive side is in defining the framework in which the search for new knowledge or for new means of prediction and control must be confined.

SOME MEMBRANE PHENOMENA FROM THE POINT OF VIEW OF INFORMATION THEORY*

HERMAN BRANSON
Howard University, Washington 1, D.C., U.S.A.

Abstract—The methods of information theory and of irreversible thermodynamics are applied to membranes. Equations are derived for the negentropy production of a membrane maintaining a concentration difference. The results are converted to bits. When applied to typical data for a nerve transporting Na^+ against a concentration gradient, the equation gives for the negentropy or information production,

$$\dot{H} = 7.3 \times 10^{13} \text{ bits/cm}^2 \text{ second.}$$

Enumeration based on $Na^+ : Cl^- : K^+ = 1 : 1 : 10$ gives a value of

$$4.3 \times 10^{13} \text{ bits/cm}^2 \text{ second.}$$

IN a classification of the significant problems of biophysics DANIELLI (1) listed four. Two of these relate intimately to membranes and their role in biological systems: cell permeability and cell electrophysiology. It is almost mandatory, then, to inquire into the behavior of membranes from the point of view of information theory. For if information theory is to have relevance to important biological problems, a coherent relation should be exhibited for membrane phenomena. This attitude was exhibited in the initial attempts to discuss biological problems within this discipline in QUASTLER's pioneering book (2).

The formidable complexity of biological membranes is a recurring theme in the immense amount of experimental data which are being accumulated. Phenomena encountered in biological membranes may range from those explainable by the assumption of simple pores of various sizes, to those requiring charged pores, and on to those necessitating a picture of the surface as possessing pores, binding sites, permeability barriers, enzymes, and transport mechanisms. It is possible, however, to ignore the details of structure and specific mechanism —as is usually the case in thermodynamics—and formulate a limited model of membrane activity satisfactory to our analysis. Thus membranes may be classified by the manner in which they react to or treat a given substance. If we schematize the membrane as separating two media in each of which the reference substance is soluble, calling one region the 'inside' and the other the 'outside', we may introduce the following notation:

Indifferent: A membrane is indifferent to a substance if the concentration of that substance is the same on both sides of the membrane at all times. Thus $C_i = C_0$, for all t.

* This work has been supported in part by the U.S. Atomic Energy Commission, AT (30–1)–892.

Responsive: A membrane is responsive to a substance if the concentrations of that substance differ on each side at the same time. But neither concentration is zero. Thus, $C_i \neq C_0$ for some t.

Exclusive: A membrane is exclusive with respect to a given substance if, for all time, the concentration on one side is finite but the concentration on the other is zero. Thus, $C_i = 0$ and $C_0 \neq 0$, or $C_i \neq 0$ and $C_0 = 0$, for all t.

The analysis of this paper will be limited to substances to which a given membrane is responsive. It must not be concluded, however, that indifferent and exclusive substances are of no biological significance. There are examples where seemingly the most important role of a membrane is its action to exclude a given substance from the internal medium or keep a component from diffusing out.

The preliminary work (2) on a responsive membrane attempted to derive by the methods of irreversible thermodynamics an expression for the negentropy production of a membrane maintaining a concentration difference. The approach rested upon the concept that the entropy is an absolute maximum at equilibrium. Hence any deviation from equilibrium would mean a decrease in the entropy. Expanding the change in entropy, ΔS, in a Taylor's series about the equilibrium point yields as the first approximation, since the first derivative terms vanish at equilibrium,

$$\Delta S = -1/2 \sum_{l,m} g_{l,m}\, a_l a_m \, . \tag{1}$$

Equation (1) was combined with some descriptive equations for the membrane and the final result

$$\frac{dH}{dt} = k\alpha \tag{2}$$

was deduced. In equation (2), H is the negentropy or the information (3), α is the rate constant governing the rapidity with which the membrane would approach uniform concentration on each side if it were not actively maintaining the concentration difference and k is Boltzmann's constant (1.380×10^{-16} ergs per degree).

Equation (2) had to be examined to determine if it is applicable to a membrane maintaining a considerable concentration difference. Its significance with respect to the relation

$$\Delta S_{\text{irr.}} = k \ln C_i/C_0 \tag{3}$$

had to be clarified (4). Equation (3) gives the irreversible production of entropy for the passage of a single particle from concentration C_i to the concentration C_0. If C_0 is greater than C_i, we have the situation postulated for the membrane, thereupon $\Delta S_{\text{irr.}}$ is negative and may be equated to information, H.

Derivation of the Rate of Production of Information

The methods of irreversible thermodynamics as presented by DeGroot (5) are applicable to effect the derivation. Following DeGroot's nomenclature, we have

$$\Delta \dot{S} = J_u X_u + J_m X_m + J_\mu X_\mu \tag{4}$$

where $\Delta\dot{S}$ is the rate of entropy production. The J's are 'fluxes' and the X's are 'forces', u is energy, m matter, and μ chemical potential. The J's are related to the X's through

$$J_u = L_{11}X_u + L_{12}X_m + L_{13}X_u.$$

Assuming an isothermal system, the X's and J's are

$$X_u = \Delta(1/T) = 0 \qquad J_u = 0$$

$$X_m = -\Delta\mu/T \qquad J_m = \frac{dm}{dt}$$

$$X_\mu = -\Delta m/T \qquad J_\mu = \frac{d\mu}{dt}$$

Whereupon, equation (4) becomes

$$\Delta\dot{S} = -J_m\,\Delta\mu/T - J_\mu\,\Delta m/T. \tag{5}$$

Substituting into equation (5),

$$\mu = \mu_0 + RT\ln a$$

where a is the activity, we get

$$\Delta\dot{S} = (R\ln a_i/a_0)\frac{dm}{dt} + R\,\Delta m\,d/dt\,(\ln a_i/a_0). \tag{6}$$

The a's refer to the activities in the two different regions.

Equation (6) is the basic relation replacing both equations (2) and (3). If we assume that the activities may be replaced by the concentrations and that we are concerned with the passage of N particles from a lower to a higher concentration, $C_0 > C_i$, then equation (6) becomes

$$\frac{dH}{dt} = \dot{H} = k\frac{dN}{dt}\ln C_0/C_i + Nk\,d/dt\,(\ln C_0/C_i).$$

If the outside is taken to be very large with no change in concentration by the addition of the particles from within,

$$\dot{H} = k\frac{dN}{dt}\ln C_0/C_i - \frac{Nk}{C_i}\frac{dC_i}{dt}. \tag{7}$$

Equation (7) gives the rate of decrease of entropy or the rate of production of negentropy or information by a membrane which is transferring material from a lower to a higher concentration where the particles leave at the rate dN/dt and the concentration within changes at the rate dC_i/dt. Thus one may look upon equation (7) as the dynamic equation describing real transport. On the other hand equation (2) describes a situation where there is no macroscopically discernible change in concentration within or without. But inasmuch as the membrane maintains a concentration difference, it is producing information at the rate given to continue the imbalance.

If the concentration within is stabilized in such a manner that for a slight change the system returns to the resting value following the equation

$$(1/C_i)\, dC_i/dt = -\alpha \tag{8}$$

where α is analogous to the rate constant appearing in equation (2), equation (7) may be written

$$\dot{H} = k\frac{dN}{dt}\ln C_0/C_i + kN\alpha. \tag{9}$$

Equations (8) and (9) seem harmonious with equation (2) interpreted as being applicable to the resting case when there is no net transport or to actual transport with the condition that $C_0 = C_i$.

The application of equation (9) to a fluctuation should follow this sequence. Initially $C_0 = C_i$, and ΔN particles jump from one solution to the other. The rate of negentropy production or the rate of production of information is $\dot{H} = \Delta H/\Delta t = k\Delta N\alpha$. At the end of this fluctuation equation (9) becomes for the *next* fluctuation $k\Delta N\alpha$

$$\dot{H} = k(\Delta N/\Delta t)\ln\left(\frac{C + \Delta C}{C - \Delta C}\right) + k\Delta N\alpha. \tag{10}$$

Suppose that the next fluctuation is the movement of ΔN particles in the direction opposite to the first fluctuation in the same interval of time, Δt. We should expect $\dot{H}$ to be the negative of its original value. Expressing the logarithm in equation (10) as $\ln\dfrac{1 + \Delta C/C}{1 - \Delta C/C}$ and making use of the relation

$$\ln\frac{1 + X}{1 - X} = 2(X + X^3/3 + \cdots), \quad \text{for} \quad X^2 < 1,$$

the equation becomes

$$\dot{H} = 2k\ \Delta N/\Delta t\ \Delta C/C + k\Delta N\alpha$$

but

$$\Delta N/\Delta t\ \Delta C/C = \Delta N/C\ \Delta C/\Delta t = -\Delta N\alpha$$

from equation (8), and since ΔC is negative in the second fluctuation.

Finally:
$$\Delta H = -k\Delta N\alpha\ \Delta t.$$

Thus the system returns to the equilibrium position on the entropy surface with an increase in entropy exactly equal to the decrease of entropy experienced in the first fluctuation.

Analysis for Charged Particles

In the derivation of the basic equations (6) and (7), the chemical potential was employed, which limits the applicability of the analysis to uncharged particles. To derive the corresponding equations for an ion, the electrochemical potential μ' replaces μ,

$$\mu' = \mu_0' + RT\ln a + ZF\phi$$

where Z is the valence, F is the Faraday, and ϕ is the potential. Substituting

and carrying through an identical analysis as followed in deriving equation (7), $\dot{H}$ becomes

$$\dot{H} = k[(dN/dt) \ln C_0/C_i + N\alpha] + (Zq/T)[dN/dt(\phi_0 - \phi_i) + N(d/dt)(\phi_0 - \phi_i)] \quad (11)$$

where q is the numerical value of the charge on the electron, when the membrane transports ions from the lower concentration to the higher concentration, $C_i \rightarrow C_0$.

Application to a Nerve

Data on the movement of ions across the nerve membrane may be substituted into equation (11) to obtain numerical values for $\dot{H}$. The knowledge of the transport of ions across the nerve membrane although quite extensive is still not complete enough to permit unequivocal choice of a model. There are two possibilities which suggest themselves for consideration. In the first, following the transport of an impulse, the nerve returns to its resting condition with respect to the concentration of Na^+ or K^+ before it can pass another impulse. The resting potential is reached at the beginning of this period. Calling this example model one, we have these data for a squid axon (6):

$[K^+]$	$C_0 = 10$	$C_i = 410$
$[Na^+]$	$C_0 = 460$	$C_i = 49$
$[Cl^-]$	$C_0 = 540$	$C_i = 40$

in units of millimoles/kg. At 300°K, $\phi_0 - \phi_i = 50$ mV with the outside positive. If the length of the recovery period is taken as 1 millisecond, the equation (11) becomes*

$$\dot{H} = k(dN/dt)\,[\ln C_0/C_i + (Zq/kT)(\phi_0 - \phi_i)]. \quad (12)$$

The first term on the right yields for the concentration gradient above

for K^+, $\dot{H} = 5.3$ bits/ion-unit time,
for Na^+, $\dot{H} = 3.3$ bits/ion-unit time,
for Cl^-, $\dot{H} = 3.7$ bits/ion-unit time.

* In applying equation (11) the second and fourth terms contribute negligibly. Examining the two terms with the data that 3.7 $\mu\mu$ moles of Na^+ enter per impulse per cm², a cylindrical nerve of 100 μ radius with unit surface area would have,

$$\text{Area} = 2\pi r l = 1 \text{ cm}^2$$

$$\text{Vol} = \pi r^2 l = 5 \times 10^{-3} \text{ cm}^3;$$

and assuming that the nerve has the density of water, the section would weigh 5×10^{-3} gm. So 49 m mole/kg would correspond to 2.5×10^5 $\mu\mu$ mole/cm², whereupon for sodium

$$\dot{H} = 2.25\, k(dN/dt) + 1.5 \times 10^{-5}\, kN \text{ per millisecond}$$

$$= 2.25\, k(dN/dt),$$

since

$$\Delta N = N \quad \text{and} \quad \Delta t \text{ is taken as 1 millisecond.}$$

The term $$\frac{Zq}{kT}(\phi_0 - \phi_i) = 1.94$$

at 300°K. Using this value, the combined rates for both the concentration and electrical terms are

$$\begin{aligned} K^+: &\quad \dot{H} = 2.5 \text{ bits/ion-unit time,} \\ Na^+: &\quad \dot{H} = 6.1 \text{ bits/ion-unit time,} \\ Cl^-: &\quad \dot{H} = 0.9 \text{ bits/ion-unit time.} \end{aligned} \tag{13}$$

These values have been arrived at by multiplying equation (11) by $\frac{\log_2 e}{k}$ to convert from k units to entropy units.

[The conversion factor was derived in a previous paper by the author in considering proteins (2). It is easily shown that $\log_y x \log_x y = 1$, which means that this conversion factor is the same as that derived by Linshitz (7) and others (8).]

On the basis of the first model the 3.7 $\mu\mu$ moles of Na^+ entering during the impulse would be extruded during the millisecond following the return to the resting potential. This time interval requires that the nerve produce information at the rate

$$\dot{H} = 9.3 \times 10^{15} \text{ ergs/°K cm}^2 \text{ sec,}$$

or

$$\dot{H} = 1.35 \times 10^{16} \text{ bits/cm}^2 \text{ sec.} \tag{14}$$

The alternative model is that the nerve does not extrude the Na^+ in so short a time. Rather the nerve passes it from within to the outside at a rate of 20 $\mu\mu$ mole/cm^2 second (9). However, the acceptance of this view does not alter equations (13). Inserting this value for Na^+ in equation (12) results in

$$\dot{H} = 7.3 \times 10^{13} \text{ bits/cm}^2 \text{ sec.} \tag{15}$$

The experimental results seem to favor this model; thus equation (15) is the more reasonable result in comparison with equation (14).*

The alternative mode of viewing the nerve according to information theory makes use of the concept that the sodium ions are chosen from the pool of ions within the nerve by some mechanism. Within the nerve the ratios are $Na^+ : Cl^- : K^+ = 1 : 1 : 10$. The mechanism chooses the Na^+ from this group, requiring $\log_2 12$ bits of information for each Na^+ selected. This value, 3.57 bits/ion, leads to

$$\dot{H} = 4.3 \times 10^{13} \text{ bits/cm}^2 \text{ sec.,} \tag{16}$$

for the nerve based on the second model which transports 20$\mu\mu$ moles cm^2 sec. of Na^+ against the concentration and electrical gradient. Assuming that the

* Dr Leroy Augenstine made the astute observation with respect to equation (14) in the discussion that it is consistent with a value of 10^3 Å^2 for the area of a protein and that on 1 cm^2 of nerve there could be 10^{13} proteins transporting one ion per millisecond. It may weaken the argument to assume that the nerve surface has that many protein molecules. The lower value 1.20×10^{13} ions/cm^2 second is consistent with any combination of rates and numbers of protein molecules responsible for transport such that the product equals this numerical value. Thus 1.20×10^{10} protein molecules transporting an ion per millisecond are suitable and require that only 0.001 of the nerve surface consists of such proteins, each of 10^3 Å^2 area.

information units, the bits, appearing in equations (15) and (16) are identical with those used in discussing the information content of the printed page, the interpretation is that a cm^2 of nerve has an enormous rate of production. The analogy of course is to a person who is called upon to separate red balls (Na^+) from white (K^+) and blue (Cl^-) balls in a box where he would reach in, pick up a ball, look at it, and if it is red it is taken out, and if not it is replaced. The nerve has some equivalent separating mechanism with the information rate of equation (15). In terms of a familiar example, taking the information content of a single printed page as 10^4 bits (9), equation (15) requires that the cm^2 of nerve surface produce information equivalent to that contained in a library of 7.3 million volumes of a thousand pages each second—this is over half the number of books in the Library of Congress! The value given by equation (15) is not inordinate, however, in comparison with the estimates of the information content of biological objects (9), where for man the value is of the order of 10^{25} bits.

The result in equation (16) may be viewed as the minimum information production necessary to effect the separation of sodium. The numerical value may be in error, for the choice could be from among more ions than the three employed. It is of interest to note that the nerve does not possess perfect coding inasmuch as it uses 1.7 times as much information to effect the separation as is required. Alternatively the information efficiency may be expressed as 59 per cent. These comparisons may be without substance because of the inadequacy of the data. The only relevant comparison may be that the physio-chemical determination of information production as summarized in equation (15) is of the same order of magnitude as the value determined by enumeration in equation (16).

REFERENCES

1. J. F. Danielli: Cytological research and biophysics. *Nature, Lond.* **162,** 753 (1948).
2. H. R. Branson: A definition of information from the thermodynamics of irreversible processes. In: *Information Theory in Biology*, ed. by H. Quastler, 25–40. University of Illinois Press, Urbana (1953).
3. L. Brillouin: *Science and Information Theory*, Academic Press, New York (1956).
4. R. C. Tolman and P. C. Fine: On the irreversible production of entropy. *Rev. Mod. Phys.* **20,** 51–77 (1948).
5. S. R. DeGroot: *Thermodynamics of Irreversible Processes*, Interscience, New York (1952).
6. R. W. Stacy, D. T. Williams, R. E. Worden, and R. O. McMorris: *Essentials of Biological and Medical Physics*, McGraw-Hill, New York (1955).
7. H. Linschitz: Information and physical entropy. In: *Information Theory in Biology*. ed. by H. Quastler, 14–15, University of Illinois Press, Urbana (1953).
8. S. M. Dancoff and H. Quastler: The information content and error rate of living things. In: *Information Theory in Biology*, ed. by H. Quastler, 263–273, University of Illinois Press, Urbana (1953).
9. A. L. Hodgkin and R. D. Keynes: Active transport. *Symp. Soc. Exp. Biol.* **8,** 423–437 (1954).

EFFICIENCY OF INFORMATION TRANSMISSION BY BIOCHEMICAL CO-FACTORS*

PETER D. KLEIN
Division of Biological and Medical Research, Argonne National Laboratory, Lemont, Illinois

Abstract—A model for the transmission of information by biochemical co-factors is described. Two points of information transfer are apparent: formation of the holo-enzyme and formation of the holo-enzyme-substrate complex. The reduction in uncertainty taking place at these points is related to the sets of compounds existing before and after these points, and the values for artificial situations calculated.

It is concluded that the artificial situation is an estimate of the minimum selection capabilities of the enzyme system.

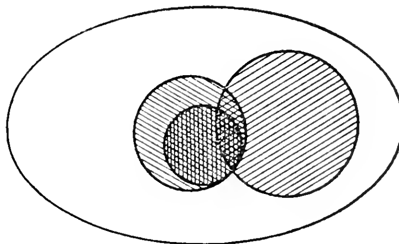
CO-FACTORS are compounds of molecular weight 100 to 2000 which participate in a host of biochemical reactions. They are not metabolised *per se*, but serve as catalysts. The moiety with which co-factors cooperate may in this instance be limited to proteins. Co-factors for a particular protein may be either exogenous (vitamins) or endogenous (hormones).

The flow diagram (Fig. 1) represents the fate of a co-factor in the organism. A particular apo-enzyme can operate on a substrate or class of substrates if provided with the suitable co-factor. In this case the co-factor is assumed to be a vitamin. Therefore, preliminary to its appearance in the cell, the compound must of necessity be ingested, absorbed and transported into the cell. Inside the cell, the compound may or may not be excreted again. Each time that it exists in a 'free' form, it has a finite possibility of leaving the site of action, including transformations which lead to the degradation of the molecule so that it cannot function. This possibility is represented by Probability Point 1. This and subsequent probability points have the following characteristics: A molecule 'passing' through this point may undergo two or more transitions; each state resulting from these transitions has a certain probability but there is no control over the state into which the molecule passes.

Next, it may be imagined that a collision between the compound and the apo-enzyme for which it may be destined takes place, leading to the formation of a complex. The formation of this complex, however, depends upon mutual exchange of information between the apo-enzyme and its co-enzyme and is indicated by Decision Point 1. For example, if the co-enzyme for cocarboxylase collides with the apo-enzyme for riboflavin, no information exchange takes place and there is no complex formation. If, however, sufficient information is exchanged, there results the formation of the holo-enzyme, or in the case of the competitive inhibitor, a pseudo-holo-enzyme. Both forms of holo-enzyme have, of course, a dissociation constant, indicated by Probability Point 2.

* This work was performed under the auspices of the U.S. Atomic Energy Commission.

FIG. 1. Pathway of utilization for a hypothetical co-factor.



CELLULAR CONSTITUENTS
APO-ENZYME ACCEPTORS
SUBSTRATE ACCEPTORS

FIG. 2. Topological considerations in evaluating information transmission by a co-factor.

The collision of the holo-enzyme with a potential substrate represents another decision point (Decision Point 2). Again, a complex will be formed if sufficient information is transferred. It is at this point that the pseudo-holo-enzyme must present the wrong information in order to be an effective inhibitor. This results in repeated cycling in the innermost loop diagrammed.

The enzyme-substrate complex has a finite probability, designated Probability Point 3, of decomposing unchanged before the reaction is catalysed to decompose the substrate into product and regenerate the holo-enzyme.

There are, then, two points in this flow sheet at which information can be exchanged; between the co-factor and the apo-enzyme and between the holo-enzyme and the substrate. These are the two points to which attention will be devoted.

At neither decision point is the decision unequivocal. There are several types of co-factor-substances which may form a complex with the apo-enzyme and several of these complexes are acceptable, though to differing degrees, to the substrate. The situation is graphically presented in the set diagram Fig. 2. The largest circle represents the class of all possible organic compounds. Let B designate these substances. A subset of B, composed of the organic substances which normally occur in cells, is designated C. Another subset of B, designated A, is comprised of substances acceptable to the apo-enzyme for complex formation. The set $A.C$† includes those compounds normally occurring in cells which are able to complex with the particular apo-enzyme considered; the set $(A–A.C)$ contains those substances which form a complex with the apo-enzyme but are not normally found in the cell.

A subset of A, designated $A.S$, contains all compounds which complex with A *and* react with the substrate. (There may be other substances in B which would react with the substrate, if complexed with a proper apo-enzyme; but these are not of concern here.) These substances which are contained in the set $C.A.S$ are the natural co-factors for the apo-enzyme-substrate pair under consideration; the substances in the set $A.S–C.A.S$ are artificial co-factors; the substances $C.A–C.A.S$ are natural antimetabolites; those in the set $A–A.S$ are artificial antimetabolites.

The information measures associated with the two decision processes can be derived from the diagram. Let $H(X)$ designate the uncertainty associated with the set X; then:

> $H(C)$ is the uncertainty of substances in a normal cell. To give this quantity meaning, we shall consider it to be the uncertainty about the nature of an organic molecule which normally collides with the apo-enzyme.
>
> $H(C.A)$ is the uncertainty concerning a substance which has formed a complex with the apo-enzyme.
>
> $H(C.A.S)$ is the uncertainty concerning the complex which has reacted with the substrate. It should be noted that in dealing with a given apo-enzyme and a given substrate (or class of substrates) all uncertainties in question are due to the co-factor.

† The set $X.Y$ contains all substances which belong to both the set X and the set Y. The set $X–Y$ contains those substances which are contained in X but not in Y; alternatively this set is designated $X–X.Y$, all substances which belong to X but not to both X and Y. The latter notation is more explicit and will be used here.

The informational performances at the decision point are related to the reduction of uncertainty, ΔH, at these points, i.e. the difference in uncertainty before and after:

$$\Delta H_{\mathrm{I}} = H(C) - H(C.A)$$
$$\Delta H_{\mathrm{II}} = H(C.A) - H(C.A.S)$$
$$\Delta H_{\mathrm{I,II}} = H(C) - H(C.A.S)$$

The comparable functions for artificial situations are given by:

$$\Delta H_{\mathrm{I}}^* = H(B) - H(B.A)$$
$$\Delta H_{\mathrm{II}}^* = H(B.A) - H(B.A.S)$$
$$\Delta H^*_{\mathrm{I,II}} = H(B) - H(B.A.S)$$

There is a fundamental difference between natural and artificial functions. The quantities $H(C)$ as well as $H(C.A)$ and $H(C.A.S)$ do not depend on the experimenter; furthermore, because of the constancy of the internal environment, they can be considered to be numbers approximating natural constants, subject to relatively small fluctuations. The function $\Delta H_{\mathrm{I,II}}$ represents the normal informational performance achieved in the particular metabolic process considered, whose average value has been placed at 9 bits (1). The quantity $H(B)$, on the other hand, is completely or partially controlled by the experimenter, who regulates the availability of substances in B; $H(B.A)$ and $H(B.A.S)$ depend on apo-enzyme, substrate *and* on the experimenter. Accordingly, the ΔH^* functions have, in general, very little interest since it is easy to make ΔH^* vanish by offering only a single co-factor which can be used, or to give it a very high value by introducing numerous compounds which are known not to react with the apo-enzyme or substrate.

A great body of data is available, however, which lends itself to an examination of the ΔH^* functions as well as $H(B.A)$ and $H(B.A.S)$ from the standpoint of the systems' responses to a series of compounds closely resembling the natural co-factors. The values obtained may be regarded as a sort of minimum residual uncertainty associated with the various systems, and they form the subject of this report.

We shall define the uncertainty functions $H(B.A)$ and $H(B.A.S)$ as:

$$H(B.A) = -\Sigma p_a \log_2 p_a$$
$$H(B.A.S) = -\Sigma p_{as} \log_2 p_{as}$$

where p_a and p_{as} are the normalized biological activities of the compounds tested in a particular system. Thus if four compounds were tested for their ability to combine with the apo-enzyme and all were found to be equally active, their p_a would each be 0.25 and $H(B.A)$ would be 2. This method of calculation takes into account the fact that with equal concentrations, equal activity may not be observed and that the information is of necessity related to the concentration required to produce a complex.

A word is in order as to the mechanics of calculation. The basic data were derived in large part from WILLIAMS and co-workers' treatise on the B vitamins (2); data on thyroxine are due mainly to the work of BRUICE, KHARASCH and

WINZLER (3). $H(B)$ was defined as the logarithm of the total number of compounds tested, both for ability to replace the natural co-factor and for those having antimetabolite activity. $H(B.A)$ was calculated from the compounds active as antimetabolites plus those having substrate activity; in the former case, the inhibition index (number of molecules of inhibitor required to overcome the action of one molecule of the true compound) was considered as the reciprocal of biological activity and suitably transformed to agree in dimension with the other data. $H(B.A.S)$ was, of course, derived from the group which showed ability to replace the natural co-factor.

RESULTS AND DISCUSSION

The H functions and ΔH_I^*, ΔH_{II}^* and $\Delta H_{I,II}^*$ are listed for a variety of compounds in Table 1. In addition, compounds for which partial data were

Table I

Compound	Organism	N	$H(B)$	$H(B.A)$	$H(B.A.S)$	ΔH_I^*	ΔH_{II}^*	$\Delta H_{I,II}^*$
Biotin	MO	33	5.04	3.01	2.67	2.03	0.34	2.37
Riboflavin	MO	40	5.32	2.91	2.61	2.41	0.33	2.71
Riboflavin	R	18	4.16	2.93	2.61	1.24	0.33	1.57
Folic acid	C	11	3.46	2.74	1.97	0.72	0.77	1.49
Folic acid	MO	51	5.67	3.97	2.55	1.70	1.42	3.12
Thiamine	R	8	3.00	1.39	0.99	1.61	0.40	2.01
Thyroxine	F	34	5.09	2.55	1.47	2.54	1.08	3.62
Thyroxine	R	43	5.43	4.28	3.22	1.15	1.06	2.21
p-Amino benzoic acid	MO	72	6.17	4.90	3.27	1.27	1.63	2.90
Biotin	R	4	2.00	—	0.13	—	—	1.87
Unsaturated fatty acids	R	10	3.32	—	1.85	—	—	1.75
Pantothenic acid	C	4	2.00	—	1.86	—	—	0.14
Vitamin D	R	10	3.32	—	2.08	—	—	1.24
Pantothenic acid	R	10	3.32	—	2.19	—	—	1.13
Nicotinamide	C	5	2.32	—	2.23	—	—	0.11
Ascorbic acid	G	10	3.32	—	2.33	—	—	0.99
Nicotinamide	D	13	3.70	—	2.66	—	—	1.04
Pyridoxine	R	11	3.34	—	2.85	—	—	0.49
Choline	R	35	5.11	—	3.12	—	—	1.99
Carotene	R	15	3.90	—	3.68	—	—	0.22
Estrogens	R	18	4.16	—	3.88	—	—	0.28
Average			3.96		2.39			1.57

Key: MO: Micro-organism F: Frog
R: Rat G: Guinea pig
C: Chick D: Dog

available are included. Fig. 3 presents some values for the first group in graphic form.

It can be seen that there is a range in $H(B.A)$ of 1.39 to 4.90 and in $H(B.A.S)$ of 0.13 to 3.88. There is also a marked tendency for ΔH_{II}^* to be smaller than ΔH_I^*, suggesting that the greater portion of the selection process is assumed by the apo-enzyme/co-enzyme complex formation, but sight must not be lost of

the fact that the area $H(B.A)-H(B.A.S)$ (representing antimetabolites) is dependent upon the number of successful inhibitors of a co-factor which have been devised.

Over-all, the mean reduction in uncertainty in terms of actual compounds is such that when confronted with fifteen compounds, assumed to be equally

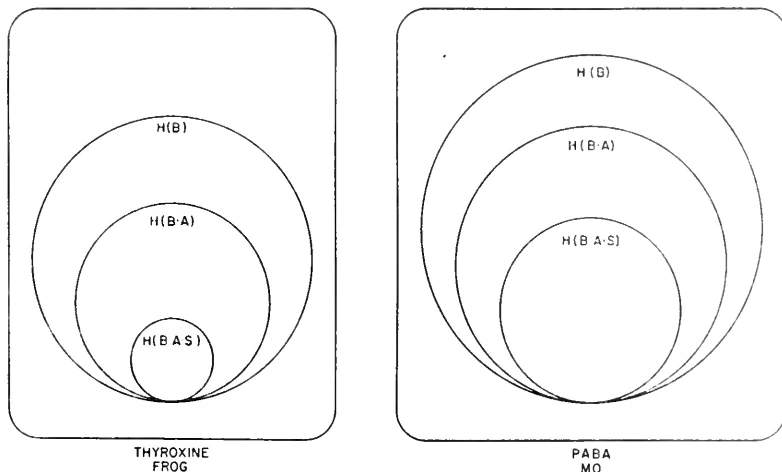


FIG. 3. Residual uncertainties associated with two co-factors.

effective, the system can weed out ten of these, leaving the equivalent of five, equally active, co-factors. Comparison of this to $H(C)$ and $H(C.A.S)$ is of course not plausible from these figures, but clearly indicates the relative chaos of the universe B, from the standpoint of the enzyme system: the assembly of letter-perfect molecules of protein or nucleic acids would be impossible under these circumstances. Nevertheless, these figures may have some interest as the minimum limits of discrimination ability by enzyme systems.

REFERENCES

1. P. F. FENTON, cited by H. QUASTLER: The specificity of biological functions. In: *Information Theory in Biology*, ed. by H. Quastler, 170–188, University of Illinois Press, Urbana (1953).
2. R. J. WILLIAMS, R. E. EAKIN, E. BEERSTECHER Jr., and W. SHIVE: *The Biochemistry of B Vitamins*. Reinhold Publishing Co., New York (1950).
3. T. C. BRUICE, N. KHARASCH, and R. J. WINZLER: A correlation of thyroxine-like activity and chemical structure. *Arch. Biochem. Biophys.* **62**, 305–317 (1956).

DISCUSSION

QUASTLER: The following interpretation of the 'residual uncertainty of co-factors' may be considered:

Given a particular apo-enzyme-substrate system, and the set B of all substances b_i which are or might be co-factors. Let c_i be the reaction rate constant of the system in the presence of the (potential) co-factor b_i. Suppose that all c_i's have been determined; it seems that there are two statistics of general interest: the average size of the c's, which characterizes the reactivity of the system in general, and the dispersion among the c's, which characterizes its specificity.

To study the specificity independently of general reactivity, we normalize the c_i's by setting their sum equal to one. We consider a function called the *tolerance* of the apo-enzyme-substrate system; the tolerance function shall have the following properties: it shall be zero if only one c_i is not zero; it shall increase with the number of substances b with non-zero c's; for a given number of substances the tolerance shall be called highest if all c_i's are equal. These postulates are satisfied by the information function:

$$\text{Value of tolerance} = -\sum_i \frac{c_i}{\sum_i c_i} \log_2 \frac{c_i}{\sum_i c_i}$$

Tentatively, one might assign a physical meaning to the tolerance as follows: the reaction rates are replaced by (suitably normalized) probabilities of a reaction following a collision; it is assumed that the system can exist in a number of mutually exclusive states (configurations) and that a reaction will occur if the system is in the proper configuration; then, the c_i's are proportional to the fractions of the time the system is in a configuration compatible with functioning with the substances b_i. The value of the tolerance is the uncertainty concerning the actual state of the apo-enzyme-substrate system.

ANTIGENIC SPECIFICITY*

BERNARD N. JAROSLOW and HENRY QUASTLER
Division of Biological and Medical Research, Argonne National Laboratory, Lemont, Illinois
and
Department of Biology, Brookhaven National Laboratory, Upton, New York

Abstract—The production of a specific antibody involves a transfer of information, and so does the specific reaction between antibody and antigen. This paper deals with the 'vocabulary' of this communication process. An antigenic determinant is considered as a 'word' of a certain number of 'letters', subject to certain constraints. It is shown that the number of 'words', the number of 'letters', and the degree of constraint can be estimated by methodical random sampling. Experimental methods suitable for this purpose are discussed. Preliminary results are given.

I. INTRODUCTION

THE question, 'How many different antigens are there?' is one that has not been explored up to now, but which arises naturally if information theory is applied. Information theory interprets the process of antibody formation as the transmission of information from the antigen to the antibody-forming mechanism, with the information then utilized and again transmitted when the antibody reacts selectively with the appropriate antigen. It is then natural to ask how much information is transmitted from antigen to antibody and vice versa. More explicitly, one will ask certain questions about the kind of information traffic between antigen and antibody—the 'vocabulary' in which this information traffic is coded, the 'alphabet' that is used to make up the words of the vocabulary. Now, information theory is not concerned with specific features of 'alphabet' and 'vocabulary' but with general properties of both, such as their sizes. The problem, 'how large is the vocabulary of information transmission between antigen and antibody,' is closely related to the question posed above. A preliminary estimate by one of us (1) has led to a rough estimate of some 125 to 500 different protein antigenic determinants, and a smaller number of different carbohydrate determinants. No attempt was made, at that time, to estimate the number of antigens of other chemical constitutions. Although these figures appear very small in the light of the specificity of immunity to the multitude of infectious agents, the antigen complexes of the organisms represent an array of many different determinants and their over-all specificity can be much larger than that of a single antigen.

The present investigation is an attempt to measure antigenic specificity. The general plan of the experiment is based on information theory; the specific methods are based on agar diffusion precipitin tests developed by Oudin (2).

* Work performed under the auspices of the U.S. Atomic Energy Commission.

II. NOMENCLATURE AND MODEL

An antigen, GN, consists of a specific determinant G and a carrier moiety N which is a macromolecule (MW > 10,000), i.e. protein, lipo-protein, glyco-protein, etc. G may consist of three or four amino acid residues, a mono- or disaccharide, an aliphatic chain, etc., with its specificity dependent upon order, size, polarity, or optical configuration of the residues. There may be a number of antigenic determinants, of the same or different specificities, on the surface of N, so that one molecule of antigen may combine with several molecules of antibody, e.g. 5 for ovalbumin to over 200 for hemocyanins (3). Combining capacity as well as antigenicity (ability to induce antibody formation) is usually proportional to the molecular weight of N.

An antibody, AB, consists of a specific combining site A and a carrier moiety B which is always a globulin, usually γ-globulin. The combining site is believed to be a chemical and/or spatial configuration which combines with the specific antigenic determinant through hydrogen bonds. The number of combining sites per antibody molecule is thought to be two or one.

The A-G reaction may be manifested in a number of ways: precipitation of a soluble antigen, agglutination of a particulate or cellular antigen, or lysis of a cellular antigen in the presence of complement.

We consider a *heterophile* reaction as one in which the reaction between G_2 and A_1 is indistinguishable from the homologous reaction of G_1 and A_1 although G_2 and G_1 come from different sources. By *cross reaction* we mean the phenomenon wherein G_2 reacts with A_1 but the strength of reaction is less than that of the homologous one (G_1 and A_1).

We can describe an antigenic determinant, G, as a 'word' of k 'letters'. By letter we mean antigenically active residues such as amino acids, mono-saccharides, etc. Let r be the size of the alphabet, i.e. the number of available letters; then

$$H\,(\text{letter}) = \alpha \cdot \log_2 r.$$

α is a constant which ranges from zero to one. Its upper limit occurs if all 'letters' occur with equal probabilities, and no two letters can ever have equivalent effects.

The average information content of a *word* averaging k letters is given by:

$$H\,(G) = \beta \cdot k \cdot H\,(\text{letter}).$$

β is a constant which ranges from zero to one. Its upper limit occurs when all letter combinations are equally probable, i.e. if there is no 'intersymbol influence'. The lower limit would obtain if there existed only one antigenically active combination of letters.

To fix the ideas on the measuring of r, H (letter), k and $H(G)$, we give the corresponding values for printed English:

$r = 26$ $\qquad$ $k \approx 4.5$

$\log_2 r = 4.7$ $\qquad$ $\beta \approx 0.6$

$\alpha \approx 0.87$ $\qquad$ $H(G) \approx 10$

$$H\,(\text{letter}) = 4.1$$

III. EXPERIMENTAL TESTS

1. *Occurrence of the Heterophile Reaction*

The incidence of heterophile reactions will depend on the number and relative frequencies of the various antigenic determinants. As nothing is known so far about relative frequencies, we assume them to be equal; this will yield a lower bound of the number of different G's. Under the assumption of equiprobability, we use the following argument (4): Let C_i and C_j be different and (as far as known) unrelated antigen complexes; let m_i and m_j denote the number of antigens in the complexes which can be differentiated and demonstrated, by a given technique, by reaction with the specific antisera S_i and S_j; let h_{ij} be the number of heterophile reactions observed; let N be the total number of different antigenic determinants which this technique will differentiate. Then, the maximum likelihood estimate $\hat{N}$ of N is given by:

$$\hat{N} = \frac{m_i \cdot m_j}{h_{ij}}$$

Assuming β to be one, we have:

$$H(G) \approx \log_2 \hat{N}$$

and

$$r^{\alpha k} \sim \hat{N}$$

This is a preliminary test of $r^{\alpha k}$.

2. *Classification of Cross Reactions*

The strength of the cross reaction presumably depends on the number of letters in common, and on the nature of these letters. We assume as a working hypothesis that the former factor is the leading one. Then, if we grade the strengths of many cross reactions, we expect to find a distribution into clearly separated groups such as strong, less strong, weak, . . . etc., cross reactions.

We suspect that the strong cross reactions are those in which the G-pair has $(k - 1)$ letters in common, the next class those with $(k - 2)$ letters, and the weakest observable reactions those with one common letter. Then, the number of distinguishable classes of strengths of cross reactions should be $(k - 1)$.

This may develop into a test of k.

3. *Ratio of Incidence of Heterophiles to Incidence of Strong Cross Reactions*

By our hypotheses, the probability of occurrence of heterophiles is the probability of having all letters in common; now, for k letters, assuming β to be one, the number N of different words is:

$$N = r^{\alpha k}.$$

Then, probability of a given word $= (1/r)^{\alpha k}$

and, probability of a heterophile $= (1/r)^{\alpha k}$.

Since we propose that in the strong cross reaction we have $(k-1)$ letters in common, the probability of this event is:

$$(1/r)^{\alpha(k-1)}.$$

There are k sets of $(k-1)$ letters in a k-letter word; hence,

$$\text{probability of a strong cross reaction} \sim \alpha\, k(1/r)^{\alpha(k-1)}$$

and

$$\frac{\text{probability (heterophile)}}{\text{probability (strong cross reaction)}} = \frac{(1/r)^{\alpha k}}{\alpha k(1/r)^{\alpha(k-1)}} = \frac{1}{\alpha k r^{\alpha}}$$

This is a test of $\alpha\, kr^{\alpha}$. One can construct similar tests for the ratio of strong to weak cross reactions, if experimental results warrant this.

Optimally we have, thus, three experimental determinations of three parameters which have computable theoretical values.

IV. METHODS

It is essential to this study that the test antigens be as nearly as possible a random sample of natural antigens. Since related organisms regularly have common antigens, we chose our sources so that no two were in the same phylum. We used GUYER (5) as our guide to classification.

Entire organisms were placed in a Waring blender for two minutes with phosphate buffer (pH 7.4) as the extracting agent. If the antigen source was of microscopic proportions it was ground in a mortar and pestle containing sterile powdered carborundum. The material was centrifuged at low speed to remove the gross particulate material and 0.2 per cent formalin was added as a preservative.

Rabbits were immunized with a series of four intravenous injections of antigen over a ten-day period. They received another series of three injections a month later and were bled three days afterwards. Serum was collected, complement was inactivated by incubation for one-half hour at 56°C and 1:10,000 merthiolate was added as a preservative.

Our test system was double diffusion agar precipitin test based on the methods of OUDIN (2) and OUCHTERLONY (6). Antiserum (0.5 ml) was placed at the bottom of a test tube (4 mm i.d.). The antiserum was topped with a layer of 1.5 ml of 1 per cent agar. After the agar had gelled, the antigen (0.5 ml) was added. Both antigen and antiserum diffused towards each other through the agar, and where they met in the proper proportions a band of precipitate became visible. For antigens with different rates of diffusion separate bands appeared. Thus, the number of bands of precipitate establish a lower limit for the number of antigens in the extract.

We should be able to differentiate between heterophile and cross reactions because the formation of a precipitate is dependent upon the relative concentrations of antiserum and antigen. In the region of great antibody excess no precipitate forms, although all the antigen is combined with antibody. In the region of large antigen excess no precipitation is seen, although all the antibody is combined with antigen. Between these two zones is a region in

which precipitate appears. Thus, in a heterophile reaction, two antigens diffusing towards each other through a field of antibody will form two lines of precipitate that become confluent on meeting. Although free antigen migrates past the diffusion front of the other, no precipitate will form because all the antibody is already combined in what is a region of antigen excess. In the case of cross reactions a different situation exists. As LANDSTEINER (7) and many others have demonstrated, where the antigenic determinants differ from one another, the heterologous antigen will precipitate only part of the antibody, leaving enough antibody in solution to form precipitate when homologous antigen is added. Thus, the peptide glycyl glycyl *glycyl* glycyl glycine (heterologous antigen) forms a precipitate with antiserum to glycyl glycyl *leucyl* glycyl glycine (homologous antigen). If this precipitate is removed and the antiserum is then mixed with the homologous antigen a precipitate will appear. The antibody that reacts with the heterologous antigen is probably 'imperfect', i.e. its specificity is less than that of the uncombined or 'avid' antibody. In our system one would expect, then, that the advancing homologous antigen would find uncombined antibody behind the front of the heterologous antigen and form a line of precipitate there. We hope that we will be able to grade the intensity of this reaction through some aberration in the band of precipitate and thus be able to classify the cross reactions as to relative strengths.

V. RESULTS

As of this writing, we have obtained some preliminary data which we present here as an indication of the limits or power of our test system.

We obtained antisera to whole body extracts of the following organisms:

Grasshopper	*Melanoplus differentialis*	phylum Arthropoda
Frog	*Rana pipiens*	phylum Chordata
Bacterium	*Escherichia coli*	sub-division Bacteria
Horse mussel	*Modiolus modiolus*	phylum Mollusca
Sea pen	*Ptilosarcus quadrangularis*	phylum Coelenterata
Giant star fish	*Pisaster giganteus*	phylum Echinodermata

In addition we tested extracts for which we had no antisera. These were:

Baker's yeast	*Saccharomyces cerevisiae*	sub-division Fungi
Earthworm	*Lumbricoides terrestris*	phylum Annelida

The results of testing each of the six antisera against each of the eight antigen extracts are shown in Table I. We observed from four to ten homologous reactions per test and two heterologous reactions. Since we have no data on the number of homologous reactions for yeast and earthworm extracts we assigned each of them a value of 7.3, which is the mean of the homologous reactions for the other six tests. Then, we have six tests, in each of which m_i was (on average) 7.3; m_j, being the sum of antigens in all other complexes, was $7 \times 7.3 = 51$; the average number of heterophile reactions was $2/6 = 0.33$. Then, if reciprocal tests are considered as independent, we obtain:

$$\hat{N} \sim \frac{7.3 \times 51}{0.33} \sim 1100$$

$$\log_2 \hat{N} \sim 10 \text{ bits}$$

*Table I. Precipitin Reactions in the Agar Diffusion Test**

Antiserum	Antigen							
	Grasshopper	Frog	*E. coli*	Horse mussel	Sea pen	Star fish	Yeast	Earthworm
Grasshopper	4	0	0	0	0	0	0	0
Frog	0	10	0	0	0	0	0	0
E. coli	0	0	4	0	0	0	1	0
Horse mussel	0	0	1	9	0	0	0	0
Sea pen	0	0	0	0	7	0	0	0
Star fish	0	0	0	0	0	8	0	0

* Each figure signifies the number of visible precipitin reactions in each test.

VI. CONCLUSION

The estimate obtained from our preliminary test is extremely crude; however, it agrees fairly well with the earlier (1953) estimate. The number of homologous reactions we observed was surprisingly low. We expect that with more potent antisera it would increase markedly. If the observed reactions are found to be cross reactions, or if repetition of the test with another set of antigens gives fewer heterologous reactions, the value for $\hat{N}$ would go up radically. Also heterophile reactions will have to be differentiated from each other to prevent a common heterophile such as the Forssman antigen or the Wasserman cardiolipid antigen from lowering the value for $\hat{N}$ by multiple appearances in the tests. On the other hand, the similarity with QUASTLER's estimate suggests that the order of magnitude of $\hat{N}$ after further experimentation will not be much greater.

The preliminary tests were made with antigen complexes. It is known that simultaneous immunization with many antigens does not produce sera of optimum potency. In later tests, it might be worthwhile to try to isolate antigens which show cross reaction or heterophilia, and use them to produce more potent antisera.

In the final analysis, we hope to obtain an estimate of the number of factors or signals that are needed to characterize an antigen as to its specificity.

REFERENCES

1. H. QUASTLER: The specificity of elementary biological functions. In: *Information Theory in Biology*, ed. by H. Quastler, 177–181, University of Illinois Press, Urbana (1953).
2. J. OUDIN: Méthode d'analyse immunochimique par précipitation spécifique en milieu gélifié. *C.R. Acad. Sci., Paris* **222,** 115–116 (1946).
3. W. C. BOYD: *Fundamentals of Immunology*, 2nd ed., p. 91. Interscience, New York (1948).
4. W. FELLER: *An Introduction to Probability Theory and its Applications*, vol. 1, p. 37, J. Wiley and Sons, New York (1950).

5. M. G. Guyer: *Animal Biology*, 4th ed., 60–79, Harper, New York (1948).
6. O. Ouchterlony: In vitro method for testing the toxin-producing capacity of diphtheria bacteria. *Acta Path. Microbiol. Scand.* **25**, 186–191 (1948).
7. K. Landsteiner: *The Specificity of Serological Reactions*, rev. ed., Harvard University Press, Cambridge, Mass. (1945).

INFORMATION CONTENT AND BIOTOPOLOGY OF THE CELL IN TERMS OF CELL ORGANELLES*

CHARLES F. EHRET
Division of Biological and Medical Research, Argonne National Laboratory, Lemont, Illinois

Abstract—The cell organelle is regarded as a characteristic structural unit of the organism that bridges the gap between the molecular and cellular levels of organization. It may arise from the nucleus as a primary organelle, or from other cytoplasmic organelles. A provisional flow-diagram is presented, according to which specifically different cell structures are derived from primary organelles by sequences of relatively simple events that involve two to five binary decisions.

I. INTRODUCTION

A MAJOR limiting factor in estimates of information content of the organism resides quite obviously in assumptions that are made regarding organization of its component parts. Thus in terms of atoms, $H(\text{man}) = 2 \times 10^{28}$, and in terms of molecules $H(\text{man}) = 3 \times 10^{25}$. The reduction in this second DANCOFF-QUASTLER estimate (1) of information content is permitted by limitations placed upon possible positions of atoms as a result of the restricted number of molecular configurations found in living systems. It follows that if many of the molecular and macromolecular configurations that are theoretically possible actually occur in only a limited number of anatomical and microanatomical (organellar) configurations, then a further reduction of this estimate by several orders of magnitude is possible. In *reductio ad absurdum* the terms might be alive or not, as AUGENSTINE has suggested (2), and the information content one bit or less; however, such a classification would be of more value to exterminators than to biologists.

The existence of such organelles as chromidia, micellae, bioplasts, etc., that occupy in all cells a functionally meaningful position between the molecular and gross anatomical levels of organization was long ago claimed by some cytologists but denied by most (3). New evidence from electron and phase contrast microscopy revives but considerably revises the earlier unifying view; it allows for characterizable 'nuclear ambassadors' each equipped with a versatile morphogenetic repertoire to roam the cytoplasm and serve the structural and functional needs of the cell. While the evidence to date is far from complete, its general implications are presented here.

II. ORGANELLES AND PRIMARY ORGANELLES

Firstly, we define and characterize *organelles* in the following manner:

1. *Organelles are elements of unitization for the macromolecular species of an organism:* they occupy niches between the macromolecular and cellular

* Work performed under the auspices of the U.S. Atomic Energy Commission.

hierarchies of organization; they are sites of biosynthetic and energy yielding cycles (4) (molecular 'chunking') *large* enough to maintain a relative degree of thermodynamic homeostasis and biochemical independence.

2. *Organelles are homologously related in two ways:* (a) *They are phylogenetically static.* Classes of organelles present the same basic appearance in all plants and animal cells: they show *synchronic stability.* Historically, the chromosome is an excellent example; presently, cytoplasmic organelles including mitochondria, flagella and cilia are even better described than are those of the nucleus. ASTBURY has proclaimed the demonstration of the basically similar patterns of cilia wherever they are found as "one of the most important microanatomical revelations of our time" (5), a statement made before DE ROBERTIS' discovery of retinal rod organization (6) cited below. Basically tubular and lamellar mitochondria are also phylogenetically ubiquitous (7–11).

(b) *They are epigenetically plastic.* Organelles react with or respond to their environment and interact with other organelles to produce novel patterns of organelle complexes and systems: they show *diachronic non-fixity.* The patterns 'develop' from similar mechanics of packing, coacervation and disintegration. (A relatedness akin to that described by THOMPSON for cells and tissues (12).) Thus the nebenkern arises by fusion of mitochondria (13), the acroblast by fusion of dictyosomes (14), the old nucleolus by fusion of young nucleoli (15), the rod sacs and rod tubules of the mammalian retina by development of the distal region of the connecting filaments, which are themselves cilia (6), the endoplasmic reticulum by fusion of spherical vesicles, and the 'prolamellar body' and lamellated grana of the plastid by fusion of lipid vesicles (16). Such earliest reactants we term 'primary organelles'.

3. *Primary organelles arise by synthesis* from molecular pools and probably never by division of pre-existing organelles.* The specific pattern has never been seen to divide in nature; it is only replicated. The essential contribution to their progeny of even such classical 'dividers' as bacterial viruses (17) and chromosomes (18) is that of a linearly ordered code *along which* new code units are synthesized, rather than that of a code which grows and 'splits'. Experiments claiming genetic continuity to centrioles (19) and to blepharoplasts (kinetosomes (20)) do not demonstrate *division* of these elements, but only continuity of topological relationships; whereas some primary organelles, microsomal to mitochondrial in size, arise from the nucleus (15, 21) or, in the case of the plant blepharoplast, *de novo* (22). Hence the organelle is not itself an 'organism' and ALTMANN's 'dictum' that all granules come from granules does not strictly apply.

4. *Organelles give rise to the gross patterns and shapes of cell parts and of whole cells* by symmetrical packing, coacervation and the formation of polarized interconnectives. They provide the periodically replicative units of brush border (23), ciliated epithelium (24, 10) and nuclear membrane (25, 36—39); the walled structures of an outer pellicle and an inner cell-mouth (26, 48); the packed units of the cirrus (27), the polarizing units of the kinety (28), and the vehicles of fluid transport 'by which structural lipids move in the cell from site of synthesis to region of lamellar growth' (16).

* This is to be distinguished from the sort of mitochondrial splitting observed between the miotic divisions during grasshopper spermatogenesis (13).

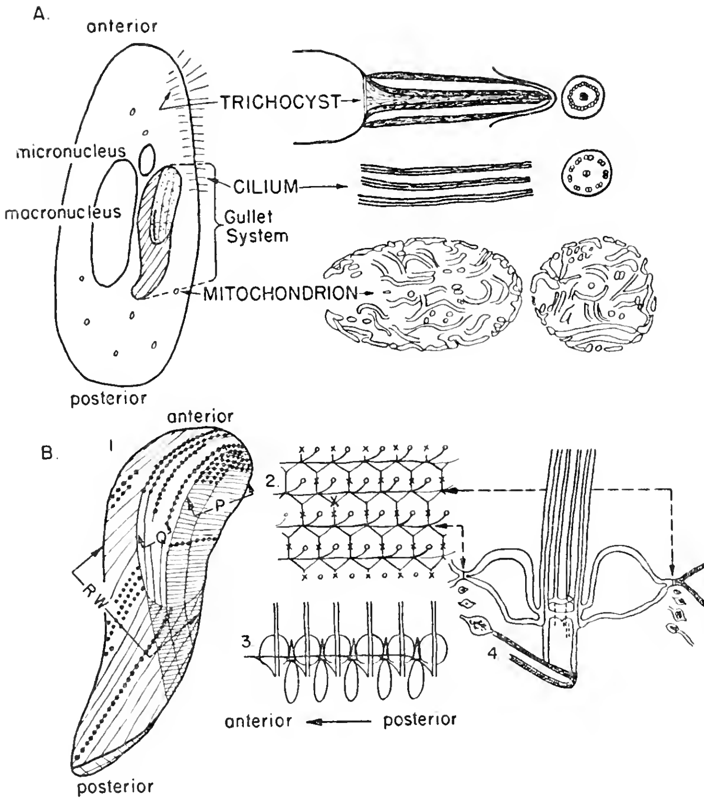


FIG. 1. Systems and complexes of organelles in *Paramecium*.

Approximate dimensions: Cilium diameter = 0.25 μ
Cilium tubule diameter = 22–30 mμ
Mitochondrion tubule diameter = 34–40 mμ
Mitochondrion tubule-lumen diameter = 8–29 mμ

Cilium to cilium distance (on centers):
in gullet system = 0.5 μ
in pellicle system = 1.8 μ
length of gullet = 20 μ

A. Orientation sketch, and diagram of some hypothetical derivatives of primary organelles.

B. Patterns of organelle packing at the free cell border.

1. Gullet (food-intake) system consists of three organelle complexes: ciliated peniculus (P), ciliated quadrulus (Q) and non-ciliated ribbed-wall (RW); close packed basal 'granule'-tracts are diagrammed.

2. Surface view of hexagonal complex of pellicle system. Ciliary loci are represented by circles, trichocyst loci by X's; kineties are formed to the right of the ciliary bases by overlapping kinetodesmal fibrils.

3. Schematized longitudinal section through pellicle system, along a single kinety; the outer elements are represented as close-packed spheres bearing cilia; the inner elements are 'coralled' out of normal packing-positions (with rare exceptions, viz. at the intersection, on 2, left center) by kinetodesmal fibrils. These inner elements are the trichocysts.

4. Cross-section of single corpuscular organelle of the hexagonal complex. Four overlapping kinetodesmal fibrils are shown under the right arrow.

The cytological and epigenetic criteria of this definition must ultimately be supplemented by cytogenetic criteria, since the organelles are certain to include such elements as some plasmagenes and all kinetosomes (20, 29), and homeostats (30). The use of the term 'organelle' in this concept has several distinct advantages. In the first place it is particularly nebulous and difficult to define at its limit values, an asset because "the boundaries of all natural units are hazy" (52); and it therefore does not lack this attribute of realism. In this way it is less mystical and arbitrary than many of the 'vital granules' of the past, and there is much room left for the contribution of the configurations at the 'hazy lower-boundary' to cell form and function, by the mechanisms discussed by WASSERMANN (31), GROSS (32) and others.

On the other hand it has sufficient traditional meaning (with or without our four-part definition) for some cytologists to agree that all of the units listed outside of brackets in Fig. 2 are 'organelles.' Our concept goes only a step further in claiming epigenetic kinships.

III. ORGANELLE DECISION-TREES

Such an unorthodox but allowable interpretation of the organelle complexes and systems of the unicellular animal *Paramecium* (shown in Fig. 1) leads to the following postulates that should be useful in describing the information content of other organisms: (a) The units of structural and functional integration in the cytoplasm are organelles ranging in size from less than 0.2 μ to 1 μ, or are composed of such organelles as fusion structures, as complexes, and as systems of complexes. (b) The organelles arise from primary organelles that in some or all cases come from the nucleus as extruded nucleoli.

According to these postulates, the primary organelle or young nucleolus starts out with a finite set of possibilities and finishes with one having been realized. We can call the point of choice a 'decision point', and discuss the events involved in terms of binary decisions. Decision points are encountered at which specialization is gained and differentiation-potential (uncertainty) is lost. The critical decision points in *Paramecium* are to remain intranuclear or become extranuclear; to remain within the fluid phase (karyoplasm or endoplasm) or to occupy an interface; to remain solitary or to pack; to form a coacervate or not. A provisional decision-tree for *Paramecium*, as well as for other animal and plant cells, is presented in Fig. 2. The differentiation of organelles in *Paramecium* insofar as it concerns the present classification can be accomplished in a sequence of 2 to 5 binary decisions; the informational performance accomplished in the sequence of decisions is not more than a few bits. Presumably, only a small number of such decisions will be needed to account for all existing differentiations in any organism. (While it is premature to assign probabilities to each decision point*, the practical methodology is nearly at hand: an animal has approximately 13,000 cilia, and about as many trichocysts; perhaps twice this number of mitochondria; a macronucleus at division is capable of numerically twice generating this entire

* Pure chance, or a probability bias imposed by the micro-environment, may not be regarded as a sufficient cause for the choice of a particular decision in every case. We do not suppose that each primary organelle contains all messages of the genome, and its specific quality may therefore predetermine some choices and exclude others.

demand in the form of primary organelles (young nucleoli) $\lesssim 0.5\ \mu$ in diameter.) One interesting aspect of this scheme that is probably not artificial is that for plant and animal cells in general, all subcategories have members, whereas for a particular cell, such as *Paramecium*, numerous vacant categories exist (e.g., 0.101 grana-eyespot). The latter condition, not the former, is the more common

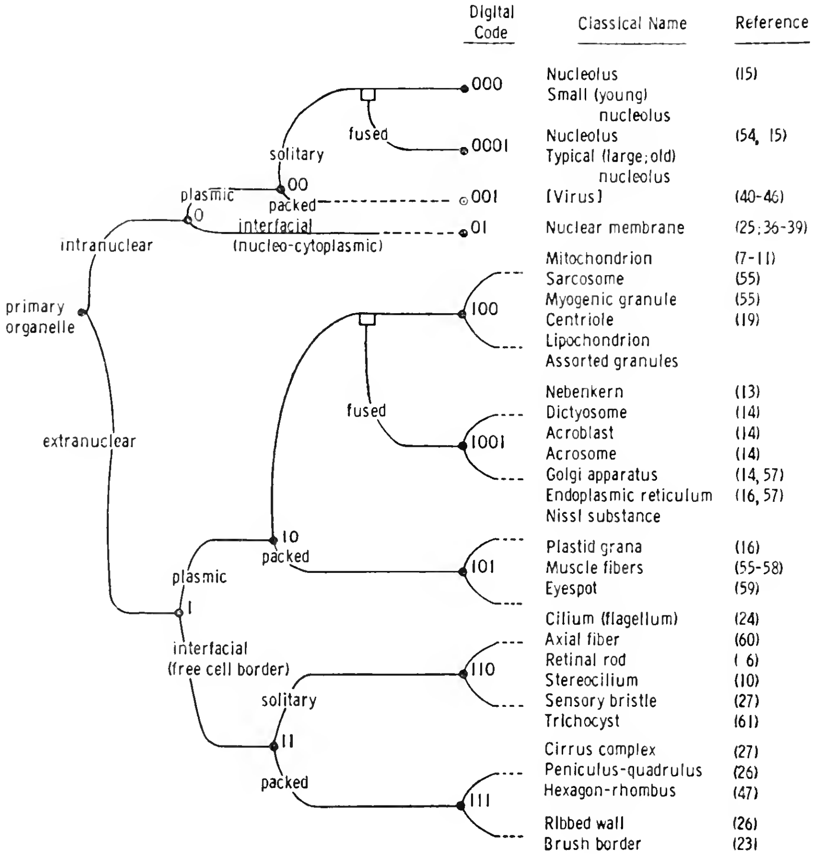


FIG. 2. Organelle Decision Tree, representing a flow-diagram of the alternate pathways and collisions available to a primary organelle in approaching minimum uncertainty in a cell.

situation in dichotomous divisions based upon a *differentia* and its negative (33). The *fundamentum divisionis* for the first two divisions is location in the cell, and for the next two divisions is location with reference to other organelles; it is evident, however, that early organeller coacervation may influence the subsequent localization of organelles, thereby upsetting the strict temporal representation of the scheme. Whether such competitive mechanisms explain null-categories in some cells, or whether we have simply failed to recognize the appropriate candidates for these categories in the cells in question remains to be seen, and of course both answers are plausible. It should be noted that the 'fusion'

decisions are represented as branch points, which may or may not be taken; thus when the nebenkern (.1001) develops from the fusion of mitochondria (.100), the mitochondrion loses potential on making this decision, but suffers no loss in failing to do so.

Admittedly several apparent and some real inconsistencies exist in this provisional flow-diagram. For example, the development of the cilium from a mitochondrion-like body has already been suggested (34) in which the cilium should appear related to, but more specified than, its supposed progenitor; the present tree only remotely relates these. This difficulty may be resolved in a common progenitor organelle (.1) at the plasmic-interfacial decision point. Another inconsistency is in the separation at 0.111 into cirrus-peniculus-hexagon complexes vs. ribbed wall and brush border, the former being ciliated, and the latter non-ciliated organelle-complexes. A more realistic decision in *Paramecium* might be to form pellicle system unit (hexagon-rhombus complex) *or* to form gullet system unit (peniculus-quadrulus-ribbed wall) dependent perhaps on location of the primary organelle at a region of pattern homogeneity (amongst hexagons or amongst rhomboidal elements) *or* at a region of pattern contrast (at the junction of hexagons and rhomboids (k_j to $k_j - 1$ in Fig. 3, discussed in the next section)). The latter condition is consistent with *Paramecium* structure, but so far has been demonstrated only in the ciliate *Stentor* (35).

Several investigators (25, 36, 37, 38) have related nuclear membrane and endoplasmic reticulum as mutual derivatives—a view not inconsistent with the present scheme in its broadest sense; in addition to the rounds of packing and fusion, some 'unravelling' may be involved in both cases. Viruses are provisionally included in the tree because they nearly satisfy our definition, because of their organelle-like ultrastructure, because of recent evidence for host relatedness (40, 41, 42) and because of frequent nucleolar involvement (43, 44, 45). If considered as particles produced by the host's gene-product synthesis that contain a small error perpetuated by error feed-back, then the *nuclear viruses* might occupy the packed (46) small-nucleolus position (.001); the *cytoplasmic viruses* would be classed with 'assorted granules' (.1000) if single, and along with 'grana' and 'fibers' (.101) if packed. It is not necessary to postulate a 'nuclear round' for the replication of either extranuclear organelles or of cytoplasmic viruses, although this may be the usual case.

IV. PRIMORDIAL GRAPH

A material basis for nucleocytoplasmic communication is therefore realized in the cell organelle, which in its primitive unspecified state provides the cell with its necessary potential of structural diversity through relatively few bits of information. A further quality related to the above (47) is that the cell's geometry of functional relations is in some respects as similar to its cytological structure as a wiring diagram is to its final construct. In other words, at the level of the cell organelle (above the macromolecular and below gross cellular dimensions), a coincidence between the functional and the structural biotopological set points of the cell exists. These relationships are represented for organelles at an intranuclear-extranuclear decision in Fig. 3a, and for functioning cytoplasmic organelle systems (48) in Fig. 3b. The intranuclear-extranuclear

decision occurs at binary fission, during macronuclear development, and possibly intermittently during interphase. In most cells during mitosis the nuclear membrane breaks down, thereby eliminating briefly the important inside-outside relationships imposed by that membrane upon particles of large size. (While *Paramecium* is an exception in that its macronuclear membrane does not 'break down', there is sufficient change to allow for the passage of particles

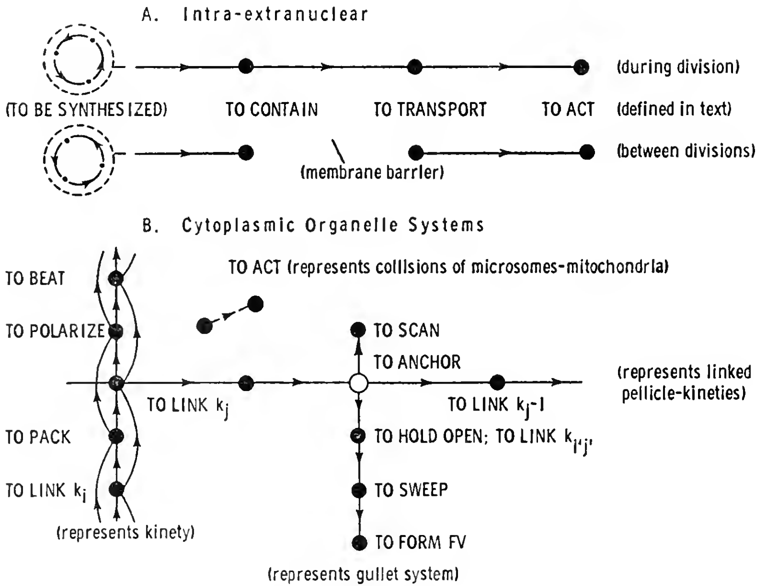


FIG. 3. Partial primordial graphs for organelles in *Paramecium*. A. Organelles engaged in communications between nucleus and cytoplasm and in construction of complexes and systems of organelles. B. Organelles engaged in systems for feeding, for locomotion and for structural integration.

up to 1 μ in diameter during fission (15).) The first function of the primary organelle is *to contain* the message introduced presumably by the chromosomal genes at time of nucleolar synthesis. Its next function is to move and *to transport*; the function shown in the figure is to move 'out of' the nucleus; this is followed by *to act* (collide, fuse, develop), an accomplishment in which the message and its vehicle are also partial power supply and building stone to make cilium, trichocyst, mitochondrion, or whatever is dictated by uncertainty loss. The decision tree (Fig. 2) is itself a partial primordial graph representing the diversity of these acts.

The cytoplasmic organelle system (Fig. 3b) operates through similarly gross functions. The function of a pellicle unit is *to beat* (row, propel), *to pack* (to fit as a block in the pellicle wall), and *to link* similar units longitudinally (k_i) and latitudinally (k_j); a column of such units (one kinety) is aligned with directional reference to each of its ciliary bases (kinetosome) and basal fibrils (kinetodesma); each kinetodesmal fibril lies to the right of its kinetosome; and each is overlapped by and in turn overlaps two or three of the fibrils antero-posteriorly

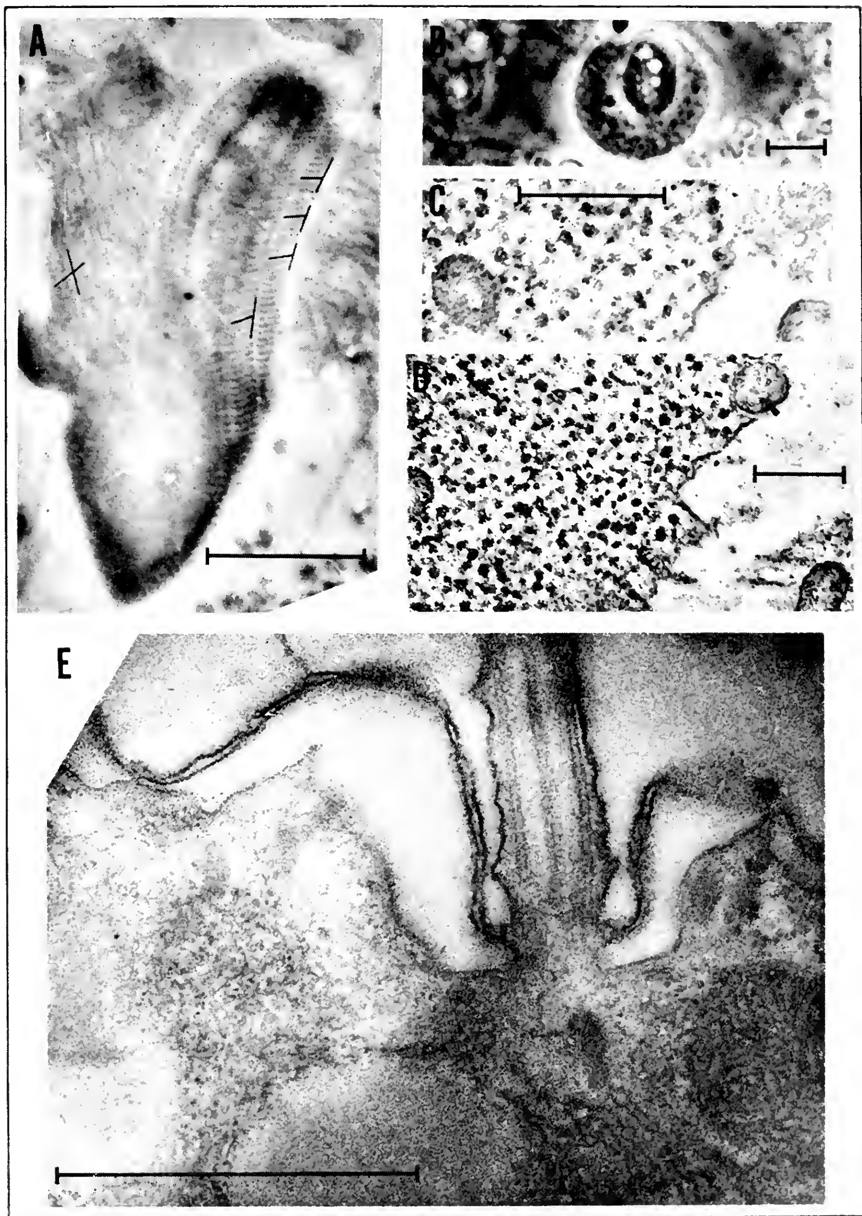


PLATE I

The appearance of organelles in *Paramecium* under phase contrast and electron microscopy. (In A and B the line represents 10 μ; in C, D, and E it represents 1 μ.)

A. Food-intake system, compression-dissected from an unfixed cell. From left to right, the non-ciliated 'granules' of the ribbed wall complex followed by four columns of the ciliated quadrulus complex and eight columns of the ciliated peniculus complex (26). Phase.

B. Macronucleus anlage in an exconjugant during extrusion of young nucleoli; net-like figure in the center is fusion product of old nucleoli (15). Phase.

C and D. Electron micrographs of a similar stage. The smaller dark bodies are chromatic elements of the nuclear matrix; the larger bodies at the left are young nucleoli, and at the right in the cytoplasm are mitochondria (15).

E. Electron micrograph of a single packing unit of the hexagonal complex of the pellicle system (48). Note the cross-sectioned kinetodesmal fibrils in the bays of cytoplasm at each side of the cilium base; at the left, a portion of a trichocyst with its golf-tee-like head; at the right a tubular mitochondrion (8, 9).

None of these photographs has been previously published. Each is a part of the work cited in parentheses, and done in collaboration with Drs E. E. POWERS and L. E. ROTH of Argonne National Laboratory.

to face p. 2

along the kinety; these units act *to polarize*; the array of kineties (or entire pellicle system) functions *to envelop* the endoplasm of the whole animal. It also functions *to anchor* in place other systems such as the food intake or gullet system. The anchorage confers a new level of asymmetry, resulting in the swimming function *to scan* or spiral; the function of the peniculus and quadrulus is *to sweep* food particles down the intake tube; their terminal cilia act *to form* food vacuoles (FV); the cilia-free ribbed wall functions to confer rigidity and *to hold open* the tube, which lies within the endoplasm; the gullet system also functions *to envelop* ($k_{i'j'}$) the endoplasm. This graph may be read in the following way: effective beating of a kinety cilium requires polarization; both functions require positioning upon the pellicle by latitudinal linking to adjacent kineties (k_j's); the previous functions require packing (a continuously surfaced pellicle). Longitudinal linkage (k_i) is required to prevent the dispersion of the surface blocks from within a kinety. Each kinety organelle may perform every function, but locally any function may be by-passed. The function *to scan* requires anchorage of the gullet organelle-complexes in position on the animal; the gullet's general function is to feed. An effective food vacuole requires that food be swept into it by the cilia of the peniculus and quadrulus; these functions require that the gullet be held open by the gullet tube-wall, which requires anchorage to the pellicle. The general function of envelopment requires the linkage of all k_i's and k_j's. (A nearly unique quality of the ciliated protozoa, which, however, should not be entirely ignored in the transformations to higher forms, is the presence of what might be called 'linkage groupings' in the cytoplasm: organelle patterns, far more complex than any known in the metazoa, appear to be as much dependent upon the previous existence of a related pattern in the cytoplasm as upon any nuclear genes (49); even *Stentor*, with its remarkable capacity to regenerate 'kineties' (actually pigmented stripes), rebuilds its mouth organelles only when a particular juncture of maximum anisotropy in the stripe pattern is available (35).) Organelle-functions are therefore given not in the terminology of the molecular level (whose necessary though not strictly pertinent relations are partially represented for the whole cell topologically in such graphs as those of the glycolytic and citric acid cycles (50)) but in the correspondingly appropriate terms of the gross operations performed.

According to this concept, the cell is entirely describable in minute detail of anatomical pattern without reference to either power or fuel. It does not matter whether the oar-like cilia are tugged by galley-slaves, gasoline engines or a creatine-phosphate-ATP system. However, the universal usage by cells of such engines places some restriction at the systems-coupling level, and probably represents a nearly unique solution of the bioenergetics problem. If the model is correct, the most complex patterns are entirely derivable by just such remarkably simple interactions as those first explicitly delineated by D'Arcy Thompson (12).

In summary, at the organelle level fundamental topological sets are recognized of two classes: those that are periodically disjoined (intranuclear from extranuclear organelles), and those that are continuously joined at non-empty intersections (cytoplasmic organelle-systems). Periodic coupling processes (such as during mitosis and nuclear membrane disappearance) occur to form non-empty intersections at all disjunctions of the first class. Below this dimensional

level and within the spheres of intracellular and extracellular molecular interaction, each set of the higher two classes is continuously capable of communication with every other set by means of diffusion and convection transport phenomena.

V. CONCLUSION

In the introduction to this conference BIGELOW suggested that in biological systems with long time-constants, for a message to be useful at the receiving point, 'all messages must be enormously complex groups of messages rather than simple ones' (51). This expression is clearly related to the limited span proposition (52), that is, that span of diversity is limited by difficulties of internal control. Therefore it is not too surprising to find that epigenetic control systems of the cell (whose internal difficulties are in the form of long time functions and thermodynamic vulnerability) solve these difficulties by the method of 'chunking' complex groups of messages into structurally and functionally unitized subsystems. That the subsystems of primary organelles appear to be phylogenetically ubiquitous might also have been predicted from the principles of biochemical evolution. But that they are structurally so alike is indeed a striking fact; although this is not to say that we should now expect the cilium of whale bronchus, the axial fiber of fern sperm, the connecting fibril of toad retina, the sweeping cilium of paramecium peniculus or the mantle cilium of a mollusc to be exactly alike. We know that such organelles are capable of antigenic distinctions even amongst the various stocks within a species (53); indeed, the mechanism of such distinctions constitutes a most crucial problem of molecular biology. That the functional and structural diagrams of an organism in terms of its organelles are topologically homeomorphic is consistent with parallel relations at other levels; in its functions between molecular and cellular levels of organization, the cell organelle fills the last gap in a complex hierarchy of unitized subsystems that characterize the organism from the atomic to the social level. The method of integrating these hierarchies and of extracting quantity of information from *any* organism that employs such mechanisms remains to be accomplished.

Acknowledgment—I wish to thank my many colleagues, particularly Howard Ducoff, E. Lawrence Powers, and Robert Schweisthal of Argonne National Laboratory for their interest and cooperation, and Henry Quastler of Brookhaven National Laboratory for his very helpful suggestions.

REFERENCES*

1. S. M. DANCOFF and H. QUASTLER: The information content and error rate of living things. In: *Information Theory in Biology*, ed. by H. Quastler, 263–273, University of Illinois Press, Urbana (1953).
2. L. AUGENSTINE: Protein structure and information content. This volume.
3. L. MONNE: Functioning of the cytoplasm. *Advances in Enzymology* **8**, 1–69 (1948).
4. D. E. GREEN: Organized enzyme systems. *J. Cell. Comp. Physiol.* **39**, Suppl. 2, 75–111 (1952).
5. W. T. ASTBURY: Introductory review: fibrous proteins and their biological significance. *Symp. Soc. Exp. Biol.* **9**, 1–9 (1955).

6. E. De Robertis: Electron microscope observations on the submicroscopic organization of the retinal rods. *J. Biophys. Biochem. Cytol.* **2,** 319–330 (1956).
7. G. E. Palade: An electron microscope study of the mitochondrial structure. *J. Histochem. Cytochem.* **1,** 188–211 (1953).
8. E. L. Powers, C. F. Ehret, and L. E. Roth: Mitochondrial structure in *Paramecium* as revealed by electron microscopy. *Biol. Bull.* **108,** 182–195 (1955).
9. E. L. Powers, C. F. Ehret, L. E. Roth, and O. T. Minick: The internal organization of mitochondria. *J. Biophys. Biochem. Cytol., Suppl.* **2,** 341–346 (1956).
10. J. Rhodin and T. Dalhamn: Electron microscopy of the tracheal ciliated mucosa in rat. *Z. Zellforsch.* **44,** 345–412 (1956).
11. K. E. Wohlfarth-Bottermann: Protistenstudien. VII Die feinstruktur der mitochondrien von *Paramecium caudatum. Z. Naturf.* **11,** 578–581 (1956).
12. W. D'Arcy Thompson: *Growth and Form,* Cambridge University Press, Cambridge, England (1952).
13. T. N. Tahmisian, E. L. Powers, and R. L. Devine: Light and electron microscope studies of morphological changes of mitochondria during spermatogenesis in the grasshopper. *J. Biophys. Biochem. Cytol., Suppl.* **2,** 325–330 (1956).
14. H. W. Beams, T. N. Tahmisian, R. L. Devine, and E. Anderson: Electron microscope studies on the dictyosomes and acroblasts in the male germ cells of the cricket. *J. Biophys. Biochem. Cytol., Suppl.* **2,** 123–128 (1956).
15. C. F. Ehret and E. L. Powers: Macronuclear and nucleolar development in *Paramecium bursaria. Exp. Cell Res.* **9,** 241–257 (1955).
16. A. J. Hodge: A possible mechanism for the morphogenesis of lamellar systems in plant cells. *J. Biophys. Biochem. Cytol.* **2,** 597–608 (1956).
17. C. Levinthal: The mechanism of DNA replication and genetic recombination in phage. *Proc. Nat. Acad. Sci., Wash.* **42,** 394–404 (1956).
18. J. H. Taylor, P. S. Woods, and W. L. Hughes: The organization and duplication of chromosomes as revealed by autoradiographic studies using tritium-labelled thymidine. *Proc. Nat. Acad. Sci., Wash.* **43,** 122–128 (1957).
19. A. F. Huettner: Continuity of the centrioles in *Drosophila melanogaster. Z. Zellforsch.* **19,** 119–135 (1933).
20. É. Chatton and J. Séguéla: La continuité génétique des formations ciliaires chez les ciliés hypotriches. Le cinétome et l'argyrome au cours de la division. *Bull. Biol,* **74.** 349–442 (1940).
21. Akira Yuasa: Critical observations on the origin of the blepharoplast and centrosome in plants. *Jap. J. Bot.* **10,** 259–284 (1939).
22. R. Lepper: The plant centrosome and the centrosome-blepharoplast homology. *Bot. Rev.* **22,** 375–417 (1956).
23. H. W. Beams, T. N. Tahmisian, and R. L. Devine: Electron microscope studies on the cells of the malpighian tubules of the grasshopper. *J. Biophys. Biochem. Cytol.* **1,** 197–202 (1955).
24. D. W. Fawcett and K. R. Porter: A study of the fine structure of ciliated epithelia. *J. Morph.* **94,** 221–281 (1954).
25. M. L. Watson: The nuclear envelope. *J. Biophys. Biochem. Cytol.* **1,** 257–270 (1955).
26. C. F. Ehret and E. L. Powers: The organization of gullet organelles in *Paramecium bursaria. J. Protozool.* **4,** 55–59 (1957).
27. L. E. Roth: Aspects of ciliary fine structure in *Euplotes patella. J. Biophys. Biochem. Cytol., Suppl.* **2,** 235–240 (1956).
28. C. B. Metz, D. R. Pitelka, and J. A. Westfall: The fibrillar systems of ciliates as revealed by the electron microscope. I—*Paramecium. Biol. Bull.* **104,** 408–425 (1953).
29. A. Lwoff: *Problems of Morphogenesis in Ciliates,* J. Wiley and Sons, New York (1950).

30. J. F. Danielli: Cytoplasmic inheritance. *Nature, Lond.* **178,** 214–215 (1956).
31. F. Wassermann: The intercellular components of connective tissue: origin, structure, and interrelationship of fibers and ground substance. *Ergebn. Anat. Entwgesch.* **35,** 240–333 (1956).
32. J. Gross: The behavior of collagen units as a model in morphogenesis. *J. Biophys. Biochem. Cytol., Suppl.* **2,** 261–274 (1956).
33. M. R. Cohen and E. Nagel: *An Introduction to Logic and Scientific Method,* Harcourt-Brace, New York (1934).
34. E. L. Powers, C. F. Ehret, and L. E. Roth: An hypothesis of mitochondrial origin in *Paramecium. Rep. Argonne National Laboratory, ANL*-5247, 55–60 (1954).
35. V. Tartar: Pattern and substance in *Stentor.* In: *Cellular Mechanisms in Differentiation and Growth,* ed. by D. Rudnick, 73–100 Princeton University Press (1956).
36. L. I. Rebhun: Electron microscopy of basophilic structures of some invertebrate oocytes. I. Periodic lamellae and the nuclear envelope. *J. Biophys. Biochem. Cytol.* **2,** 93–104 (1956).
37. H. Gay: Nucleo-cytoplasmic relations in salivary gland cells of *Drosophila. Proc. Nat. Acad. Sci., Wash.* **41,** 370–375 (1955).
38. H. Swift: The fine structure of annulate lamellae. *J. Biophys. Biochem. Cytol., Suppl.* **2,** 415–418 (1956).
39. G. D. Pappas: The fine structure of the nuclear envelope of amoeba proteus. *J. Biophys. Biochem. Cytol., Suppl.* **2,** 431–434 (1956).
40. A. Garen and N. D. Zinder: Radiological evidence for partial genetic homology between bacteriophage and host bacteria. *Virology* **1,** 347–376 (1955).
41. A. Lwoff: Control and interrelations of metabolic and viral diseases of bacteria. *Harvey Lectures (Ser. L),* 1954–1955, 92–111, Academic Press, New York (1956).
42. J. H. Northrop and J. S. Murphy: Appearance of new phage types and new lysogenic strains after adaptation of lysogenic *B. megatherium* to ammonium sulfate culture medium. *J. Gen. Physiol.* **39,** 607–624 (1956).
43. G. N. Calkins: The life history of *Cytoryctes variolae,* Guarnieri. *J. Med. Res.* **11,** 136–172 (1904).
44. W. R. Duryee: Seminar on transmission studies on renal adenocarcinoma of the frog. *J. Franklin Inst.* **261,** 377–379 (1956).
45. T. H. Dunnebacke: Cytopathic changes associated with poliomyelitis infections in human amnion cells. *Virology* **2,** 811–819 (1956).
46. C. Morgan, H. Calderon, H. Rose, and D. Moore: Structure and development of viruses observed in the electron microscope. IV—Viruses of the R1-APC group. *J. Biophys Biochem. Cytol.* **2,** 351–360 (1956).
47. N. Rashevsky: Life, information theory and topology. *Bull. Math. Biophys.* **17,** 229–235 (1955).
48. C. F. Ehret and E. L. Powers: The systems and complexes of primary organelles in *Paramecium. J. Protozool.,* **3** *Suppl.* **3,** 5 (1956).
49. P. B. Weisz: Morphogenesis in protozoa. *Quart. Rev. Biol.* **29,** 207–229 (1954).
50. N. Rashevsky: Contributions to topological biology: some considerations on the primordial graph and on some possible transformations. *Bull. Math. Biophys.* **18,** 113–128 (1956).
51. J. N. Bigelow: Unpublished.
52. H. Quastler: Unitization in biological systems. *Essays on Biological Unitization,* Report No. R-52, Control Systems Laboratory, University of Illinois, Urbana (1953).
53. J. R. Preer: Gel diffusion studies on the serotypes of *Paramecium.* Jordan Hall Dedication Program, Indiana University (1956).
54. E. Borysko and F. B. Bang: Electron microscope observations on nucleolar structure. *J. Appl. Phys.* **23,** 163 (1952).
55. H. J. Weinstein: An electron microscope study of cardiac muscle. *Exp. Cell. Res.* **7,** 130–146 (1954).

56. A. MOSCONA: Cytoplasmic granules in myogenic cells. *Exp. Cell Res.* **9**, 377–380 (1955).
57. G. E. PALADE: The endoplasmic reticulum. *J. Biophys. Biochem. Cytol., Suppl.* **2**, 85–98 (1956).
58. A. J. HODGE: The fine structure of striated muscle. *J. Biophys. Biochem. Cytol., Suppl.* **2**, 131–142 (1956).
59. J. J. WOLKEN: A molecular morphology of *Euglena gracilis*. *J. Protozool.* **3**, 211–221 (1956).
60. B. A. AFZELIUS: The fine structure of the sea urchin spermatozoa as revealed by the electron microscope. *Z. Zellforsch.* **42**, 134–148 (1955).
61. B. P. POTTS: Electron microscope observations on trichocysts. *Biochim. Biophys. Acta* **16**, 464–470 (1955).

* The last eight references are not specifically discussed in the text, but are referred to as examples in the Decision-Tree diagram, Fig. 2. Many of the cytological references cited concur with the data rather than with the interpretations of the authors.

QUANTIFICATION OF PERFORMANCE IN A LOGICAL TASK WITH UNCERTAINTY

A. RAPOPORT
Mental Health Research Institute, University of Michigan

Abstract—Tasks to which information theory has been applied characteristically do not involve 'reasoning', i.e. the drawing of inferences. The present paper explores the possibility of applying information theory to measuring performance of logical tasks. We note at once that any task in which a *necessary* conclusion must be reached from given information has *formally speaking* no information content. From the information-theoretical point of view, therefore, no information is gained in the process of solving a purely mathematical or logical problem, no matter how 'complex'.

There are problems, however, in which in addition to the making of inferences, information must be obtained in the process of solution. Success of solution can be measured by the rate of obtaining such information and by the degree of completeness with which it is utilized. Assuming complete utilization at each step, the efficiency of solution depends on the efficiency with which information is obtained. A classical example is the coin-weighing problem in which a deviant coin and the direction of its deviation must be determined in the fewest possible weighings. Information theory provides not only the minimum number of weighings for such a problem but also a method for constructing the best 'strategy'.

In the present paper a particular logical task with uncertainty is discussed from the information-theoretical point of view. It is shown that the construction of an information-getting strategy depends very strongly on the instructions given the subjects and on the inferences which the subjects make from the instructions. Thus the practical problem of quantifying the performance of a logical task carries within it certain ambiguities which must be resolved if information theory is to be of use in psychological tests based on such tasks.

INFORMATION theory is mainly concerned with a quantity called the amount of uncertainty associated with a situation in which choices or guesses are made. This uncertainty can be viewed as a measure of ignorance. For example, we are the more ignorant of the value about to be assumed by a random variable with a discrete domain, the more values it can assume and the more nearly equi-probable these values are.

Defining every situation of ignorance is a set of postulates with a subjective flavor. *Somebody* is ignorant. At least this is the case in real situations involving subjects whose state of ignorance is to be inferred. It may be argued from certain philosophical points of view that this intrusion of subjective concepts is unsatisfactory, and attempts should be made to circumvent them or to eradicate them altogether. I don't want to take sides on this question, but only to point to some of its manifestations by way of indicating its persistence. The question has been raised in connection with the foundations of probability theory. There the attempts to circumvent the subjective element have given rise to the so-called 'objectivist school', which sought to define probabilities of events 'objectively' in terms of the relative frequencies of the events. Opposition to

'subjectivist' notions are also, I think, at the root of many philosophical objections to quantum theories.

The question of where to draw the line between subjective and objective frames of reference also arises in connection with attempts to link information theory with thermodynamics so as to make it useful in theoretical molecular biology. Whatever the merits of the attempts may be, this area of investigation does bring out rather pointedly the necessity of examining the possibly 'subjective' postulates underlying the description of the situations studied. Indeed, the philosophical question 'What is a subjective postulate?' comes to the forefront in the region where thermodynamics, quantum theory and information theory, meet—a triple point.

There is an area of possible application of information theory, however, where clearly subjective postulates are not only unavoidable but central. This is the area of psychology. Psychology is a study of behavior. Since psychology has gradually outgrown the austere positivistic restrictions of strict behaviorism, it has become respectable again to include into psychological theory considerations based on how the situation looks from inside the subject. To be sure, these matters must somehow be inferred from overt behavior, but once they are inferred there is no reason why these 'subjective variables', for example, subjective probabilities, utility functions, and so on, cannot enter as parameters in a theory. Indeed, if these parameters are determinable and stable, they serve to 'objectify' the subjective and thus contribute to the success of psychology as a science.

Information theory was applied from its very inception to psychological investigations. These applications have often been criticized. The grounds for criticism have been many, but a recurrent theme has been the failure of many psychologists to realize that information theory is worthless without an underlying set of postulates for each situation. Just as the application of probability theory to any situation necessitates the determination of a 'sample space', that is, a set of elementary events with a priori assignment of probabilities or a probability distribution function, so is the case with information theory.

Yet it was shown by de Finetti (1), Savage (2), and others that a rigorous theory of probability could be constructed backwards. That is to say, beginning with certain preferences of individuals for certain outcomes as reflected in their choices of actions under uncertainty, a set of subjective probabilities of events could be inferred, provided certain 'rationality criteria' of behavior were satisfied by the individuals. The question to what degree such rationality criteria are in actuality satisfied is another question which has led to many interesting investigations in their own right; so it is not entirely an unfortunate one. It must in any case be admitted that the 'subjective probability' of an event can in principle be defined, and thus statements such as, 'The Democrats will with probability 0.6 win in 1960,' are not wholly devoid of operational sense, provided the expressed 'subjective' probability is inferrable by explicit rules from observed behavior and enjoys a certain stability. Such assertions have no sense in the conceptual framework of the objectivist school, since the election of 1960 is a unique event whose 'probability' cannot be deduced from a frequency of occurrence.

The operational definition of subjective probability introduces probability

theory into psychology in a significant way, not as a mere appendage to statistics. I think the situation is similar in the case of information theory. Instead of lamenting the ambiguity of the 'universe of discourse' in psychological situations, which stands in the way of a straight-forward application of information theory, we may well seek to infer the universe of discourse as it looks from inside the subject. This is, indeed, a central task of the psychologist, and it is improper for him to shun it.

It remains true, however (and it is just as it should have been), that the early psychological experiments based on information-theoretical considerations were constructed in such a way as to eliminate idiosyncratic subjective probabilities. In memory tasks one starts with some set of thoroughly randomized and presumably equalized stimuli. One introduces redundancies in terms of actual biases of occurrence-frequencies objectively determined. The same techniques prevail in experiments in which the capacity of the individual as a channel is measured. These are all attempts to translate into experimental psychology situations occurring in communication engineering. The human being is studied as a piece of communication apparatus. I believe this strategy to be entirely correct as far as it has gone. I am sure, however, that its limitations are apparent most of all to the investigators who pursue it. Somewhere along the line a transition must be made which will allow the application of information theory to psychology as distinct from psychophysics. In other words, the perceptual world of the subject must eventually become a focus of interest. There is no reason why information theory should not become a useful tool in such investigations.

Characteristically, the tasks just described (such as rote learning, multiple choice responses, and so on) do not involve the deductive process. Indeed, from the formal information-theoretical point of view, the results of information theory are not applicable to a deductive process, because there is no 'uncertainty' in such processes. The solution of a mathematical problem, no matter how complex, yields no information from the information-theoretical point of view. From either the common-sense or the psychological point of view such a conclusion seems bizarre. The information-theorist can, of course, argue that his technical definition of information departs from common-sense and psychological notions of what constitutes information, and he is technically correct. Yet it might be instructive to try to bring the two concepts of 'information' into closer agreement.

Formally speaking, no information is gained in the solution of a purely mathematical or logical problem, because the solution is implicitly contained in the already known conditions. But the solution is not initially known to the subject. Is there a way to measure the extent of his ignorance? There might be, if we are willing to abandon the omniscient position from which the solution is seen as a necessary inference, hence of zero uncertainty, and enter the perceptive of the cognitive field of the subject, to whom only a range of possible solutions and, perhaps, associated subjective probabilities present themselves. But how does one get into this perceptive field? Obviously by observing the subject's behavior. But how does one make inferences from the observations to what that perceptive field may be? It would be gratifying to be able to say that for every configuration of the cognitive field, there is a specific behavior pattern,

but unfortunately this is not the case, as will be seen in the following example.

Suppose we present the subject with the famous coin-weighing problem. One of twelve coins is of odd weight. It is required to determine the coin and whether it is lighter or heavier than the rest in a minimum number of weighings on a balance using only the coins for weights.

Information theory not only reveals that three weighings are necessary and sufficient but also indicates the strategy. Obviously there are $\log_2 24 = 4.59$ bits of information (uncertainty) in the problem. A weighing can yield a maximum of $\log_2 3 = 1.59$ bits. Therefore, at least three weighings are necessary and may be sufficient. Further analysis shows that the first weighing can yield the full 1.59 bits and that only if four coins are weighed against four. The second weighing must involve six coins chosen in such a way that the three outcomes have probabilities 3/8, 3/8, 2/8, which yields 1.55 bits. The third weighing will therefore involve three coins, that is, 1.59 bits, in three fourths of the cases and two coins or 1 bit in one fourth of the cases, i.e. an average of 1.45 bits. The total information is 4.59, exactly equal to the initial uncertainty.

Now the uncertainty in the problem as it is presented is clearly perceived. At least it is easy to recognize that there are initially twenty-four possibilities. It takes some effort to determine the remaining uncertainty after each weighing, but it is none too difficult to do so. We may therefore suppose that in most instances the 'uncertainty' of the problem is perceived by a fairly intelligent subject correctly, that is, in accordance with the 'objective' assignment of uncertainty. However, it is by no means true that the majority of subjects proceed to the solution in the optimal way. That is, they cannot deduce the 'correct' strategy, even when they perceived the 'actual' amount of uncertainty in the problem.

It appears, therefore, that it is too much to expect to be able to deduce the subject's personal evaluation of uncertainty from his strategy in the solution of a problem in which both the deductive process and resolution of actual uncertainty must operate. However, this circumstance only reveals the situations to be more 'psychological' than they appear in the light of the personal evaluation of uncertainty. Not only is this evaluation personal but also the choice of strategy is, and the latter is by no means always optimal relative to the uncertainty perceived. We are reminded of a similar difficulty in the psychology of decisions in which subjective estimates of probabilities and subjective utility functions are intimately intertwined.

As pointed out, ours is a similar problem. Assuming that the solution of a logical task with uncertainty will be determined by two 'subjective' characteristics, namely, (a) the amount of uncertainty perceived by the subject at each step, and (b) his preference of strategy for a given amount of perceived uncertainty, then our problem is to determine these subjective characteristics in the course of a solution of a problem. It should be mentioned that some obvious techniques for determining subjective uncertainty are in most cases unusable. If, for example, the solving process is interrupted to ask the subject what he does or does not know, the subject may through these questions become aware of relations he had not been aware of or he may doubt some assumptions he had been making correctly but with insufficient justification.

I will now describe a task which has been adapted to an analysis of the problem-solving process in such a situation.*

The subject is faced with a board on which nine numbered light bulbs are arranged in a circle, at the center of which is a tenth bulb. Each of the peripheral bulbs may be lit by an adjacent button. Moreover, relays are so arranged inside the apparatus that lighting of certain lights may result in the lighting of other lights following a constant 'synaptic delay' of three seconds. A 'problem' is a programming in the apparatus so that certain causal relations are established among the lights. These causal relations are *only partially* represented by arrows on a chart attached to the mounting board. Examples are given in Figs. 1 and 2.

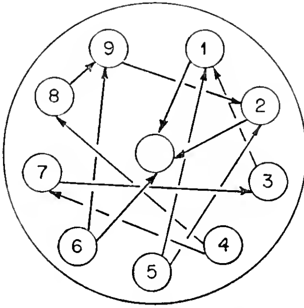


FIG. 1. Problem 2 on PSI

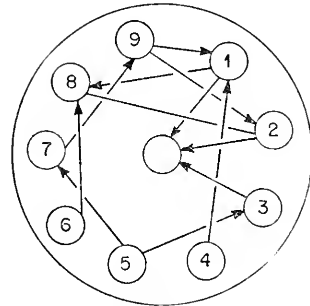


FIG. 2. Problem 3 on PSI

The point of the problem is that the meanings of the arrows on the chart are ambiguous. An arrow from A to B may mean that A is necessary to light B or sufficient, or both, or that A inhibits B. The subject's task is to obtain sufficient information about these relations, by pushing any button he chooses, to be able to cause the center bulb to be lit by manipulating buttons 4, 5, and 6 only. We will refer to these as the circled buttons.

There is a unique solution to each problem, consisting of a certain sequence of pushes of the circled buttons or of their combinations. For example, the solution to problem 2 (Fig. 1) is the pushing of buttons 4, 6, 5, 6 in the successive time periods. The solution to Problem 3 (Fig. 2) is 5, 0, 45, 6, 45.

There are a number of 'rational' approaches to the problem. Let us begin by making a chart of the connections indicated by the arrows. Figs. 1 and 2 are formally equivalent (as linear graphs) to Figs. 3 and 4. However, many characteristics of the problems are visually displayed in Figs. 3 and 4, which immediately suggest various lines of attack. These charts display what the

* The apparatus to be described, 'PSI', based on the isomorphism of certain networks of relays and the calculus of propositions (previously discovered by SHANNON (4) and by MCCULLOCH and PITTS (5)) was developed in Chicago by R. JOHN, J. G. MILLER, S. MOLNAR and H. J. A. RIMOLDI. The adaptation of the instrument to experiments of the type described is largely due to JOHN (3), who has listed a great number of performance parameters to be observed in the problem solving process. Of these the so-called 'inferential lag' (defined below) seems to me of particular importance. John's terminology and definitions differ somewhat from those in this paper, but the basic ideas (as yet unpublished) have been the point of departure for the present analysis.

subject does not know. For example, in Fig. 3, the convergence of arrows leading from lights 1, 2, and 6 upon C (the center light) induces the following questions: Which combinations of 1, 2, and 6 are necessary or sufficient or both to light C? Is one or more of them an inhibitor? If so, which one? Similar ambiguities are apparent at the other three junctures, that is at lights

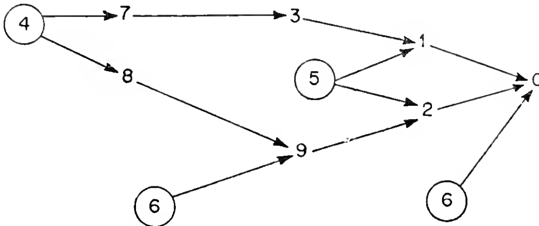


Fig. 3. Problem 2 displayed in time sequence

1, 2, and 9. A single arrow leading to a light presents no ambiguity. This is in consequence of the condition explained to the subject that each arrow 'means' something, hence a single arrow can mean only 'necessary and sufficient', otherwise its presence or absence would make no difference.

The subject can now ask specific questions. He can ask, for example, a question about each converging juncture. The question can pertain to the

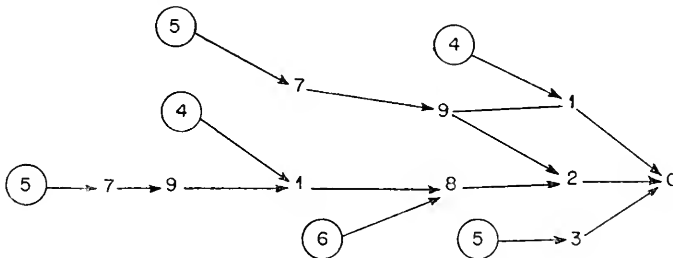


Fig. 4. Problem 3 displayed in time sequence

meaning of the arrows or to the combinations necessary or sufficient to light the bulb on which the arrows of the juncture converge. In the first case, he will be labeling arrows, in the second case the lights. Or he may proceed in a different way. Noting that all the possibilities of solution are displayed by the circled buttons in their proper time sequence, he may ask how each button is involved in the solution when its turn comes, by being pushed or *not* by being pushed.

Note that each of these perceptions of the problem implies a different 'information content'. According to one scheme, one seeks a 'yes' or 'no' answer to every non-null combination at a juncture. There are sixteen such combinations in all in both problems, hence sixteen bits of uncertainty if the 'yes's' and 'no's' are assumed independent. They are not independent, but this interdependence can be arrived at *a priori* only by deduction, which may or may not be made. Thus the uncertainty of the situation depends on the

state of mind of the subject. According to another scheme (labeling arrows), one can assign the values to the arrows in eighteen different ways at the triple juncture† and in four different ways at each of the three double junctures. Hence there are 18×64 different ways. This gives a little over ten bits of uncertainty. According to the last scheme, one has to decide whether to push or not to push each of the circled buttons in the time period when they appear on the chart. Here Problem 2 seems to have four bits of uncertainty and Problem 3 seems to have six bits. Clearly, the amounts of uncertainty associated with each scheme are different, but so are the yields of each trial, because one counts the yield in different kinds of statements, which have different *a priori* probabilities of being true.

One can push the analysis still further and thus reduce the information content of each problem by utilizing the rule that each problem has a unique solution. In this analysis the 'sample space' would be all possible problems having unique solutions involving the circled buttons at the proper time periods. Several of such problems would 'map' on each solution, and since the number of problems mapping on each solution are not equal, neither are the probabilities of the respective solutions. The value 4 bits for Problem 2 is a consequence of the equi-probability of all sixteen solutions (strictly speaking fifteen, barring the null solution where no button is pushed). If the solutions are not equi-probable the information content is correspondingly reduced.

This calculation is extremely tedious and has not been carried out. It is mentioned only to stress the general idea that the information content of the PSI problems depends significantly on the 'sample space' according to which probabilities are assigned. This sample space is presumably chosen (perhaps unconsciously) by the subject; hence the amount of uncertainty in the problem is a 'subjective' quantity, difficult to ascertain but in principle inferrable from a thoroughgoing analysis of the problem solving process.

One sees thus that even pursuing a far-reaching analysis and assuming perfect memory, it is not easy to derive the best strategy in the sense of minimizing exploratory trials. When one takes into consideration the ambiguities present in the subject's mind, who may not even have the convenient visual representation of the time sequence in his mind's eye, one realizes that far more psychology than can be formally treated by information theory at this time is involved in the problem.

Nevertheless, it is possible to cast the problem into information theoretical terms. One hopes, at any rate, that the concepts of information theory can be extended to cover situations where the subject's perception of the problem is an important unknown, That is, information theory may help formulate such situations in quantitative and analytic language. We have attempted to do so in the following way. We record the successive trials. Each trial must yield at least one of the sixteen 'crude facts', i.e. combinations of lights at each juncture

† In view of the rule that each arrow must have a meaning, the number of ways values can be assigned to the arrows equals the number of distinct irreducible disjunctions among the subsets of the arrows. Thus for three converging arrows there are seven non-null subsets (i.e. 'disjunctions' involving only one subset), three disjunctions among the singles involving two singles, three among the doubles involving two doubles, three involving a single and a double, one involving all three doubles, and one involving all three singles, eighteen in all.

which do or do not light the node of the juncture. Some trials yield more than one fact, but some yield no *new* facts. The fact yield is also recorded, counting the facts which the subject had the opportunity to observe, but not counting the inferences which he could have made.

Combinations of these facts make possible inferences about the meanings of the arrows at each juncture. For example, in Fig. 4, the facts '1 does not light 8' and '6 does not light 8' allow the inference that the arrows at this juncture should be labeled as shown in Fig. 5.

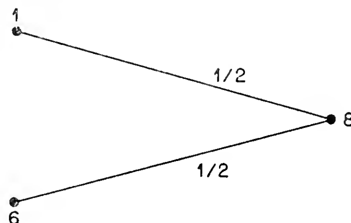


FIG. 5.

These possible juncture inferences are also recorded, and thus their rate of accumulation. Finally, the juncture inferences collate into the bits of information directly related to the solution of the problem—whether or not to push the circled buttons involved in the successive time periods.

When all this information is available (by inference, of course), there is no formal uncertainty left in the problem. However, in most cases the problem is not yet solved. The extra trials made by the subject, who has the solution available by inference, constitute the 'inferential lag'. We thus have various possible measures of subjective uncertainty over and above the 'objective' measures. The most obvious difference is revealed in the repetitions of trials (ordinary failure to record information obtained). Next we have the explicitly redundant trials, that is, those which while being new trials yield no new facts. Next the inferential lag already mentioned. All these can be measured both in time units and in numbers of trials.

The apparatus and the analysis of the problem solving process offer many opportunities for elaborate experimental designs, but they all hang on the question of how 'standard' these tasks are. In other words one needs to answer the question of whether there is a level of performance on each problem characteristic of a given subject, so that the variance in performance in a population of subjects can be adequately accounted for by a variance of some inherent ability.

Although this question has not yet been answered definitively, there are indications of a certain stability of performance. A set of experiments was performed at the Mental Health Research Institute, University of Michigan, in which the 'subject' in each case was a group of three students who solved the problems cooperatively by discussing each move and by coming to unanimous decisions on which move to make next. Eight such groups solved Problem 2 and then went on to solve Problem 3. The average number of moves for Problem 2 was about thirteen and for Problem 3 about nineteen. This is a first indication of the relative difficulty of the problems. That this difference is real is indicated

by the fact that *all eight* groups increased the number of moves from Problem 2 to Problem 3. When the groups were rank-ordered on their performance on Problem 2, the rank order was preserved (with just one reversal of two consecutive groups) on Problem 3. Another set of eight groups was given Problem 2 and then Problem 3 with a money incentive to minimize the number of moves on the latter. Under these conditions again all eight groups *decreased* the number of moves (averaging only eight on Problem 3). In spite of the radically changed situation, the rank order of these eight groups was again preserved from Problem 2 to Problem 3 (again with a reversal of only one pair of consecutive groups).

When groups are rank-ordered according to time of solution, no discernible correlation appears from one problem to another. These results point to a possible stable relation between the complexity of the problem and the effectiveness of solution strategy adopted by each of our trios of subjects. The lack of correlation in *rates* of performance points to possible extraneous effects such as the nature of the discussion process itself. At any rate the fact that the most prominent regularities are found in the performances as measured by the number of moves raises the hope that these regularities are the reflections of the uncertainty content of the problems as perceived by the subjects. It is noteworthy that, while on the level of observable crude facts and on the level of inferences about the meanings of the arrows, the two problems have the same uncertainty contents (about sixteen and ten bits each), on the level of major inference involving the circled buttons, Problem 2 has four bits of uncertainty while Problem 3 has six. The approximately 50 per cent increase in the average number of moves from Problem 2 to Problem 3 may well be a reflection of the increase in uncertainty on that level. Whatever the case may be, the results warrant further experimentation with a view of establishing the expected level of performance of a given subject on a given problem, once the set of uncertainties on various levels of observation and inference characteristic of the problem and certain factors of strategy efficiency characteristic of the subject are known. It is evident that the number of various problems which can be programmed into the PSI apparatus is astronomical.

REFERENCES

1. B. DE FINETTI: La Prévision: ses lois logiques, ses sources subjectives. *Ann. Inst. Henri Poincaré* **7**, 1–68 (1937).
2. L. J. SAVAGE: *The Foundations of Statistics*, J. Wiley and Sons, New York (1954).
3. E. R. JOHN: *Contributions to the study of the problem-solving process*, Preprint No. 1, Mental Health Research Institute, University of Michigan, Ann Arbor (1956).
4. C. E. SHANNON: A symbolic analysis of relay and switching circuits. *Trans. Amer. Inst. Elect. Eng.* **57**, 713–723 (1938).
5. W. S. MCCULLOCH and W. PITTS: A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **5**, 115–133 (1943).

PART IV

DESTRUCTION OF INFORMATION BY IONIZING RADIATION

THE disorganization of highly ordered macromolecules of biological importance by the action of ionizing radiation is a field of study owning a half-century of history, a tremendous literature, and possibly a somewhat feeble accomplishment in terms of clear and unexceptionable conclusions. With the development of information theory, and its subsequent application to biological systems, there appears to be substantial basis for cherishing the hope that it may constitute a valuable tool in the analysis of the experimental results of radiobiology and their translation into knowledge concerning biological phenomena. The present section of the symposium is dedicated to this goal. The first two papers, by GORDY and by PLATZMAN and FRANCK, explore different aspects of the interpretation of physical and chemical effects of ionizing radiation on proteins and related substances; for without some measure of fundamental physical insight into the mechanisms of this action, the utilization of information theory in radiobiology would appear unlikely to emerge from an ineffectual state of pleasant vagueness. In the third paper, by MOROWITZ, positive steps are taken in the analysis of some relationships between information theory and ionizing-radiation action. The following two short papers, which stem from discussion by KOCH and by AUGENSTINE, are devoted to the almost perennial question of the role of sulfur-bonding in radiobiology, as was also, to a large extent, the further discussion at the symposium, part of which is summarized in the final pages of the section. It is disquieting to have to record that the views on this perplexing problem are still seriously discordant.

R. L. P.

ELECTRON SPIN RESONANCE IN THE STUDY OF RADIATION DAMAGE*

WALTER GORDY

Department of Physics, Duke University, Durham, North Carolina

Abstract—It has been demonstrated by a Duke University microwave group that the electron spin resonance of the resulting unpaired electron can give specific information about the radiation damage in proteins, nucleic acids, and many other biologically significant chemicals. The structures of their electron resonances show that free radicals of various types are formed from the different amino acids and simpler peptides by ionizing radiations. However, in numerous proteins only two structural patterns are obtained, either separately or in combination. One of these is like the common pattern obtained for cysteine, cystine, and glutathione and is believed to arise from an unpaired electron (electron hole) on the protein sulfur. The other pattern (obtained alone in proteins which have no sulfur) is a doublet characteristic of the interaction of the electron spin with the spin of a single proton. The latter appears to arise from an electron on a carbonyl oxygen interacting with a proton of the hydrogen bridge, or possibly on a —CH— of the peptide chain which has lost an R side group. There is no evidence that the ionizing radiation breaks the polypeptide backbone structure of the proteins. The results seem to require that an electron hole or vacancy created at a given location in the protein molecule can migrate to other locations where it has lower energy.

I. INTRODUCTION

YESTERDAY evening when coming over from the airport I discovered that I was in the car with a biologist. After making this discovery, about half way over, I asked my new acquaintance what it is that the biologists expect of the physicists, what help—if any—we physicists can be to them. He told me that we could give them better instruments. What they need as biologists, he said, are newer and better instruments to see into things. He made no mention of information or theory. I didn't ask him whether we were to bring the instruments or just send them by mail. Nevertheless, I think that a physical instrument which brings information out of biological things should be accepted as a ticket of admission to a discussion of information theory in biology, especially one held under the auspices of physicists!

The instrument which I offer as an admission ticket was not invented by me. Electron magnetic resonance was discovered in 1945 by a Russian scientist, ZAVOISKY (1). Nor can I claim to be the first to apply electron resonance to the study of radiation damage. That, I believe, was first accomplished by HUTCHISON (2), who in 1949 detected *F*-center resonance in neutron-irradiated alkali halides. Our group at Duke University, we are proud to say, was among the first to show the applicability of electron magnetic resonance in the study of biological substances, and the first, we think, to detect such resonances in irradiated proteins. COMBRISSON and UEBERSFELD (3), independently of our

* This research was supported by the United States Air Force through the Air Force Office of Scientific Research ARDC contract No. AF 18(600)-497.

work, found resonances in certain amino acids. Their results did not agree with ours, except with those for glycine.

Our group has now obtained electron spin resonances of scores of biological substances which have been subjected to ionizing radiation. These include amino acids (4), peptides (5), fatty acids (6), nucleic acids (7), various proteins (4, 8), enzymes (8), hormones (9), and vitamins (9). Some of these results we think we understand, at least partially; others we do not pretend to understand. This does not discourage us, however. Some twenty to thirty years were required for obtaining reasonably definitive interpretations of X-ray diffraction patterns of a few of the simpler proteins. Nevertheless, it must have been apparent from the first that these patterns contained a wealth of information which would eventually be decoded by the persistent scientist. In electron spin resonance we now have a direct method for studying radiation damage which is comparable to the X-ray diffraction method for the study of structures. It is, in fact, a specific for such studies, for it 'sees' not the normal biological matter but the radicals, or broken pieces of molecules torn apart by ionizing radiations.

Descriptions of microwave spectrometers for observation of electron magnetic resonances are available (10, 11). Such spectrometers can now be obtained commercially. Descriptions of theoretical methods and applications to chemical and biochemical problems are given in recent publications (11, 12, 13, 14, 15, 16).

In the observation of electron magnetic resonance the sample to be investigated is placed in a microwave cavity at a point where the magnetic component of the microwave radiation is strongest. The cavity containing the sample is so placed in a d.c. magnetic field that the lines of the d.c. field lie perpendicular to the magnetic component of the microwave radiation. When the d.c. field is adjusted to the proper strength for resonance, microwave radiation will be absorbed. The value of the field for resonance is:

$$H = \frac{h}{g}\frac{\nu}{\beta} \tag{1}$$

Numerically,

$$H\,(\text{gauss}) = 0.7143\nu\,(\text{Mc/sec})/g \tag{2}$$

where g is the spectroscopic splitting factor for the paramagnetic species. It is found that for practically all organic free radicals, including those produced in solids by ionizing radiation, the g value is very close, within a fraction of a per cent, to the g factor for the free electron spin, 2.0023. This comes about because possible orbital moments are largely averaged out by the motion of the unpaired electron, or by the spreading out over a number of atoms (delocalization) of its molecular orbital. The persistent observation of a g factor near that of the free electron spin has led to the designation of this resonance as electron spin resonance.

In the vector model, the electron spin vector would precess about the direction of the applied field H. Quantum mechanically there are only two stable orientations for this precessing vector, which represents an average or the 'expectation value' for the electron spin momentum. These correspond to the two observable components, $+\frac{1}{2}$ and $-\frac{1}{2}$, of the electron spin vector along a fixed direction in space. Because of the interaction of the magnetic moment of the spinning

electron with H, the potential energy of the electron is slightly greater for one of the orientations than for the other. The difference in energy for the two orientations is equal to the microwave quantum energy $h\nu$ which will induce the spin vector to flip over from one orientation to the other. The classical Larmor precessional frequency of the electron spin vector about the direction of H is equal to the absorbed microwave frequency. Thus the precessing electron is in tune with, or at resonance with, the microwave radiation.

In normal organic matter about us, the electrons are all—or nearly all—in the lowest orbital levels, with the maximum limit of two electrons in each molecular orbital. According to the Pauli principle, two electrons can share an orbital only if their spins are aligned in an antiparallel manner. If it is assumed that the spin vector of one electron flips over in an imposed field, that of its orbital mate must flip in the opposite direction at the same time, thus preventing any observable absorption or emission of radiation. To produce an observable electron spin resonance in normal organic matter, one must by some means lift electrons out of the completely filled orbitals of the ground level. Strong ionizing quanta, such as those of x-rays, can eject electrons from ground molecular orbitals with sufficient energy to free them entirely from the parent molecule. If a molecule loses a single electron through ionizing irradiation, the ionized molecule—if it holds together—will have a single unpaired electron in one of its orbitals. This electron is now free to flip over in an external field without the opposite flipping of a partner. The singly ionized molecule is thus paramagnetic and can execute electron spin resonance. Furthermore, the electron which is knocked away from one molecule may become attached to a neighboring molecule and thus convert it into a negatively charged radical. Since the latter molecule is presumed to have all its bonding orbitals filled, the new arrival must go into an orbital of higher energy and remain unpaired. Thus resonance of electrons on negatively charged molecules might likewise be detected. If the electron is ejected with sufficient energy it may, of course, ionize several molecules before coming under the control of a particular molecule. The end result is the same, however, except that a single quantum has, in effect, been able to ionize more than one molecule. Two types of charged radicals are thus produced. If the barrier to the return passage of the electron between the molecules is high, as is the case in most organic solids, a sufficiently high concentration of charged radicals can be built up in this way to give a detectable electron spin resonance. The molecules may be small ones such as the amino acids or long-chain macromolecules such as the proteins or nucleic acids. The only requirement is that the separated electrons cannot easily become paired again, i.e. that the radicals produced by ionizing radiation have a lifetime sufficiently long for a detectable quantity to be built up.

II. NATURE OF INFORMATION CONTENT IN ELECTRON SPIN RESONANCE

If the spin of an odd electron of a radical were entirely free from perturbing influence of its environment, its resonance would be a single, sharp, isotropic line with a g factor of 2.0023. Not much information is contained in such a simple signal, although one could measure the lifetime of the radical from

its rate of decay. Also, the very fact that electrons could achieve such freedom within an organic solid might itself be classed as desirable information. Fortunately, however, the electron resonance signals are often rich with information about the environment of the unpaired electrons. Our problem is to decode their messages. There are at least three important sources of information in these resonances. The first is the hyperfine structure arising from interactions of the electron spin moment with magnetic moments of various nuclei around or near the unpaired electron. The second is the small residual spin-orbit interaction which in some instances makes the g factor slightly anisotropic and different from the free spin value of 2.0023. The third is the information which can be obtained from the line widths and shapes. The most important of these sources is the nuclear hyperfine structure.

Most instruments used for detection of electron spin resonance plot the intensity of absorption at a particular frequency as a function of d.c. magnetic field. The appearance of the plot depends upon the instrument as well as upon the actual, intrinsic shape of the resonance. I shall not discuss possible variations in the actual line-shapes, but shall here assume that the resonances have gaussian shape when the intensity of absorption at a constant frequency is plotted as a function of d.c. magnetic field. A high-fidelity receiver and recorder (or cathode ray scope) would reproduce closely the actual shape of the resonance curve, as shown in Fig. 1(a). The high-fidelity systems are not, however, the most sensitive systems. The most sensitive methods of detection employ modulation of the resonance relative to the observation frequency. A narrow-band amplifier is tuned either to the modulation frequency or to a higher harmonic of this frequency. If one uses a frequency modulation which is very small as compared to the width of the resonance and a phase-sensitive amplifier tuned to the modulation frequency, a curve like that in Fig. 1(b) is obtained. This curve represents the first derivative of the actual line-shape. If one uses such a receiver and tunes to the second harmonic of the modulation frequency, a curve like that in Fig. 1(c) is obtained. This curve represents the second derivative of the actual line-shape. Both the first and second derivative curves are commonly employed in display of electron spin resonances. In interpretation of the curves it is desirable to know what method of detection has been employed, especially when there are structural components incompletely resolved. In the illustrations which follow we shall sometimes use first and sometimes second derivative displays.

This brief description of the appearance of the signals and the simplified theory of the structure of the resonance given below will, I hope, make it possible for you, whether you are a biologist, chemist, physicist, or hybrid, to share with us some of the fun of trying to decode the complex microwave messages which we have been receiving from biological substances. You will be able, I hope, to decide for yourself what is definitely proved by the resonances, what is strongly suggested but not proved, and what is merely hinted.

1. *Nuclear Hyperfine Structure*

The hydrogen nucleus, with a relatively large magnetic moment, 2.79 nm, and nuclear spin of $\frac{1}{2}$, is abundant in all organic matter. The only other nucleus with a non-zero spin abundantly found in biochemicals is N^{14} ($I = 1$ and

$\mu_I = 0.40$ nm). Carbon, oxygen, and sulfur are of course also prominent constituents of biochemical matter, but their most abundant isotopes have zero nuclear spins and hence cannot interact with the electron spin. In strong resonances one might detect effects caused by C^{13} (spin $\frac{1}{2}$ and natural abundance 1.12 per cent) or S^{33} (spin 3/2 and natural abundance 0.74 per cent). For some

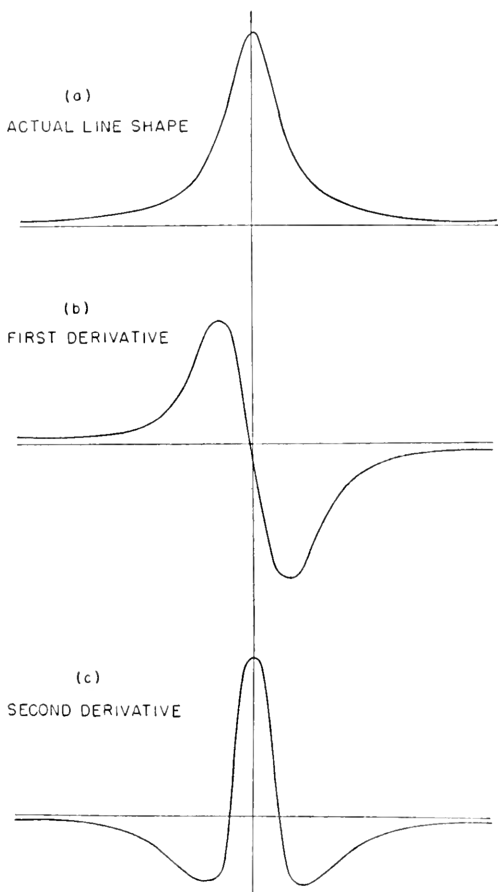


FIG. 1. Appearance of resonance signals as detected in various ways: (a) High fidelity. (b) First derivative curve obtained by small modulation of the resonance with a phase-sensitive receiver tuned to the modulating frequency. (c) Second derivative curve obtained by small modulation of the resonance with phase-sensitive receiver tuned to twice the modulation frequency.

substances one can obtain samples concentrated with C^{13}, S^{33} or O^{17}. Hyperfine structure of their nuclei thus obtained will greatly augment the information obtained from proton hyperfine structure, but it is fortunate for these studies that C^{13} is not the more abundant isotope of carbon. If hyperfine structure from all the nuclei were present at one time, the resulting pattern would often be unresolvable and its decoding thus more uncertain. As it is, there is seldom

any ambiguity about the identity of the nucleus which gives rise to the nuclear hyperfine structure of electron resonances in irradiated organic matter. Usually, it must be hydrogen. By substitution of deuterium for hydrogen, one should often be able to learn which hydrogens give rise to a particular splitting.

When the electron spin resonances of organic radicals are observed in the microwave region at frequencies of 30,000 Mc/sec, the corresponding magnetic field required is 10,700 gauss. A magnetic field of such strength is usually sufficient to produce the Paschen-Back effect, in which the $\boldsymbol{I} \cdot \boldsymbol{S}$ coupling is broken down and both $\boldsymbol{I}$ and $\boldsymbol{S}$ precess about the direction of $\boldsymbol{H}$. Under these conditions the resonance frequencies of the various components at constant field strength H_0 can be expressed as:

$$h\nu = g\beta H_0 + \sum_i A_i m_i \tag{3}$$

where A_i is the coupling constant of the electron for a particular nucleus i with spin I_i and where the magnetic quantum numbers have the values:

$$m_i = I_i, \quad I_i - 1, \cdots -I_i. \tag{4}$$

Usually the resonances are observed at a fixed frequency, ν_0, by variation of the d.c. magnetic field. The resonant field strengths for the various hyperfine components are then:

$$H = H_0 + \frac{1}{g\beta} \sum_i A_i m_i \tag{5}$$

$$= H_0 + \sum_i \Delta H_i m_i \tag{6}$$

The summation is again taken over all the coupling nuclei for each combination of the magnetic quantum numbers. All orientations of a given nucleus (all values of its m) are equally probable and independent of those of the other nuclei. In this expression $H_0 = h\nu_0/g\beta$ is the resonant field strength for the central component of the structure at the observation frequency, ν_0, or that for resonance if there were no nuclear perturbation; ΔH_i is the component separation (in magnetic field units) caused by a particular nucleus i. Obviously, $\Delta H_i = A_i/g\beta$. In these applications we can set g as equal to 2.00 and write:

$$\Delta H_i \text{ (gauss)} = A_i \text{ (Mc)}/2.80. \tag{7}$$

If all the coupling nuclei in a given free radical have the same coupling to the electron spin, one can define

$$T = \sum_i I_i, \tag{8}$$

and

$$m_T = T,\ T - 1,\ T - 2,\ \cdots,\ -T, \tag{9}$$

and can write equation (6) in the simpler form:

$$H = H_0 + (\Delta H)M_T \tag{10}$$

There will be $(2T + 1)$ components corresponding to the different values of M_T. This simplification is often possible in organic free radicals in solids where the coupling nuclei are all hydrogens. It is apparent that, where all

the equally coupling nuclei have the same spin, $T = nI$, and the total number N of components of the multiplet will be:

$$N = 2nI + 1 \tag{11}$$

or

$$n = \frac{N - 1}{2I} \tag{12}$$

Thus n equally coupling hydrogens ($I = \frac{1}{2}$) gives $n + 1$ components. The intensities of the components are proportional to the number of different combinations of the m_i's which give the same sum $\sum_i m_i$ or same value of M_T. Because all the $+\frac{1}{2}$ and $-\frac{1}{2}$ orientations of n hydrogens are equally probable and mutually independent, the intensities of a multiplet formed by equally coupling hydrogens will be gaussian.

The interaction constant A_i of the electron spin with the moment of a particular nucleus may contain both an isotropic and an anisotropic component. The isotropic component, the Fermi term, is independent of the orientation of the sample in the magnetic field and arises from the non-vanishing probability density, $\psi_0 \psi_0^*$, of the electronic wave function at the nucleus in question. Since only the s atomic orbitals are non-vanishing at the nucleus (radius $r = 0$), the presence of an isotropic coupling term for a particular atom in a molecule generally indicates s character in the bonding orbitals of that atom.

For an unpaired electron occupying wholly an s orbital of a particular atom, the coupling to the nucleus of that atom arises entirely from the non-vanishing density $\psi_0 \psi_0^*$ at the nucleus and has the value (17):

$$A_s = \frac{16}{3} \beta\beta_I g_I \psi_0 \psi_0^* = \frac{8}{3} \frac{hcg_I R\alpha^2 Z^3}{n^3} \tag{13}$$

where β is the Bohr magneton; β_I, the nuclear magneton; g_I, the g factor (μ_I/I) of the nucleus; h, Planck's constant; c, the velocity of light; R, the Rydberg constant; α, the fine structure constant; Z, the effective nuclear charge; and n, the total quantum number. For atomic hydrogen in the ground state, A is known to be 1420 Mc/sec. This value with equation (7) gives $\Delta H_i = 507$ gauss as the expected splitting for the atomic hydrogen doublet for the strong-field case ($H \gg \Delta H_i$). The non-isotropic components are zero because of the spherical symmetry of the s orbital. Thus the isotropic coupling to the nucleus of a particular atom gives a good measure of the s orbital contribution of that atom to the molecular wave function of the odd electron in a free radical.

An electron at a fixed distance from a nucleus i with non-zero spin will experience a magnetic field component arising from the magnetic moment of the nucleus. If the spin vectors of both the electron and the nucleus precess about the direction of an applied field H (this corresponds to the strong-field, Paschen-Back case), the non-vanishing field component at the electron, ΔH, caused by the nucleus, will lie along H and will have the value:

$$(\Delta H) = \left(\frac{m}{I}\right) \mu_I \beta_I \left(\frac{1}{r^3}\right) (3 \cos^2 \theta - 1) \tag{14}$$

where I, m and μ_I are the spin, magnetic quantum number, and magnetic moment (in nm units) of the nucleus, β_I is the nuclear magneton, r is the radius vector from the nucleus to the electron, and θ is the angle between r and H. Although the nucleus may be regarded as located at a fixed point within the molecule or crystal, the electron definitely cannot be so regarded. Hence, to find the averaged or effective $(\Delta H)_{\text{eff}}$ acting on the electron in a molecular orbital ψ, we must average the above quantity over the orbital ψ. Thus

$$(\Delta H)_{\text{eff}} = \left(\frac{m}{I}\right)\mu_I\beta_I \int \psi\left(\frac{1}{r^3}\right)(3\cos^2\theta - 1)\,\psi^*\,d\tau \quad . \tag{15}$$

Since the coordinates are separable, we can write this equation as:

$$(\Delta H)_{\text{eff}} = \left(\frac{m}{I}\right)\mu_I\beta_I\left\langle\frac{1}{r^3}\right\rangle_{\text{Av}}\langle 3\cos^2\theta - 1\rangle_{\text{Av}}, \tag{16}$$

where

$$\left\langle\frac{1}{r^3}\right\rangle_{\text{Av}} = \int \psi_r\left(\frac{1}{r^3}\right)\psi_r^*\,d\tau$$

and

$$\langle 3\cos^2\theta - 1\rangle_{\text{Av}} = \int \psi_\theta(3\cos^2\theta - 1)\psi_\theta^*\,d\tau \quad .$$

To attack such a problem one can assume, as is usually done in other calculations of molecular orbitals, that ψ is a linear combination of atomic orbitals, ψ_a, ψ_b, etc. We then readily get a part of the solution for we already know, at least to a fair approximation, $\left\langle\frac{1}{r^3}\right\rangle_{\text{Av}}$ and $\langle 3\cos^2\theta - 1\rangle_{\text{Av}}$ for electrons in various kinds of atomic orbitals. Expressions for these to various degrees of approximation together with coupling constants actually measured are available in standard texts on atomic spectra (17, 18). There is more to the problem than this, however. Although an overlap or cross term of the form $\psi_a(1/r^3)(3\cos^2\theta - 1)\psi_b^*$ may possibly be neglected, an electron in an atomic orbital of atom B might have a significant interaction with the nucleus of an adjacent atom A. It is thus necessary to include terms of the form:

$$\int \psi_b\left(\frac{1}{r_{ab}^3}\right)(3\cos^2\theta_{ab} - 1)\psi_b\,d\tau, \tag{17}$$

where r_{ab} and θ_{ab} are the coordinates of an electron on atom B referred to the nucleus of atom A as the origin. The values of these terms are sensitive functions of the hybridization of the atomic orbitals and of the direction of the projections of the major lobes of the hybridized orbitals. As we get greater skill in the procedure, these orientation-dependent coupling terms should give additional information about orbitals of radicals. Expressed in convenient numerical units equation (16) becomes:

$$\Delta H\,(\text{in gauss}) = 5.05\,\frac{m\mu_I}{I}\left\langle\frac{1}{r^3}\right\rangle_{\text{Av}}\langle 3\cos^2\theta - 1\rangle_{\text{Av}}, \tag{18}$$

where μ_I is in nm units and r is in Å.

In simple cases where single crystals can be prepared, it should be possible to measure $\langle 1/r^3 \rangle_{Av}$ for the interaction of an electron in atomic orbital of atom B interacting with the nucleus of another atom A. Such applications are made later in the discussion. If $\langle 1/r^3 \rangle_{Av}^{-\frac{1}{3}}$ is greater than the interatomic distance, the electron may be in a hybridized orbital of B which projects away from A. If it is less than the atomic distance, the electron may be in a hybridized orbital which projects toward A. In some instances $\langle 1/r^3 \rangle_{Av}$ may be so large that the field of the electron at the nucleus may be greater than the applied field. The nucleus would not then necessarily precess about the direction of H, and the above formula would not hold for all values of θ. It should still hold when θ equals zero or 90°, for then the field of the electron at the location of the nucleus would have, on the average, the same direction as H. If the cloud of the electron is symmetrical about the bond axis between A and B, the angle θ would, in effect, measure the orientation of the bond axis in the field H. For this case $\langle 3 \cos^2 \theta - 1 \rangle_{Av}$ equals 2 for $\theta = 0$ (bond axis parallel to H), and $\langle 3 \cos^2 \theta - 1 \rangle_{Av}$ equals -1 for $\theta = 90°$ (bond axis perpendicular to H). Thus the ΔH should have twice the value for θ equal to zero as that for θ equal to 90°. The dipole-dipole interaction of the electron with the nucleus averages to zero when the electron is entirely outside the nucleus and is moving in such a manner that its averaged density achieves spherical symmetry about the nucleus during its lifetime in a spin state.

Nuclear hyperfine structure of any type becomes independent of the magnetic field strength after the field becomes sufficiently strong to achieve the Paschen-Back case, which is assumed in the above treatment. Thus nuclear hyperfine structure can be readily distinguished from the splitting which arises from anisotropy in the g factor, discussed below, if measurements are made at two or more frequencies, both with strong fields. Although the direct-dipole type interaction with the nucleus varies with orientation in the field, it does not vary with the magnitude of the field after the strong field case is reached.

Figure 2 shows the type of hyperfine structure theoretically predicted for the strong field case for various radicals with equally coupling nuclei having spins of $\frac{1}{2}$ (H or F, for example). Figure 3 illustrates a few cases where the coupling of one or two of the nuclei differs from that of the others. It is apparent that these cases are easily distinguishable.

2. *Residual Spin-Orbit Coupling*

If the odd-electron density of a radical is largely concentrated on a non-s orbital of a single atom of a radical or is shared mainly by only two atoms, as it is on the —N—N— group of diphenyl picryl hydrazyl (DPPH), effects of spin-orbit interaction are not entirely negligible. The orbital momentum will be oriented by the strong electrical force of the chemical bond and will not be free to precess about the applied field. Bond-oriented orbital components will give rise to an observable anisotropy in the magnetic susceptibility and thus in the observed g factor. If the odd electron wave function is symmetric about a particular bond as in DPPH, the observed g factor will reflect this symmetry: if all such bonds in a given sample were oriented along the applied H, the $g_{\parallel}$ factor would differ from the $g_{\perp}$ observed when the bonds are all oriented

perpenidcular to H. For an arbitrary orientation θ of the bond axis with H, the observed g_θ factor would have the value:

$$g_\theta = \sqrt{g_{\parallel}^2 \cos^2 \theta + g_{\perp}^2 \sin^2 \theta} \tag{19}$$

In a sample in which the bond angles have arbitrary orientations in the field H such as would be true in a powder or polycrystalline sample, the

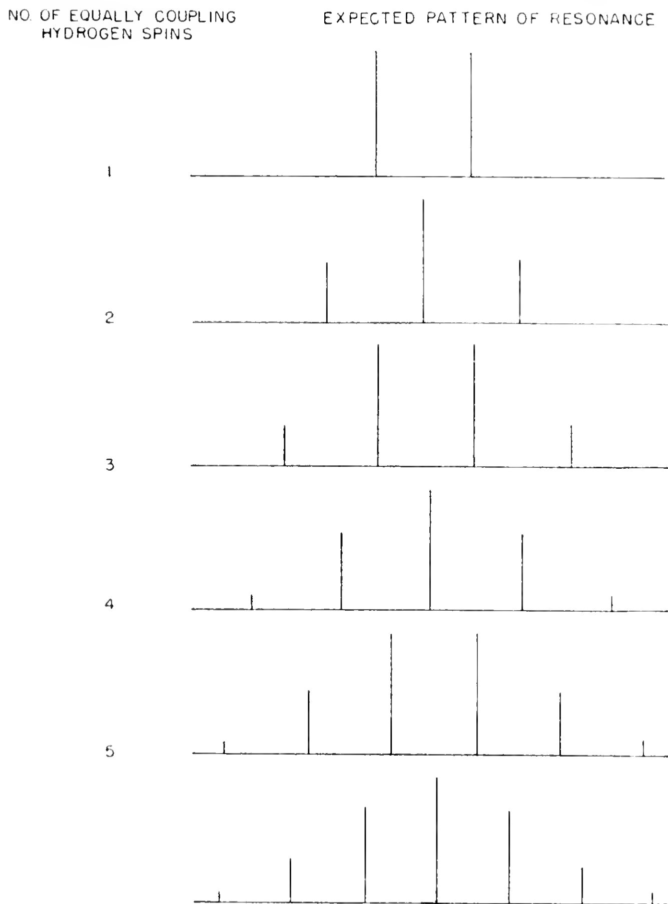


FIG. 2. Types of hyperfine structure predicted for strong-field case for various radicals having different numbers of equally coupling hydrogens or other nuclei of spin $\frac{1}{2}$.

resonance absorption would spread over all values of the field intermediate between that corresponding to the resonance value for $g_{\parallel}$ and $g_{\perp}$. The $g_{\perp}$ would apply for any orientation of H in a plane perpendicular to the bond axis, whereas the $g_{\parallel}$ value would apply only for H along the bond axis. Thus for random orientations in the polycrystalline samples, the $g_{\perp}$ value has the greater weight, and the resonance has an asymmetric form with the highest

peak corresponding to the $g_{\perp}$ value. Such a resonance will have a shoulder or shelf on one side with the edge of the shoulder corresponding to $g_{\parallel}$. First derivatives of an asymmetric resonance arising from an anisotropic g factor in X-irradiated cystine are shown in Fig. 4 for three different observation frequencies. That the apparent structure in these curves is due to anisotropy

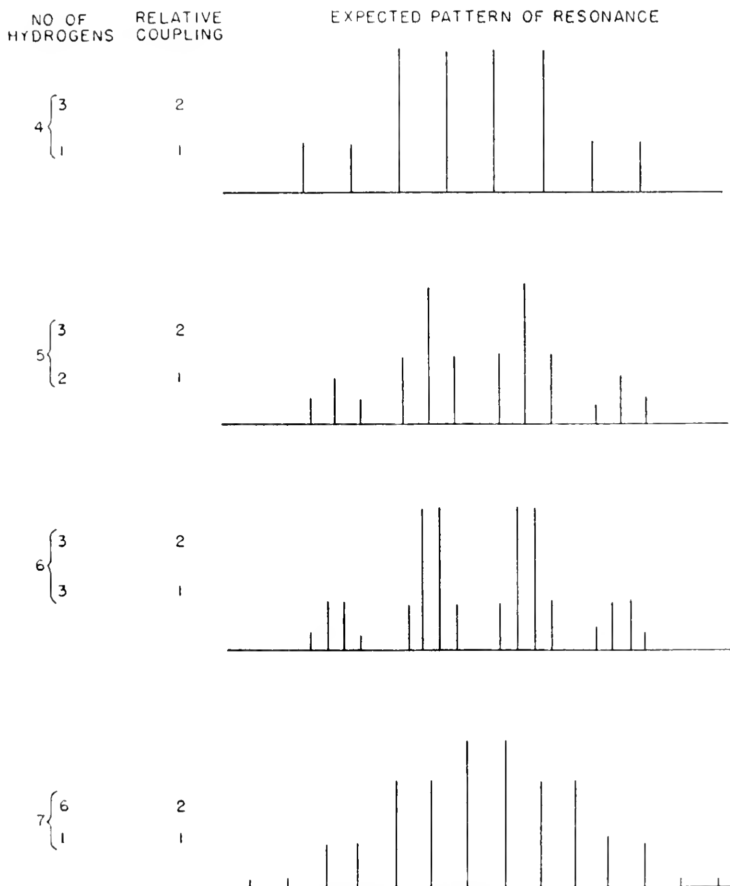


Fig. 3. Some illustrative theoretical hyperfine patterns of radicals with two sets of H nuclei (or other nuclei of spin ½). All nuclei of one set have the same coupling, but those of the two sets differ as indicated.

in the g factor has been established by measurements on a single crystal of cystine at different orientations in the field (19). Such curves show some differences depending upon amount of modulation, variations in natural line widths, degree of anisotropy in g, as well as variations in observation frequency or H value. Nevertheless, there should always be a bend point in the derivative curves corresponding to the H for resonance at $g_{\perp}$ and a lesser one for $g_{\parallel}$. This fortunate circumstance allows measurement of $g_{\parallel}$ and $g_{\perp}$ even in polycrystalline samples. The identity of these bend points, if in doubt, can usually

be established by variation of the modulation amplitude and observation frequency. The outermost bend points will in general correspond to $g_{\parallel}$ and $g_{\perp}$.

III. FREE RADICALS IN IRRADIATED AMINO ACIDS AND SIMPLE PEPTIDES

The work of our group at Duke University has revealed that the isotropic *s*-orbital contributions of hydrogen atoms in aliphatic hydrocarbon radicals are very significant and that they give rise to hyperfine structure in the spin resonance of these radicals which is frequently of the order of 100, and sometimes as much as 200, gauss. This coupling is an order of magnitude greater than that generally found for the aromatic ringed radicals (14, 15, 20) which can be prepared chemically and observed in solution. Furthermore, the first measurements indicated, and later work on single crystals (21) confirmed, that the isotropic *s*-orbital coupling to the hydrogen nuclei in aliphatic hydrocarbon radicals is generally much greater than the orientation-dependent, dipole-dipole component. This very fortunate circumstance makes possible detection and often identification of the aliphatic hydrocarbon radicals made by irradiation of solid matter in the polycrystalline powder and even in impure biological solids. In other words, it seems possible with microwave spectroscopy to 'fingerprint' many of the common radicals produced within solid matter by irradiation. I need not emphasize the usefulness of such a set of fingerprints for the study of radiation damage.

There are two important factors which we believe to be mainly responsible for the reduction of the anisotropic nuclear coupling in hydrocarbon radicals. One of these is the spreading of the odd electron density over a large molecular orbital so that there is no appreciable fraction of the total density near a particular nucleus. The other is the twisting, turning, tunneling, tumbling, or other motion of the radicals, or their parts, within the solid cages in which they are trapped. The first is generally more important for large radicals than for small ones, and the latter is generally more important for room temperature and elevated temperatures than for lower ones.

These properties of aliphatic free radicals and their remarkably long lifetime within solids were not predicted by theory. The conclusions were forced upon us from the experimental evidence for them. Furthermore, this pronounced isotropic interaction through the *s*-orbitals immediately gives much information about the electronic wave functions and structure of hydrocarbon radicals. The large coupling to the H nuclei in the CH_3 radical (total spread of quartet 70 gauss) indicates that this radical is not planar. Amazingly, the characteristic pattern of the ethyl free radical, C_2H_5, is a symmetrical sextet (or approximately so) of about 130 gauss spread. This indicates equivalent, or nearly equivalent, coupling to the electron spin of all five protons.

Fig. 5 illustrates some characteristic hyperfine patterns of hydrocarbon radicals produced by x-irradiation of some simple peptides. Compare these with the theoretical patterns for different numbers of equally coupling protons in Fig. 2. Similar patterns have been obtained by irradiation of amino acids (4) and other compounds (6, 22, 23) with x-rays and with ultraviolet light (24).

Figs. 6 and 7 illustrate somewhat more complex resonances.

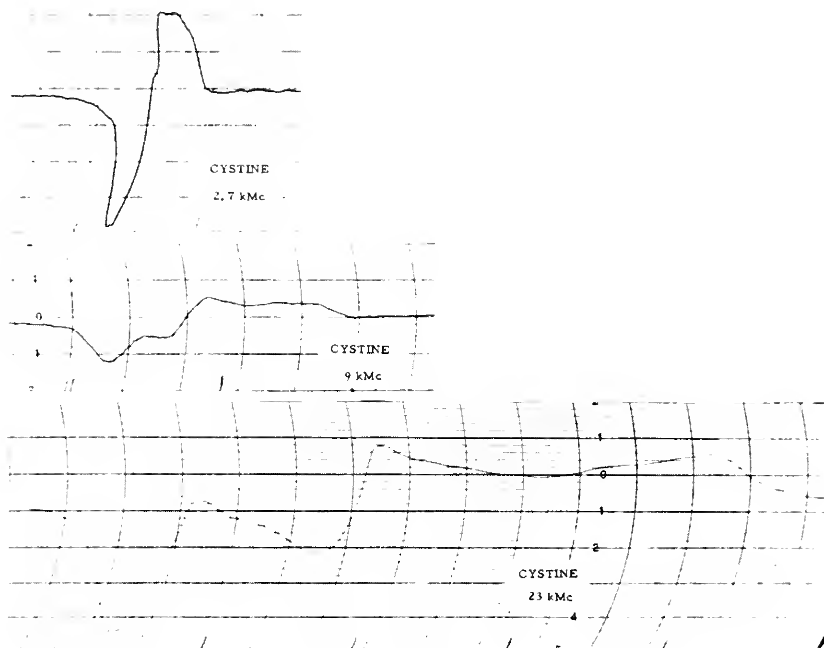


FIG. 4. First derivative curves at different frequencies of powdered cystine after x-irradiation in a vacuum. The markers at the base are 68 gauss apart. The top curve for 2.7 kMc requires a magnetic field of 960 gauss; the middle curve at 9 kMc, one of 3200 gauss; the bottom curve at 23 kMc, one of 8200 gauss.

to face p. 250

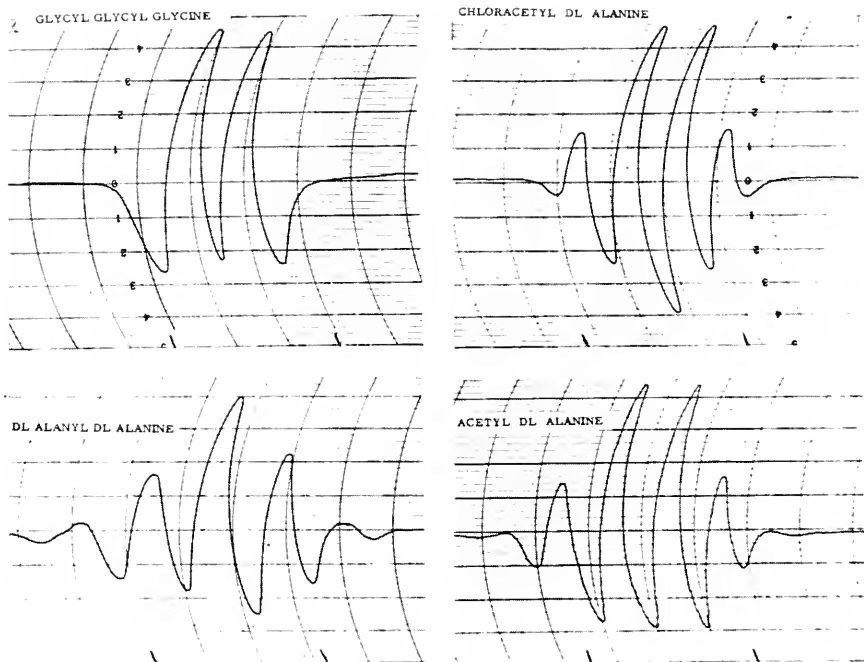


FIG. 5. Some illustrative patterns of resonances of x-irradiated peptides (second derivative curves). The markers at the base are spaced 68 gauss apart. The observation frequency is 9 kMc. From G. McCORMICK and W. GORDY (5).)

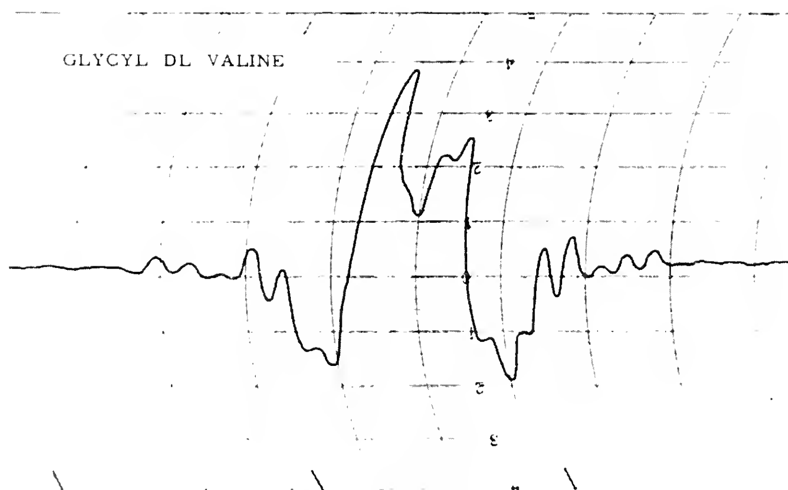


FIG. 6. Hyperfine pattern (second derivative curve) of the radical produced by x-irradiation of glycyl DL-valine. Marker spacing is 68 gauss. Observation frequency, 9 kMc. (From G. McCORMICK and W. GORDY (5).)

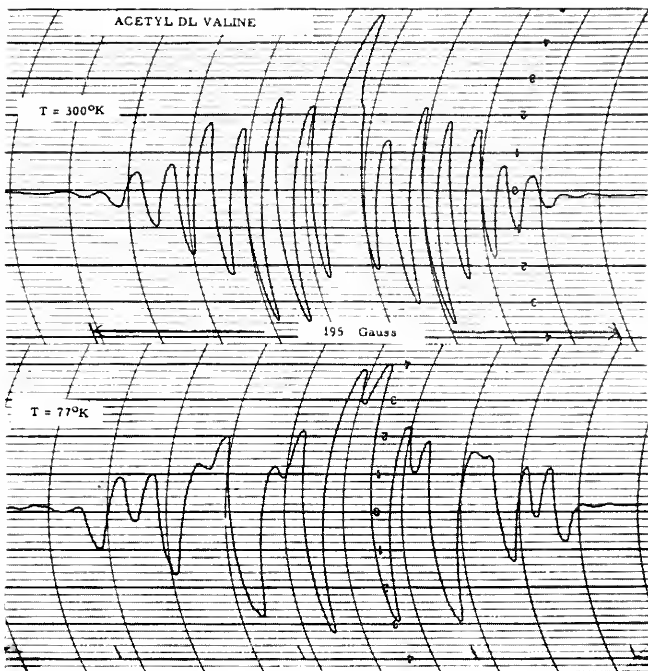


FIG. 7. Hyperfine pattern (second derivative curve) of the radical produced by x-irradiation of acetyl DL-valine. Marker spacing, 68 gauss. Observation frequency, 9 kMc. (From G. McCORMICK and W. GORDY (5).)

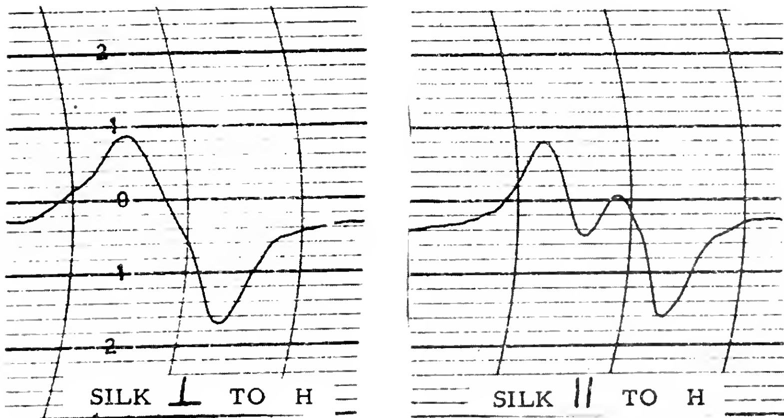


FIG. 8. Resonances (first derivative curves) obtained for x-irradiated silk with strands oriented parallel and perpendicular to the magnetic field. The observation frequency is 23 kMc. (From W. GORDY and H. SHIELDS (8).)

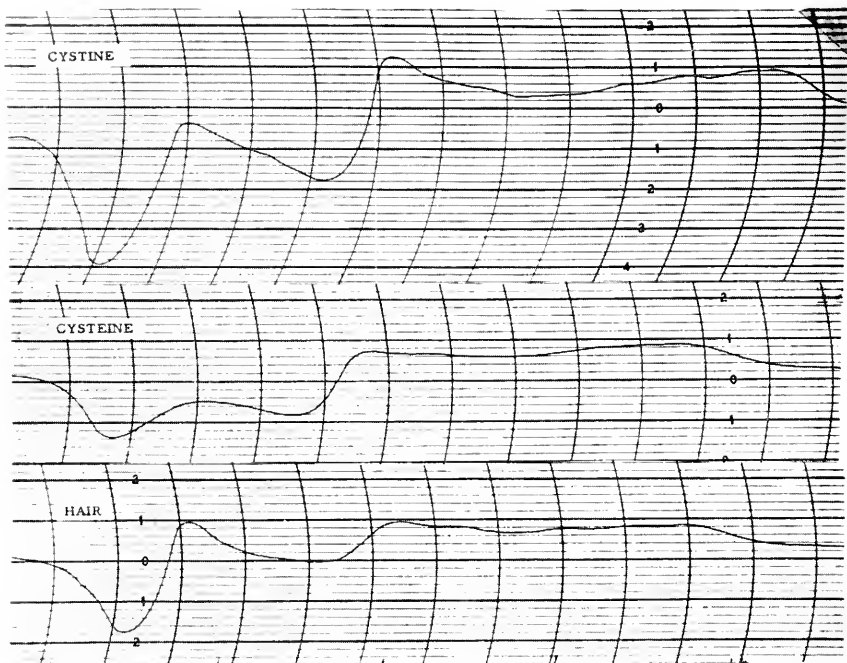


FIG. 9. Resonances (first derivative curves) of x-irradiated hair compared with similar resonances for cystine and cysteine. Marker spacing at base, 68 gauss. Observation frequency, 23 kMc. (From W. GORDY and H. SHIELDS (8).)

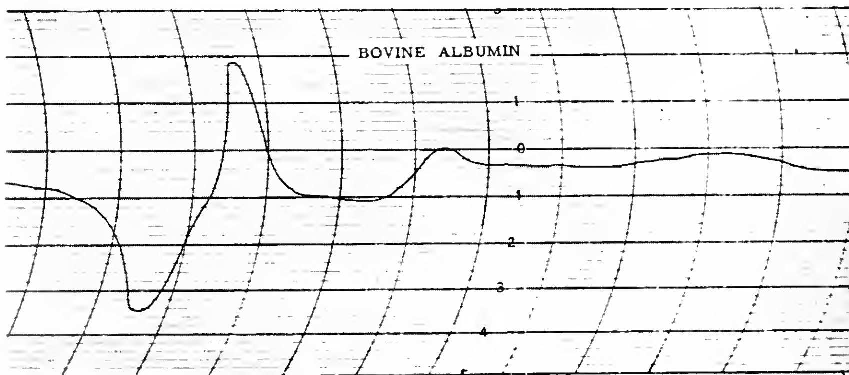


FIG. 10. Resonance (first derivative curve) of bovine albumin which represents a combination pattern of the glycyl-glycine (or silk) doublet and cysteine (or hair) resonance. Observed at 23 kMc. (From W. GORDY and H. SHIELDS (8).)

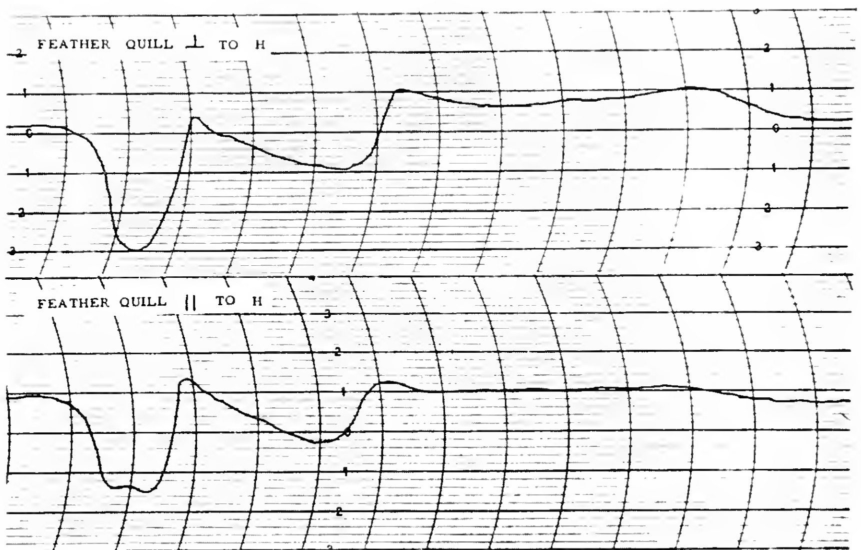


FIG. 11. Resonances (first derivative curve) of x-irradiated feather quill (of a goose) at 23 kMc for parallel and perpendicular orientation in the magnetic field. (From W. GORDY and H. SHIELDS (8).)

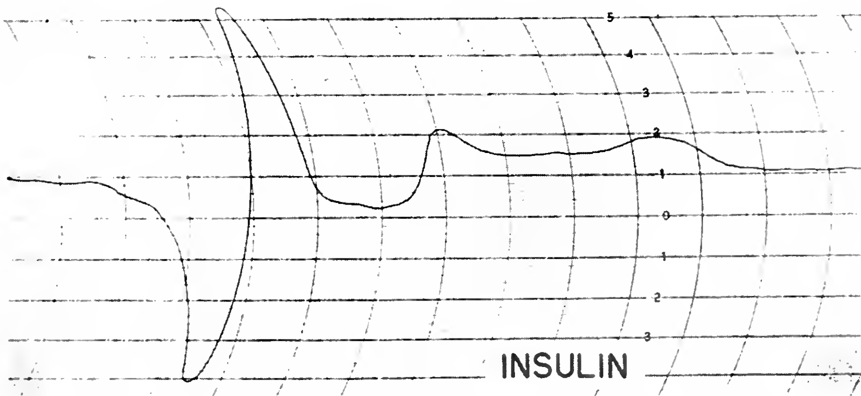
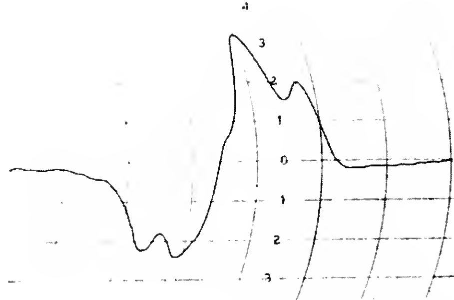


FIG. 12. Resonance (first derivative curve) of x-irradiated insulin observed at 23 kMc. (From W. GORDY and H. SHIELDS (8).)

Cholesterol, $C_{27}H_{45}OH$

X-rayed in vacuum
(indicating the top curve)



X-rayed in air
(indicating the bottom curve)

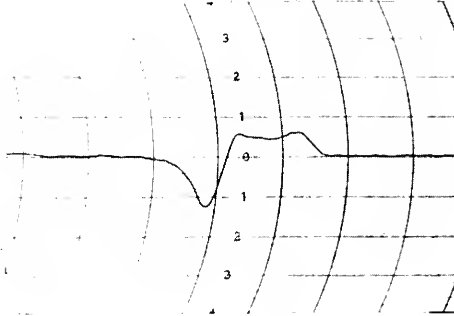


FIG. 13. Resonance (first derivative curve) of cholesterol at 2.7 kMc, x-irradiated in a vacuum (upper figure) and in air (lower figure). (From H. N. REXROAD and W. GORDY (24).)

There are far too many radicals already observed in irradiated amino acids and peptides to discuss them here. I should like to mention one more, however. The pattern of Fig. 7 for acetyl valine consists mainly of a set of nine symmetrical doublets spread over 200 gauss. There is another resonance near the center of the group which I ignore for the present discussion. Seemingly, the nine doublets must arise from eight equally-coupling protons and a ninth with coupling only about half as much as each of the eight at room temperature, and only about a fourth as much at liquid air temperature. This pattern requires an almost unimaginable radical. The odd electron must spread two-fifths of its total density in 1*s* orbitals of the eight equivalent hydrogens. This indicates a radical with a high concentration of hydrogens. It is difficult to design a radical with eight equally coupling hydrogens, especially with a ninth coupling differently. The $(CH_3)_3C$ radical would have nine equally coupling hydrogens which would be expected to give a hyperfine spread of the order of 200 gauss. If we should assume that one of the hydrogens in $(CH_3)_3C$ is replaced by a group RH with only one coupling hydrogen (such as OH) and one which does not noticeably disturb the coupling of the other two, we would have a radical which might account for the acetyl valine pattern of nine doublets.

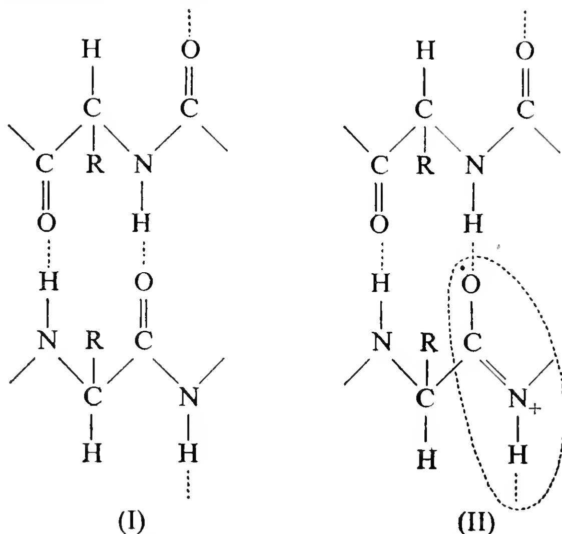
IV. RADIATION DAMAGE IN PROTEINS

In contrast to the varied hyperfine patterns found for the resonances of the x-irradiated amino acids and simple peptides, we have found mainly (but not exclusively) two patterns either singly or in combination for numerous proteins. One of these patterns consists of a simple doublet arising from interaction of the odd electron spin with a single proton spin. The other pattern is a field-dependent one like that of powdered or polycrystalline cystine, cysteine, or glutathione. Fig. 8 illustrates the first type; Fig. 9, the second; and Fig. 10 is a combination of the two patterns.

In our first papers on electron resonances in irradiated proteins (4), we suggested that the doublet pattern in the proteins might arise from an odd electron localized mainly on an oxygen joined by a hydrogen bridge as indicated in Structure II.

Model I represents a structural section of the unirradiated β-keratin protein. The doublet structure, we thought, might arise from dipole-dipole interaction of the electron spin with the proton of the hydrogen bridge. Partly to test this hypothesis, H. W. SHIELDS and the author (25) have made observations on strands of irradiated silk directed along the applied magnetic field, and also on strands directed perpendicular to it. It is known from infrared and x-ray studies (26) that hydrogen bridges in silk lie approximately in a plane perpendicular to the direction of the silk strands. If we assume, for simplicity, that the odd electron density is symmetrically localized on the oxygen, the θ of equation (18) would measure the angle of the O—H axis with the magnetic field. Hence, when the silk strands are along the applied field, θ equals 90° for all hydrogen bridges, and the doublet splitting is the same for all radicals of the silk. Under these conditions one would expect a clear resolution of the doublet. When the silk strands are perpendicular to the applied field, the

hydrogen bridges have all orientations with the field from 0 to 180°. With this arrangement one would expect the individual components of the doublet to be broader and less well resolved than for the parallel case. These features are not completely in accord with the observed results on silk. The doublet splitting for $\theta = 0$ (parallel case) is found to be approximately 25 gauss, somewhat larger than that previously estimated from the polycrystalline material, and also significantly larger than that expected for the hydrogen bonded model.



Furthermore, the separation of the doublet seems to be greater for the parallel case.

It should be appreciated that what is proved for silk is simply that the radical formed is one in which the odd electron interacts with one and only one proton, and that this interaction is at least partly anisotropic. Later we hope to obtain more specific evidence from deuterium substitution in glycyl glycine, which appears to have the same doublet as that for silk.

Irradiated feather quill gives a composite pattern of a doublet and the cysteine-like resonance. However, the doublet is not as widely spaced as that for silk and is not resolved for a polyoriented sample. It has been found (19) that the strong component to the left of the cysteine-like resonance in feather quill (Fig. 11) is partially resolved into a doublet when the feather quill is arranged parallel to the applied magnetic field, whereas it has only about half the width of the unresolved resonance for the perpendicular orientation. Presumably the structure of the feather quill is that of the alpha helix of PAULING and COREY (27), with the helix axis along that of the quill. Interestingly, the cysteine-like component of the resonance is not orientation-dependent. We believe for reasons given later that this situation indicates that the S—S or the C—S bonds of the quill have many different orientations with respect to the quill axis.

A resonance found to be prominent in x-irradiated proteins which contain sulfur is like that of cystine, shown in Fig. 4. Biological substances such as hair (Fig. 9), hoof, horn, and feather have this as the predominant if not the

only pattern, despite the fact that the cystine or cysteine content is only a few percent. The fact that the same pattern, but one very different from any so far obtained from non-sulfur compounds, is observed for many sulfur-containing proteins and for cysteine, cystine, and glutathione convinces us that the odd electron giving these resonances is essentially localized on sulfur. Whether it is on a single sulfur or is shared by two sulfurs of the S—S link, as originally suggested, remains a question to be answered by later work. That the odd electron is localized mainly on one or two atoms is borne out by the large amount of residual spin orbit coupling evidenced by the anisotropy in the observed *g* factor, as already explained.

Because cysteine with only —SH sulfur gives the same type of resonance as cystine with —SS— sulfur, it is uncertain whether the electron which gives rise to the characteristic resonance of Figs. 4 and 9 is on a single S or is shared between two sulfurs to form a 'three-electron bond'. When the plus charge accompanying the odd character arrives at the S of the —SH of cysteine, it would probably 'shock' off either the naked proton to leave the neutral free radical RS·, or the H atom to leave RS^+, where R represents the part of the cysteine exclusive of the SH group. In the latter case, the H atom would escape through the lattice or react with something. (We have been unable to detect the free hydrogen radical at room temperature in any irradiated substances.) We do not know at this time which if either of these two events occurs. Interestingly, RS^+ is not a free radical, and no resonance would be detected for this case until further events had transpired. At room temperature, however, the molecules may flop about sufficiently to allow the RS^+ to react with the —SH of a neighbor and to release another H and form the same charged disulfide radical which has been postulated for irradiated cystine. The common patterns of cystine and cysteine might be thus explained. I should say, however, that the two patterns although alike are not identical: the resonance pattern of cysteine has a slightly greater over-all width than that of cystine, a variation which we believe arises from the difference in dielectric medium. If the radicals were different—if one were RS· and the other were $R \cdot (SS)^+ \cdot R$ —a much greater difference would be expected.

If the resonance in irradiated cysteine arises from RS· mentioned above, the resonance of cystine must arise from the same radical, which would result first from ionization of the cystine molecule and later from rupture of the S—S bond to leave RS· and RS^+. There is no evident mechanism by which the positive charge could disrupt the S—S bond other than the initial 'shock' of the sudden arrival of the charge. Such 'shock' effects can be anticipated from the Franck–Condon principle (28). They would hardly be expected to break the S—S link, because its potential curve would be lowered and its bond length shortened by the removal of an anti-bonding electron. The positive charge would have no preference for either sulfur; and, if the S—S bond holds, the odd electron would be shared equally by both sulfurs to form an additional half-bond. The Franck–Rabinowitch caging principle (28) would also tend to prevent the breaking of the S—S link by the 'shock' effect. The two S atoms are in a sense caged and hindered from flying apart by the large inert R groups to which they are attached. Any 'shock' energy would probably be dissipated as vibrational energy throughout the whole dimeric molecule. Such a charged

link would of course tend to attract other agents such as O_2 or H_2O which might later sever the bond or an electron which would restore the normal S—S link.

Although we are not yet certain whether the cystine or cysteine-like resonance arises from radicals of the type R—(S · · · S)$^+$—R or RS·, we are inclined to favor the latter. It would seem that the neutral radical would enjoy the longer life and hence be the more probable one to be detected. Furthermore, in the RSH compounds the formation of the RS· radical would require the simpler process. With the present information we are inclined to believe that R—(S · · · S)$^+$—R is the primary radical formed by ionization of the disulfide compounds but that the healing of the molecule through capture of an electron or later rupture of the charged link, probably by attraction of other groups, or molecules, may occur so rapidly that this charged radical is not the one detected, but rather the neutral radical R—S. However, our interpretations are still tentative. Because we consider the question an important one we are continuing to investigate it experimentally. Studies using S^{33} can clear up this uncertainty. What already seems established is that the odd electron giving rise to the pattern is essentially localized on the sulfur.

The large anisotropy in the g factor for the cystine-type resonance suggests the potential usefulness of this resonance for obtaining structural information about the proteins. Studies by SHIELDS and the author on single crystals of cystine (19) showed a resonance simpler and much narrower than that for polycrystalline cystine, and one which shifted position sensitively with orientation in the magnetic field. After this observation the same crystal was crushed up and found to give the resonance pattern characteristic of polycrystalline cystine, shown in Fig. 4. Observations (19) on strands of hair and on feather quill at various orientations in the d.c. magnetic field showed only the polycrystalline type of cystine resonance for all orientations. It is significant, we think, that the cystine-like resonance in these proteins is not orientation-dependent, for that fact gives convincing proof that the bonds to sulfur, either the C—S or S—S links, in the keratins are randomly oriented (in contrast to hydrogen bridges). We have also made measurements (19) on stretched and unstretched hair and found no significant variation in its cystine-like resonance pattern. In all cases it is like that of the polycrystalline cystine.

The resonance of x-irradiated insulin may exhibit a third type of protein resonance (cf. Fig. 12). It has the characteristic sulfur or cystine-like pattern but with a relatively sharp resonance superimposed (at the left of the pattern). Although it has the same g factor—that of the free electron spin—this component to the left seems too sharp to be classified as an unresolved doublet like that of feather quill or silk. Possibly this sharp component of the insulin resonance may arise from an electron trapped in one of the unsaturated ringed structures known to be on the side chains of this protein. The ringed structure may act as a sink or trap for the odd electron produced by ionizing radiation. We have found a similarly sharp resonance (29) for x-irradiated polystyrene, where the odd electron observed is believed to be trapped in the aromatic rings attached to the backbone structure.

Considering the varied patterns which we found for the resonances of the different amino acids and simple peptides, it was at first surprising to us that

the proteins gave such simple patterns with the same few features, described above, repeating so often either singly or together. We were forced to conclude that the electron hole or vacancy created by an ionizing quantum or particle at any given locality in the protein can move through the polypeptide chain until it reaches one of a few traps or sinks where it becomes lodged. One such low-energy trap we believe is sulfur. Both —SH and —S—S— groups are effective traps. Possibly the unsaturated rings of certain side chains are an important trap.

Furthermore, we must postulate that there are effective traps for the electrons knocked away in the ionization process since these do not always seem to be able to return readily to fill the hole. Because they have not given recognizable resonances, we do not speculate on the negative traps. For most of them, the resonances may be too broad for detection.

V. PROTECTIVE MECHANISMS

I am sure that there are many who have suspected that some proteins when ionized can hold together and conduct the electron hole to certain side-chain groups such as the sulfur link. I think that I have heard Professor E. C. Pollard, of Yale, and members of his group express such views. However, from my brief and sketchy acquaintance with the literature in this field I surmise that this question has been a highly debatable one. In the microwave resonances we have a new and perhaps more direct type of evidence in favor of the migration of the electron holes to certain side-chain groups.

Now that there is new evidence for effective resistance to the breaking of the polypeptide backbone of the proteins by ionization, it is interesting to speculate on the reasons why this is true. If one of the electrons of a localized, covalent bond between two atoms were suddenly removed, the two atoms might—according to Franck–Condon principle—become dissociated while trying to adjust to the new and shallower potential curve with the longer equilibrium distance commensurate with the 'one-electron bond'. It might be supposed that the Franck–Rabinowitch caging would help to prevent any two atoms of a protein chain faced with such an emergency from becoming dissociated. However, the evidence which we have obtained for the migration of the electron vacancy to a sink in the side chains indicates that a particular bond of the polypeptide chain does not have to face the Franck–Condon catastrophe because the bonds are not in a strict sense localized. If we imagine that charge density equivalent to a single electron is removed completely from the localized region of two adjacent atoms along the main chain, we must, at the same time imagine that this charge density is restored quickly, before the atoms have time to move significantly apart, by the flow of electronic charge from a side chain group such as the S—S link. It might be better to think of the ionization as taking place only at these sites where the electron vacancy is detected. A molecular chain or polymer which can conduct a hole out to a non-essential side-chain sink or to a point where a simple recapture of an electron restores the *status quo* has, in effect, a built-in, remarkably effective method of self-protection from radiation damage. Such polymers have a high survival value in a world where ionizing radiations are ever present.

Although our measurements were made in dry—reasonably dry—samples, it seems likely that the same transfer of an electron hole to low-energy sites, such as the side-chain sulfur, would take place in the proteins of living systems. The better mobility of charges in the more fluid systems should only speed up the recapture of an electron and hence the recovery of the system. Of course the attack on the charged radicals such as —(S—S)$^+$— by molecules like H_2O would also be speeded up in the living systems, but in the living systems the electron recovery might well be the more rapid. Even if a break in the S—S bond should occur, this might be less damaging and more easily healed than a break in the polypeptide trunk line.

We seem to be proposing here a self-protective mechanism which would prevent almost any radiation damage whatever to proteins. This is not true for several reasons, one of which is that not all proteins have —S—S— links in their side chains. There are other traps for the 'hole' where bonds are probably broken as postulated for silk, or for the sulfhydryl group, where the hydrogen atom or proton is believed to be freed. A free hydrogen atom could cause trouble in the living system, even though it could be temporarily spared from the S—H group of the protein. Moreover, not all damage to proteins in the living systems is due to the direct ionization of the protein which we have been discussing here. Much of the damage (30) is thought to be done by radicals such as H, OH, and OOH produced by radiation in the inter-penetrating fluid, which later attack and damage the protein. These are the so-called indirect effects.

About the time of our initial experiment on the proteins, a very significant experiment of an entirely different kind was in progress by ELDJARN, PIHL, and SHAPIRO (31) which indicated that the indirect effects are probably not as significant as had been previously thought, and that a high degree of protection could be achieved by previously converting the —SH groups in proteins to —S—S— links. Their experiments are of a chemical nature and employ tagged sulfur (S^{35}) in cysteamine ($NH_2C_2H_4SH$). I shall not attempt to give the details of their experiments but merely to connect their results with ours. The interdependence of the two apparently different types of results has been pointed out in an interesting paper by EHRENBERG and ZIMMER (32). Our results indicate that any ionization of a protein which contains S—H groups would always tend to dissociate the —SH group through the migration of the 'hole' or positive charge to the S. Because of the large cross-section of the proteins there would be a large release of H atoms by this mechanism unless there were many competing —S—S— links or other traps in the protein to protect the —SH. The experiment of ELDJARN *et al.* would seem to 'protect' the —SH group by first destroying it! By carrying the hydrogen away peacefully in a harmless molecule they prevent its being released by the irradiation as a damaging free radical. Later, after the upheaval is past, it can be restored peacefully if needed.

Our results, as well as those of ELDJARN *et al.*, suggest that some agents may exert their protective effects by becoming temporarily attached through a chemical bond to the protein or other thing which they protect. Cystine, glutathione, or other agent which gives up electrons easily is needed for protection against the damaging effects of positive holes. Cysteine, glutathione

in the reduced form, or other —SH agents may exert their protective effects by forming an —S—S— link with an —SH of the protected molecule, as ELDJARN *et al.* proved for cysteamine. Electron sinks which collect the electrons knocked out of the holes, and thus prevent them from causing damaging reactions, would also be protective agents. The most desirable electron storage tank would be a molecule which would accept the electron without itself becoming dissociated, would hold it loosely, and would give it up easily when it was needed elsewhere.

I should like to add that the electron sources (traps for electron holes) attached to side chains are not necessarily restricted to protective action against direct hits: they may also protect from some of the indirect effects. Certain free radicals produced in the medium around the protein might exert their damage simply by stealing away an electron from some point in the protein. This would of course be replaced by an electron borrowed from the protective group, just as if the electron had been removed by irradiation. The effects of ionized O_2 or H_2O—if there are such things—would be, I suppose, to ionize the protein when they came near it. An OH radical might react with the protein molecule, or it might simply ionize the protein and form OH^-. I do not know which would happen in a particular case. I simply wish to illustrate a possible unrecognized protective mechanism against indirect action of the radiation. Some specialists on radiation effects evidently believe that the damage to the protein of the cells is due mainly to the indirect effect of radicals produced in the medium around the protein molecules and that the protective action of such agents as cystine or glutathione is entirely the elimination of these radicals before they get to the protein. I do not mean to imply that such effects and the mechanism proposed to protect against them are not very important. What seems clear is that protection is also needed against direct hits as well, and if possible against those radicals or charges produced in the medium which survive long enough to reach the protein. Because of the ability of the protein to transfer a charge, it now seems possible to provide this type of protection too. In fact it has already been achieved in some measure by ELDJARN *et al.*

The protective mechanism which I have proposed is strikingly related to enzyme activity. Pollard's group at Yale, and perhaps others, have been making experiments which show, I believe, that a single hit in a large enzyme molecule by an ionizing particle is enough to destroy the enzyme activity of that molecule. This is not strange if the sensitive sites for the enzyme activity are synonymous with the sinks or sources for electrons about which we have been talking, and the ability of the enzyme molecule to conduct a hole or excitation is required for enzyme action.

I like to think of the protective agents which are described here as enzymes which prevent reactions. I know, of course, that the normal function of an enzyme is to cause reactions. Some enzymes, so I understand, exert their catalytic action by accepting spare electrons for a time and giving them up again later. In the vivid language of Professor Henry Eyring, they take over the unnecessary children (electrons) during the divorce proceedings and give them back after the remarriages have taken place. One kind of 'protective enzyme' supplies children to prevent divorces (broken bonds) and then later recovers children indistinguishable from those given up (electrons all). Another

kind provides temporary abode for the disrupted children to prevent their disturbing the neighbors. Our living systems probably have already built in both types of protective agents in sufficient quantity to provide reasonably good protection from ionizing radiation encountered in normal living of the past. For the future we may need to add some.

I have not space to discuss radiation damage to other substances—such as fatty acids, nucleic acids, and hormones—for which our group has obtained many spin resonance data similar to that described here. I have not space to discuss the important effects of oxygen on radiation damage to molecules, about which we have obtained information from spin resonance, of the type shown in Fig. 13. I hope that I have described enough of the results to convince you—and the biologist who rode with me in the car—that microwave electron-spin resonance is an important new way of 'seeing into' biological things.

ACKNOWLEDGEMENTS

The experimental results I have discussed here were obtained largely by my students—W. B. ARD, H. W. SHIELDS, C. G. MCCORMICK, H. N. REXROAD, A. VAN ROGGEN and C. F. LUCK. Others now active in the project are J. H. HADLEY, F. W. PATTEN, and DR. I. MIYAGAWA. For their valuable assistance I am grateful. Our research is supported by the Office of Scientific Research, U.S. Air Force, and the Office of Ordnance Research, U.S. Army.

Without implying that he is to any degree responsible for any misinterpretations I may have made, I wish to acknowledge with thanks some stimulating and enlightening discussions with Dr. JAMES FRANCK during the preparation of this paper.

REFERENCES

1. E. ZAVOISKY: Paramagnetic relaxation of liquid solutions for perpendicular fields. *J. Phys., Moscow* **9**, 211–216 (1945);
   E. ZAVOISKY: Spin magnetic resonance in the decimetre-wave region. *J. Phys., Moscow* **10**, 197–198 (1946).
2. C. A. HUTCHISON: Paramagnetic resonance absorption in crystals colored by irradiation. *Phys. Rev.* **75**, 1769–1770 (1949).
3. J. COMBRISSON, and J. UEBERSFELD: Detection de la Résonance Paramagnétique dans Certaines Substances Organiques Irradiées. *C.R. Acad. Sci., Paris* **238**, 1397–1983 (1954);
   W. GORDY: Quadrupole couplings, dipole moments and the chemical bond. *Disc. Faraday Soc.* **19**, 14–29 (1955).
4. W. GORDY, W. B. ARD, and H. SHIELDS: Microwave spectroscopy of biological substances. I Paramagnetic resonance in x-irradiated amino acids and proteins. *Proc. Nat. Acad. Sci., Wash.* **41**, 983–996 (1955).
5. G. MCCORMICK, and W. GORDY: Electron-spin resonance in x-irradiated peptides. *Bull. Amer. Phys. Soc.* **1**, 200 (1956).
6. W. GORDY, W. B. ARD, and H. SHIELDS: Microwave spectroscopy of biological substances. II Paramagnetic resonance in x-irradiated carboxylic and hydroxy acids. *Proc. Nat. Acad. Sci., Wash.* **41**, 996–1004 (1955).
7. H. SHIELDS, and W. GORDY: Electron-spin resonance of x-irradiated nucleic acids. *Bull. Amer. Phys. Soc.* **1**, 267 (1956).
8. W. GORDY, and H. SHIELDS: Electron-spin resonance in x-irradiated proteins. *Bull. Amer. Phys. Soc.* **1**, 199 (1956).
9. H. N. REXROAD, and W. GORDY: Electron-spin resonance in x-irradiated hormones and vitamins. *Bull. Amer. Phys. Soc.* **1**, 200 (1956).

10. W. GORDY, W. V. SMITH, and R. TRAMBARULO: *Microwave Spectroscopy*, J. Wiley and Sons, New York (1953).
11. D. J. E. INGRAM: *Spectroscopy at Radio and Microwave Frequencies*, Butterworth, London (1955).
12. B. BLEANEY and K. W. H. STEVENS: Paramagnetic Resonance I. *Rep. Progr. Phys* **16**, 108–159 (1953).
13. K. D. BOWERS and J. OWEN: Paramagnetic Resonance II. *Rep. Progr. Phys.* **18**, 304–373 (1955).
14. W. GORDY: Microwave and Radiofrequency Spectroscopy. In: *Chemical Applications of Spectroscopy*, ed. by W. West, chap. 3, Interscience Publishers, New York (1956).
15. J. E. WERTZ: Nuclear and electronic spin resonance. *Chem. Rev.* **55**, 829–955 (1955).
16. W. GORDY: Microwave spectroscopy. In: *Handbuch der Physik*, ed. by S. Flügge, Vol. 28, 1–78, Springer-Verlag, Heidelberg (1957).
17. N. F. RAMSEY: *Nuclear Moments*, J. Wiley and Sons, New York (1953).
18. H. E. WHITE: *Introduction to Atomic Spectra*, McGraw-Hill, New York (1934);
E. U. CONDON and G. H. SHORTLEY: *Theory of Atomic Spectra*, Macmillan, New York (1935).
19. H. SHIELDS and W. GORDY: unpublished data.
20. H. S. JARRETT and G. J. SLOAN: Paramagnetic resonance absorption in triphenylmethyl and dimesitylmethyl. *J. Chem. Phys.* **22**, 1783–1784 (1954);
G. E. PAKE, S. I. WEISSMAN, and J. TOWNSEND: Paramagnetic resonance of free radicals. *Disc. Faraday Soc.* **19**, 147–158 (1955).
21. A. VAN ROGGEN, L. VAN ROGGEN, and W. GORDY: Electron-spin resonance in x-irradiated single crystals of amino acids. *Bull. Amer. Phys. Soc.* **1**, 266–267 (1956).
22. C. G. MCCORMICK and W. GORDY: Microwave investigations of radiation effects in solids: Methyl and ethyl compounds of Sn, Zn and Hg. *J. Amer. Chem. Soc.* **78**, 3243–3246 (1956).
23. C. F. LUCK and W. GORDY: Effects of x-irradiation upon some substances in solid state: Simple alcohols, amines, amides and mercaptans. *J. Amer. Chem. Soc.* **78**, 3240–3243 (1956).
24. H. N. REXROAD and W. GORDY: Electron spin resonance of radicals produced by ultraviolet irradiation. *Bull. Amer. Phys. Soc.* **2**, 227 (1957).
25. H. W. SHIELDS and W. GORDY: details to be published.
26. B. W. LOW: The structure and configuration of amino acids, peptides and proteins. In: *The Proteins*, ed. by H. Neurath and K. Bailey, Vol. IA, chap. 4, Academic Press, New York (1953).
27. L. PAULING and R. B. COREY: Two hydrogen-bonded spiral configurations of the polypeptide chain. *J. Amer. Chem. Soc.* **72**, 5349 (1950).
28. J. FRANCK and R. PLATZMAN: Physical principles underlying photochemical, radiation-chemical, and radiobiological reactions. In: *Radiation Biology*, ed. by A. Hollaender, Vol. I, Chap. 3 McGraw-Hill, New York (1954).
29. H. W. SHIELDS, W. B. ARD, and W. GORDY: Electron-spin resonance in x-irradiated plastics and synthetic fibers. *Bull. Amer. Phys. Soc.* **2**, 227 (1957).
30. W. M. DALE: Basic radiation biochemistry, Chap. 4; and E. S. G. Barron: The effect of x-rays on systems of biological importance, Chap. 5, In: *Radiation Biology*, ed. by A. Hollaender, Vol. I, McGraw-Hill, New York (1954).
31. L. ELDJARN, A. PIHL, and B. SHAPIRO: Cysteamine-Cystamine: On the mechanism for the protective action against ionizing radiation. In: *Proceedings of the International Conference on the Peaceful Uses of Atomic Energy*, Vol. II, 335–342, United Nations, New York (1956).
32. L. EHRENBERG and K. G. ZIMMER: Die Bedeutung der Mikrowellen-spektroskopie für die Strahlen-biologie. *Hereditas* **42**, 515–521 (1956).

A PHYSICAL MECHANISM FOR THE INACTIVATION OF PROTEINS BY IONIZING RADIATION

ROBERT PLATZMAN and JAMES FRANCK
Department of Physics, Purdue University, Lafayette, Indiana and Research Institutes (Fels Fund), The University of Chicago, Chicago, Illinois

Abstract—The extraordinary sensitivity of living systems and certain of their components to ionizing radiation must stem, at least in part, from a great sensitivity of individual molecular or macromolecular species. Analysis of the interaction of such species and swiftly moving charged atomic particles shows that the initial events of energy transfer cannot be responsible for this sensitivity, but that events immediately subsequent to *ionization* acts definitely can be. This is because the time scale for the production of new electric charges is so short as to evoke a violent reaction of the medium, a reaction which is related intimately to the dielectric behavior at very great frequencies. Such behavior is not as yet fully explored experimentally, for many of the frequencies concerned lie between the readily accessible infrared and the microwave regions, but the known dielectric properties of highly polar systems like protein and nucleic-acid macromolecules do disclose the existence of regions of strong dielectric absorption, which are to be identified with dipole oscillations and rotations of polar atoms and molecular groupings. The wave of polarization to which a sudden production of electric charge in the interior of the macromolecule gives rise must cause profound degradation of the molecular organization. This may be viewed as resulting from rupture of many weak polar bonds (such as hydrogen bonds) which maintain the intricate organization and which are involved in the above-mentioned dielectric absorption, the ruptures being essentially simultaneous. The dynamic effect on the molecular structure therefore is without parallel in any other variety of action presently accessible to experimental study—for example, thermal, chemical, or electrochemical action, all of which are in the present context essentially adiabatic in character. The mechanism clearly explains at once the striking difference in sensitivity of the media to ionizing radiation and to ultraviolet light. An approximate quantitative analysis suggests that inactivation of common proteins by a single ionization act is unlikely, but rather that several may be required. Since the effects of the ionizations in a particle track or electron spur are additive (these events being virtually simultaneous), the familiar influence of spatial correlations of ionization is qualitatively explained. The greater radiation sensitivity at elevated temperatures is another obvious consequence. Other predictions of the theory are a dependence of radiation sensitivity upon molecular anisotropy, and a wide variation in the injury to identical molecules exposed to ionizing radiation.

I. INTRODUCTION

LIVING systems embody two distinct varieties of intricate organization. One is the complex static structure of the macromolecules which are essential components of cells; the other is dynamic and is manifested in the delicate organization whereby the functions of the cell and of the organism are achieved.

Living systems are extraordinarily sensitive to ionizing radiation. This is perhaps the most striking single result of experiments in radiobiology and has been emphasized repeatedly (1, 2, 3, 4). It is customary, and plausible, to identify

this great sensitivity with a disruption of complex organization, but it is not known which of the two varieties is so highly susceptible. Indeed, both are likely ultimately to be involved.

In the case of the functional organization, many proposals have been made concerning the initial point of attack. Thus, the destruction or transformation of sulfur-containing groups, of critical enzymes, and of various other essential constituents present in small or in trace amounts, or the production of powerful poisons, have been implicated. The answer is unlikely to be unique, and from its pursuit, which must involve biological questions of the highest complexity, most of the contributions of radiobiology to the science of biology probably will devolve.

The degradation of crucial macromolecules by ionizing radiation is, on the other hand, amenable to *in vitro* experimentation and to purely physico-chemical theoretical analysis. It is the purpose of this paper to examine the possible explanations for such disruption from a simple physical point of view, and to present and investigate one mechanism which is based realistically upon physical and chemical principles and is in full accord, at least qualitatively, with the results of experiment.

A paramount experimental fact is the exceedingly great sensitivity of these macromolecules to ionizing radiation compared to their sensitivity to ultraviolet light. This fact is without parallel in the radiation chemistry of simple organic or inorganic systems, and no clue is provided by the conventional theory of the interaction of swiftly moving charged atomic particles with simple molecules. It is therefore imperative to reanalyze this interaction, with specific regard to the character of the absorbing medium.

The primary processes through which the radiation affects the medium cannot differ qualitatively from those in a simple molecular system. In the case of proteins, for example, the very weak bonds which bind the polypeptides together, and even the peptide bonds, must be essentially without influence on the optical dispersive properties of the medium,* and hence the varieties of energy transfer *from* charged particles, their statistical distribution, and even their spatial distribution will differ only slightly from the corresponding quantities for a simple mixture of amino acids having the same over-all composition. (This statement is not correct with respect to the energy dissipated by moderation of the subexcitation electrons, but that portion of the energy transfer cannot alone be responsible for the great sensitivity.)

Many of the events immediately subsequent to primary absorption of the incident radiation, on the other hand, must differ strikingly from those in a simple system. The reason (5) is that the course of such events is determined by the dielectric properties of the medium, and these, in contrast to the behavior at high (optical) frequencies, are profoundly different for a condensed system composed of highly polar molecules. Ionization in nonpolar substances is usually followed more or less quickly by recombination, so that the chemical consequences of absorption of ionizing radiation are very similar to chemical changes induced by ultraviolet light of appropriate frequency. In a system

* This is even approximately true for the near-ultraviolet absorption spectrum, which is more sensitive to such bonds than the excitations of greater energy that dominate the phenomena of energy transfer from ionizing radiations.

having a great dielectric constant, however, most or all of the initial recombination is inhibited (6), and the chemical effects of ionization stem from interaction with the medium of spatially separated electric charges. This interaction is intimately related to the processes of dielectric absorption characteristic of the medium. It may be noted parenthetically that nonpolar macromolecules, such as many plastics, display utterly different dielectric behavior—in particular showing far less dielectric absorption—and, therefore, that discussions of mechanisms in radiobiology based upon analogies with such polymers, however attractive and despite their current vogue, are perilous, to say the least.

II. CONSEQUENCES OF IONIZATION IN A POLAR MEDIUM

Before treating the particular case of biologically important macromolecules it will be useful to consider briefly the general consequences of an ionization act in a condensed polar medium. By polar medium is meant one with a high value of the static dielectric constant ϵ_s. The condition $\epsilon_s \gg 1$ implies (with only one exception, which is not germane*) the existence of dipolar molecules, and these must always possess regions of strong dielectric absorption at greater frequencies than those at which the dipoles relax. This dielectric absorption embraces all of the familiar infrared absorption, and more: it includes the region from about 30 μ to 1 mm, which at the present time is not accessible with commercial instruments and is therefore virtually unexplored (although a few investigations of simple molecules have been made, particularly in recent years, and more are now under way). The dispersive properties of proteins, for example, are almost completely unknown in this spectral region. But it is certain that such absorption must be common and intense among complex polar substances, for only those resonances arising from strong bonds and small masses lie in the 1 to 20-μ region; weaker bonds and greater masses entail absorption at greater wavelengths.

The production of an electric charge in a medium will ultimately induce a strong polarization similar to that produced, say, in a condenser filled with the same substance. This polarization will not grow uniformly, however, but rather in several stages, each increase in polarization occurring at a time corresponding in order of magnitude to the reciprocal of its characteristic frequency (7); the entire spectrum of dielectric absorption is, of course, involved. The total energy transferred to the medium as the result of an ionization act, excluding the kinetic energy of the ejected electron, can be divided into three parts: that involved in the polarization about the positive ion, a similar quantity released about the electron, or the negative ion which it produces, and finally (and usually much later) the thermal recombination energy of positive and negative ion. The last of these is small (in the case of liquid water, for example, it amounts only to a few per cent of the total) and may be ignored.

It merits emphasis that an ionization act *per se* does not usually cause a molecular dissociation process. Analysis of results from mass-spectrographic investigations shows that stable parent ions are the rule, and that the dissociation of ions, which is usually the main subject of such studies, results from additional

* This refers to substances, such as certain semiconductors, in which intense electronic absorption at comparatively low frequencies gives rise to a high value of ϵ_s.

energy communicated in conjunction with the ionization act in the form of electronic excitation energy of the parent ion. In the vapor phase this additional energy may have time to be concentrated in a particular degree of freedom, but in condensed phases it will ordinarily be dissipated by internal conversion and thermal conduction, leaving the parent ion in its ground electronic state, which is usually stable. Hence the simple identification of an ionization act with splitting of a molecule, which is so common in the literature of radiation chemistry and radiobiology, must be viewed with skepticism in so far as condensed phases are concerned. Such rupture may indeed occur, however, during the growth of polarization about a freshly formed electric charge. It is then very much a consequence of the interaction of the ion with its environment, and in the case of valence-bond breakage imposes special energy requirements with regard to both availability and mode of communication. These requirements may be satisfied, for example, in instances where dissociation would lead to much greater localization of the charge, and therefore to greater polarization energy.

An important property of the energy transferred to the medium by virtue of the growing polarization about a freshly produced electric charge is that most of it is transferred in initial stages, during which the electric field strength is very great. This follows from the Born formula

$$\Delta E = (e^2/2R)[(1/\epsilon(t_1)) - (1/\epsilon(t_2))] \qquad (1)$$

giving the difference in self-energy of the electric field about a charge of magnitude e outside the sphere of radius R, between instants when the effective dielectric constant has the values $\epsilon(t_1)$ and $\epsilon(t_2)$. Since the electronic polarization ($\omega \sim 10^{16}$ sec^{-1}) is effective virtually instantaneously, the initial value of ϵ is approximately equal to the square of the ordinary refractive index n, or about 1.5. By the time that ϵ has increased to (say) 5, most of the total energy $(e^2/2R)(1/n^2 - 1/\epsilon_s)$ will have been dissipated. (Paradoxically, if $\epsilon_s \gg 1$ the behavior under discussion is nearly independent of the magnitude of ϵ_s.) This argument has the important consequence that a major portion of energy transferred to a polar medium by virtue of an ionization act will be communicated, in a very short time, to degrees of freedom associated with weak polar ('secondary') bonds, and that the region receiving this energy will be considerably more extensive than that affected by a slow change in electric field intensity. The total energy so communicated will be of the order of magnitude of 100 kcal/mole for each (electronic) charge produced, but will depend upon the 'size' of the positive ion, or, in general, upon the structure at a molecular level. If the bonds affected are very weak, a substantial fraction of them will be broken, so that the corresponding amount of energy cannot properly be said to have been converted to heat: a portion of it truly has been used to 'melt' a certain structure; obviously, subsequent 'resolidification' often will occur and then will release part or all of the energy as heat.

The fate of the ejected electron is similar. It will progress a distance more than sufficient for the developing polarization to inhibit initial recombination (6), and eventually will attach itself to a negative-ion forming group such as OH, if one is available, or to some other entity capable of binding it. The net

energy released in such an attachment process is small, apart from the contribution from polarization, and may be positive or negative, since the energy evolved in negative-ion formation (electron affinity) must, if it is substantial, compensate for the energy required to rupture a chemical bond. (This is a consequence of the fact that electron affinities of molecules always are small; the only large electron affinities are those of certain atoms and radicals, which by their nature must be present in bound states.) Thus the effect on the medium will be very much like that for the positive ion.

III. PHYSICAL CONSEQUENCES OF IONIZATION IN PROTEINS

The ejection of an electron in an ionization act by even a fairly slow secondary electron is an exceedingly quick process and may be considered to have duration of at most 10^{-15} sec. The response of a highly polar medium to such an event has been analyzed qualitatively above. After subtraction of the electronic polarization, equation (1) shows that a total amount of energy approximately equal to $e^2/2n^2R$ will be dissipated during the several subsequent stages of polarization; here R denotes a distance of the order of magnitude of the mean separation of polar molecular groups, and for proteins must be only slightly greater than atomic dimensions. A value of R of about 2 Å thus corresponds to an energy dissipation of 60 kcal/mole. If all of this energy were expended in dissociating secondary bonds, which may be considered to have dissociation energies of approximately 5 kcal/mole, on the average, a rupture of some twelve secondary bonds would be expected. A more detailed analysis for the particular case of proteins, which leads to the same conclusion, will now be presented. Although it is again based upon the Born formula, which applies strictly only to a continuous dielectric, the error caused by neglect of molecular inhomogeneity will not alter the result in order of magnitude.

The development of polarization ensuant to electric charge localization in the medium may be divided into four stages. These stages, although distinct in character, are by no means without mutual effects, but such interactions can be disregarded in the present analysis, which is only semi-quantitative.

1. *Electronic Polarization*—This effect, in contrast to the others, is strongly coupled to the physical processes which lead to localization of the charge, and indeed to the initial ionization act itself (8). Its influence on the secondary-bond structure probably is negligible.

2. *'Infrared', or True Vibrational Polarization*—The polarization resulting from degrees of freedom corresponding to the characteristic infrared oscillations is developed during a period corresponding to the longest wavelength of such oscillations, or about 3×10^{-14} sec. With a plausible value of 1.7 for the dielectric constant after this stage of polarization, equation (1) yields an energy dissipation of 11 kcal/mole. Thus at most two secondary bonds can be broken, and more probably none are.

3. *Secondary-Bond Polarization*—It has already been emphasized that proteins and similar substances must possess regions of intense dielectric absorption at frequencies between the accessible infrared and the radiofrequency or even microwave regions. Part, or the whole of this absorption, which has yet to be investigated experimentally, stems from the highly polar

secondary bonds (hydrogen bonds of various sorts, salt linkages, etc.) upon which in large measure maintenance of the structure of the macromolecule depends. The polarization corresponding to this absorption is established during the period of from roughly 10^{-13} to 10^{-12} sec following charge localization. Information concerning the magnitude of the dielectric constant subsequent to such polarization apparently is lacking, but fortunately, according to equation (1), the exact value is comparatively unimportant provided that it is much greater than unity, which is certainly the case. (For water it is approximately five.) Thus an energy of about 35 kcal/mole is released during this stage of polarization. It would be erroneous, however, to suppose that this amount comprises the total effect. The Born energy of polarization is the electrostatic energy difference between unpolarized and polarized states (free energy), and this net diminution in energy includes a *positive* energy of the various degrees of freedom (active coordinates of secondary bonds) equal in magnitude to this net energy. Thus about 35 kcal/mole are dissipated to the medium and 35 kcal/mole reside in the 'bonds' as potential energy resulting from deformation and cleavage. A maximum of about fourteen secondary linkages may therefore be broken.

4. *Orientation Polarization*—This is the type of process usually considered in studies (9) of dielectric relaxation of proteins (and, regrettably, often imagined to be the only variety of dielectric absorption). It occurs at far greater wavelengths than the preceding types (e.g. at relaxation times of order of magnitude 10^{-7} sec) and is without influence on the secondary bonds. (Thus, these electric waves do not denature proteins, whereas intense irradiation in the 20 to 50-μ region would very likely do so.)

To summarize, energy sufficient to dissociate about sixteen secondary linkages will be released within an extremely short time interval after localization of an electric charge of magnitude e in a protein molecule. Not all of this energy need be used in bond rupture: a portion will be communicated to heat—for example, to quantum oscillations, both primary and secondary, and also to waves of long wavelength. But since the major interaction is with the secondary-bond degrees of freedom, it is likely that the actual number of broken bonds is a substantial fraction of the maximum number. A conservative estimate would be ten.

It is obvious that this consequence of ionization* will have a profound influence on the structure. It is now universally believed that, as first proposed by MIRSKY and PAULING (10), the organization of the macromolecules is achieved and sustained by a very great number of secondary bonds, and that the primary-bond structure is identical in native and denatured states. Modern elaborations have refined details while retaining the basic ideas of Mirsky and Pauling. (Thus LUMRY and EYRING (11) distinguish between several different arrangements of secondary bonds—for example, in the states of reversible and irreversible denaturation—and propose the useful term *conformation* changes for these variations.) In some respects denaturation may be viewed as a quasi phase-transition, on a submacroscopic level. Although isolated secondary bonds are continually being opened in random fashion by thermal energy, each bond

* The essential idea was described briefly in a previous publication (5).

is normally re-formed in the same configuration and the structure maintained by the constraints imposed by neighboring bonds; only if a number of disturbances overlap will there be a chance that closure of the bonds occurs in improper fashion, and that the disorder becomes irreparable. The model makes possible a satisfactory interpretation of thermodynamic and even kinetic data for thermal denaturation of many proteins (12). It immediately suggests that explanation of the great radiation sensitivity of proteins must be sought in a means of communicating energy from a swiftly moving charged particle to the secondary bonds. Direct energy transfer is negligible, for the coupling is too small (6). The process here advanced provides the required mechanism. Although the analysis given above is admittedly crude (a circumstance for which the lack of relevant and important information on protein structure is chiefly responsible) it is certainly not speculative: it is based upon well established physical principles which are, perhaps, unfamiliar in their present implication.

The *simultaneous* cleavage* of approximately ten secondary bonds following charge localization constitutes a violent perturbation of the protein structure, but probably does not suffice to denature most proteins, at least at ordinary temperatures. This conclusion is suggested by an examination of representative data (Table I) from analysis of thermal-inactivation kinetics, taken from the

Table I. Critical Number of Hydrogen-Bond Ruptures From Thermal Inactivation Rates

	Molecular weight (W)	$\Delta H\ddagger$	$\Delta S\ddagger$	N_1	N_2	W/N_1
Insulin	12,000	35.6	23.8	7	2	1700
Trypsin	24,000	40.2	44.7	8	4	3000
Pepsin	37,000	55.6	113	11	9	3400
Peroxidase (milk)	40,000	185	466	37	39	1100
Ovalbumin	43,000	132	316	26	26	1700
Hemoglobin	68,000	75.6	153	15	13	4500
Yeast invertase	120,000	110	263	22	22	5500

(based chiefly upon reference (12)). $\Delta H\ddagger$ is the enthalpy of activation, in kcal/mole, $\Delta S\ddagger$ is the entropy of activation, in cal/mole deg, and the values of N are calculated by: $N_1 = \Delta H\ddagger/5$, $N_2 = \Delta S\ddagger/12$.

work of STEARN (12). Stearn proposed a calculation of N, the number of secondary bonds which have been ruptured in the activated complex (i.e. the critical number for disordering of the conformation), by assuming an average energy requirement of 5 kcal/mole (N_1), or, alternatively, an average entropy increase of 12 cal/mole deg (N_2). The values of N_1 and N_2 so calculated from velocities of thermal denaturation are in impressive accord with one another,

* The simultaneity of secondary-bond cleavage, which plays a decisive role in the mechanism here proposed, has not been accorded much attention in radiobiology heretofore. It necessarily underlies much of the thinking about mechanisms in thermal denaturation, at least implicitly, and has been invoked, for example, in connection with a model of chemical denaturation by KAUZMANN (13).

except in the case of the smaller proteins, for which the energy and entropy requirements probably do not correspond well with the averages assumed. However, an entropy increase of 12 cal/mole deg is much greater than would be expected for simple cleavage of a secondary bond. This suggests the entirely reasonable conclusion that unfolding occurs when there are broken, not *any* N secondary bonds, but a particular selection of N of them. Clearly, the selection must be a very special one, embracing bonds at certain decisive locations. (Because of cooperative effects this would be true even if all of the bonds were equivalent in their stabilizing action, which is unlikely to be the case.) In the case of secondary-bond rupture following ionization, the bonds affected are more or less localized, and therefore less effective, on the average, than the numbers N listed in the table. Hence the required number of ruptured bonds for ionizing radiation, N_i, must substantially exceed N. Since N is in the neighborhood of ten for even the smallest enzyme molecules, it is evident that the effect of a single electronic charge is almost, but not quite violent enough to inactivate a typical, small protein macromolecule. Even the combined effects of the positive and negative charges, if they are localized in the same molecule, which must usually be the case, may be expected to be just 'subcritical' (except, possibly, in the case of the smallest molecules).

This conclusion leads immediately to the following important consequences.

1. *Variation in the Response of Various Proteins*—Because of differences in structural features among proteins of comparable size, the effectiveness of one or two charges may have wide variation. Furthermore, N_i would be expected to increase with the molecular volume, but not necessarily in a simple way. (Note that N in Table I shows a definite correlation with molecular weight (W), but that W/N is by no means constant.) In all likelihood N_i/N would also exhibit interesting differences.

2. *Effect of Temperature on Radiation Sensitivity*—In cases in which a single electric charge (or a pair of them) is subcritical, its effect may be critical at elevated temperatures, because of the augmented probability that the ambient thermal disorder can supplement the radiation effect and bring it past the threshold for denaturation. This explains, qualitatively, the pronounced thermal sensitivity which has been observed for some inactivation cross sections (14, 15) and which apparently has not received a satisfactory interpretation heretofore (14).

3. *Effect of Anisotropy on Radiation Sensitivity*—Anisotropy of the structure (at the microscopic level) may contribute greatly to the radiation sensitivity. An extreme example would be DNA, which is stabilized by numerous secondary bonds having a degree of freedom for oscillation with an almost common *direction*. Abrupt production of electric charge would rupture many of these bonds simultaneously, causing a portion of the structure to collapse. It is entirely possible that the great radiation sensitivity of DNA, which is found in a variety of experiments (16), may have its origin, at least to some extent, in this effect. The predicted role of molecular anisotropy might be tested experimentally, since proteins and allied molecules exhibit interesting differences in this characteristic.

4. *Collective Effect of Individual Activations on Radiation Sensitivity*—Although a single electric charge may be insufficient to effect denaturation, a very small number of them would suffice. This points to the importance of

the analysis of spatial correlations of charge production* (also called 'spur' or 'cluster' distribution), a subject that has not yet been brought to a quantitative basis.† (It should be emphasized that the bond ruptures caused by positive and negative charges arising from a single spur are all essentially simultaneous.)

If accurate information concerning these correlations were available, the path would be open for study of N_i, and it could be anticipated that both N_i and the ratio N_i/N would prove helpful in the study of protein structure. A closely related subject is the dependence of denaturation efficiency on the so-called density of ionization; the mechanism predicts that as this parameter increases, the effectiveness first rises (as more of the charges are formed in proximity to one another) but ultimately declines, on an energy basis (as the number of bonds ruptured in a molecule exceeds the minimum number required for unfolding). This is indeed observed in many types of experiment, although the rising or the declining portion of the dependence may be enhanced or suppressed in individual cases, depending upon the specific effect. Such behavior should be distinguished from the corresponding one of simple radiation-chemical systems (18, 19), which stem from secondary chemical reactions occurring subsequent to the primary processes; the distinction is not trivial. (In the case of complex biological systems, such as whole cells, the dependence, although it often appears to be similar and may be closely related, must clearly have a far more complex origin (20).) It may be noted that for large protein molecules the disorganization about even a densely ionized track may be insufficiently extensive to produce denaturation. Hence there would be anticipated some thermal sensitivity, although in general less than in the case of sparsely ionizing radiations. Combination of the disorder produced by several localized electric charges is by no means the only possible kind of collective effect, for contributions may be made by excitation events and even by energy transfer to valence-bond and secondary-bond oscillations from subexcitation electrons, both of which must by themselves be minor influences. Changes in certain molecular properties may indeed demand such collective action. For example, permanent dissociation of a valence bond following an excitation act is very unlikely in proteins, but if a dissociative excitation and an ionization should occur close together, the secondary-bond breakage caused by the ionization would prevent healing of the rupture. Subsequent thermal action would then denature and *fragment* the molecule. It is a suggestive possibility

* Following LEA (1), many investigators have inferred from their experimental data that inactivation is accomplished by a single 'average' primary ionization. This is in rough agreement with the general conclusion reached above, but it is not a quantitative statement. Most analyses of experimental data currently being offered appear to be insufficiently detailed and accurate to refine it.

† The proposal that Auger cascades may have an important role in the chemical and biological effectiveness of ionizing radiation (17) is highly relevant to the conclusion that a single electronic charge will in general be subcritical, for each cascade must unquestionably result in destruction of the secondary-bond structure on an extensive scale. (One factor is shown by equation (1): the polarization energy is proportional to the *square* of the electric charge.) In this connection it may be mentioned that the detailed calculations in the paper cited apply only to heavy-particle irradiation; for fast electrons and gamma-rays the yield of Auger cascades is very much greater, being of the order of magnitude of a few per cent of all ionization events, in proteins.

that interchain disulfide bonds may be sensitive regions for such an effect, especially in view of the prominent contribution made by cystine absorption to inactivation of proteins by ultraviolet light (21, 22), but there is no convincing evidence that disulfide-bond cleavage is a major factor in protein denaturation by ionizing radiation. Even the connection with ultraviolet inactivation is ambiguous, because of the different character of excitation produced by charged particles (cf. *infra*).

5. *Effect of the Environment on Radiation Sensitivity*—Since the external environment of the protein can and does participate in the structural stabilization of the molecule, it may alter the effectiveness of the various possible disturbances; the temperature effect already discussed is an instance of this. For example, the medium can contribute externally and internally attached water molecules, various interacting ions, and even chemical influences, and the altered array of secondary bonds may clearly respond differently to the disturbances caused by irradiation. After irradiation and resultant unfolding the imposed forces may impede further unfolding and may, indeed, with the help of thermal agitation, promote healing of the disorganization. On the other hand they may under certain circumstances enhance the radiation sensitivity. This accounts in a general way for pH and other solvent effects. There is *in principle* no simple way to correlate such solvent influences with their effects in ordinary thermal or biochemical inactivation, since the response to sudden charge localization is completely different in character from that involved in such phenomena.

6. *Spectrum of Radiation Injury*—The previous considerations show clearly that in a system of identical protein molecules exposed to any variety of ionizing radiation, a broad range of effects on the molecules must occur. This variability has its origin in (a), variations in the disturbance following localization of a single charge, owing to both the intrinsic variability of the effect of the charge at a given position, and to its localization at different possible sites (e.g. in the interior or on the periphery of the molecule); and (b), in variations in the cooperative effects discussed above, which can differ in number, degree, and proximity (extent of overlapping of regions of charge-induced disorder is obviously a cardinal factor). (Thus N_i certainly is not unique.) The consequence is a wide range of change in properties, different molecules exhibiting qualitative as well as quantitative differences. This spectrum of radiation injury is manifest when appropriate measures are taken to detect it, and the suspicion arises that the common conclusion from irradiation experiments that proteins are either inactivated or unaffected cannot possibly be general, and may often be an oversimplification or even an artifact imposed either by the conditions of an experiment or its interpretation. That thermal denaturation is not a unique transformation is, of course, elementary knowledge; on the basis of the present analysis it appears likely that radiation denaturation may cover an even broader range. Ample proof of the spectrum of radiation injury is provided by the work of FRICKE (23, 24). Thus partial unfolding of the main chains, in addition to denaturation, is indicated by changes in optical rotation, serological response, and $\Delta H\ddagger$, and there is evidence for a small amount of fragmentation, with formation of a variety of products of lower molecular weight. The so-called 'after-effect', a diminished thermal stability of irradiated proteins, is simply

interpreted as stemming from the portion of the spectrum that is subcritical. (BUTLER has shown (3, 16) that DNA is also more sensitive to thermal destruction after irradiation.) FRICKE has even specifically resolved the thermally labile component into a number of fractions with differing thermal response (23). In the case of ovalbumin irradiated with gamma-rays he found the denatured product to be less degraded than the thermally denatured product; this is as expected, since large-scale unfolding can only occur thermally. Another manifestation of the spectrum of radiation injury is the differing reaction to post-irradiative environment that is occasionally observed. This phenomenon, of which the after-effect is a special case, is related to the effect of radiative environment, discussed above, but it clearly involves a later phase of the injury—in particular, partial damage will have been stabilized by closure of many hydrogen bonds, although often in an incorrect way. (This can be inferred from the very low values of heats of denaturation, which show that in thermal denaturation most of the bonds do form again after unfolding.) Such disordered molecules may be further altered by certain external influences and may be restored, at least in part, by others. It has been remarked (15) that a dependence of the inactivation of irradiated hemoglobin (and other proteins) on the pH of the solvent in which they are dissolved *after* irradiation is anomalous, but according to the views set forth here such a dependence is not unexpected.

In the above discussion the term 'localized electric charge' was used in place of 'ionization act' to denote the center of the polarization wave. The motion of an electron vacancy produced by ionization in a protein has been the subject of much conjecture, but a cogent analysis has yet to be made. Although it is certainly true that in (for example) a simple, isolated organic molecule, the precise designation of an original site of ejection of a valence electron has little meaning, this cannot be taken as proof that an electron vacancy has unlimited ability to migrate along the skeleton of a protein or similar macromolecule. One reason is the low degree of symmetry of the molecule, and its greatly differing regions of potential. Another, which often is overlooked, is the influence of the external polarization on this migration. The electronic part of the polarization sets in almost immediately at ionization, and the various low-frequency varieties follow as discussed above. All of them severely limit the mobility of both positive and negative charges. It therefore appears unlikely that an electron vacancy can cross a secondary linkage, or possibly, indeed, even a peptide bond. In the case that several electron vacancies are produced within the same molecule, whatever motion may be possible must enhance the potency of the effect for disordering, for the coulomb repulsion will tend to separate the final sites of localization, thus preventing diminution in effectiveness by too great confinement.

It should be emphasized that the mechanism developed in this paper applies strictly only to the effect of radiation on an isolated macromolecule, a somewhat hypothetical situation approximated in experimental work on 'dry' proteins (1, 4). Immediate effects of the environment have also been touched upon. For proteins in solution, or in living cells, indirect effects, of a simple or complex chemical nature, must also contribute to the observed behavior, and no generalizations regarding the relative potency of the two, except that neither is likely to be negligible, seem warranted.

It may also be noted that the chemical effects may be reversed, but that the disorganization caused by localization in a macromolecule of several freshly created electric charges cannot be; hence protection from or cure of radiation damage *at the molecular level* cannot possibly be complete, even in principle.

IV. THE ROLE OF EXCITATION

Absorption of ionizing radiation leading to the formation of a certain number of ion pairs must also produce a comparable number of electronically excited molecules. This is true for the effects of primary charged particles as well as for secondary ones, and is an elementary consequence of electromagnetic theory. Indeed, the ratio of total number of excited to total number of ionized molecules is, except for slow electrons, simply related to familiar optical properties of the absorbing matter, and the available evidence shows that this ratio is unlikely to depart from unity by more than a factor of about 2, even allowing for the disturbing effects of slow electrons. The ratio is known accurately, at present, only for the noble gases, for which it is 0.4. For all molecular systems it must be greater.

Just as in the case of ionization, which was discussed above, excitation—whether produced by absorption of ionizing radiation or of ultraviolet light—does not itself 'break bonds'. The initial acts of energy transfer are all* ones in which energy is communicated to the *electronic* systems of molecules; subsequent rearrangement of atomic positions may then result in dissociation. For polyatomic molecules the probability that bond rupture will follow excitation is by no means unity, and may be quite small.

In molecules like amino acids, polypeptides, and proteins, excitation commonly is followed by dissociation or by internal conversion, but only very rarely by luminescence (5). In general, experimental work (which has usually been restricted, for practical reasons, to wavelengths greater than about 2200 Å) indicates small quantum yields for inactivation, of order of magnitude 10^{-2} to 10^{-3}. Analysis of the absorption processes has not progressed to the stage of identifying them either with dissociation or with internal conversion, but the following explanations for the low efficiency seem attractive. In the case of dissociation, that is, cleavage of a primary (valence) bond, the secondary-bond structure may prove capable of sustaining the conformation, at least until activation energy becomes available for healing the rupture. (Thus the cage effect is enhanced.) This proposal is supported by the fact that dissociation by moderate or long-wavelength ultraviolet radiation does not provide much energy in excess of the bond dissociation energy; thus at 2200 Å, not more than several hydrogen bonds could be broken. The structure should therefore remain otherwise intact, with closure of the bond a much more likely ultimate result than denaturation. Internal conversion, on the other hand, releases a substantial quantity of energy to oscillational modes, but the coupling is chiefly with *valence* oscillations (C—H, C—C, etc.); by the time the energy reaches the secondary bonds it will have been dissipated too extensively to have much effect.†

* The only direct bond breakage arises from momentum transfer to atoms from swiftly moving particles, in so-called nuclear collisions (17). This is usually a minor effect.

† However, individual internal conversion processes may be responsible for isomerization, and thus for such biological phenomena as gene mutation.

Another factor which diminishes the effectiveness of excitation by ultraviolet light is the spatial isolation of the individual absorption events. Thus the secondary bonds ruptured as a consequence of a single internal conversion process may heal before serious unfolding occurs. In the case of excitation by charged particles, however, the excited molecules are often produced in close proximity (and simultaneity) to other activations, and hence must undoubtedly contribute to the disorganizing action. Such collective effects have already been discussed.

One characteristic of ionizing radiation which always should be kept in mind is the difference in nature of the excited molecules produced from those that have been studied photochemically: they correspond, for the most part, to radiation in the vacuum ultraviolet region, where most of the optical transition probability invariably lies. Little is known of polyatomic molecules with regard to optical phenomena and to processes following excitation in this spectral region. However, such radiation may have far greater potency than the readily accessible ultraviolet, for either in dissociation or in internal conversion processes, it always releases sufficient kinetic energy to break many more secondary bonds. There is indeed some experimental indication of this in the rise of quantum yields at the shortest wavelengths studied (25, 22). Thus the role of excitation in radiobiology probably is greater than usually is (cf., e.g., (14)) supposed.

V. CONCLUSION

The mechanism considered here provides a realistic physical basis for understanding the remarkable fact that a polar macromolecule of molecular weight as great as 10^5 can be inactivated by only a few ionization acts, and it is capable of explaining qualitatively a variety of experimental results. It replaces the notion that ionizing radiation acts merely by breaking chemical bonds directly, which, apart from its superficiality, does not actually explain denaturation at all. No attempt has been made in the present paper to analyze in detail the myriad data on numerous kinds of radiation effect for varying quality and quantity of radiation, varying environment, etc. Indeed, further development must await, in most points, the further elucidation of protein structure, especially in its dependence upon the secondary-bond configuration. In particular, the number, disposition, and mutual dynamical behavior of the secondary bonds, as well as the character of their large-scale stabilizing action, must be more fully understood. At some future stage of development radiation studies may provide a valuable tool in advancing this knowledge, for the action of ionization, as described here, is completely different in character from other types of attack which are investigated, such as heat, salts and other chemically inert solutes, and chemical agents, all of which act essentially adiabatically at the atomic level. In essence it has been demonstrated that the marvellous stabilizing action manifested in natural polar macromolecules is intrinsically ineffectual against the nonadiabatic disturbance of an ionization act.

REFERENCES

1. D. E. LEA: *Actions of Radiations on Living Cells,* Cambridge University Press, Cambridge, England (1946).
2. L. H. GRAY: The initiation and development of cellular damage by ionizing radiations. *Brit. J. Radiol.* 26, 609–618 (1953).

3. J. A. V. Butler: The action of ionizing radiations on biological materials. Facts and theories. *Radiat. Res.* **4**, 20–32 (1956).
4. E. Pollard: Primary ionization as a test of molecular organization. In: *Advances in Biological and Medical Physics*, ed. by J. H. Lawrence and C. A. Tobias, **3**, 153–189 (1953).
5. J. Franck and R. Platzman: Physical principles underlying photochemical, radiation-chemical, and radiobiological reactions. *Radiation Biology*, ed. by A. Hollaender, Vol. I, Chap. 3, McGraw-Hill, New York (1954).
6. R. L. Platzman: Energy transfer from secondary electrons to matter. In: *Basic Mechanisms in Radiobiology*, 22–37, Nat. Res. Council (U.S.) Publ. No. 305 (1953).
7. R. Platzman and J. Franck: Theory of the absorption spectra of halide ions in solution. In: *L. Farkas Memorial Volume*, 21–36, Research Council (Israel), Jerusalem (1952).
8. R. L. Platzman: Influences of details of electronic binding on penetration phenomena. In: *Symposium on Radiobiology*, ed. by J. J. Nickson, Chap. 9 (cf. §VIII), J. Wiley and Sons, New York (1952).
9. J. T. Edsall: The size, shape and hydration of protein molecules. In: *The Proteins*, ed. by H. Neurath and K. Bailey, Vol. I-B, 698–712, Academic Press, New York (1953).
10. A. E. Mirsky and L. Pauling: On the structure of native, denatured, and coagulated proteins. *Proc. Nat. Acad. Sci., Wash.* **22**, 439–447 (1936).
11. R. Lumry and H. Eyring: Conformation changes of proteins. *J. Phys. Chem.* **58**, 110–120 (1954).
12. A. E. Stearn: Kinetics of biological reactions with special reference to enzymic processes. *Advances in Enzymology*, ed. by F. F. Nord, **9**, 25–74 (1949).
13. W. Kauzmann: Denaturation of proteins and enzymes. In: *The Mechanism of Enzyme Action*, ed. by W. D. McElroy and B. Glass, 70–110, Johns Hopkins Press, Baltimore (1954).
14. E. C. Pollard, W. R. Guild, F. Hutchinson, and R. B. Setlow: The direct action of ionizing radiation on enzymes and antigens. In: *Progress in Biophysics*, ed. by J. A. V. Butler and J. P. Randall, **5**, 72–108, Pergamon Press, London and New York (1955).
15. R. Setlow: Radiation studies of proteins and enzymes. *Ann. N.Y. Acad. Sci.* **59**, 471–483 (1955).
16. J. A. V. Butler: Effects of x-rays and radiomimetic agents on nucleic acids and nucleoproteins. In: *Ciba Foundation Symposium on Ionizing Radiations and Cell Metabolism*, ed. by G. E. W. Wolstenholme and C. M. O'Connor, 59–69, Churchill, London (1956).
17. R. L. Platzman: On the primary processes in radiation chemistry and biology. In: *Symposium on Radiobiology*, ed. by J. J. Nickson, Chap. 7, J. Wiley and Sons, New York (1952).
18. E. J. Hart, W. J. Ramler, and S. R. Rocklin: Chemical yields of ionizing particles in aqueous solutions: Effect of energy of protons and deuterons. *Radiat. Res.* **4**, 378–393 (1956).
19. R. H. Schuler and A. O. Allen: Radiation chemistry studies with cyclotron beams of variable energy—yields in aerated ferrous sulfate solution. *J. Amer. Chem. Soc.* **79**, 1565–1572 (1957).
20. R. E. Zirkle: The radiobiological importance of linear energy transfer. In: *Radiation Biology*, ed. by A. Hollaender, Vol. I, Chap. 6, McGraw-Hill, New York (1954).
21. R. Setlow: A relation between cystine content and ultraviolet sensitivity of proteins. *Biochim. Biophys. Acta* **16**, 444–445 (1955).
22. R. Setlow and B. Doyle: The action of monochromatic ultraviolet light on proteins. *Biochim. Biophys. Acta* **24**, 27–41 (1957).
23. H. Fricke: Kinetics of thermal denaturation of x-rayed egg albumin. *J. Phys. Chem.* **56**, 789–795 (1952).
24. H. Fricke, C. Leone, and W. Landmann: The role of molecular degradation in the loss of serological activity in gamma-irradiated ovalbumin. *Radiat. Res.* **7**, 316 (1957).
25. A. D. McLaren: Photochemistry of enzymes, proteins, and viruses. *Advances in Enzymology*, ed. by F. F. Nord, **9**, 75–170 (1949).

INFORMATION AND INACTIVATION OF BIOLOGICAL MATERIAL*

HAROLD J. MOROWITZ
Department of Biophysics, Yale University, New Haven, Connecticut

Abstract—An analysis of target theory has been carried out in terms of the language of information theory. Certain results suggest that radiation and thermal inactivation experiments can be used to set limits on the values of information content of biological structures. A group of such limits has been discussed, as well as a suggestion for using 'radioactive suicide' experiments to evaluate information content.

INFORMATION theory provides a discipline for quantifying order and specificity in biological structures. Ionizing radiation and heat provide more or less random methods of disordering biological structures. Therefore, we may anticipate that information theory and studies of the biological effects of heat and ionizing radiation may in some way complement each other. In particular, if we can make some quantitative statements about the amount of disordering necessary for loss of biological function, we are then able to say something about how much order is involved in specifying the system.

The concept of target volume has an analogue in the representation of a structure in terms of a series of symbols. If inactivation curves are exponential and the target volume is less than the volume of the structure, we may conclude that part of the structure (the critical target) has an information density higher than the rest of the structure. That is, a subset of symbols in the array require much closer specification than the rest of the array. If no energy is transferred and there are b symbols in the subset, the target volume will be bV/M, where V is the total volume of the structure and M is the total number of symbols of equal volume needed to specify the structure. If there is energy transfer with high efficiency over g atoms, the volume will be of the order of bg^3V/M, assuming no overlap of partial volumes.

In this paper we shall be concerned with those biological materials that can be dried, taken to low temperatures and then returned to a functional state without appreciable loss of activity. This class of materials includes enzymes, viruses, spores, and transforming principle. In these entities we may conclude that the information is contained in the structure. Several methods have been used to evaluate the information content of these resting systems.

We shall outline briefly two methods that have been used to evaluate the information content of biological materials. In both methods it is assumed that the atomic composition and volume are known. The volume may then be divided into a number of elementary atomic volumes. To specify the

* Supported by a Research Grant from the John A. Hartford Foundation.

system completely, we need to state, in some pre-arranged sequence, which type of atom is present in each elementary volume and the bonds between that atom and its six nearest neighbors. Our specification then consists of a message giving the appropriate symbol to each elementary volume. To calculate the average information per symbol, we consider the probability p_{ij} of having the ith type of atom in the jth bonding state. The average information is then given by

$$H = -\sum_{ij} p_{ij} \log_2 p_{ij} \tag{1}$$

If the p_i's are assumed from the average composition of dry bacteria (hydrogen, 52.2 mole per cent; carbon, 29.9 mole per cent; nitrogen, 7.6 mole per cent; oxygen, 5.8 mole per cent; phosphorus, 2.9 mole per cent; sulfur, 1.6 mole per cent), and if we assume that all the types of bonding configurations have equal *a priori* probabilities, we can then calculate that H is of the order of 4.0 bits per atom. Since the different bond configurations have rather different probabilities, our figure is high and 3 bits would probably be a more realistic estimate.

An alternative but equivalent method of finding the information content is to assume that all states of the system have equal *a priori* probability. If there are N possible states and L of these are biologically functional, the probability that the system is in a functional state is L/N and the information is given by

$$H = -\log_2 L/N = \log_2 N - \log_2 L \tag{2}$$

If the system must be completely specified, L equals one and H takes on its maximum value, $\log_2 N$. We may then calculate N from the number of permutations of the atoms in the elementary atomic volumes and the permutations of the bond states (1). This leads to the same average information content per atom as the previous treatment.

However, from a point of view of biology, we would like to know the actual value of H rather than H_{max}. If we consider a large collection of spores or viruses or enzymes in contact with a thermal reservoir at temperature T and allow the system to come to thermal equilibrium over a long time, we may regard the collection as a Gibbsian ensemble, and the ratio of the final activity to the initial activity is a measure of the *a priori* probability of finding the system in a functional state. In general the activity decreases with time in an exponential fashion. In all the experiments that have been carried out, the sample is too small and thermal equilibrium is never reached. This enables us to determine a lower limit of the information content, but the limit is too low to be of any practical use. For example, dry *Bacillus subtilis* spores show an exponential inactivation over twelve decades. There is no indication that the system is nearing equilibrium so we may conclude that the *a priori* probability of finding the system in a living state is less than 10^{-13}. H is then greater than $-\log_2 10^{-13}$ or greater than 42. Since the upper limit (based on $L = 1$) for this system is of the order of 10^{11} bits, the thermal data do not help very much in bracketing the figure. Experimentally it is not feasible to carry thermal inactivation studies below an activity of 10^{-12} because of difficulties in the

sample size and the assay in the presence of all the inactivated material.

It may be noted in passing that the consideration of the system in terms of a Gibbsian ensemble may provide some insight into the origin of life or the *a priori* probability of a biologically functional structure arising *de novo*.

In considering the information aspects of ionizing radiation, we shall confine ourselves to anhydrous systems and consider only the direct action of radiation. We must then consider the effect of a primary ionization in altering the structure of biological molecules. Present evidence indicates that primary ionizations occur in a random fashion along the track of the fast charged particle. However, the subsequent events are much less clear. It is difficult to make quantitative statements about the probability of the energy being transferred from the site of the original ionization to an energy sink in the material. For purposes of developing the theory, we shall first make the simplest possible assumption that the result of a primary ionization is a bond break, or rearrangement of bonds at the site of the ionization. Many structures are inactivated by a single ionization within the structure. If the previous hypothesis applies, such structures have an information content close to H_{max}, since L must be unity if any rearrangement destroys the functional integrity of the structure. It should be remembered that H_{max} is the formal upper limit if the calculation is based on atomic specification. It would be possible to start from other points of view, such as monomer specification, functional unit specification, or genetic specification, and arrive at different values of an H function for use in subsequent analysis.

However, there are many indications that the simple assumption made above is not valid. For tobacco mosaic virus (2), the target volume is about half the total volume of the particle, yet the infectious unit is presumably the RNA which is only six percent of the total volume. Many enzymes show a target volume equal to the physical volume of the molecule (3), yet recent evidence suggests that several amino acids can be removed from the enzyme without loss of activity (4). It is difficult to see why bond rearrangements in these amino acids should lead to loss of function. Some enzymes show a target volume larger than the physical size of the molecule.

These factors indicate energy transfer from the site of the ionization to an energy sink within the molecule. Recent studies by GORDY (5) and SETLOW (6) tend to suggest that sulfur–sulfur bonds are the ionization sinks in protein. If we assume that this is the case and that the energy of a primary ionization is transferred with a high efficiency to these bonds, then we can arrive at a minimum value of the information content of molecules which contain these bonds and are inactivated by a single ionization. Since about one in every 400 of the atoms is involved in an S–S bond, we may conclude that $MH/400$ per atom is a crude estimate of the minimum information necessary to specify the structure of M atoms. In this case radiation experiments would enable us to set a lower limit to the information content.

Consider next a structure which requires several ionizations to cause an inactivation. If there are M atoms in the structure, each ionization may transform the system from its original state to any one of the MB neighboring states, where B is the average number of ways in which each atom can be bonded to its neighbors. If on the average x hits are required to inactivate the structure,

then L must be at least $\frac{(MB)^x}{x!}$. H is then given by

$$H = H_{\max} - x \log_2 MB + \log_2 x! \tag{3}$$

It was argued above from equation (1) that there are about 3 bits per atom, so that $H_{\max}$ is a number of the order of $3M$; thus if x is small in comparison to M, a structure requiring multiple hits will still have an information content close to $H_{\max}$, in cases where no energy transfer is assumed.

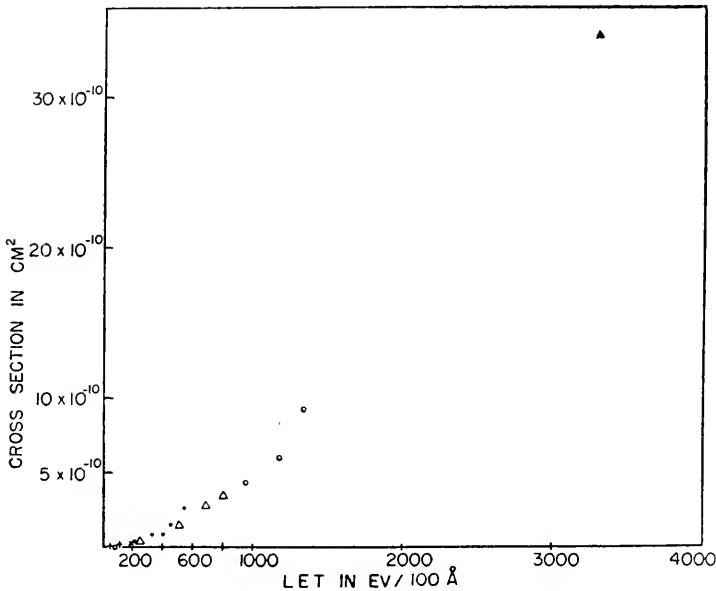


Fig. 1

There are cases in which x may be an appreciable fraction of M. If one plots cross-section as a function of linear energy transfer (LET), the shape of the curve gives an indication of target thickness and number of ionizations required. If it can be shown that x ionizations are required in a distance A for an inactivation, we can divide the structure into sub-structures of volume A^3 in which case H is given by

$$H = 3A^3m - \log_2 \frac{(A^3mB)^x}{x!} \tag{4}$$

where m is the number of atoms per unit volume. If x is an appreciable fraction of A^3m the substructure may have an information content smaller than the maximum value. The information content of the entire structure, which is the sum of that for the substructures, will be correspondingly smaller than in the case of complete specification.

Let us now consider the specific case of the irradiation of spores of *Bacillus subtilis* with fast charged particles. At all values of LET studied the inactivation curves are exponential functions of the dose. Fig. 1 shows the curve obtained

by Mr. J. Edward Donnellan of the Yale Biophysics Department, and indicates inactivation cross-section for colony formation as a function of LET. From the electron irradiations of PROCTOR *et al.* (7), the target volume for these spores is of the order of 10^{-17} cm^3. However, this seems to be the volume of the substructure with the highest information density. For we see that as we increase the LET the radiation rapidly becomes more efficient in causing inactivation. What we are doing in increasing LET is to increase the probability of several ionizations in a given substructure of the spore. Since the cross-section then rises so dramatically, we must conclude that targets of lower information density than the one originally inactivated at low ion densities are now coming into play. Since the curves are exponential, the multiple ionizations in any substructure must be coming from the same fast charged particle. Under these circumstances, x must of necessity be small compared with M, and the secondary targets must still retain an information content near H_{max}, if we ignore energy transfer.

We may conclude, in general, that any large structure which is capable of being inactivated by the passage of a single fast charged particle through that structure probably has an information content which is an appreciable fraction of H_{max}. In general, if information can be transferred with high efficiency over g atoms, H is probably greater than $H_{max}/(2g)^3$.

There seems to be a possibility of reducing energy transfer and thus getting a better estimate of information content. When enzymes are irradiated with fast charged particles and the experiments are carried out at different temperatures, the target cross-section is found to be an increasing function of temperature (3). The possibility exists that energy transfer is being reduced at the low temperatures, and data taken in this range might provide a better index of the actual information content. However, considerations of this type demand a thoroughgoing analysis of the physics of the situation.

Another method of random disordering exists which might provide an even more powerful tool for the elucidation of information content. It has been recently shown (8) that viruses labeled with P^{32} lose activity on standing, and the rate of loss is associated with the amount of P^{32} incorporated. Now the decay of a radioactive atom incorporated in a biological structure, and the consequent transmutation of the atom, represents a random disordering. If the labeling is random, the rate of decay should provide a measure of the fraction of atoms of the labeled type which require precise specification in order for the structure to be functional. Such an information evaluation should be possible for phosphorus, sulfur, hydrogen, carbon, sodium and calcium. Thus we may inquire about a complicated structure like a spore: how many of the phosphorus atoms in the structure are required to specify a functional unit? Experiments and calculations of this type should serve to limit the value of the information content of biological structures.

REFERENCES

1. H. J. MOROWITZ: Some order–disorder considerations in living systems. *Bull. Math. Biophys.* **17**, 81–86 (1955).
2. E. C. POLLARD and A. E. DIMOND: The inactivation of tobacco mosaic virus by ionizing radiation. *Phytopathology* **46**, 214–218 (1956).

3. E. C. POLLARD, W. R. GUILD, F. HUTCHINSON, and R. B. SETLOW: The direct action of ionizing radiation on enzymes and antigens. In: *Progress in Biophysics*, **5**, 72–108, Pergamon Press, London and New York (1956).
4. C. B. ANFINSEN: The limited digestion of ribonuclease with pepsin. *J. Biol. Chem.* **221**, 405–412 (1956).
5. W. GORDY, W. B. ARD, and H. SHIELDS: Microwave spectroscopy of biological substances. *Proc. Nat. Acad. Sci., Wash.* **41**, 983–1003 (1955).
6. R. B. SETLOW: A relation between cystine content and ultraviolet sensitivity of proteins. *Biochim. Biophys. Acta* **16**, 444–445 (1955).
7. B. E. PROCTOR, S. A. GOLDBLITH, E. M. OBERLE, and W. C. MILLER, JR.: Radiosensitivity of *Bacillus subtilis* under different environmental conditions. *Radiat. Res.* **3**, 295–303 (1955).
8. G. STENT and C. A. FUERST: Inactivation of bacteriophage by decay of incorporated radioactive phosphorus. *J. Gen. Physiol.* **38**, 444–457 (1955).

DISCUSSION

QUASTLER: Dr. MOROWITZ's analysis of the informational aspects of radiation effects, and his concept of information density, are very important developments. As a matter of fact, I believe them to be so important that even small differences in interpretation are worth mentioning, and this is the reason for making the following comments.

To rephrase the situation: consider a structure (message) consisting of a distinct substructures (words) of b elements (letters) each. Let H' be the information content per letter and T' the information measure of constraints between letters. Then:

$$H'' = \text{information content per word} = b(H' - T')$$

$$H''' = \text{information content of message} = a(H'' - T'')$$

(where T'' represents the informational aspects of constraints between words) and

$$H'''/ab = \text{information density in bits per letter.}$$

If the 'letters' are atoms in living matter, then I suspect the constraints T' and T'' to be quite considerable and to reduce H'''/ab to rather less than three bits per atom.

We introduce noise of such a character that a single noise event results in the functional destruction of a single letter, and examine the functional value of the message after it has suffered a known number of noise events. If a single event destroys the functional value, then all letters are functionally essential and the functional information density is H'''/ab. If the number of noise events needed averages more than one, then the informational density must be less than maximum, and this can occur in two entirely different situations.

(I) A number a_0 of the letters in each word (or a number b_0 of the words) is either irrelevant or can be reconstructed, provided every one of the $a - a_0$ essential letters (or $b - b_0$ essential words) is intact. Then the functional information density is

$$\frac{H'''}{ab}\,\frac{a - a_0}{a} \quad \text{or} \quad \frac{H'''}{ab}\,\frac{b - b_0}{b}$$

and a single event *can* cause loss of function but does so only with probability $(1 - a_0/a)$ or $(1 - b_0/b)$, respectively. This is the situation where the target size is less than the total size of the structure.

(II) Up to a_r letters in every word (or up to b_r words) can be destroyed without loss of function—and these letters or words do not belong to a predetermined sub-set but can be *any* letters (or words). This is the case with error-correcting codes; in this case no single event can cause loss of function, but $a_r + 1$ events will every time. The functional information density is again less than maximum, being

$$\frac{H'''}{ab}\,\frac{a - a_r}{a} \quad \text{or} \quad \frac{H'''}{ab}\,\frac{b - b_r}{b}$$

but it is reduced by the presence of redundant information—not by irrelevant information as in the first case. These two situations must be sharply distinguished.

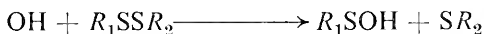
With atoms and molecules, the error-correcting mechanism can be a cage effect or something of the kind. This could be the situation where loss of function is caused by clusters of events, i.e. passage of a particle of high linear energy transfer. It may be suspected that the protective effect of redundancy of chemical structure extends only over regions of rather limited sizes which would imply that the reduction of information density could be rather substantial.

THE ABSENCE OF RADIATION-INDUCED DISULFIDE INTERCHANGES*

ARTHUR L. KOCH †
Division of Biological and Medical Research, Argonne National Laboratory, Lemont, Illinois,
and
Department of Biochemistry, University of Florida, Gainesville, Florida

Abstract—Mixtures of cystine and its *bis*-dinitrophenyl derivative were: (a) irradiated as dry films, (b) treated in aqueous solution with Fenton's reagent, and (c) irradiated in aqueous solution. In none of these cases could any of the interchange product, mono-dinitrophenyl cystine, be detected. It is therefore inferred that disulfide interchange is not a primary cause of protein denaturation or enzyme inactivation by ionizing radiations.

As a consequence of treatment of proteins and protein solutions with large doses of ionizing radiations, denaturation, as assessed by decreased solubility at the isoelectric point, frequently occurs (1). The involvement of sulfur linkages in these solubility changes as well as in the concomitant loss of biological activity has been frequently suggested. The importance of the oxidation of existing thiol groups to disulfides is well documented (2), but another possibility is that polymerization results from disulfide interchange, in a manner similar to that postulated by HUGGINS, TAPLEY, and JENSEN (3) to account for gel formation in the presence of urea. In their case the initiator of the chain reaction was assumed to be a sulfhydryl group exposed by the unfolding of the protein. This unfolding results from the breaking of intramolecular hydrogen bonds by the urea. One could, however, conceive of chain reactions initiated by the free radicals produced by ionizing radiations. For example, hydroxyl radicals produced by the action of the radiation on water could react to produce a sulfenic acid and a sulfide radical, which could then react further:



$$OH + R_1SSR_2 \longrightarrow R_1SOH + SR_2$$

$$SR_2 + R_3SSR_4 \longrightarrow R_2SSR_4 + SR_3 \ .$$

Such a mechanism appears attractive in view of the properties of sulfur compounds as presented by CALVIN (4, 5). Further reactions of R_1SOH could also lead to polymerization as a consequence of dismutation to the sulfide and sulfinic acid.

In order to investigate these possibilities, a model system was studied. The system chosen was that utilized by RYLE and SANGER (6). These authors

* This work performed under the auspices of the U.S. Atomic Energy Commission.

† Present address: Department of Biochemistry, University of Florida, Gainesville, Florida.

were particularly interested in the possibility of interchange under strong *aqueous* acid conditions that are present during protein hydrolysis. In our laboratory this system has been used to study the effects of strong *anhydrous* acid media (7). In the former case interchange was found and this could be repressed with thiol compounds; in the latter case no interchange occurs.

The system consists of a mixture of *bis*-2,4-dinitrophenyl-L-cystine (*bis*-DNP cystine) and L-cystine. If the interchange occurs, the reaction product, mono-2,4-dinitrophenyl-L-cystine (mono-DNP cystine) may be readily measured by removing the *bis* compound by acid ether extraction, or by chromatography in a solvent system consisting of aqueous 5 per cent Na_2HPO_4 overlaid with isoamyl alcohol. The spots can be visualized by observation with near ultra-violet light. By a combination of these two techniques additional sensitivity can be obtained.

Dry Irradiation

Using this system it was quickly established that even with doses as large as 3×10^7 r of Co^{60} gamma rays no detectable interchange product was produced in the radiation of dry films of mixed cystine and its *bis*-dinitrophenyl derivative. As very small amounts could be detected by the combination of the extraction and chromatographic techniques, it is felt that disulfide interchange cannot be of importance in the denaturation of dry protein samples, as certainly much less than one interchange per 1000 disulfide bonds could have been detected.

Effect of OH *Radicals*

In aqueous solution the experimental situation is quite different. We first investigated the effects of hydroxyl radicals by themselves. Experiments (Table I) with Fenton's reagent (a mixture of Fe^{++} and H_2O_2 prepared as described

Table I. The Absence of Interchange Produced by OH *Radicals*

The complete system contains 1×10^{-3} M cystine; 1.25×10^{-4} M *bis*-DNP cystine; 1.2×10^{-2} M phosphate buffer, pH 7.9; 5×10^{-5} M Fe^{++} (as $FeSO_4$); 5×10^{-3} M H_2O_2. For analysis, 0.5 ml aliquots were added to 2.0 ml of 1 N HCl. The solution was then exhaustively extracted with ether and read at 350 mμ in the Beckman spectrophotometer.

	Increase in optical density	
	5.5 hours	48 hours
Complete	0.048	0.052
Minus Fe^{++}	0.062	0.052
Minus H_2O_2*	0.004	0.084
Minus Fe^{++} and H_2O_2*	−0.006	0.100
If interchange is complete (calculated)		0.489

* The interchange observed at 48 hours in absence of H_2O_2 is caused by thiol compounds produced by the hydrolysis of the disulfide (6).

by COLLINSON *et al.* (8)) did not give any increase in the content of non-ether soluble chromophoric material as compared with a control containing H_2O_2 alone. Even at the end of forty-eight hours no increase was apparent, although ribonuclease would have been destroyed completely at the end of thirty minutes (8). The increase in optical density in the H_2O_2 control is small but significant (20 per cent of theoretical at the end of forty-eight hours), and probably represents oxidation to ether insoluble materials such as dialkyl sulfoxides and cysteic acids.

Irradiation in Aqueous Solution

When aqueous solutions of mixtures of cystine and its *bis*-dinitrophenyl derivative were irradiated with 1×10^7 r, the results obtained were equivocal because of the influence of side reactions causing change in the chromophoric moiety.

Possibly this effect is akin to the well known photo-destruction of dinitrophenyl derivatives generally. That such a process is occurring follows from the observations that the samples irradiated with 6×10^6 r of Co^{60} gamma rays were less intensely colored than the non-irradiated controls; the optical density at 350 mμ was reduced to one third that of the controls. The bulk of the 350 mμ absorbing material after irradiation was insoluble in ether. This product was clearly not the interchange product, because it possessed the wrong spectrum (high absorption at 250 mμ, with only a shoulder at 380 mμ, whereas the wavelength for the maximum absorption of the mono-dinitrophenyl cystine under these conditions is 360 mμ). In addition, the same material was produced (in higher yield) in the control containing no cystine.

An attempt was made to lower the dose to a point where the above side reaction would not obscure the possible interchange reaction (see Table II).

Table II. The Effect of Co^{60} *Gamma-ray Irradiation of* bis-*DNP Cystine*

The complete system contains 1×10^{-3} M cystine; 1.25×10^{-4} M *bis*-DNP cystine; 1.2×10^{-2} M versene buffer, pH 7.9; where indicated, 5×10^{-4} M *N*-ethyl maleimide (NEMI) and 5×10^{-5} M *p*-chloro mercuri-benzoate (PCMB). The samples were irradiated with 4×10^5 r of Co^{60} gamma-rays in a 60-minute period.

	Optical density		
	Before ether extraction	After ether extraction	Component with $R_F = 0.6$
Unirradiated	0.510	0.040	0
Complete	0.220	0.171	0
Plus NEMI	0.255	0.198	0
Plus PCMB	0.281	0.209	0
Minus cystine	0.155	0.106	0

After a dose of 4×10^5 r no colored material could be detected with the R_F of 0.6, which is the R_F of mono-dinitrophenyl cystine, although significant

amounts of the reactants were still present. In this experiment the dose was delivered in a one-hour period. During this time the effect of the spontaneous interchange catalyzed by thiols produced by hydrolysis of the disulfides is small (6). In the presence of versene, which prevents metal-catalyzed oxidation by molecular oxygen, the interchange is greater because of greater persistence of thiol. For this reason, controls were included containing thiol binding reagents which completely block thiol catalyzed interchanges.

These preliminary experiments indicate that interchange of disulfide bonds is not a prominent feature of radiation-induced denaturation. Further work will be required to assess the role of disulfide linkages in secondary aspects of denaturation. In addition, further work should be carried out using a disulfide interchange indicator that is not itself influenced by irradiation in aqueous solution and thus affords a more sensitive assay for interchange in aqueous media.

Acknowledgement—The author acknowledges with gratitude the continued interest and valuable advice of his colleagues Howard S. Ducoff, George A. Sacher, and Antreen Pfau.

REFERENCES

1. H. FRICKE: The denaturation of proteins by high frequency radiation. *Cold Spr. Harb. Symp. Quant. Biol.* **6**, 164–170 (1938).
2. E. S. G. BARRON and P. FINKELSTEIN: Studies on the mechanism of action of ionizing radiations—X. Effect of x-rays on some physicochemical properties of proteins. *Arch. Biochem. Biophys.* **41**, 212–232 (1952).
3. C. HUGGINS, D. F. TAPLEY, and E. V. JENSEN: Sulphydryl-disulphide relationships in the induction of gels in proteins by urea. *Nature, Lond.* **167**, 592–593 (1951).
4. M. CALVIN: Mercaptans and disulfides, some physics, chemistry, and speculation. In: *Glutathione*, ed. by S. Colowick *et al.*, 3–26. Academic Press, New York (1954).
5. J. A. BARLTROP, P. M. HAYES, and M. CALVIN: The chemistry of 1,2-dithiolane (trimethylene disulfide) as a model for the primary quantum conversion act in photosynthesis. *J. Amer. Chem. Soc.* **76**, 4348–4367 (1954).
6. A. P. RYLE and F. SANGER: Disulphide interchange reactions. *Biochem. J.* **60**, 535–540 (1955).
7. A. L. KOCH, W. A. LAMONT, and J. J. KATZ: The effects of anhydrous strong acid on ribonuclease and lysozyme. *Arch. Biochem. Biophys.* **63**, 106–117 (1956).
8. E. COLLINSON, F. W. DAINTON, and B. HOLMES: Inactivation of ribonuclease in dilute aqueous solution. Inactivation by hydroxyl radicals. *Nature, Lond.* **165**, 267–269 (1950).

A PROPOSED MECHANISM OF PROTEIN INACTIVATION*

L. G. AUGENSTINE
Brookhaven National Laboratory, Upton, Long Island, New York

Abstract—An hypothesis dealing with the role of disulfide bonds in protein inactivation by physical agents has been discussed with reference to material presented at this conference. It is proposed that the critical effect becomes localized at a 'weak-link', causing first the rupture of a disulfide bond, followed by the breaking of neighboring intramolecular bonds and finally the rupture of a second disulfide bond. Much of the evidence upon which these postulates are based is reviewed. The manner in which this model defines a target volume is indicated and alternative methods of disulfide splitting are discussed.

THE author has previously proposed (1) an hypothesis dealing with the general problems of protein inactivation and the importance of disulfide bonds in maintaining protein structure. This hypothesis was originally presented to account for heat denaturation data. It has since been extended in an attempt to account for inactivation by ultraviolet light and by ionizing radiation ('direct effect') (2, 3, 4). The model is a special case of more general ones proposed by MIRSKY and PAULING (5), LUMRY and EYRING (6) and PLATZMAN and FRANCK (7), and would depend for its accomplishment upon physical processes similar to those described by the latter authors. It is to be emphasized that this scheme is not advanced as the only mechanism whereby protein inactivation can occur, but rather as the *most likely*.

It is proposed that the critical effect of the physical agents mentioned is not to cause indiscriminate molecular disorganization. Rather, their primary effect becomes preferentially localized at certain points in the molecule†. Further, certain of these points (collectively called the 'weak-link') are involved in processes which are characteristic of all proteins and which lead to inactivation. These processes can be characterized as occurring in three distinct steps.

1. The breaking of an S—S bond;
2. The breaking of a variable number of neighboring intramolecular bonds (e.g. H-bonds); and
3. The rupture of a second S—S bond.

Step 1 requires about 20 kcal/mole, or 0.9 eV per molecule (8), and a negligible entropy factor, while an appreciable entropy increase is associated with step 2. Step 3 allows the spontaneous formation of a structure incompatible with further activity. Although irreversibility could result from the

* Research carried out at Brookhaven National Laboratory under the auspices of the U.S. Atomic Energy Commission.

† For instance, PLATZMAN (4, p. 19) has pointed out that 'the most stable position for a migrating electron vacancy to become localized is at a site that can be crudely identified with the atom of lowest ionization potential'.

formation of a single new bond (4)*, it is probably produced most often by the spontaneous breaking of a large number of intramolecular bonds which is accompanied by a very large increase in entropy. Steps 1 and 2 would constitute the activated state of physical denaturation, i.e., reversible inactivation, whereas the rupture of the second S—S bond, step 3, allows irreversible inactivation to proceed. The large entropy change often found to be associated with irreversible denaturation indicates that a partial unfolding of the molecule usually occurs, and therefore the 'weak-link' is probably involved in latching the molecule together. Once the second S—S bond has ruptured, the degree of denaturation will depend both upon the extent to which unfolding proceeds and upon subsequent reactions of the newly exposed groups of the altered molecule. The number and the location of the intramolecular bonds involved in step 2 is thought to be essentially invariable for a given protein, and to depend upon its particular structure in the region of the weak link. It is postulated that a variable number of bonds is involved in step 2 since the enthalpy of denaturation activation varies widely between different proteins (8, 1).

These generalizations are consistent with a variety of experimental findings, many of which have been discussed elsewhere (1,2,3), and therefore will be only mentioned briefly.

(a) Data reported by a number of investigators give a value for the free energy of denaturation activation of $\Delta F^* = 24.8 \pm 1.5$ kcal/mole (or 1.1 eV/molecule), which is slightly in excess of that required to rupture an S—S bond (8).

(b) Mild denaturation is reversible, whereas violent denaturation is not.

(c) Following activation an additional 20 kcal/mole (for trypsin) is necessary to initiate a large entropy change (about four or five times that for ΔS^* of activation). This is thought to be a large configurational change.

(d) An average of two to three cysteine equivalents per insulin molecule corresponds to a fifty per cent reduction in its biological activity (the present hypothesis would predict three cysteine equivalents per molecule, i.e. two for each reversibly inactivated and four for each irreversibly inactivated molecule).

(e) The appearance of the full protein sulhydryl titer is invariably associated with complete loss in activity.

(f) Disulfide bonds are likely involved in the latching together of large segments of the insulin molecule, since reoxidation of the reduced insulin molecule causes aggregation.

(g) The ultraviolet action spectra for the inactivation of trypsin and ribonuclease, both of which have high cystine contents, are peaked at a wavelength corresponding to maximum cystine absorption, and the quantum efficiency is strongly correlated to their cystine content (9), (10); however, SETLOW and DOYLE (10) found that the action spectra for gramicidin and aldolase, which had little or no cystine, roughly paralleled the molecular absorption spectra. They concluded that although there must be more than one inactivation mechanism, a quantum absorbed by cystine could be as much as twenty-five times as effective in producing inactivation as one absorbed by an aromatic

* For instance, inactivation due to freezing and drying, which is apparently not accompanied by a gross opening of the molecule (5), may depend upon such a process.

amino acid, and as much as fifteen times as effective as a quantum which split a peptide bond.

(h) The electron spin resonance measurements reported by GORDY (11) indicate that irradiation of proteins usually converts some of the disulfides into free radicals.

(i) The native configuration of the ribonuclease molecule can be greatly disrupted (as indicated by viscosity measurements) by destroying its H-bonds with urea without destroying its function; however, as soon as the disulfides are oxidized activity is lost immediately (12).

(j) The recent findings of LEONE (13) are in excellent agreement with the two main aspects of the hypothesis. First, he found that the antigenic properties of γ-irradiated serum albumin were the same whether an average of 9 or 90 eV per molecule had been absorbed indicating that irradiation caused this protein to unfold in a characteristic fashion. Second, the single S—S bond of serum albumin was likely involved since ultracentrifugation patterns of the irradiated material contained only monomers, dimers and small amounts of di- and tripeptides, with no evidence of larger aggregates. SETLOW and DOYLE (10) also noted smaller fragments produced by ultraviolet irradiation of trypsin, but found that the degradation components produced by a wavelength very favorable to a cystine effect were more prominent and homogeneous than those produced by a less favorable wavelength. They interpreted this as evidence for more than one inactivation mechanism.

(k) Studies of the inactivation of protein monolayers (2,3,14,15) yielded results which were consistent with the model; however, the data could only disprove but not prove the hypothesis. For example, molecules in compressed monolayers show reduced inactivation from both surface forces (14) and irradiation (15). This was expected, since the scheme proposed here would predict that an external force, such as the monolayer film pressure, should lower the probability of the second S—S bond being ruptured; and even if step 3 occurred, an external pressure should be able to maintain molecular structure sufficiently intact so that restitution would be enhanced. However, it was estimated that the proposed mechanism might account for no more than two-thirds of the inactivation observed.

Some proteins, such as ovalbumin (16) and serum albumin, contain fewer than two S—S bonds. In such proteins other bonds which (i) have comparatively small rupture energies and (ii) are involved in latching large segments of the molecule together, would assume the functions of the cystine in this scheme.

The present model provides a specification of the 'target volume' for irradiation inactivation. Associated with each atom is a probability that energy will be absorbed and migrate to the weak link in amounts sufficient to rupture that structure completely. The sum of these probabilities over the whole molecule gives the 'effective target volume'. Thus, the actual physical target need not have sharp boundaries (see also discussions by LEA (17), BURTON (18) and SETLOW and DOYLE (10) of target elements having probabilities other than 0 or 1).

The probabilities, and thus the 'size' of the target volume, will depend upon a number of factors. For instance, the possibility—discussed by PLATZMAN and FRANCK (7)—of complementary effects between thermal and absorbed

energies suggests that the target volume should decrease as the temperature at which protein is irradiated is decreased. Consistent with this is the fact that the x-irradiation cross-section of phage Tl has been found to be a linear function of the irradiating temperature (19); also the inactivation of trypsin ultraviolet-irradiated at 300°K is about three times as great as at 90°K (10). The target volume should also decrease as the quantum of energy absorbed is decreased: the inactivation cross-section of bovine serum albumin bombarded with very low energy electrons was found to increase with increasing β energy and a measurable cross-section was obtained with particle energies as low as 10 ev (20).

The recent results and interpretations of YALOW (21) are particularly pertinent to the hypothesis discussed here. Her irradiations of insulin, serum albumin and cystine indicated that disulfide bonds are reduced both under conditions where direct and indirect effects should predominate. However, she proposed that the splitting occurred between the C and S (leaving an S—S—C configuration) rather than between the sulfurs. Although the data cited previously appear to indicate a splitting of the S—S bond, most of those same data would be compatible with a reduction of the C—S bond instead. To select between the two possibilities may be difficult, since the energy required to split a given bond in a compound such as cystine may be quite different than that required in a protein; as LUMRY and EYRING (6) point out, various of the intramolecular bond angles of proteins may be distorted in order to effect structural compromises which minimize free energy. However, YALOW (21) has pointed out that the production of a C—S—S· radical is probably more consistent with Gordy's findings than the other alternative.

The failure of KOCH (22) to detect radiation-induced disulfide interchanges either in solution or the dry state does not disprove the hypothesis proposed here. The dosages they used (up to 3×10^7 r) are much larger than those required by other workers (3×10^4 r) to liberate sulfur groups (21, 23) from similar compounds. These results indicate that although the splitting of disulfide bonds may well be critically involved in protein inactivation, seven per cent or less* of the liberated —SH groups undergo interchange.

REFERENCES

1. L. AUGENSTINE: Structural interpretations of denaturation data. In: *Information Theory in Biology*, ed. by H. Quastler, 119–124, University of Illinois Press, Urbana (1953).
2. L. AUGENSTINE and R. RAY: Trypsin monolayers at the air-water interface. III. Structural postulates on inactivation. *J. Phys. Chem.* **61**, 1385–1388 (1957).
3. L. AUGENSTINE: Trypsin monolayers at the water-air interface. Ph.D. thesis, University of Illinois (1956).
4. E. C. POLLARD: The direct effect of radiation on proteins, viruses, and other large molecules. In: *Biochemical Aspects of Basic Mechanisms in Radiobiology*, ed. by Harvey M. Patt, Nat. Res. Council Washington Publ. no. **367**, 1–29 (1954).

* The $1/e$ dose, D^*, for a spherical molecule of mol. wt. 240 and density 1.35 is 2×10^9 r. A dose of 3×10^7 r could potentially rupture 1.5 per cent of the S—S bonds (from $dN/N = dN/D^* = (3 \times 10^7)/(2 \times 10^9)$). If one interchange per 1000 disulfides could have been detected (22) but none was found, then seven per cent (1/15) or less of the 'hit' molecules were eventually involved in disulfide interchanges.

5. A. Mirsky and L. Pauling: On the structure of native, denatured and coagulated proteins. *Proc. Nat. Acad. Sci., Wash.* **22,** 439–447 (1936).
6. R. Lumry and H. Eyring: Conformation changes of proteins. *J. Phys. Chem.* **58,** 110–120 (1954).
7. R. Platzman and J. Franck: A physical mechanism for the inactivation of proteins by ionizing radiation. This volume.
8. A. Stearn: Kinetics of biological reactions with special reference to enzymic processes. *Advances in Enzymology* ed. by F. F. Nord, **9,** 25–74 (1949).
9. R. Setlow: A relation between cystine content and UV sensitivity of proteins. *Biochim. Biophys. Acta* **16,** 444–445 (1955).
10. R. Setlow and B. Doyle: The action of monochromatic ultraviolet light on proteins. *Biochim. Biophys. Acta* **24,** 27–41 (1957).
11. W. Gordy: Electron spin resonance in the study of radiation damage. This volume.
12. C. Anfinsen and R. Redfield: Protein structure in relation to function and biosynthesis. *Advances in Protein Chemistry* **11,** 1–100 (1956).
13. C. Leone: personal communication.
14. R. Ray and L. Augenstine: Trypsin monolayers at the air-water interface. I. Film characteristics and the recovery of enzymatic activity. *J. Phys. Chem.* **60,** 1193–1199 (1956).
15. L. Augenstine and R. Ray: Trypsin monolayers at the air-water interface. II. Effect of x-ray and UV radiation upon enzymatic activity. *J. Phys. Chem.* **61,** 1380–1384 (1957).
16. G. Tristram: Amino acid composition of purified proteins. *Advances in Protein Chemistry* **5,** 84–148 (1949).
17. D. Lea: *Actions of Radiations on Living Cells,* Cambridge University Press, Cambridge, England (1946).
18. M. Burton: Elementary Chemical processes in radiobiological reactions. In: *Symposium on Radiobiology,* ed. by J. J. Nickson, 117–138, J. Wiley and Sons, New York (1952).
19. C. Bachofer, C. F. Ehret, S. Mayer, and E. L. Powers: The influence of temperature upon the inactivation of a bacterial virus by x-rays. *Proc. Nat. Acad. Sci., Wash.* **39,** 744–750 (1953).
20. F. Hutchinson: Energy requirements for the inactivation of bovine serum albumin by radiation. *Radiat. Res.* **1,** 43–52 (1954).
21. R. Yalow: Cleavage of the disulfide bond as a result of the indirect or direct effect of ionizing radiation. National Biophysics Conference, Abstract No. 118 (1957).
22. A. L. Koch: The absence of radiation-induced disulfide interchanges. This volume.
23. W. Dale and J. Davies: The liberation of hydrogen sulfide by x-radiation from cysteine and glutathione. *Biochem. J.* **48,** 129–132 (1951).

DISCUSSION

Platt: This is a tangential comment which may be relevant to the paper by Professor Gordy and that by Professor Platzman and Professor Franck. It is that Dr. Meyerson* has recently done some extremely beautiful work on species of various mass fragments that break up during their flight time in the mass spectrometer, and by isotopic substitution experiments, checking the species that come out, he has proved that when isopropylbenzene is bombarded with 50-volt electrons, the side group is changed to a cyclopropyl group in which all three carbons are equivalent. He has also proved that if a molecule like toluene is bombarded, the decay in flight indicates the existence of a species of the form of a tropylium ion, a C_7 ion in which again all the carbons are equivalent. Similar results are obtained down to zero excess electron energy.

* P. N. Rylander, S. Meyerson, and H. M. Grubb, "I. The Cationated and Cyclopropane Ring" *J. Amer. Chem. Soc.* **78,** 5799–5802 (1956): Organic ions in the gas phase. II. The tropylium ion. *J. Amer. Chem. Soc.* **79,** 842–846 (1957). See also later papers of this series by same authors in *J. Chem. Phys.* and *J. Phys. Chem.*

The result, I think, is that one should probably not speak about traditional valence bonds in traditional directions in a molecular species which has been ionized by ions or radiation. The notion of the classical valence bond is peculiar to a closed shell system, that is, an electronically closed shell with a filled lower energy band and with a large energy gap between the highest filled state and the lowest empty state. I think that it might not be possible to attribute the electrons whose spins Professor GORDY detects to a particular site in the original primitive molecule, because these bonds may have been completely rearranged. Also it may not be possible to attribute Professor PLATZMAN's type of damage to a particular section of the molecule, if this sort of thing, cyclopropyl or tropylium ion, is very common. There may be a large section of the molecule which is doing a merry-go-round of interchangeable carbon atoms before the system settles down.

PLATZMAN: I should like to record my skepticism as to the ubiquitous and decisive role of disulfide-bond cleavage in radiobiology. This is not to question the participation of such breakage in denaturation by ionizing radiation or any other agent: since disulfide bonds make an important contribution to the structural stability of many proteins they must certainly be involved in structural breakdown. Dr. AUGENSTINE has, indeed, attempted here to describe the relationship between the contribution of secondary bonds and that of disulfide bonds. However, the argument—frequently heard in recent years—that the great sensitivity of a protein molecule to ionizing radiation can be understood in terms of migration of an electron vacancy produced almost anywhere in the molecule to a disulfide bond, with resultant cleavage of that bond and unfolding of the molecule, is open to most serious objections. In the first place, such long-range migration is unsupported by independent evidence and, as indicated in my paper, is likewise unsupported by physical principles. The fact that electron holes are observed to move freely in certain nonmetallic crystals is of doubtful relevance because of the different dielectric properties of such crystals, and because of their periodic structure. Moreover, even though Professor GORDY's proof of the existence of free valences at sulfur atoms is impressive (although the precise number of such radicals in relation to the radiation dose is still uncertain), a simple causal connection between formation of the radicals and inactivation of the protein has yet to be established, and, indeed, may not exist at all. It is quite possible that they are a secondary factor in denaturation, however conspicuous they may be in the paramagnetic–resonance spectrum. Furthermore, the logical link between disulfide-bond cleavage and electron-vacancy migration is also unproven. A simpler and plausible mechanism for a strong sulfur-atom signal is given by the action of subexcitation electrons: that these can attack the disulfide bonds effectively even though the latter are present in low concentration is strongly suggested by the small dissociation energy of such bonds and also by the marked red displacement of the absorption spectrum of cystine in relation to the spectra of most other amino acids.

GORDY: I certainly think that Professor PLATZMAN's suggestion is worthy of consideration. I try to keep the sulphur-sulphur bonds in my brain open when discussing these complex systems.

PART V

AGING AND RADIATION DAMAGE

A FEATURE of our times is that people are now living long enough so that the problems and diseases of the aged have become an important medical speciality, and at the same time we are, of necessity, embarking on the development of civil and military technology which generates radioactivity, an agent which, uncontrolled, will contribute to shortening our lives. There is evidence that these two attributes of this age are more than incidentally related.

It is well for us to remember that the biological effects of radiation are not new, for the same radiation by which BECQUEREL discovered radioactivity very soon thereafter burned his person. An understanding of these effects has come slowly. The relationship between aging and radiation damage has been dormant in the literature for a long time and has come to prominence only recently.

The first papers on the effect of radiation on life span were published by W. P. DAVEY (1, 2) in 1917 and 1919. The care exercised in dosimetry and in showing that the observed effect is due to the x-rays and not to some experimental artifact was most remarkable for the time at which this work was done. Davey found that the life span of the beetle *Tribolium confusum* was shortened by large amounts of x-rays and lengthened slightly by small amounts. The first result seems to be well established today. The second result is still frequently reported.

GOWEN and STADLER (3) in 1952 found an increased life span for male *Drosophila melanogaster* given 2500 r, although the life span of the female was decreased. The effect appeared in LORENTZ's data (4) on the LAF_1 mouse and inbred guinea pigs receiving 0.11 r per day. He did not consider this statistically significant, although SACHER (5) later stated that the effect is significant and that it had been confirmed by himself and D. GRAHN. GOWEN (6) found a shortening of the life span in male mice from ten distinct inbred strains—even for small single doses of x-rays. However, for female mice he found an increase in life span for doses up to 320 r. But the number of litters produced was reduced even for small doses. The explanation given was that the semi-sterility induced by x-rays reduced the hazards of pregnancy. For low doses, it was argued, this more than compensated for the somatic x-ray damage.

It is not generally accepted that there is a stimulation due to x-rays. Probably cases where this seems to occur can be explained as an artifact, perhaps following GOWEN's explanation. At any rate further research on this point is well justified.

In 1937 RUSS and SCOTT (7) published a report on the biological effects of continuous gamma irradiation. They found the significant features known today, namely, that there is a cumulative permanent damage reflected by a death rate higher than that of the controls, sterility or semi-sterility, high infant and prenatal mortality of progeny from both male and female irradiated parents. They confirmed these results in 1939 and specifically called attention to accelerated aging in the irradiated rats (8).

The invention of the nuclear reactor in 1942 added immensely to the industrial and laboratory hazards of radiation and to the concern for evaluating these hazards. HENSHAW (9) in 1944 again called attention to the similarity between the pathology of aging and the pathology of radiation damage. SACHER (10)

in 1950 and BRUES and SACHER (11) in 1952 considered radiation injury and lethality and normal aging from the same point of view and gave an analysis in terms of survival curves.

In 1953 H. A. BLAIR (12) emphasized this relationship and extended the notion to internal emitters. He pointed out that the shortening of life, even with bone seekers such as Po, Pu and Ra, is not attributable solely to bone pathology since other tissues are also damaged in a way similar to total body irradiation. Blair's remark is based on observations by BOYD *et al.* (13) that tissue changes were of the type produced in rats by 550 r whole-body irradiation. In 1954 FURTH, UPTON, CHRISTENBERRY, BENEDICT, and MOSHMAN (14) called attention to this relationship in the case of LAF_1 mice exposed to atomic bomb radiation. In the same year UPTON, FURTH, and CHRISTENBERRY (15) made the same observation with regard to late effects from thermal-neutron irradiation of *RF* mice.

The similarity between aging and radiation damage is paralleled by chemical carcinogens. CLOUDMAN, HAMILTON, CLAYTON, and BRUES (16) reported that mice painted with a carcinogenic agent (methylcholanthrene) exhibited a life shortening not to be explained by a single pathology. They indicated an analogy between life shortening from hydrocarbons and total-body irradiation.

RUSSELL (17) has recently found that the increased prenatal and infant mortality of offspring from irradiated parents continues throughout life and is reflected by a reduced average life. He studied only the offspring of male mice irradiated by neutrons from an atomic weapon. Presumably, the effect is general and applies to offspring of both male and female animals subjected to any ionizing radiation. The relation of this work to that of RUSS and SCOTT is clear and the need for detailed study is paramount.

Some may feel that establishing a relation between two unexplained effects gets one nowhere. However, as PLATT (18) points out, such a relationship may help one effect to explain the other.

The concept of premature aging as a measure of damage from various deleterious agents seems to be well enough established for practical use in understanding the nature of biological damage. Information theory may well have a contribution to make to the elucidation of these problems of our times which are so important from so many points of view.

H. P. Y.

REFERENCES

1. W. P. DAVEY: The effect of X-rays on the length of life of *Tribolium confusum*. *Gen. Elect. Rev.* **20**, 174–182 (1917). (An article nearly identical to this is to be found as follows: W. P. DAVEY: The effect of X-rays on the length of life of *Tribolium confusum*. *J. Exp. Zool.* **22**, 573–592 (1917).)
2. W. P. DAVEY: Prolongation of life of *Tribolium confusum* apparently due to small doses of X-rays. *Gen. Elect. Rev.* **22**, 479–483 (1919).
3. J. W. GOWEN and J. STADLER: Irradiation effects on viability of *Drosophila melanogaster*. *Genetics* **37**, 586–587 (1952).
4. E. LORENTZ: Effects of long-continued total-body gamma irradiation on mice, guinea pigs, and rabbits. In: *Biological Effects of External X and Gamma Radiation*, 226–242, ed. by R. E. Zirkle, *National Nuclear Energy Series* **IV-22B**, McGraw-Hill, New York (1954).
5. G. A. SACHER: On the statistical nature of mortality, with especial references to chronic radiation mortality. *Radiology* **67**, 250–258 (1956).
6. J. W. GOWEN: private communication (1957).
7. S. RUSS and G. M. SCOTT: Some biological effects of continuous gamma irradiation, with a note on protection. *Brit. J. Radiol.* **10**, 619–629 (1937).

8. S. Russ and G. M. Scott: Biological effects of gamma irradiation. *Brit. J. Radiol.* **12,** 440–441 (1939).
9. P. S. Henshaw: Experimental roentgen injury. IV Effects of repeated small doses of X-rays on blood picture, tissue morphology and life span in mice. *J. Nat. Cancer Inst.* **4,** 513–522 (1944).
10. G. A. Sacher: The survival of mice under duration-of-life exposure to X-rays at various dose rates. Argonne Nat. Lab. Report No. CH-3900 (1950).
11. A. M. Brues and G. A. Sacher: Analysis of mammalian radiation injury and lethality. In : *Symposium on Radiobiology*, 441–465, ed. by J. J. Nickson, J. Wiley and Sons, New York (1952).
12. H. A. Blair: The shortening of life span by injected radium, polonium and plutonium. University of Rochester Atomic Energy Project, Document No. UR-274 (1953).
13. G. A. Boyd, H. E. Silberstein, R. M. Fink, A. Frenkel, W. L. Minto, R. G. Metcalf, G. Casarett, and G. M. Suter: Pilot studies on the intravenous lethal dosage of polonium, plutonium, and radium in rats. In: *Biological Studies with Polonium, Radium and Plutonium*, ed. by R. M. Fink, Chap. 7, *National Nuclear Energy Series* VI.–3. McGraw-Hill, New York (1950).
14. J. Furth, A. C. Upton, K. W. Christenberry, W. H. Benedict, and J. Moshman: Some late effects in mice of ionizing radiation from an experimental nuclear detonation. *Radiology* **63,** 562–570 (1954).
15. A. C. Upton, J. Furth, and K. W. Christenberry: Late effects of thermal neutron irradiation in mice. *J. Cancer Res.* **14,** 682–690 (1954).
16. A. M. Cloudman, K. A. Hamilton, R. S. Clayton, and A. M. Brues: Effects of combined local treatment with radioactive and chemical carcinogens. *J. Nat. Cancer Inst.* **15,** 1077–1083 (1955).
17. W. L. Russell: Shortening of life in the offspring of male mice exposed to neutron radiation from an atomic bomb. *Proc. Nat. Acad. Sci., Wash.* **43,** 324–329 (1957).
18. J. R. Platt: Functional geometry and the determination of pattern in mosaic receptors. This volume.

A STUDY OF AGING, THERMAL KILLING, AND RADIATION DAMAGE BY INFORMATION THEORY

HUBERT P. YOCKEY
Health Physics Division,
Oak Ridge National Laboratory, Oak Ridge, Tennessee

Abstract—The information theoretic formalism developed in the author's paper in Part I has been applied to the calculation of survival curves. The results have been compared for a variety of organisms ranging from viruses to mammals. The deleterious agent varied in the magnitude of its quantum energy from thermal and chemical energies to several Mev.

It should be emphasized that this article says very little about models. In general the accepted model of the organism is taken and its behavior is calculated from those features of the model which have the aspects of a communication system.

In spite of the complexity and variety of the organisms and the range of energy in the lethal agent and without ad hoc assumptions pertaining to models, we were able to account for the main features and some details of the survivorship curves. Many additional experiments will be suggested; some of them have been pointed out and some predictions have been made.

The idea of the storage and transfer of information and its destruction by deleterious agents seems to link the material discussed. Much of this material may otherwise seem unrelated or only vaguely so.

I. INTRODUCTION

THE study of the survivorship curve has contributed to the quantification of some essential but otherwise qualitative notions in biology. The effect of various insults such as ionizing radiation, ultraviolet light, temperature, disease, chemicals, and so forth, is very often measured by survivorship of a suitable test organism. On the other hand, the experimenter may be interested in the survival response of a particular organism as a function of maturation, nutrition, strain difference, or the like, and may use some convenient agent as a test stimulus.

Survivorship does not contain all we feel intuitively is involved in the concepts of 'vigor' or 'fitness' but it does contain much of what can be defined and measured in an unequivocal and operational way that is associated with those ideas. These facts, together with the application in evaluating quantitatively hazards to man, make this subject one of great practical and theoretical interest.

Information theory is peculiarly well qualified to provide a mathematical treatment of these matters. The survivorship curve is a property of the ensemble of organisms rather than of the individual. It reflects the generalized decay of the organization of a system. The central thesis of this paper is that aging, thermal killing, and radiation damage reflect essentially the same action, namely, the destruction of the information content of the cell. The ideas discussed in the author's previous article in this volume will be applied to the calculation of

haploid survivorship, diploid survivorship, and to the role of equivocation in the germ line.

II. SURVIVORSHIP FOR THE HAPLOID CASE

OTTO RAHN (1, 2) was the first to suggest in two pioneer papers published in 1929 and 1930 that the genetic structure is the sensitive element in the cell for radiation damage, thermal killing, and the action of some chemicals. He later reviewed the data on disinfectant action and confirmed this opinion (3). LEA (4) was a strong supporter of this idea and used it in his development of the target theory. This is generally an accepted view today (5) although the role of gene mutations and chromosome aberrations is a matter of debate (6, 7).

In a previous article in this volume we showed that this notion follows directly from the application of information theory to the current conception of the storage and transfer of genetical information in the cell and the synthesis of proteins. It is therefore of interest to continue the argument and attempt to calculate survivorship curves. We also showed that error will exist in the genetical information of all real organisms. The organism will live and multiply according to DANCOFF's principle (8), in spite of these errors. We argued previously that there must be a distribution of message entropy values among the elements of the ensemble of organisms. Suppose that the number of errors in the genetical information is increased as a result of, say, radiation. Those elements of the ensemble near the lethal limit will succumb even though they were quite viable before irradiation.

This is a notion peculiar to information theory. The communications analogy is SHANNON's channel capacity theorem (9). This theorem shows that if a channel has a capacity C, it is possible, by proper coding, to send information at rate C or less through the channel *with as small a frequency of errors as desired.* Thus, though the noise level in the channel will affect the channel capacity C, it will not prevent nearly perfect transmission of information. This can be assured by proper coding. As long as this limit C is not exceeded, it is impossible for the recipient to know the noise level in the channel or information source.

With these points in mind, we now return to the suggestions of RAHN and LEA, keeping further in mind the idea of WATSON and CRICK that a mutation is a change in the order of nucleotide bases in DNA (or some other information-bearing molecule). We have proposed that the action of radiation or other deleterious agent at the molecular level is such that the nucleotide pair mimes some other nucleotide pair insofar as protein synthesis and replication are concerned (10). The action of radiation may therefore be thought of as causing lethality through gene mutation by decreasing the message entropy of some members of the ensemble below the lethal limit. This is essentially the suggestion of RAHN and of LEA phrased in the language of information theory, and it follows from the argument given in the previous article in this volume. On this basis we may proceed to calculate the force of mortality on the ensemble.

The distribution of message entropy in the ensemble will be represented by a probability distribution $\rho(H, \lambda)$, where λ is a measure of the magnitude of the deleterious agent and the initial distribution is $\rho(H, 0)$. This distribution will vary with the genetic character of the ensemble of organisms. It can probably be derived from first principles, at least for simple cases, when more

is known about storage and transfer of information in organisms. For the present it will be necessary to make some simple assumptions, however.

It was proposed previously (10) that death occurs when the value of H decays below some limit H_d. Let l be the number of organisms in the population representing the ensemble. The probability per unit λ of leaving the population $(1/l)(dl/d\lambda)$ is called the force of mortality. The force of mortality will be the probability density unit per H at H_d after exposure λ multiplied by the rate of decrease of H per unit λ at H_d.

$$\frac{1}{l}\frac{dl}{d\lambda} = \rho(H_d, \lambda)\left.\frac{dH}{d\lambda}\right|_{H_d} \tag{1}$$

The value of $\rho(H, \lambda)$ varies continuously with λ; no organisms leave the microensemble which at $\lambda = 0$ lies between H and $H + dH$.

$$\rho(H, 0)\left.\frac{dH}{d\lambda}\right|_{\lambda=0} d\lambda = \rho(H_d, \lambda)\left.\frac{dH}{d\lambda}\right|_{H_d} d\lambda \tag{2}$$

The relation of $\rho(H_d, \lambda)$ to $\rho(H, 0)$ is as follows:

$$\rho(H_d, \lambda) = \rho(H, 0)\,\frac{H_0 - H + \frac{1}{4}\sum_{i,j} p(i)\log_2 p_i(j)}{H_0 + H_d + \frac{1}{4}\sum_{i,j} p(i)\log_2 p_i'(j)} \tag{3}$$

where $p_i'(j)$ is the value corresponding to H_d.
Equation (1) may be written

$$\frac{1}{l}\left(\frac{dl}{d\lambda}\right) d\lambda = \rho(H_d, \lambda)\left[H_0 - H_d + \frac{1}{4}\sum_{i,j} p(i)\log_2 p_i'(j)\right] J(\lambda)\, d\lambda \tag{4}$$

In many cases the action of the deleterious agent will be of the first order so that $J(\lambda) = J_0$, a constant. Let us assume that $\rho(H, 0)$ is of such shape that $\rho(H_d, \lambda)$ is a constant.

Equation (4) may be integrated:

$$\log_e l/l_0 = \rho(H_d, \lambda)\left[H_0 - H_d + \frac{1}{4}\sum_{i,j} p(i)\log_2 p_i'(j)\right] J_0\lambda \tag{5}$$

Equation (5) represents haploid survivorship as a function of λ for many types of destructive influences under many experimental conditions—but not for all influences, or conditions, or haploid organisms. T. ALPER (11) found the rate of inactivation by gamma rays of dysentery phage S13 to increase with increasing dose at 130 rad/min. At 5.3 rad/min the survival curves departed markedly from the exponential form although that form was found when catalase was present. WATSON (12) also reported the same phenomenon with phage T2. ALPER (13) later showed that the gas treatment of phage could result in departure from the exponential form. A number of cases of a non-exponential inactivation curve for viruses are discussed by LURIA in a recent review (6).

GATES (14) discussed the deviations from an exponential curve for ultraviolet irradiations of bacteria. Recent work by URETZ (15) has shown that

inactivation of haploid yeast is exponential for x-rays and sigmoid for ultraviolet. ANDERSON (16) has irradiated two biochemical mutants of strain B of *E. coli* with x-rays, namely, the streptomycin dependent strain and the purine-less strain. An exponential survival curve is obtained in oxygen while a sigmoid curve is obtained in nitrogen for each strain. HOLLAENDER, BAKER, and ANDERSON (17) have discussed the effect of oxygen and other chemicals on the x-ray sensitivity for mutation production and survival.

HOLLAENDER and STAPLETON (18) have shown that many types of survival curves may be obtained ranging from exponential to a very pronounced sigmoid shape depending on the experimental conditions.

STAPLETON, SBARRA, and HOLLAENDER (7, 19) have studied the nutritional aspects of survival of bacteria from ionizing radiation. They showed that the B/r strain of *E. coli* grown on a complete medium such as nutrient broth exhibited radiation-induced requirements for nutritional factors. They presented some evidence showing that such bacteria are not stable auxotrophic mutants.

The dependence of survivorship on nutritional factors is explained by ZIRKLE and TOBIAS (20) from hit-theory concepts. They state that: 'Accordingly, the number n of essential sites in the haploid chromosome set might vary with the composition of the medium; in general, one would expect that the richer the medium the fewer would be the observed number of essential sites. On the other hand, if the 'inactivation' of a 'site' is not a mutation, but a gross change in chromosome state or configuration, the number of sites would be independent of the composition of the medium.'

The interpretation of these results given by the authors quoted loses some force since essentially the same results are found for viruses by FRIEDEWALD and ANDERSON (21), by LURIA and EXNER (22), and by DALE (23). The explanation offered by LURIA and EXNER is based on a two-fold action of the radiation, a direct and an indirect effect.

The case where two deleterious influences operate simultaneously is interesting. WOOD (24, 25) has studied the x-ray survival of haploid yeast as a function of temperature. The curves show a distinct tendency to be concave downward for temperatures between 45°C and 55°C. He finds a 'softening' or 'memory' of exposure to temperature and x-rays for the action of the other. URETZ (15) finds very little 'softening' in his study of the action of x-rays and ultraviolet on haploid yeast. We are not aware of a study of the ultraviolet survival as a function of temperature, although such data would be of importance to complete knowledge of these effects.

GRAY (26) has pointed out recently that a view is gaining general acceptance that a site may be inactivated by a single fast electron, but not by the absorption of a single photon. The site mentioned by Gray is interpreted as a nucleotide pair in the present paper. The action of the deleterious effect may be, partly at least, to throw the nucleotide pair into an excited tautomeric form. In such a form it may be more easily damaged by a successive interaction. The extent to which this occurs may very well depend on the chemical environment. At any rate, for the present purposes, there is reason to believe that $J(\lambda)$ may be represented by a polynomial in λ. The higher order terms represent higher order reactions.

In that event, it is possible to begin to understand why an exponential

survival is obtained under some conditions, but not under others. The shape of the survival curve depends on both the environment and on the genetic character of the organism. Thus, it may be possible to obtain some separation between purely biological and purely physical or chemical phenomena—in so far as such a separation has meaning—by means of the present theory. The function $\rho(H, 0)$ is related to the distribution of genetical information in the ensemble and so is characteristic of the biology of these problems. $J(\lambda)$ represents the interaction of radiation and matter, and will be determined by the physics and chemistry of the situation. It is in this regard that the current controversy on the role of direct and indirect action bears on the present theory.

Not all haploid organisms exhibit the exponential survivorship curve. NYBOM (27) reported sigmoid X-ray survival curves for the green algae, *Chlamydomonas eugametos*, *C. moewusii*, and *C. reinhardi*. Genetic experiments involving tetrad analysis indicate that the haploid character of these organisms is reasonably certain. JACOBSON (28) has studied *C. reinhardi* in some detail and finds a sigmoid survival curve. The mathematical nature of this curve is such that it does not correspond with target theory calculations. This and other cases of sigmoid survival curves in haploid organisms will be discussed in the next section.

III. SURVIVORSHIP FOR DIPLOID

LEA (4) rejected the gene mutation suggestion for 'killing of organisms other than bacteria or viruses' in favor of the view that chromosome aberrations are the main cause for lethal effects in polyploid tissue. He did this partly on the ground that a 'recessive lethal mutation in a diploid cell will not be lethal unless it is in the X chromosome of the male, owing to the presence of a normal allelomorph in the same cell'.

ZIRKLE and TOBIAS (20) retained the recessive lethal mutation hypothesis in their study of X-ray survival curves in yeast. It was shown by TOBIAS and STEPKA (29) that irradiated diploid yeast exhibits an inheritable increase in radiosensitivity presumably because of an increased load of recessive lethal mutations. MORTIMER and TOBIAS (30) obtained direct experimental evidence for X-ray induced recessive lethal mutations by demonstrating a reduction in the fraction of germinating spores produced by X-ray exposed diploid yeast cells.

MORTIMER (31) obtained further evidence for the existence of recessive lethal mutations in studies of the conjugation of yeast cells of opposite mating type. See results shown in Fig. 1. MORTIMER argued, as LEA had, that the viability of zygotes should be unaltered because of the presence of the normal allelomorph and that therefore recessive lethal mutations could not be responsible for all the radiation damage.

Chromosome aberrations clearly represent an increase in the equivocation in the genetic information. It is difficult to see how this is to be calculated at the present writing. It is unclear also what their role is in insults milder than damage due to ionizing radiation, such as thermal killing and aging. SACHER (32) has called attention to the need for cytological investigation of the part chromosome aberrations play in the development of late effects from radiation damage. RUSSELL (33) has argued that chromosomal aberrations probably have little to do with radiation hazards to man. At any rate, one may argue that

recessive lethal gene mutations play an important role in the lethality of diploid cells. Since it is possible to present a calculation of the equivocation due to this process let us calculate the survivorship curve according to this notion and see how it compares with experiment.

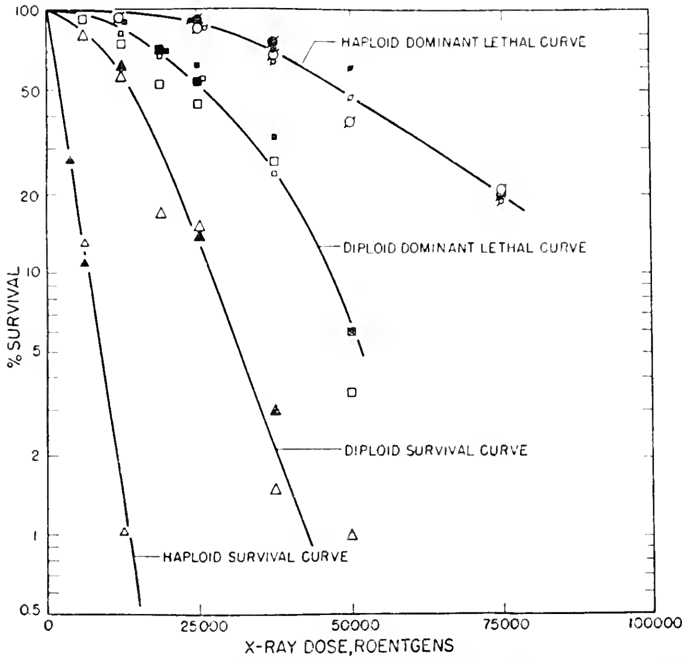


FIG. 1. Percent survival for yeast with one irradiated parent. Haploid × haploid cross (oo), haploid × diploid cross (oO), etc. The first symbol represents a cell of the α-mating type, the second one of a-mating type. A filled letter o designates the irradiated parent. Haploid dominant lethal curve: ȯ, ●o; ●̇, o●; ọ, ●O; Ȯ, O●. Diploid dominant lethal curve: □, ●O; ■, O●; □̇, ●o; ■o●. Diploid survival curve: △, ●—$\alpha\alpha$ diploid, ▲, ●—$\alpha\alpha$ diploid. Haploid survival curve: ▲, ●—α haploid; ▲, ●—α haploid. (From ref. (61)).

The decay of the correct read-off probability is given by equation (10) of my paper in Part I:

$$\frac{d}{d\lambda} p_i(j) = -J(\lambda)\, p_i(j) + \tfrac{1}{4} J(\lambda) \tag{6}$$

In the diploid case $J(\lambda)$ cannot contain a constant term because of the protection afforded by the unaffected allelomorph. Therefore $dp_i(j)/d\lambda$ must depend at least linearly on λ. The polynomial for $J(\lambda)$ is in this case, where J_1 is a constant:

$$J(\lambda) = J_1 \lambda \tag{7}$$

Substitute this function in equation (4)

$$\frac{1}{I}\frac{dI}{d\lambda} = \rho(H_d, \lambda)\left[H_0 - H_d + \tfrac{1}{4}\sum_{i,j} p(i) \log_2 p_i'(j)\right] J_1 \lambda \tag{8}$$

The assumption of Section II concerning the nature of $\rho(H_d, \lambda)$ is retained and the expression for the survival curve is:

$$\log_e l/l_0 = \rho(H_d, \lambda)\left[H_0 - H_d + \frac{1}{4} \sum_{i,j} p(i) \log_2 p'_i(j)\right] \frac{J_1\lambda^2}{2} \tag{9}$$

That some survivorship curves have wholly or substantially the form of equation (9) for normal aging where λ is the time can be shown for a wide variety of organisms. Some examples are shown in Fig. 2 and also one in

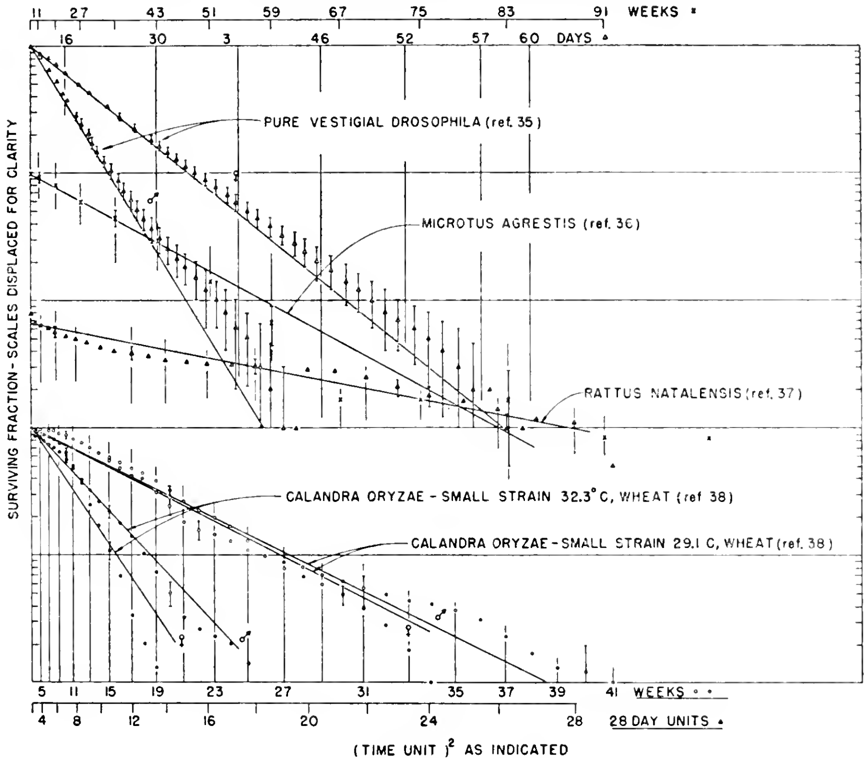


FIG. 2. Survivorship curves for two insects and two mammals plotted against the square of the age (see equation (9)). Straight lines for pure vestigial *Drosophila* (35) obtained from maximum likelihood estimate. Other straight lines are for comparison only. (See text.)

Fig. 3. The data for pure vestigial *Drosophila* are from PEARL and PARKER (35) and represent a life table. The animals are kept under ideal conditions and the number which die in certain time intervals is recorded. All data on Fig. 2 and Fig. 3 are obtained this way. The curve has been fitted by these authors to a function of the following form where a, b, c, d, e are positive constants.

$$\log l = e^{a\lambda}(b - c\lambda + d\lambda^2 - e\lambda^3) \tag{10}$$

PEARL and PARKER were aware that this description involves too many constants and that irrelevant statistical fluctuations are preserved by equation (10).

LESLIE and RANSON (36) fitted their data on the vole to a function of the form of equation (9). They give a conventional χ^2 analysis to justify the hypothesis. The χ^2 analysis cannot be applied to data obtained as these life tables are obtained since the points are not statistically independent. The random variable is the time of death of each animal, not the number alive at

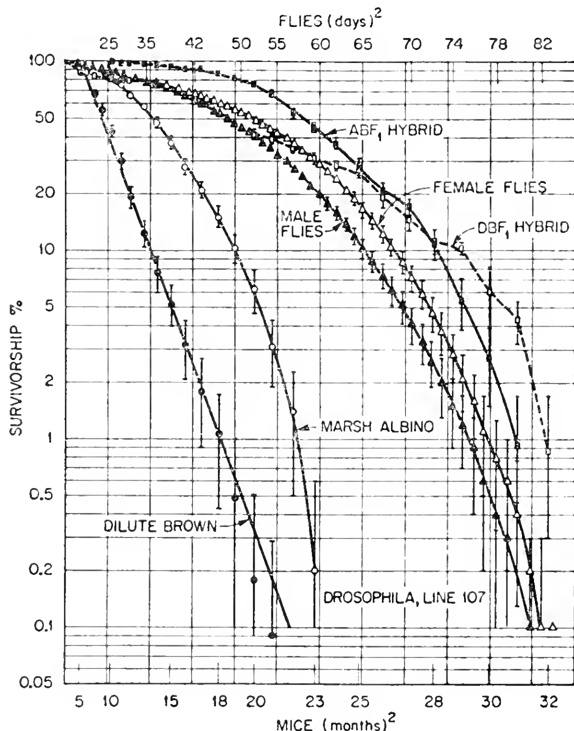


FIG. 3. Normal aging survivorship of certain strains of mice (52, 53) and *Drosophila melanogaster* (35). Note that the abscissa is plotted as age squared to show that the dilute brown strain follows equation (9) but that others do not.

a given interval of time. We are unable to justify our hypothesis in any objective mathematical way. The standard errors given are calculated as follows where l_i is the number in the i interval and l_{i+1} the number in the $i+1$ interval

$$\pm\sqrt{\frac{(l_i - l_{i+1})}{l_i}}\, l_{i+1} \tag{11}$$

The points in Fig. 2 lie very near to a straight line and so it is plausible, at least, that equation (9) represents the normal aging survivorship for some organisms.

If the destruction of genetical information is the feature common to the action of the deleterious agents discussed in this article then the survivorship curve should be relatively insensitive to the character of the agent except insofar as reflected by the form of $J(\lambda)$. That such is indeed the case is shown

by the results obtained for x-ray and thermal killing of diploid yeast and other organisms.

T. H. WOOD (24, 25, 39) has reported data on the x-ray survival and thermal killing of yeast, since repeated and verified by URETZ (15). The data for diploid yeast are given in Fig. 4 and λ is the time at the indicated temperature or the

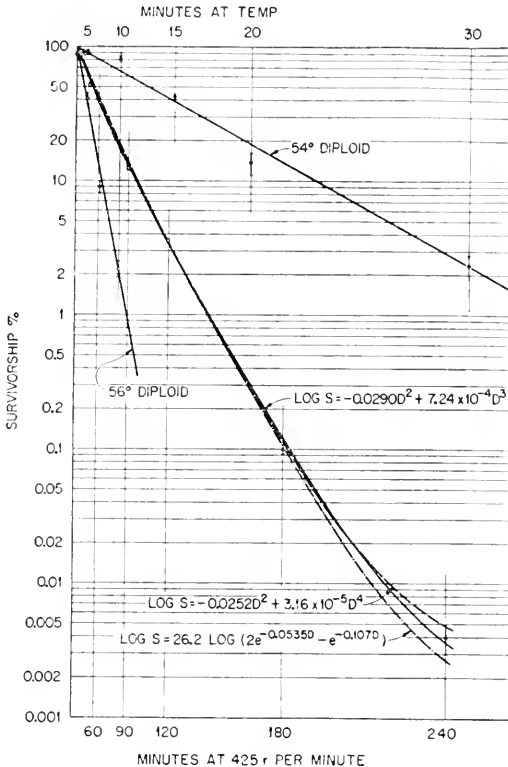


FIG. 4. Thermal and radiation killing of yeast from Wood (25, 39). Note that the abscissa is plotted as the square of time at temperature indicated or at 425 r per min. Straight lines drawn for comparison only.

x-ray dose as the case may be. The data have been fitted, using KIMBALL's method (40), to several curves and the results are shown in Table I.

The function $\log I/I_0 = n \log [2e^{-k_d\lambda} - e^{-2k_d\lambda}]$ was derived from hit-theory by LURIA and DULBECCO (41) and used by ZIRKLE and TOBIAS (20) and by WOOD (39). We have fitted WOOD's data allowing k_d to be determined internally by the diploid data and obtain the values shown. When the haploid value $nk_h = 2.49 \times 10^{-4}\, r^{-1}$ is imposed on the diploid data $n = 21.3$; $\chi^2 = 5.1$; $P > 0.5$. WOOD obtains graphically the slightly different values shown in Table I. It is clear from Fig. 4 and Table I that the term in λ^2 represents substantially, if not completely, the behavior of the survival curve as a function of dose. The fact that the fit may be made satisfactory by including a small term in either λ^3 or λ^4 supports this conclusion.

WOOD's survivorship curves for thermal killing in yeast (25) are also given in Fig. 4. The term in λ^2 again substantially represents the behavior of the survivorship curve. The curve retains its form when the temperature is changed. The χ^2 test shows in each case a very poor fit to λ^2 but this presumably reflects

Table I. Goodness of Fit for Wood's X-ray Survival of Diploid Yeast

Function	Constants	χ^2	P*
$\log I/I_0 = -a\lambda^2$	$a = 0.022$	60.	<0.001
$\log I/I_0 = -a\lambda^2 + b\lambda^3$	$a = 0.029$; $b = 7.2 \times 10^{-4}$	5.5	≈ 0.5
$\log I/I_0 = -a\lambda^2 + b\lambda^4$	$a = 0.025$; $b = 3.2 \times 10^{-5}$	8.3	≈ 0.2
$\log I/I_0 = n \log [2e^{-k_d\lambda} - e^{-2k_d\lambda}]$	$n = 26.3$; $nk_d = 2.20 \times 10^{-4}\ \mathrm{r}^{-1}$ (this paper) $n = 30$; $nk_d = 2.41 \times 10^{-4}\ \mathrm{r}^{-1}$ (from Wood (39))	3.9	>0.5

* 7 degrees of freedom

the existence of small higher order terms as in the X-ray case. Attention is also called to the aging of the grain beetle *Calandra oryzae* at 32.3°C and at 29.1°C shown in Fig. 2. The survivorship curve again retains the same shape, changing only the coefficient of λ^2.

The sensitization to thermal killing of *Paramecium caudatum* following X-irradiation was first reported by GIESE and HEATH (42). They found a slow recovery effect, requiring several days. This parallels the earlier discovery of GIESE and CROSSMAN (43) of sensitization to thermal killing by ultraviolet and by visible light in the presence of photodynamic dyes (44).

BALDWIN (45) has pointed out the similarities between thermal killing and killing by X-rays for the hymenopterous insect *Dahlbominus fuscipennis*. The immediate consequence of both insults is a coma from which the insect may recover to die later of delayed effects. Aging decreases the tolerance for both temperature and X-rays. The dose-survivorship curve is not given accurately but it has roughly the same shape for each agent. The diploid females are more resistant than the haploid males. These observations parallel those of WOOD on yeast.

It was mentioned briefly in the section on haploid organisms that NYBOM (27) has reported sigmoid X-ray survival curves for three species of green algae, *Chlamydomonas eugametas*, *C. Moewusii*, and *C. reinhardi*. JACOBSON (28) has studied *C. reinhardi* in some detail and shows that the X-ray survivorship curve fits accurately an equation of the form of equation (9). He points out that this can 'be explained by a redundancy of genetic information.' CLARK and HERR (46) irradiated the haploid male and diploid female of *Habrobracon*

uglandis at three stages of growth in air and nitrogen. Rough survivorship curves are given using eclosion as the criterion of survival. The haploid males apparently do not exhibit an exponential survival. This point should be studied further but it may be that the male haploid insects also exhibit a redundancy in the genetic information similar to *Chlamydomonas* in spite of their haploid character. The argument used to derive equation (9) applies in these cases as well as in the diploid case.

So far in the discussion it has been argued that deviations from the ideal form of the survivorship curves were due to the deleterious agent and $\rho(H, 0)$ was regarded as having the same form. The fact that many organisms, particularly hybrids, do not exhibit the survivorship curves corresponding to equation (9) is shown in Fig. 3. This behavior is closely associated with the genetic constitution of the organisms. There are a number of facts which support this conclusion.

Consider the survivorship curves of vestigial *Drosophila melanogaster* in Fig. 2 which differs from the wild type, whose survivorship curve is shown in Fig. 3 by a single gene. The same general effect has been reported by CLARKE and SMITH (47) for *Drosophila subobscura*. The hybrids between two inbred lines designated 'B' and 'K' exhibit a life span essentially double that of the parent inbred strains. The data are not sufficiently extensive to determine the mathematical form of the survivorship curve, but the inbred strains seem to have roughly the same type as the vestigial *Drosophila melanogaster* of PEARL and PARKER shown in Fig. 2 while the hybrid has the same form as the wild type shown in Fig. 3. This effect is also shown by mice. The survivorship curves for normal aging are given in Fig. 3 for two hybrid strains and for the hybrids of each with the C57 strain.

It is therefore a very plausible conclusion that the survivorship curve is very sensitive to the genetical character of the ensemble and that the change of shape can be ascribed to the form of $\rho(H, 0)$.

This function $\rho(H, 0)$ plays a role in information theory not unlike the equation of state in thermodynamics. We are at liberty to admit many types of probability distribution in H but it must be the same one in a given ensemble of organisms for all experiments. That this is the case is illustrated for mice by the resemblance between the survivorship curves for gamma- and x-irradiation and those for normal aging. The purposes of most of the work in this field, particularly in the case of acute killing, are served by obtaining an LD_{50}. The results are ordinarily reported by probit analysis and the life table is not given. We will not attempt to review the very extensive literature, which was not developed for the present purpose. Rather we will quote one experiment which involved a very large number of mice and which has been extensively studied and reported (48, 49, 50). In Fig. 5 the acute killing from atomic bomb radiation as a function of dose is shown from CRONKITE *et al.* (51), on LAF_1 mice. This is to be compared with the normal aging survivorship curve obtained from the controls. All curves on this figure are normalized by being passed through the 3 per cent survivorship point, after the custom of PEARL. The data of MURRAY and HOFFMAN (52) and MURRAY (53) giving normal aging life tables for hybrid and in-bred mice are also shown. The agreement, of course, is not exact but the curves for gamma-ray acute lethality

agree with the normal aging curves as well as these curves agree with each other.

There are several interesting details which should be pointed out. The gamma-ray data show a remarkable collinearity with aging data below the 10 to 20 per cent survivorship value but rise above the aging curve to a much sharper 'knee'. This effect is probably due to recovery, a phenomenon which is associated with radiation damage but not with aging. CRONKITE *et al.* noted

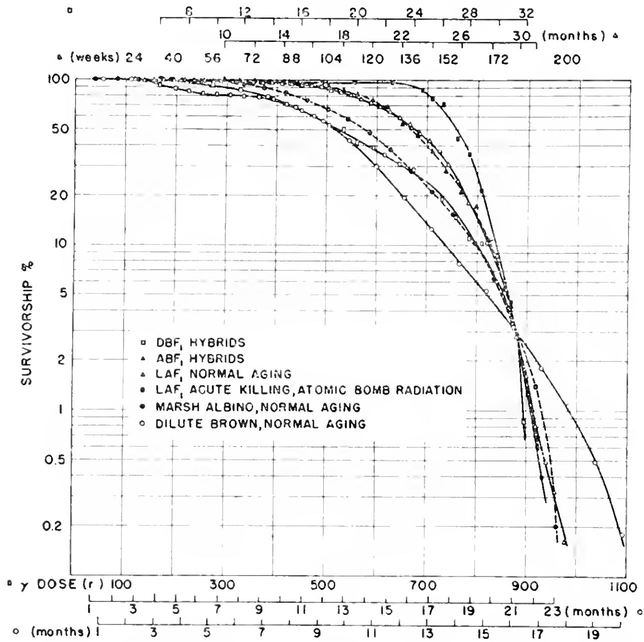


FIG. 5. Normalized survivorship of certain strains of mice for normal aging and acute radiation killing for LAF_1 (51, 52, 53). LAF_1 normal aging curve is due to A. C. Upton and A. W. Kimball, personal communication (see also (48)). Radiation dose and age are plotted linearly.

a small mortality below about 600 r and took pains to establish its reality as a radiation effect. This feature is also present in the aging curves—a bit more pronounced, to be sure, since there is no recovery. This feature of these curves, incidentally, fits very well with the present theory, or any theory which relates aging and radiation damage, but is otherwise quite a puzzle. These mice were presumed to be as nearly identical as possible and so should have been killed by nearly the same dose. (A discussion of a possible molecular basis for recovery has been given (10). This effect is left out of this article for simplicity.)

COLE, NOWELL, and ELLIS (54) have reported the late survivorship of LAF_1 mice protected from 800 r of 250 kVP x-rays by spleen homogenate. The survivorship curve has the same shape but is shifted so that the mean age at death is eight months earlier than the unirradiated control. This shows that even if a certain organ system, in this case the hematopoietic system, is so aided in repair that its role as a cause of death is greatly modified, nevertheless, the survivorship curve has the same shape.

The discussion given above indicates that even though it is necessary to assume $\rho(H, 0)$ of different shapes for different ensembles of organisms, once chosen, the same shape is required to represent survivorship data with little regard for the nature of the deleterious agent causing mortality.

The case for loss of information content by action of chemical mutagens or carcinogens is less clear than that for radiation. Radiation is no respecter of local chemical detail—it sees only an electron gas held together by positive charges. Side reactions complicate the experimental problems in obtaining good survivorship curves for the chemical radiomimetics. In fact, we know of no such curves at all, although some data of this sort are discussed by RAHN (3). Nevertheless, there is good reason to believe there is considerable miming (55, 56, 57) of radiation effects, in general, and of aging in particular.

This is illustrated by a paper by CLOUDMAN, HAMILTON, CLAYTON and BRUES (58). They studied malignant tumors and survival in CF-1 female mice painted with methylcholanthrene, irradiated with P^{32} β particles, and with these two insults in combination. A striking decrease in life span was found in those mice painted with methylcholanthrene. This was not due to any single pathologic state, not even to the pulmonary tumors generated in these animals. Survival was also shortened in those mice which did not have pulmonary tumors. The authors had the impression that life was shortened in a general way similar to the life shortening effects of total-body irradiation.

The carcinogenic effects of the two agents used in the experiments were approximately additive. This observation as well as the life shortening and the carcinogenesis correlates very well with the view that these effects are manifestations of the destruction of genetic information in the somatic cells. Since exposure to such chemicals is probably on the increase there is a practical as well as a theoretical reason for pursuing this matter further.

IV. THE ROLE OF EQUIVOCATION IN THE GERM LINE

It was said in my previous article in this volume that the ideas developed there should be applicable both to the germ line and to the somatic line. In this section we shall consider the effect of equivocation on the ability of the germ line to transmit specificity. We do not have available as much experimental material as that which pertains to the somatic line but there are several experiments which are very good and are very germane to the phase of information theory in biology discussed in this section.

It should be remembered that there are a number of error correction methods peculiar to the germ line. Among these are fertilization or conjugation and the selection value of the independent existence of cells in the germ line. The germ line may therefore be expected to exhibit a recovery from damage to a degree not found in the soma.

An experiment in which the germ line is propagated parthenogenically and so resembles very much the somatic line has been reported by LANSING (59, 60). He studied the effect of parental age on the survivorship of two species of the rotifer, namely, *Euchlanis triquetra*, which lives normally about a week, and *Philodina citrina*, which survives normally nearly a month.

The method of the experiment was to observe the survivorship curve for

a series of generations each produced from eggs laid on a given day in the life of the parent. LANSING called such a series an 'orthoclone'. An orthoclone obtained from a senile stage in the life of the rotifer was designated as an old orthoclone or a 'geriaclone' whereas an orthoclone from adolescent organisms was called a young orthoclone or a 'pediaclone'.

For each species it was found that the geriaclone could be followed to extinction in a few generations. In the case of *Philodina citrina* even the six-day orthoclone died out in the seventeenth generation. It was observed that the longevity of the five-day orthoclone tends to increase. The maximum life span of that orthoclone was not found but appeared to be indefinite.

It was found for each species of rotifer that the life shortening could be reversed by starting a pediaclone as an off shoot from a geriaclone. The limit to the ability to lengthen life seemed to be the fact that egg production does not appear until about the fifth day for *Philodina citrina* and about the fourth day for *Euchlanis triquetra*.

The number of animals used to establish a life table was sixty, a number too small to avoid considerable fluctuations. However, the curves shown in LANSING's papers give the impression that the shape of the survival curves is maintained. This feature seems to be in common with data discussed above in Section III, and in particular with the work of FURTH *et al.* (48), on the late effects of ionizing radiation on mice.

The decline and extinction of viability in the germ line is accounted for in the present theory by the accumulation of equivocation in the gene code as it is transmitted in the germ line. The recovery is regarded as being due to selection and propagation of that portion of the ensemble with a relatively low amount of equivocation.

The explanation offered by LANSING is quite different from that given here. He attributes his results to a transmissible factor which appears at cessation of growth. In particular, his assertion that the factor is non-genic appears to contradict the point of view adopted here. Actually there is no contradiction with the latter assertion since LANSING was undoubtedly thinking of genetic factors in terms of the ideas concerning the gene current at the time of writing, and indeed today. However, as LANSING notes, 'it is striking that the experimental observations on the primitive rotifer as well as conclusions derived therefrom are entirely compatible with conclusions drawn from mammalian experiments.' The feature of these and other organisms which is the same is the chemical composition of the genetic material and for this and other reasons it seems to me that an explanation for so ubiquitous a phenomenon as aging must be related to the genome.

The germ line provides an opportunity to study the error correction function of conjugation. The most extensive data relating damage in the germ line from one parent or both to survival seem to be due to MORTIMER (31, 61). He obtained survivorship curves for yeast zygotes formed by the conjugation of cells of opposite mating types. The following crosses were obtained: haploid × haploid (oo), haploid × diploid (oO), diploid × haploid (Oo), and diploid × diploid (OO). In the symbolism used a capital O represents a diploid cell, a lower case o represents a haploid cell; a filled letter (●o) irradiation. The first symbol indicates the α-mating, the second the *a*-mating type.

The survival curve for each haploid type in Fig. 1 is of the usual exponential form, equation (5). According to the discussion in Section II above, this is to be understood as the full expression of recessive lethal mutations. The survival curve for diploid exhibits the sigmoid shape whether the irradiation is done before or after conjugation (31). Note that the abscissa in Fig. 6 is

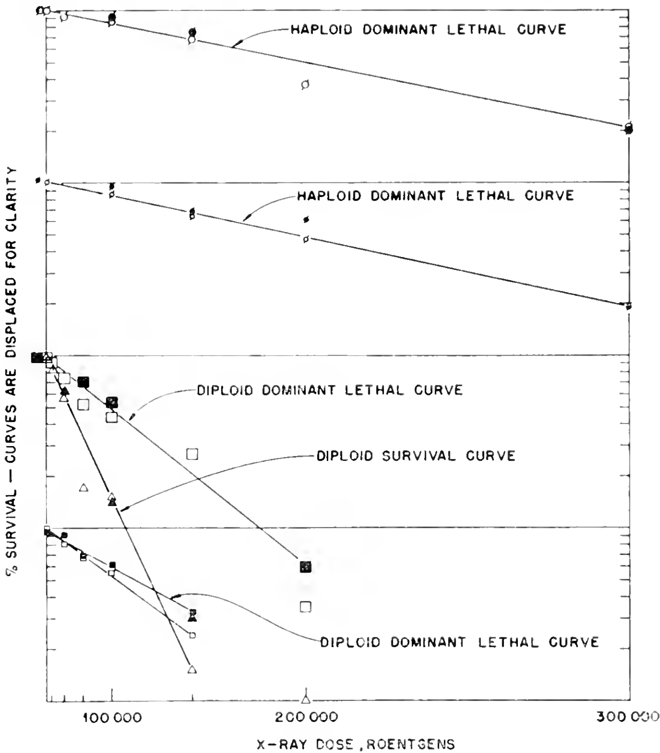


FIG. 6. Survivorship for yeast with one irradiated parent. Data from ref. (61). Haploid × haploid cross (oo), haploid × diploid cross (oO), etc. The first symbol represents a cell of the α-mating type, the second one of a-mating type. A filled letter o designates the irradiated parent. Haploid dominant lethal curve: o, ●o; ●, o●; O, ●O; O, O●. Diploid dominant lethal curve: □, ●O; ■, O●; □, ●o; ■, o●. Diploid survival curve: △, ●—αα diploid, ▲, ●—αα diploid. Haploid survival curve: ▲, ●—α haploid; ▲, ●—α haploid. Note that abscissa is the square of the dose.

the square of the dose. The straight lines have been drawn for comparison purposes only, but, as in Figs. 2, 3 and 4, it is clear that log I/I_0 is well represented by λ^2.

According to the discussion of Section III, this is to be understood as indicating that an error is not expressed in the diploid cell except when errors are paired in the two sets of chromosomes. This follows from the dependence of log I/I_0 on λ^2.

The shielding of errors by a normal allele is seen to be very effective in the (●o) or the (o●) case so that survival is very much greater than for the

irradiated diploid (●●). This shielding seems to be complete for errors due to first order damage; that is, damage such that $J(\lambda) = J_0$ in equation (6). If this shielding were not complete a term in the first power of λ would be apparent in Fig. 6.

Thus far the application of information theory to the currently accepted model of the diploid cell succeeds very well. We find features we expect and do not find ones we do not expect. OWEN and MORTIMER's data on the dominant lethal survival enable one to study second order effects ordinarily submerged in those of the first order discussed in the paragraphs above.

Figure 6 shows that there is a very small dominant lethal expression of damage but only when the damage is of the second order in the irradiated parent. That is, a single error or group of errors is shielded but pairs of errors or pairs of groups of errors are expressed, to some degree at least. This is evident since log I/I_0 behaves as λ^2. It is to be expected that this higher order damage exists in the haploid and in the diploid (●●) but cannot be observed because of the lethality due to first order damage.

The survivorship has the same λ^2 behavior for the other ploidies, but a curious feature is that this higher order damage is expressed to a greater degree in the higher ploidies, contrary to expectation. Perhaps this is a model-sensitive phenomenon (as higher order phenomena often are). If that is so further experimentation may tell us more about polyploidy.

It was pointed out in Section II that $J(\lambda)$ is related to the interaction of radiation and matter. This indicates that repeating OWEN and MORTIMER's experiments with other deleterious agents may be very fruitful. For example, URETZ (15) has shown that the ultraviolet survivorship of haploid yeast is sigmoidal. If this means in the case of haploid survival $J(\lambda) = J_1\lambda$, these errors will probably be shielded in the zygote. We choose the next higher term $J(\lambda) = J_2\lambda^2$ so that log I/I_0 may be expected to behave as λ^3. Higher powers in λ may be found in the expansion of $J(\lambda)$ depending on the effectiveness of shielding of recessive lethal mutations in the zygote.

These results should apply to organisms other than yeast and in particular to the survivorship of F_1 progeny in mice. F_1 progeny with one irradiated parent should have a shorter life span than the unirradiated parents. F_1 progeny with two irradiated parents should have a still shorter life span. That this is at least partly the case is shown by recently reported results by RUSSELL (34). He reports a life shortening in the offspring of male mice exposed to neutron irradiation from a nuclear detonation. The dose was rather low; the highest to the parent was 186 rep, but only two such offspring were obtained. Rather small numbers of individuals were obtained from other parents also so that the estimate of the magnitude of the effect is rough, although its existence seems to be established. The life shortening seemed to be of the same order of magnitude in the father as in the offspring, however.

WALLACE (62) has reported work on *Drosophila* in which he has irradiated several populations for as many as 150 generations. His criterion of viability is survival from egg to emergence and this work refers only to the second chromosome. He finds that the fitness of a population does not necessarily continue to decrease under the influence of radiation.

These experiments together can be understood from the point of view

developed in my previous article in this volume without *ad hoc* assumptions. Furthermore, certain interesting predictions can be made.

LANSING's remarkable work on the rotifer is a particularly interesting beginning to understanding the problems discussed in this article. If aging, thermal killing, and radiation damage are really aspects of the destruction of information content then there should be, as discussed above, a reciprocity between the respective agents. It would be particularly interesting to know if LANSING's results could be obtained by suitable x-, gamma- or ultraviolet-irradiation, or also by a thermal or chemical treatment. These organisms should be well adapted to this type of research.

Among the diploid organisms, of course, mice and *Drosophila* are of paramount importance. It would be extremely pertinent to look for the same reciprocity in this material. In addition, one should expect it to be possible, given a strain of one of these animals with a rectangular survivorship curve, see Fig. 5, to change it by irradiation to one of the type corresponding to equation (9) in several generations.

Acknowledgements—It is a pleasure to acknowledge the help of Dr. A. W. Kimball in matters concerning the statistics of data analysis. The analysis of T. H. Wood's data was greatly aided by his sending his exact experimental values. Thanks are due also to Dr. Arthur C. Upton and Dr. A. W. Kimball for furnishing the numerical data on the LAF_1 normal aging survivorship. I am very grateful to R. K. Mortimer for the use of his data.

REFERENCES

1. O. RAHN: The size of bacteria as the cause of the logarithmic order of death. *J. Gen. Physiol.* **13,** 179–205 (1929).
2. O. RAHN: The order of death of organisms larger than bacteria. *J. Gen. Physiol.* **14,** 315–337 (1930).
3. O. RAHN: Death of bacteria by chemical agents. *Biodynamica* **5,** 1–116 (1945).
4. D. E. LEA: *Actions of Radiations on Living Cells*, Cambridge University Press, Cambridge, England (1946).
5. M. R. ZELLE, and A. HOLLAENDER: Effects of radiation on bacteria. In: *Radiation Biology*, ed. by A. Hollaender, Chap. 10, Vol. II, McGraw-Hill, New York (1955).
6. A. HOLLAENDER (ed.): Vol. I, *Radiation biology*, McGraw-Hill, New York (1954).
7. A. KELNER, W. D. BELLAMY, G. E. STAPLETON, and M. R. ZELLE: Symposium on radiation effects on cells and bacteria. *Bact. Rev.* **19,** 22–44 (1955).
8. S. M. DANCOFF, and H. QUASTLER: The information content and error rate of living things. In: *Information Theory in Biology*, 263–270, ed. by H. Quastler, University of Illinois Press, Urbana (1953).
9. C. E. SHANNON, and W. WEAVER: *The Mathematical Theory of Communication*, University of Illinois Press, Urbana (1949).
10. H. P. YOCKEY: An application of information theory to the physics of tissue damage. *Radiat. Res.* **5,** 146–155 (1956).
11. T. ALPER: Hydrogen peroxide and the indirect effect of ionizing radiations. *Nature, Lond.* **162,** 615–616 (1948).
12. J. WATSON: The properties of x-ray inactivated bacteriophage. II Inactivation by indirect effects. *J. Bact.* **63,** 473–485 (1952).

13. T. ALPER: Bacteriophage as indicator in radiation chemistry. *Radiat. Res.* **2,** 119–134 (1955).
14. F. L. GATES: A study of the bacteriocidal action of ultraviolet light. *J. Gen. Physiol.* **13,** 231–260 (1929).
15. R. B. URETZ: Additivity of x-rays and ultraviolet light in the inactivation of haploid and diploid yeast. *Radiat. Res.* **2,** 240–252 (1955).
16. E. H. ANDERSON: The effect of oxygen on mutation by x-rays. *Proc. Nat. Acad. Sci., Wash.* **37,** 340–349 (1951).
17. A. HOLLAENDER, W. K. BAKER, and E. H. ANDERSON: Effect of oxygen tension and certain chemicals on the x-ray sensitivity of mutational production and survival. *Cold Spr. Harb. Symp. Quant. Biol.* **16,** 315–326 (1951).
18. A. HOLLAENDER, and G. E. STAPLETON: Fundamental aspects of radiation protection from a microbiological point of view. *Physiol. Rev.* **33,** 77–84 (1953).
19. G. E. STAPLETON, A. J. SBARA, and A. HOLLAENDER: Some nutritional aspects of bacterial recovery from ionizing radiations. *J. Bact.* **70,** 7–14 (1955).
20. R. E. ZIRKLE, and C. A. TOBIAS: Effects of ploidy and linear energy transfer on radiobiological survival curves. *Arch. Biochem. Biophys.* **47,** 282–305 (1953).
21. W. F. FRIEDEWALD, and R. S. ANDERSON: Factors influencing the inactivation of the rabbit papilloma virus by x-rays. *Proc. Soc. Exp. Biol., N.Y.* **45,** 713–715 (1940).
22. S. E. LURIA, and F. M. EXNER: The inactivation of bacteriophage by x-rays—influence of the medium. *Proc. Nat. Acad. Sci., Wash.* **27,** 370–375 (1941).
23. W. M. DALE: The effect of x-rays on enzymes. *Biochem. J.* **34,** 1367–1373 (1940).
W. M. DALE: The effect of x-rays on the conjugated protein d-amino acid oxidase. *Biochem. J.* 36, 80–85 (1942).
24. T. H. WOOD: Influence of temperature and phase state on x-ray sensitivity of yeast. *Arch. Biochem. Biophys.* **52,** 157–174 (1954).
25. T. H. WOOD: Lethal effects of high and low temperature on unicellular organisms. *Advances in Biological and Medical Physics* **4,** 119–165 (1956).
26. L. H. GRAY: Cellular radiobiology. *Annual Review of Nuclear Science* **6,** 353–422 (1956).
27. N. NYBOM: Some experiences from mutation experiments in *Chlamydomonas. Hereditas* **39,** 317–324 (1953).
28. B. S. JACOBSON: Reversible and irreversible effects of x-rays on vegetative survival in *Chlamydomonas reinhardi.* Ph.D. thesis, Univ. California, Berkeley (1957).
B. S. Jacobson: Evidence for recovery from x-ray damage in *Chlamydomonas. Radiat. Res.* **7,** 394–406, (1957).
29. C. TOBIAS, and B. STEPKA: University of California Radiation Laboratory Report No. 1922, (1952).
30. R. K. MORTIMER, and C. TOBIAS: Evidence for x-ray induced recessive lethal mutations in yeast. *Science* **118,** 517–518 (1953).
31. R. K. MORTIMER: Evidence for two types of x-ray lethal damage in *Saccharomyces cerevisiae. Radiat. Res.* **2,** 361–368 (1955).
32. G. A. SACHER: On the statistical nature of mortality, with especial reference to chronic radiation mortality. *Radiology* **67,** 250–257 (1956).
33. W. L. RUSSELL: Genetic effects in mammals. In: *Radiation Biology,* ed. by A. Hollaender, Chap. 12, Vol. I, McGraw-Hill, New York (1954).
34. W. L. RUSSELL: Shortening of life in the offspring of male mice exposed to neutron radiation from an atomic bomb. *Proc. Nat. Acad. Sci., Wash.* **43,** 324–329 (1957).
35. R. PEARL, and S. L. PARKER: Experimental studies on the duration of life, IX. New life table for *Drosophila. Amer. Nat.* **58,** 71–82 (1924).
36. P. H. LESLIE, and R. M. RANSON: The mortality, fertility and rate of natural increase of the vole (*Microtus agrestis*) as observed in the laboratory. *J. Anim. Ecol.* **9,** 27–52 (1940).

37. W. D. OLIFF: The mortality, fecundity and intrinsic rate of natural increase of the multi-mammate mouse, *Rattus (Mastomys) natalensis* (Smith) in the laboratory. *J. Anim. Ecol.* **22**, 217–226 (1953).
38. L. C. BIRCH: Experimental background to the study of the distribution and abundance of insects. *Ecology* **34**, 698–711 (1953).
39. T. H. WOOD: Reproducibility of x-ray survival curves for yeast cells. *Proc. Soc. Exp. Biol., N.Y.* **84**, 446–452 (1953).
40. A. W. KIMBALL: The fitting of multi-hit survival curves. *Biometrics* **9**, 201–211 (1953).
41. S. E. LURIA, and R. DULBECCO: Genetic recombinations leading to production of active bacteriophage from ultraviolet inactivated bacteriophage particles. *Genetics* **34**, 93–125 (1949).
42. A. C. GIESE, and D. C. HEATH: Sensitization to heat by x-rays. *J. Gen. Physiol.* **31**, 249–258 (1947).
43. A. C. GIESE, and E. B. CROSSMAN: The action spectrum for sensitization to heat with ultraviolet light. *J. Gen. Physiol.* **29**, 79–87 (1945).
44. A. C. GIESE, and E. B. CROSSMAN: Sensitization of cells to heat by visible light in presence of photodynamic dyes. *J. Gen. Physiol.* **29**, 193–202 (1946).
45. W. F. BALDWIN: Similarities in killing by heat and by x-radiation in the insect *Dahlbominus fuscipennis* (Zett.). *Radiat. Res.* **5**, 46–51 (1956).
46. A. M. CLARK, and E. B. HERR Jr.: The effect of certain gases on the radiosensitivity of *Habrobracon* during development. *Radiat. Res.* **2**, 538–543 (1955).
47. J. M. CLARKE, and M. J. SMITH: The genetics and cytology of *Drosophila subobscura*. XI Hybrid vigor and longevity. *J. Genet.* **53**, 172–180 (1955).
48. J. FURTH, A. C. UPTON, K. W. CHRISTENBERRY, W. H. BENEDICT, and J. MOSHMAN: Some late effects in mice of ionizing radiation from an experimental nuclear detonation. *Radiology* **63**, 562–570 (1954).
49. A. C. UPTON, and J. FURTH: Induction of pituitary tumors by means of ionizing irradiation. *Proc. Soc. Exp. Biol., N.Y.* **84**, 255–257 (1953).
50. A. C. UPTON, and J. FURTH: Spontaneous and radiation induced pituitary adenomas of mice. *J. Nat. Cancer Inst.* **15**, 1005–1015 (1955).
51. E. P. CRONKITE, V. P. BOND, W. H. CHAPMAN, and R. H. LEE: Biological effects of atomic bomb gamma radiation. *Science* **122**, 148–150 (1955).
52. W. S. MURRAY, and J. G. HOFFMAN: Physiological age as a basis for the comparison of strains of mice subject to spontaneous mammary carcinoma. *Cancer Res.* **1**, 298–307 (1941).
53. W. S. MURRAY: Studies on the inheritance of mammary carcinoma in the mouse. *Cancer Res.* **1**, 123–129 (1941).
54. L. J. COLE, P. C. NOWELL, and M. E. ELLIS: Incidence of neoplasms and other late lesions in mice protected against lethal x-ray doses by spleen homogenate. *J. Nat. Cancer Inst.* **17**, 435–445 (1956).
55. B. WALLACE: Dominant lethals and sex-linked lethals induced by nitrogen mustard. *Genetics* **36**, 364–373 (1951).
56. S. POMPER, and K. C. ATWOOD: Radiation studies on fungi. In: *Radiation Biology*, ed. by A. Hollaender, Chap. 11, Vol. II, McGraw-Hill, New York (1955).
57. O. G. FAHMY, and M. J. FAHMY: Chemical and radiation mutagenesis in *Drosophila melanogaster*. In: *Progress in Radiobiology*, 205-210. Proceedings of the Fourth International Conference on Radiobiology, Oliver and Boyd, Edinburgh and London (1956).
58. A. M. CLOUDMAN, K. A. HAMILTON, R. S. CLAYTON, A. M. BRUES: Effects of combined local treatment with radioactive and chemical carcinogens. *J. Nat. Cancer Inst.* **15**, 1077–1084 (1955).
59. A. I. LANSING: A transmissible, cumulative, and reversible factor in aging. *J. Geront.* **2**, 228–239 (1947).

60. A. I. LANSING: Evidence for aging as a consequence of growth cessation. *Proc. Nat. Acad. Sci., Wash.* **34,** 304–310 (1948).
61. M. E. OWEN and R. K. MORTIMER: Dominant lethality induced by x-rays in haploid and diploid *Saccharomyces cerevisiae. Nature, Lond.* **177**, 625–626 (1956).
62. B. WALLACE: Studies on irradiated populations of *Drosophila melanogaster.* In: *Progress in Radiobiology*, 201–204, Proceedings of the Fourth International Conference on Radiobiology, Oliver and Boyd, Edinburgh and London (1956).

ENTROPIC CONTRIBUTIONS TO MORTALITY AND AGING*

GEORGE A. SACHER
Division of Biological and Medical Research,
Argonne National Laboratory, Lemont, Illinois

Abstract—All dynamic physiologic processes are attended by fluctuations. The magnitude of these fluctuations is determined by the inherent regulatory capacity of the specific process and by the magnitude of random disturbances arising both in the environment and within the organism. A system with these characteristics has, in a given environment, a determinate probability of failure per unit time. As a consequence of the ubiquitous random component in physiologic performance, a population of individuals that are indistinguishable by any combination of physiologic measurements will nevertheless manifest time-survival and dosage-survival curves with finite dispersions. This is illustrated by means of a one-dimensional model system subjected to a stationary Gaussian random noise disturbance. In real biological populations, there is a component of variance between individuals. This can be taken into account by a straightforward generalization of the basic equations for homogeneous populations.

In this approach, aging is interpreted as a secular change in the values of the parameters of the regulatory mechanisms. These secular changes are ultimately due to irreversible changes in permanent or self-reproducing macromolecules. The rate of such irreversible change is in turn dependent in part on the magnitude of local fluctuations away from ideal steady state conditions for biochemical syntheses. There are thus two aspects to the stability of organisms—the probability of mortality per unit time and the rate of increase of this probability with time (age). Both are intimately dependent on the fluctuation characteristics of physiologic performances.

I. INTRODUCTION

THIS paper discusses mortality and aging insofar as they depend on certain statistical characteristics of organisms and populations. These characteristics, which may be subsumed under the closely related concepts of fluctuation, entropy, and information, have their origin in the dynamic nature of physiologic processes. Much of the current methodology for the analysis of survival curves is founded on the theory that the observed distributions of survival are due to the existence of a distribution of sensitivities in the populations tested. The present discussion is intended to emphasize the statistical nature of the mortality process *within* the individual, or in populations of indistinguishable individuals. Only those aspects of behavior are considered that have to do with the establishment and preservation of the steady state of physiologic function, and that can be described by a set of fixed relations among a finite, and in fact quite small, number of physiologic processes. Implicit in this approach is the conception of physiologic process as functional unit rather than as ultimate enzymatic reaction-step.

* Work performed under the auspices of the U.S. Atomic Energy Commission.

II. PHYSIOLOGIC REGULATIONS

The ability to maintain the physiologic steady state in the face of an unfavorable environment is called homeostasis (1). A number of quantitative indices of homeostatic capacity are in use. In ecological studies the tolerable range of an environmental variable, such as ambient temperature or salinity (of sea waters), is widely employed (2). Resistance to transient stresses is a more common measure in experimental physiology. If the response can be followed continuously, measures such as the amplitude of displacement of function, and rate of return to normal may be obtained. The above may be referred to as *determinate measures* of homeostatic capacity, for they reflect the fact

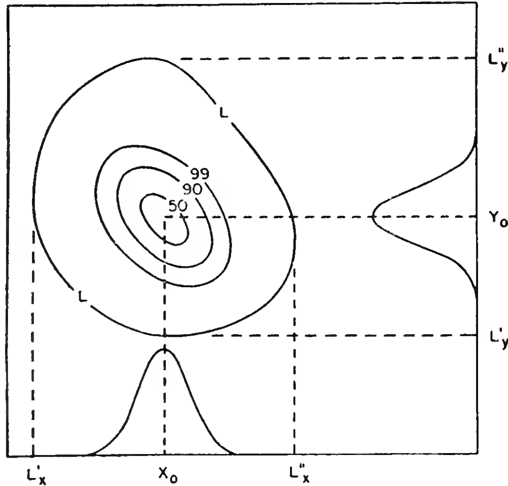


FIG. 1. Schematic representation in two dimensions of the probability distribution of physiologic states and their relation to the boundary delimiting viable from non-viable states. The probability distribution is indicated by elliptical contours of equal probability. The delimiting boundary (L–L), called the *lethal bound*, is indicated as a sharp line, and is treated as a precise value in the derivation of equation 15. A more realistic representation is given in Fig. 2.

that the homeostatic mechanism, even if it functions perfectly and without error, has but a finite regulatory capacity, set by the physical limitations of the mechanism.

In summary, there is a closed region in the physiologic configuration space within which some degree of stable physiologic function may persist, and beyond which stable function is impossible. This is indicated schematically in Fig. 1. The boundary surface of this region will be called the lethal bound, and denoted by L.

The quantitative properties of the lethal bound differ for different physiologic processes. In the case of white blood cells, the lethal bound on the low side, either for number of circulating cells or for number of proliferative cells, is only a small fraction of the normal level. Similarly the lethal range on the high side is considerably above normal levels. The boundary values for erythrocytes lie somewhat closer to the normal values. Blood glucose is rather more

sharply limiting on the low side than on the high. Blood pH must be held to very close tolerances on both sides of normal.

III. SOME PROPERTIES OF FLUCTUATIONS

From a consideration of the components of variation within and between individuals under different environmental conditions, it can be inferred that the observed variation can be attributed to (a) random fluctuations in the common environment that have a uniform influence on all animals maintained therein and (b) independent random fluctuations of each animal that must originate either within the animal or in local fluctuations of the environment that are independent for each animal. The magnitude of the fluctuation arising within the animal due to internal random noise is reasonably well known in a few experimental situations of a psychophysical or neurophysiological nature. Except for certain obvious aspects, such as temperature and humidity, the nature and properties of the environmental random variables is for the most part unknown. More significant perhaps than the purely environmental random variables are these that might be classified as organism-environment interactions. Such relations as pathogenicity, parasitism, dominance-submission, predator-prey relationships, etc., are in this class, and make contributions to the variability of individual performance that defy estimation. For the present purpose, the fact that intra-individual fluctuations exist is sufficient: the question of their nature can be deferred.

IV. FLUCTUATION AND THE PROBABILITY OF MORTALITY

The set of physiologic processes can be written formally as

$$X_i = \phi_i\,[\alpha_{ij}, A_{ij}, X_j, t] \qquad (i, j = 1, \ldots, n) \tag{1}$$

where the X_i denote the set of physiologic variables, the α_{ij} denote internal parameters, and the A_{ij} external parameters.

The *state* of an individual at a given moment is specified by the values at that moment of the n physiologic variables X_i. One can conceive of, and in principle construct, a population made up of indistinguishable individuals, in which the value of each internal parameter for every member lies within an arbitrarily small range. In such a population under constant environmental conditions the time-average of any function of the physiologic variables, X_i, for one individual is equal to the average over the population of the same function of the X_i at any moment in time.

If we locate a frequency distribution of physiologic states in the configuration space, the result is as seen in Fig. 1. The contours enclosing percentages of the distribution (such as 50, 90, 99) are drawn approximately in accord with the supposition that the bivariate distribution of states is Gaussian, and the situation is roughly in scale for a 'healthy' population, i.e. there is only a small probability of observing states near the lethal bound.

Since contact with the lethal bound removes an individual from the population, the distribution of states must be modified in the neighborhood of the boundary. Furthermore, the frequency distribution of states is not by itself

enough to permit a calculation of the probability per unit time that fluctuations will reach L. To answer these questions, we turn to a consideration of the dynamic nature of the fluctuation process within the individual. The complete description of a fluctuation process is given by specifying its correlation function, which is in one dimension

$$\rho(\tau) = \langle x(t)x(t+\tau)\rangle_{av}/\langle x^2(t)\rangle_{av} \tag{2}$$

where x is a deviation from the mean,

$$X = x_0 + x \tag{3}$$

The correlation function is a measure of the degree to which a fluctuation present at time t persists at a later time $\tau + t$, averaged over all values of t. The nature of $\rho(\tau)$ depends on the nature of the system. A process that obeys the differential equation

$$\frac{dx}{dt} + \beta x = 0 \tag{4}$$

returns to equilibrium as

$$x = x_0\, e^{-\beta t} \tag{5}$$

Corresponding to this, if a stationary pure random Gaussian noise source $f(t)$ is applied,

$$\frac{dx}{dt} + \beta x = f(t) \tag{6}$$

the resulting correlation function is

$$\rho(\tau) = e^{-\beta\tau} \tag{7}$$

It can be shown (3) that if the correlation function in one dimension is given by equation (7), then the fluctuation process is Markoffian and is completely described by the joint probability distribution

$$W_2(x_1 x_2 t) = \frac{1}{2\pi\sigma^2\,(1-\rho^2)^{\frac{1}{2}}} \times \exp\left[-\frac{1}{2\sigma^2(1-\rho^2)}\left\{x_1^2 + x_2^2 - 2\rho x_1 x_2\right\}\right] \tag{8}$$

This gives the joint distribution of observations of x separated by time t, where ρ is defined by equation (7). The variance, σ^2, of the distribution of x satisfies

$$\sigma^2 = D/\beta \tag{9}$$

where $4D$ is the (constant) *spectral density* of the random noise (white noise) source. The conditional probability distribution, which describes the distribution of x_2 when x_1 is fixed, is

$$P(x_0/x, t) = \frac{1}{[2\pi\sigma^2(1-\rho^2)]^{\frac{1}{2}}} \times \exp\left[-(x-\bar{x})^2/2\sigma^2(1-\rho^2)\right] \tag{10}$$

where σ^2 and ρ are defined as above and

$$\bar{x} = x_0 e^{-\beta t} \tag{11}$$

As $t \to \infty$, this becomes the stationary distribution of fluctuations,

$$P(x) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} \exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{12}$$

This distribution is Gaussian because of the linearity of mechanism specified by equation (4). Fluctuation processes are not in general Gaussian if the dynamical equations are non-linear.

Equation (12) gives the stationary frequency distribution of fluctuations over the entire x-axis. We now introduce the condition that there is at X_a an *absorbing barrier* at a distance λ from the mean

$$\lambda = X_a - x_0$$

An individual remains in the distribution only as long as the path described by his fluctuation process remains in the region $x < \lambda$. If this situation prevails for a time sufficiently longer than the relaxation time of fluctuations, $1/\beta$, a stationary distribution is again established and there will be a stationary probability q per unit time that the path will intersect $x = \lambda$. This 'absorption rate' is the mortality rate for the model fluctuation process. The stationary frequency distribution in the presence of an absorbing barrier may be obtained from equation (12) by the following argument.

In the steady state there is a stationary diffusion current, j, into the barrier. The desired frequency distribution $Q(x)$ must satisfy the steady state differential equation for diffusion in the presence of a force field (4). This equation is, in the notation of equations (6) and (9),

$$j = KQ(x) - \beta\sigma^2 \frac{\partial}{\partial x} Q(x) \tag{13}$$

where $K = -\beta x$ is the restoring force.

A solution of equation (13) satisfying the boundary condition

$$Q(x) = 0 \qquad (x = \lambda)$$

is

$$Q(x) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} \left[e^{-\frac{x^2}{2\sigma^2}} - e^{-\frac{(x-2\lambda)^2}{2\sigma^2}} \right] \tag{14}$$

The mortality rate, q, is equal to the diffusion current, j, normalized by the area under the distribution, $Q(x)$, so that from equation (13) we find the constant mortality rate to be

$$q = \frac{1}{\int Q(x)\, dx} \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \frac{\beta\lambda}{\sigma} e^{-\frac{\lambda^2}{2\sigma^2}} \tag{15}$$

A rigorous discussion of this class of stochastic processes (4) indicates that the validity of equation (15) is subject to the limitation

$$q \ll \beta$$

This restriction is met if we have $\lambda > 3\sigma$ (as is the case in the applications considered). The normalizing integral in equation (15) is then not appreciably less than unity, so equation (15) for the mortality rate reduces to

$$q = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \frac{\beta\lambda}{\sigma} e^{-\frac{\lambda^2}{2\sigma^2}}, \qquad (\lambda > 3\sigma) \tag{15a}$$

Equation (15) gives the dependence of the mortality rate on the parameters β, λ and σ (or β, λ and D) in the stationary state of a system specified by equation (6) and subject to a stationary random force function with spectral density $4D$. Although this model is too simple and artificial to be an adequate description of an actual mortality process, it should be noted that equations equivalent to equation (4) give an approximate description of a number of different physiologic mechanisms.

Equation (15) can be extended to the case of time-dependent mortality rates, as they are observed in animal populations, if the parameters are sufficiently slowly changing functions of time, so that stationariness of the fluctuation process is preserved. This is a reasonable assumption with regard to the life tables of animal populations in their normal environments. It is also considered for the purpose of this discussion that the fixed and the random components of environmental forces are stationary throughout life.

Experimental data on homeostatic capacity for a variety of mechanisms as a function of age indicate that this capacity diminishes during adult life (5). We therefore expect a steady decrease in the value of β. Since $\sigma^2 = D/\beta$, the value of σ will be increased by a decrease in β. The observed dispersion of physiologic variables does not increase markedly with age. This may imply that the recovery constant does not diminish much during the life span, but it may also be due to the effect of the distribution of parameters in the population, since it can be estimated that about half the total variance in a typical outbred population is variance between members, and this variance is reduced by selection, for as mortality proceeds in a heterogeneous population the subpopulations with the more disadvantageous parameter values will experience heavier mortality and thus be preferentially eliminated from the surviving population.

We have also examined (6) one simple mathematical model of a homeostatic mechanism that introduces a plausible type of non-linearity of recovery. In this model there arises a relation between the location of the mean state and the value of the recovery constant. In the notation used here, λ and β would decrease concomitantly.

The methodological difficulty in the study of mortality processes is that mortality data are not by themselves sufficient for the unique determination of their parameters, even in the simplest cases. In earlier treatments (7) the expedient was therefore adopted of assuming that the mean state, λ, is the only parameter that changes with age. There is abundant evidence that the mean values of physiologic variables change with age (8). Advantage was also taken of the fact that changes in mean physiologic state with age are usually small in degree. This justified taking the linear term of the expansion of λ^2 about the initial value λ_0,

$$\lambda^2 = \lambda_0^2 + 2\lambda_0 \,\Delta\lambda \tag{16}$$

Then taking logarithms and collecting the constant terms into lumped constants the relation was obtained

$$G(t) = \log q(t) = a + b\,\Delta\lambda \tag{17}$$

where $G(t)$ is called the Gompertzian.

The linear approximation to λ^2 is satisfactory for a first-order description of the relation of injury to age and to dosage of agents that cause permanent injury, such as x-rays (7).

The fact that β also tends to decrease with age does not alter the generalization made previously (7) that the Gompertzian is a linear measure of mean physiologic state. The entire exponent

$$\frac{\lambda^2}{2\sigma^2} = \frac{\lambda^2\beta}{2D}$$

can be expanded, yielding an expression of the form

$$\frac{\lambda^2}{2\sigma^2} = a_0 + a_1\,\Delta\lambda + a_2\,\Delta\beta \tag{18}$$

Furthermore if the mechanism depends on several variables, the same expansion procedure again yields an exponent term that is a linear function of the displacements of all of the parameters. Thus, *within the range of parameter values that occur in the course of natural aging*, the Gompertzian is an approximate linear measure of the mean physiologic state.

V. DESCRIPTION OF THE n-DIMENSIONAL FLUCTUATION PROCESS

The consideration of the general n-dimensional case will take as its starting point the empirical description of the n-variate process in terms of its moments. The observational data consist of a large number m of sets of observations on one individual or on m indistinguishable individuals, where each set is a measurement of each of n variables at a given time. The *first moments* are the n mean values,

$$x_{i0} = \frac{1}{m}\sum_{j=1}^{m} X_{ij} \tag{19}$$

The second central moments are the covariances

$$v_{ik} = \frac{\sum_{j=1}^{m} (X_{ij} - x_{i0})(X_{kj} - x_{k0})}{m} \tag{20}$$

The covariances are related to the standard deviations and correlation coefficients as

$$v_{ik} = \sigma_i\,\sigma_k\,\rho_{ik} \tag{21}$$

where σ_i are the standard deviations and ρ_{ik} is the *total correlation coefficient* between the ith and kth variables. The covariance matrix

$$\underline{\mathrm{V}} = \begin{Bmatrix} v_{11} & \cdots & v_{1n} \\ & \cdots & \\ v_{n1} & \cdots & v_{nn} \end{Bmatrix} \tag{22}$$

is a non-negative quadratic form, as is the correlation matrix

$$\underline{R} = \begin{Bmatrix} \rho_{11} & \cdots & \rho_{1n} \\ & \cdots & \\ \rho_{n1} & \cdots & \rho_{nn} \end{Bmatrix} \tag{23}$$

Given the covariance matrix $\underline{V}$, the frequency distribution of the displacements in n dimensions is determined. In the case that $\underline{V}$ is positive-definite, so that the rank is equal to the order n, the distribution is (9)

$$p(x_1, \cdots, x_n) = \frac{1}{(2\pi)^{n/2}(V)^{1/2}} \exp\left[-\frac{1}{2V}\sum_{i,k}^{n} V_{ik}x_ix_k\right] \tag{24}$$

where V is the determinant of $\underline{V}$ and V_{ik} is the cofactor of v_{ik} in $\underline{V}$. The coefficients, V_{ik}/V in the exponent of equation 24 are terms in the inverse of the covariance matrix,

$$\frac{1}{V}\{V_{ik}\} = \underline{V}^{-1} = \underline{\Lambda} = \{\lambda_{ik}\} \tag{25}$$

If $\underline{V}$ is positive semi-definite, the rank r is less than the order n, and the frequency distribution is an r-dimensional distribution in r independent linear functions of the x_i (9)

$$y_\lambda = \sum_{i=1}^{n} \alpha_{i\lambda}x_i \qquad (\lambda = 1, \cdots, r) \tag{26}$$

or

$$\underline{y} = \underline{Ax} \tag{27}$$

and the moment matrix for y is

$$\underline{M} = \underline{AVA'} \tag{28}$$

The frequency distribution for y is then

$$P(y_1, \cdots, y_r) = \frac{1}{(2\pi)^{n/2}(M)^{1/2}} \exp\left[-\frac{1}{2M}\sum_{j,\lambda}^{r} M_{j\lambda}y_jy_\lambda\right] \tag{29}$$

The case that the rank of the matrix of covariances is less than the order is frequently encountered in the initial description of biological systems in terms of the variables of direct observation.

VI. FLUCTUATION, ENTROPY, AND INFORMATION

The entropy of a system at equilibrium has, in classical thermodynamics, a precise value, $S_0(x_{10}, \cdots, x_{n0})$ where the x_{i0} are the values of the state variables at equilibrium. However, if thermal agitation or other disturbance causes small displacements of the state variables, the entropy decreases by an amount, ΔS. Expanding ΔS in a Taylor series in terms of the displacements, x_i (10), we obtain

$$\Delta S = \sum_i \frac{\partial S_0}{\partial x_{i0}} x_i + \frac{1}{2}\sum_{i,j} \frac{\partial^2 S_0}{\partial x_{i0}\partial x_{j0}} x_ix_j + \cdots \tag{30}$$

plus higher order partials. At equilibrium the first partial is equal to zero, so

$$\Delta S = \tfrac{1}{2}\sum_{i,j} \frac{\partial^2 S_0}{\partial x_{i0}\partial x_{j0}} x_i x_j \tag{31}$$

$$= -\tfrac{1}{2}\sum_{i,j} S_{ij}{}^0 x_i x_j \tag{32}$$

where $S_{ij}{}^0$ is a positive-definite matrix.
There is a formal equivalence between the $S_{ij}{}^0$ and the λ_{ij} defined by equation (24),

$$\underline{\mathrm{S}}^0 = k\underline{\Lambda} \tag{33}$$

where k is Boltzmann's constant. Thus the λ_{ij}, which we may term the partial coefficients of the frequency distribution of fluctuations, are proportional to the coefficients $S_{ij}{}^0$ of the quadratic form for the mean entropy decrease due to fluctuation in the system. The physiological systems that are under consideration and are not completely described by a small number of variables, and accordingly the complete fluctuation distribution and fluctuation entropy would not be estimated. However, the S_{ij}, or the λ_{ij}, are additive, so an initially incomplete description can be completed as knowledge of the system increases.

From the definition of entropy by Boltzmann

$$S = k \int p(x) \log p(x)\, dx \tag{34}$$

it follows that the fluctuation entropy coefficient $S_{ij}{}^0$ in equation (32) can be written

$$S_{ij}{}^0 = -\int \cdots \int \frac{\partial^2 \log p}{\partial x_{0i}\, \partial x_{0j}}\, p\, dx_1 \cdots dx_n \tag{35}$$

where $p = p(x_1, \cdots, x_n)$.

In one dimension this reduces to

$$S^0 = -\int \frac{\partial^2 \log p}{\partial x_0{}^2}\, p\, dx \tag{36}$$

This is identical with the definition of information given by R. A. Fisher (11). The equivalence continues to hold in the n-dimensional case. It should be noted, however, that the Fisher information is a defined quantity, whereas the $S_{ij}{}^0$ are terms in an approximation formula.

There is a close relationship between information theory and the analysis of fluctuation processes as can be made evident in terms of the equivalences brought out above. Where there are distinct classes the information is

$$H = -\Sigma\, p_i \log p_i \tag{37}$$

In the one-dimensional continuous case we write

$$H(x) = -\int p(x) \log p(x)\, dx \tag{38}$$

In a large number of cases, the representation of $\log p(x)$ in terms of three terms of a Taylor series is a good or even an exact description. The expression for the information then becomes

$$H(x) = -\int \left[p(x) \log p(x_0) + \tfrac{1}{2}x^2\, p(x)\, \frac{\partial^2}{\partial x_0{}^2} \log p(x_0) \right] dx \tag{39}$$

The information function is thereby resolved into separate terms for the expected values and for the deviations from expectation. The analysis of fluctuation processes falls into the latter class.

The formal equivalences between fluctuation entropy and Fisher information does not imply complete equivalence of the concepts. The theory of entropy fluctuations deals with the stationary fluctuation process in a single individual or in a group of indistinguishable individuals, where in either case the ergodicity condition is satisfied. There is no such restriction on the applicability of the Fisher information. The case of non-ergodic populations (individual differences in parameters) can be covered by obvious generalizations of the fluctuation theory, so this distinction is not a permanent one.

Determination of the Lethal Bound

Thus far in the presentation the existence of the lethal boundary surface has been a postulated property of physiologic mechanisms. In terms of the linear models of fluctuation processes that have been discussed the lethal bound is of necessity an arbitrarily assumed property, for a continuous linear process by its nature has no failure point. The escape from this unsatisfactory situation is by way of a more thorough mathematical analysis of homeostatic properties. The lethal boundary has a natural interpretation as a 'divide' on a potential surface (compare with Fig. 2). When it is possible to discuss the homeostatic processes as non-linear systems with multiple equilibria, the lethal bound, and also the boundaries between different viable steady states, will appear as necessary topological properties of the physiologic mechanisms. We have under way some investigations of simple non-linear stochastic mortality models, and the early results are quite interesting (6).

VII. ENTROPIC CONTRIBUTION TO THE AGING PROCESS

Brief consideration was given above to the direction of change of homeostatic parameters with age. This section will deal with the influence of physiologic fluctuation on the rate of aging.

It is an intuitive judgment that physiologic steady states of organisms tend to maximize the efficiency of physiologic function in the environments to which the organisms are fitted. The approach to greatest efficiency is presumably by means of natural selection operating on the genetically controllable thermodynamic properties of enzymes. The characteristics of physiologic performance, and in particular the values of the phenomenological rate constants are ultimately dependent on the activities, specificities and stabilities of the constituent enzymes. Thus, to give an account of the age changes in the values of the phenomenological parameters one must turn to the consideration of function at the biochemical level. The rate at which irreversible change occurs in a biological system will be discussed for three situations:

(a) as a function of temperature, independent of metabolic activity;
(b) as a function of metabolic activity in an undisturbed steady-state;
(c) in a steady-state disturbed by fluctuations much greater than thermal, i.e. by the fluctuations of physiologic state discussed above.

The analysis of irreversible molecular changes as a function of temperature

is a part of the general theory of absolute rate processes, and is also the object of a great deal of experimental work, particularly on proteins. It is discussed in another paper in this volume (12).

It has been suggested by a number of investigators that the rate of aging is a function of the level of metabolic activity. In evidence of this is the relation

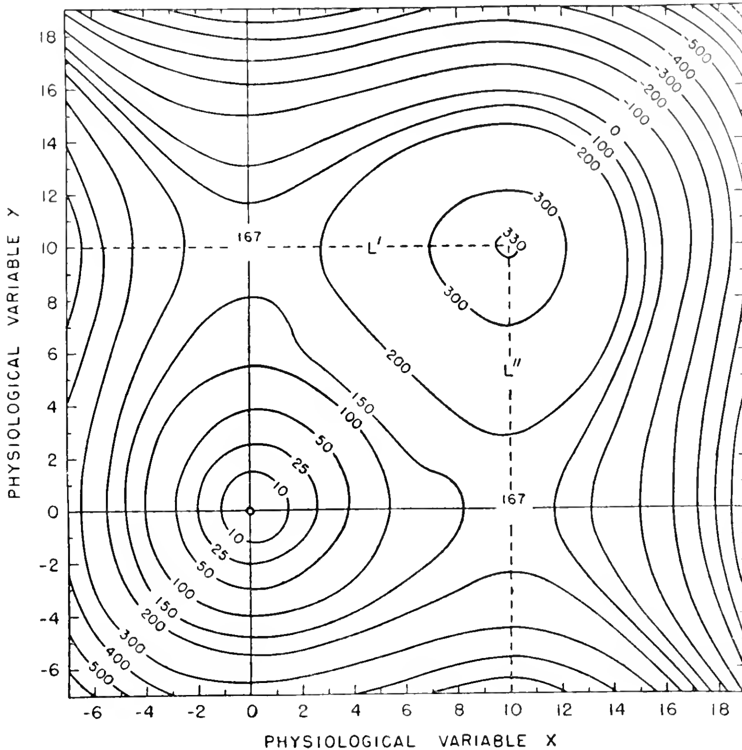


FIG. 2. Interpretation of the probability distribution of fluctuation and of the lethal bound in terms of a *potential* which is a function of the physiologic state variables (6). The solid curves are the isopotential contours. The dashed line L'–L'' is the *lethal bound.* This is the highest point ('divide') on the potential surface in any direction from the steady state, O. Different parts of the lethal bound can be at different potentials. If the potential is markedly lower in one part of the divide, most escapes will occur through this pass. Such preferential directions of escape may be identified with the occurrence of specific disease conditions. The contour lines are isopotential contours of the potential function

$$E = a_1x^2 - b_1x^3 + a_2y^2 - b_2y^3$$

A stochastic mortality process has already been investigated for the one-dimensional case of the cubic potential (6).

between metabolic rate (or body size) and life expectation among the Mammalia, and the dependence of life expectation on environmental temperature in cold-blooded forms. However, the relation is not a simple one. Birds, with body temperatures as much as 5°C higher than mammals (13), tend to have life

expectations (14) considerably greater than mammals (15) of equal body size and metabolic rate. Some primates (16) outlive carnivores or herbivores of equivalent body size by a considerable margin. This is obvious in the case of man compared with lower animals. In a study on grasshoppers, in which growth rate and life expectation were investigated as functions of temperature (17), the Arrhenius coefficient for growth rate was $\mu = 18{,}400$ cal and that for mean death rate was $\mu = 6300$ cal. Though both are temperature-dependent, the difference in μ values suggests that survival is not directly dependent on metabolic rate.

An association between the level of metabolic activity and the rate of accumulation of irreversible molecular change is certainly to be expected on physico-chemical grounds. The presence of poisons in the environment, and the ever-present possibility of incorrect reactions, imply the existence of a non-zero error rate per molecular event.

Finally we consider the influence of macroscopic fluctuations in physiologic state on the rate of accumulation of irreversible molecular changes. The calculation of the error rate due to fluctuation for a particular biochemical reaction would require a more detailed specification of the fluctuation process than is envisaged in the previous development, which dealt with a comparatively small number of important physiologic functions. This fluctuation in state of the organism as a whole would certainly play a part, but it would be necessary to consider in addition the independent fluctuations of small regions. These would usually have little immediate influence on the state of the whole organism, but they would be significant for the probabilities of irreversible change within the regions. The consideration of the problem of local fluctuations cannot be undertaken here.

It is presumed that the physiological steady state condition is one in which, through the action of natural selection, the ratio of incorrect to correct reactions is a minimum. This minimum rate is the metabolic error rate $\bar{\varepsilon}_M$, defined above. Deviations from the steady state in any direction bring about conditions in which the probability of incorrect reactions increases. This component of the error rate is called the fluctuation error rate, $\bar{\varepsilon}_F$. The fluctuation error rate would then in general be a monotone increasing function of the displacement, and the simplest assumption is that this function is a quadratic.
In one dimension this is

$$\varepsilon_F = mx^2 \tag{40}$$

where x is the displacement from the steady state. Then, for the one-dimensional model process discussed above, with stationary distribution of displacements given by equation (12), the mean error incidence per unit time is

$$\bar{\varepsilon}_F = \frac{m}{(2\pi\sigma^2)^{\frac{1}{2}}} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2\sigma^2}} dx \tag{41}$$

We find

$$\bar{\varepsilon}_F = m\sigma^2 \tag{42}$$

This is not a solution of the problem, for the evaluation of m cannot yet be carried out. However, the essential point for the present discussion is that the fluctuation error rate is an increasing function of m and of the dispersion of displacements, σ.

The mortality rate for the same model system,

$$q = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \frac{\beta\lambda}{\sigma} e^{-\frac{\lambda^2}{2\sigma^2}}, (\lambda > 3\sigma) \tag{15a}$$

is also an increasing function of σ^2, for the exponential factor in equation (15) varies much more strongly with σ than does the constant factor $\frac{1}{\sigma}$. The effect of accumulating errors will be to reduce λ and β and to increase σ (see above, Section IV). All of these changes tend to increase the mortality rate as age increases. Therefore it is concluded from this qualitative discussion that the *mortality rate* of different species, and the *rate of increase of mortality rate* with age for the same species are positively correlated. There are too many uncertainties to permit a statement of the functional relation between these quantities at present. However, we have here a possible basis for the relative constancy in the form of the life table for species as widely different as fruit fly, mouse and man.

The total error rate includes all three terms discussed above

$$\bar{\varepsilon}_{\text{total}} = \bar{\varepsilon}_T + \bar{\varepsilon}_M + \bar{\varepsilon}_F \tag{43}$$

where the subscripts denote temperature, metabolic rate and fluctuation, respectively. The existence of contributions to the error rate arising from background ionizing radiations and other environmental noxae must also be acknowledged. Perhaps the best viewpoint is that the physical basis for each term demonstrably exists, but we do not know the absolute contribution of any of them. This will be a major experimental problem.

All of these contributions arise when the environment and the population are in a steady state of fluctuation. The course of aging is also influenced to an important extent by very large disturbances that occur infrequently in the lifetime of the individual. Illness and crippling accident are examples, but changes of nutrition, etc., have equally important effects, as do also insults such as adventitious poisoning. The unique nature of these events requires that they be treated historically rather than on the basis of statistical uniformity of occurrence. Under experimental conditions it can be shown that exposure of a population to ionizing radiations leaves a permanent residue of injury (7). JONES (18) has demonstrated that human sub-populations selected on the basis of a history of given diseases have a permanent increase in their mortality at later ages. Some writers have attributed aging in general to the action of such major disturbances. Against this position it can be argued that the large common factor in the aging of human or animal populations points to an agency that acts with comparative uniformity on all members of the population and within each individual over the course of life. This is compatible with the statistical uniformities that appear in the summation of a large number of small independent events as proposed herein.

REFERENCES

1. W. B. CANNON: Organization and physiological homeostasis. *Physiol. Rev.* **9**, 399–431 (1929).
2. C. L. PROSSER: Physiological variation in animals. *Biol. Rev.* **30**, 229–262 (1955).

3. M. C. WANG and G. E. UHLENBECK: On the theory of Brownian motion II. *Rev. Mod. Phys.* **17**, 323–342 (1945).
4. S. CHANDRASEKHAR: Stochastic problems in physics and astronomy. *Rev. Mod. Phys.* **15**, 1–89 (1943).
5. W. B. CANNON: Aging of homeostatic mechanisms. In: *Problems of Aging*, 2nd ed. ed. by E. V. Cowdry, Williams and Wilkins, Baltimore (1939).
6. G. A. SACHER: Generation of a model dosage—survival curve from a stochastic hypothesis (abstract). *Radiat. Res.* **6**, 447 (1957).
7. G. A. SACHER: On the statistical nature of mortality, with especial reference to chronic radiation mortality. *Radiology* **67**, 250–257 (1956).
8. M. BURGER: *Altern und Krankheit*, 2nd ed., G. Thieme, Leipzig (1954).
   ed. by A. I. LANSING: *Problems of Aging*, 3rd ed. Williams and Wilkins, Baltimore (1952).
   L. BINET and F. BOURLIÈRE: *Précis de gérontologie*, Masson et Cie, Paris (1955).
9. H. CRAMER: *Mathematical Methods of Statistics*, Princeton University Press (1946).
10. P. EPSTEIN: *Textbook of Thermodynamics*, J. Wiley and Sons, New York (1937).
11. R. A. FISHER: Statistical information. In: *Information Theory in Biology*, 12–13, ed. by H. Quastler, University of Illinois Press, Urbana (1953).
    R. A. FISHER: Probability, likelihood and quantity of information in the logic of uncertain inference. *Proc. Roy. Soc.* (*A*) **146**, 1–8 (1934).
    M. G. KENDALL: *The Advanced Theory of Statistics*, Vol. II, Chas. Griffin and Co., London (1946).
12. H. P. YOCKEY: A study of aging, thermal killing and radiation damage by information theory. This volume.
13. S. BRODY: *Bioenergetics and Growth*, Reinhold Publishing Co., New York (1945).
14. S. S. FLOWER: Contribution to our knowledge of the duration of life in vertebrate animals. V. Mammals. *Proc. Zool. Soc. Lond.* **1931**, 145–234 (1931).
    F. BOURLIÉRE: *Natural History of Mammals*, Knopf, New York (1954).
15. S. S. FLOWER: Contributions to our knowledge of the duration of life in vertebrate animals. Birds. *Proc. Zool. Soc. Lond.*, **1925**, 1365–1422 (1925).
16. E. P. WALKER: *The Monkey Book*, Macmillan, New York (1954).
    R. M. YERKES and A. W. YERKES: *The Great Apes*, Yale University Press, New Haven (1929).
17. J. R. PARKER: *Bull. Univ. Mont. Agric. Res. Sta.*, no. 223 (1930).
18. H. B. JONES: A special consideration of the aging process, disease and life expectancy. In: *Advances in Biological and Medical Physics*, ed. by J. H. Lawrence and C. A. Tobias, **4**, 281–337, Academic Press, New York (1956).

A QUANTITATIVE DESCRIPTION OF LATENT INJURY FROM IONIZING RADIATION*

H. A. Blair

Department of Radiation Biology, University of Rochester School of Medicine and Dentistry

Abstract—A group of hypotheses previously discussed by the writer to account for the kinetics of radiation injury in mammals is reviewed. That radiation injury is proportional to dose, is partly irreversible, and that irreversible injury adds to new acute injury to produce lethality, appear to be valid. Recovery is not a single process for the whole animal but proceeds at different rates in different regions. The lethal threshold diminishes, presumably to zero, in old animals but not in proportion to life expectancy throughout adult life; rather it changes more slowly at first and then more rapidly. Irreversibility of injury differs with different radiations. With x- or gamma-rays it appears to be a similar fraction with doses smaller than about 100 r, but increases with larger brief single doses. The data in general are not sufficiently extensive and accurate to test hypotheses critically.

Over the past several years (1, 2, 3, 4, 5) I have discussed the adequacy of certain hypotheses to provide an empirical mathematical description of radiation injury and its effect on the duration of life. These hypotheses have been fairly successful in outlining a broad picture of radiation injury, in correlating many of the data and in suggesting critical experiments. It has become obvious, however, that they are deficient in some details and require amplification or revision. I propose at this time to discuss those changes in these hypotheses which appear to be necessary and also to point out some of the areas in which the data are inadequate to form the basis of quantitative correlations.

The hypotheses in question are as follows:

(a) The total injury produced by ionizing radiation is proportional to the dose.
(b) This injury is reparable in part and irreparable in part.
(c) Recovery from reparable injury occurs at a rate proportional to its magnitude.
(d) In consequence of (a) and (b), irreparable injury accumulates in proportion to total dose.
(e) Reparable and irreparable injury add in all proportions and death occurs when their sum attains a level which is proportional to the remaining life expectancy.

The injury defined here is a latent form observable at present only in terms of additional radiation dose. With acute exposures this injury has largely disappeared in most species before the clinical syndrome of radiation sickness has fully developed. There is presumably a quantitative causal relationship

* This paper is based on work performed under contract with the United States Atomic Energy Commission at the University of Rochester Atomic Energy Project, Rochester, New York.

between this latent injury and the clinical syndrome, but this has not yet been established. The advantages of developing non-lethal methods for detecting latent injury will be mentioned later.

It should also be remembered in what follows that the minimal lethal injury to an animal may not initially be manifest clinically at all, and that death occurs only after many days during which clinical signs develop. Because most mammals, if they are going to die from irradiation, do so within three or four weeks, it is customary to describe the lethal dose as that one which will kill one-half the experimental group within thirty days and to designate it LD_{50} or LD_{50} 30 days. The events which happen between the time of exposure, when presumably a lethal threshold for primary injury must be reached, or exceeded, if death is to occur, and actual death, are outside the scope of this discussion.

The LD_{50} for most mammals using whole body exposure is within the range of 400 to 800 roentgens for the young adult.

In accord with the above hypotheses, the rate of development of injury I under exposure at constant dose rate γ is

$$\frac{dI}{dt} = A\gamma - \beta(I - \alpha\gamma t) \qquad (1)$$

in which β is the rate of recovery per unit injury and A and α are constants.

Integration of equation (1) gives for the level of injury after exposure for time t

$$I = \frac{(A - \alpha)}{\beta} \gamma (1 - e^{-\beta t}) + \alpha\gamma t \qquad (2)$$

If the time of exposure is sufficiently short that no significant recovery occurs during exposure, as is usual in determining the acute median lethal dose or LD_{50}, $e^{-\beta t}$ may be replaced by $1 - \beta t$ so that equation (2) becomes

$$I = A\gamma t = A\alpha$$

α being the total dose.

If, now, α is the LD_{50}, the injury I is the lethal injury and according to postulate (e)

$$I = A\alpha = S_0 - S \qquad (3)$$

in which S_0 is the normal life expectancy of the animal and S is its age at radiation death. The constant of proportionality associated with $S_0 - S$ is taken arbitrarily as unity

For animals irradiated at daily constant rates for periods of some months $e^{-\beta t}$ may be neglected. This reduces equation (2) to

$$I = \frac{(A - \alpha)}{\beta} \gamma + \alpha\gamma t \qquad (4)$$

or, on using equation (3), to

$$\frac{S_0 - S}{\gamma} = \frac{A - \alpha}{\beta} + \alpha t \qquad (5)$$

Because nearly all chronic radiation experiments are begun on the young adult animal and also because postulate (e) cannot possibly be valid in very young animals in which the lethal dose rises instead of diminishes with age,

it is convenient to measure S_0 and S from the beginning of irradiation so that t in equation (5) is replaceable by S to give

$$\frac{S_0 - S}{\gamma} = \frac{A - \alpha}{\beta} + \alpha S \tag{6}$$

This equation represents existing (1, 2) data on chronic irradiation of mammals well within their possible errors. Such errors may be large in long-term experiments owing to infections and other accidents. An example of the fit is given in Fig. 1 for the data in Table I.

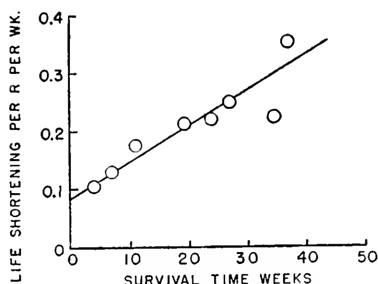


FIG. 1. Data by HENSHAW (17) on chronic irradiation of mice plotted according to equation (6). The data are given in Table 1. They cover a wider range than most. The scatter in such experiments frequently increases as the duration of the experiment gets long.

Table I. Data by HENSHAW (17) *on Mice Irradiated Chronically Five Days per Week until Death*

No. of animals	Daily dose (r)	Survival time (S) (weeks)	$S_0 - S$ observed	Total dose (r)	$\frac{S_0 - S}{\gamma}$	$\frac{\gamma(A-\alpha)}{\beta}$ (weeks)	$\alpha\sigma$ (weeks)	$S_0 - S$ calculated
10	0	45.8	0			0	0	0
15	5	37.0	8.8	925	.352	2	6.2	8.2
15	10	34.6	11.2	1730	.224	4.1	11.7	15.8
15	15	27	18.8	2025	.250	6.1	13.7	19.8
14	20	23.8	22.0	2380	.220	8.1	16.1	24.2
14	25	19.3	26.5	2410	.212	10.1	16.3	26.4
10	40	11	34.8	2200	.174	16.2	14.9	31.1
10	60	7	38.8	2100	.129	24.3	12.1	36.4
10	80	4	41.8	1600	.105	32.4	10.8	43.1

It is obvious now, however, that this equation should fail, providing all the other postulates are valid, because of the inaccuracy of postulate (e) even in the region of adult ages. Actually this postulate could have been written: 'LD_{50} diminishes in proportion to life expectancy.' Consequently it can be tested directly by measuring LD_{50} as a function of life expectancy.

In Fig. 2 are plotted LD_{50} data on Rochester rats (6) as a function of age. It will be seen that LD_{50} increases with age in young animals, is maximal in young adults, then declines slowly with age. As was mentioned above, postulate

(e) could possibly apply only to the adult stage. It should be noted that this curve is not convertible into a LD_{50}-life expectancy relation because, owing to mortality among the animals, the sample at each succeeding age is different from those going before. Those dying early have the shortest life expectancies and presumably the lowest LD_{50}'s, although this latter point cannot be proven directly.

In the legend to Fig. 2 are also given the days of life expectancy for the adult data only. These are fairly linear but do not extrapolate to $LD_{50} = 0$, when $S_0 - S = 0$, but to about 300 r. Presumably later points will diverge

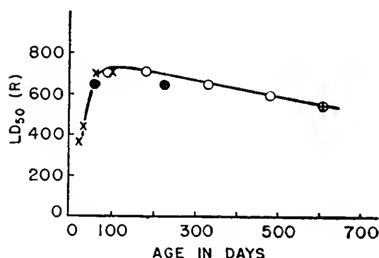


FIG. 2. Median lethal dose in roentgens as a function of age in rats of the Rochester strain (6). The animals at different ages are not directly comparable because, for example, of a group selected at 100 days, only about two-thirds survive to 500 days. The actual median survival times of control animals for the groups irradiated at 5, 11 and 16 months, respectively, are 450, 375 and 330 days after the time of irradiation. Therefore, life expectancy does not decrease as rapidly as the age of selection increases.

toward zero. Because it requires maintenance of animals for about three years to obtain data for a point at the advanced ages, it may be some time before the curves of Fig. 2 are well determined even in short-lived animals. However, GRAHN and associates (7) at Argonne National Laboratory have shown in mice that the lethal dose as measured by repeated daily doses decreases rapidly from middle age with an apparent tendency toward zero at old age.

At the present time it is not possible to state the situation more clearly than that the lethal threshold in the adult is some diminishing function of life expectancy, not a linear function throughout as required by postulate (e). This can be expressed also as

$$LD_{50} = F(S_0 - S) \tag{7}$$

and this as yet undetermined function should replace $S_0 - S$ in equation (6). However, there is considerable indirect evidence which will be discussed later that

$$LD_{50} = k(S_0 - S) \tag{8}$$

in fairly close approximation, k being a constant for values of $S_0 - S$ up to 20 per cent of S_0. It is important to establish the form of equation (7) in several species so that estimates can be made of variation of LD_{50} with age in man.

It is not clear whether equation (6) fits chronic data because equation (7) is sufficiently linear in the region in which most of the data lie (shortening of life span by one-half or less) or because of some other compensatory factor. In

any case putting k from equation (8) equal to unity, as is done in equation (6), may modify the constants A and α. This possibility should be considered when comparing the numerical values of these constants in different species and considering their absolute values.

The constants β and α/A of equation (1) and their variations with age, if any, can be determined directly. According to equation (1) the injury I, when exposure is stopped, should be repaired exponentially between its initial value and its irreversible residual. This repair was first studied in mammals by HAGEN and SIMMONS (8) using the rat. They assumed exponential repair to

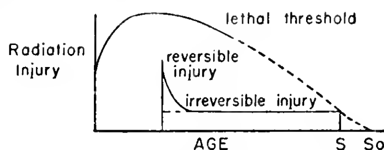


FIG. 3. A schematic representation of the LD_{50}, or lethal injury threshold, for an individual animal with a representation of injury from a single exposure. Data discussed in the text indicate that the irreversible injury is the same, and remains constant independent of the adult age at which it is laid down. This threshold curve cannot be measured directly owing to the change by death of the sample as the age of selection is made older and older. The related curve which is measurable is that of LD_{50} as a function of remaining life expectancy.

zero. Usually in making this determination a single substantial sub-lethal dose is given to a large group of animals followed by test doses to sub-groups at increasing intervals. Actually, the residual injury should be determined separately by testing one group after all repair presumably has taken place as illustrated in Fig. 3. The test doses less the residual in roentgens should then demonstrate simple exponential repair according to equation (1). However there is still another complicating factor in that it has been demonstrated recently that all parts of the animal do not recover at the same rate. CARSTEN and NOONAN (9), for example, have shown that following irradiation of the rat abdomen alone, recovery occurs with a half time between one and two days whereas it occurs in the animal with abdomen shielded with a half-time of about five days, and in the whole animal in about one week according to HAGEN and SIMMONS (8). Possibly, however, the strain used by Carsten and Noonan would demonstrate whole-body recovery at the same rate as abdomen-shielded recovery. In any case we are confronted with the fact that partial body recovery, at least for some tissues, is different from that for whole body. We do not yet know whether recovery of the parts is independent of whether or not other parts have been irradiated. It is fairly certain, however, that faster abdominal recovery can occur following whole-body irradiation, and probably accounts for the several observations (10, 11, 12) that recovery during the first day, as measured by whole-body test doses, is considerably faster than backward exponential extrapolation of later recovery. Strictly then equation (1), in some cases, if not in all, should be written with two or more constants β and the data analysed appropriately.

It will be seen that the existence of more than one constant, β, will not alter the form of equations (4) to (6) but the apparent value of β determined from

chronic data will have no exact counterpart in recovery measured directly by test doses.

According to the hypotheses, recovery should not exceed the irreversible component $\alpha\gamma t$. The data confirm that at least after some months following exposure the irreversible component is demonstrable as a decrease of LD_{50} and it is ultimately demonstrable as a decrease in life span. Nevertheless, there are some data on recovery showing that in the first few weeks test doses for lethality may attain or even exceed values for animals not previously irradiated. The natural conclusion from these data is that recovery may be complete or even more than complete in that an apparent tolerance to radiation is developed. Owing to a number of factors, the nature of this apparent transient complete or over-recovery is not clear. One of the factors is that if the experiments are done on young animals which have not attained maximal LD_{50} (Fig. 2), increase of LD_{50} during recovery will obviously make recovery appear greater than it really is. This defect may not be obviated by comparison with controls at each stage of the experiment, because LD_{50} may increase differently with age in the irradiated and control groups.

Another disturbing factor is that fast recovery of the abdominal region will make the earlier part of the recovery curve fall faster than is appropriate to the remainder of the body and the later part of the recovery curve will be lower, because, after the abdomen has recovered considerably, the dose required to kill will be greater than it would be if the whole body were recovering together. This factor will tend to obscure an irreversible remainder until all recovery has proceeded as far as it will.

Another possibility is that the animal may develop a transient physiological reaction to acute radiation injury which temporarily raises the lethal threshold for a second dose.

For all these reasons the irreversibility of radiation injury probably cannot be evaluated properly until at least several weeks after a substantial dose

The question of whether parameters in biological systems are age dependent should always be raised. In the case of recovery, for the reasons given above, evaluation of the constants or constant β is difficult by direct measurement. Nevertheless if the unanalysed recovery curve itself is similar at different ages this is an indication that the constants have not varied. HURSH and CASARETT (13) have shown in the rat that the recovery curve at 546 days (beyond middle age) is similar to that at 107 days (young adult). More study should be given this problem, but at present there is no indication that the rate of recovery is age-dependent.

The problems of whether irreversible injury is the same per unit injury at all ages, whether it slowly diminishes or increases, whether it gives rise to shortening of life because it is identical with ordinary aging or because it promotes ordinary aging, and whether it can be altered in any way, once laid down, are of considerable interest with respect to the setting of permissible levels for human exposure. If, for example, the irreversibility of radiation injury could be reduced the consequences of exposure would be reduced similarly.

Referring to Fig. 3 the indications at present, though far from complete, suggest that irreversible injury once laid down remains at constant level,

as depicted, until it intersects the curve of diminishing lethal threshold to cause the animal to die prematurely.

One indication of this has been obtained by BAXTER (14) in fruit flies. These flies normally live for about fifty days and lose half their life span if exposed to 75,000 r in a single dose. They die on day 26 approximately whether irradiated on day 1, day 25, or any day in between. Presumably recovery is very rapid in this species, and the irreversible component has the same effect when laid down at any time which is early enough in life to allow the whole potential life-shortening to be made manifest.

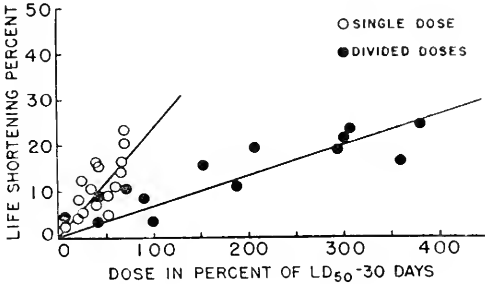


FIG. 4. Life shortening in per cent of normal span as a function of LD_{50} for rodents exposed to single doses or divided doses of x- or gamma-radiation. In the case of divided doses the radiation was stopped sufficiently long before death, or was at a sufficiently low daily level, that life shortening was caused only by irreversible injury, all, or nearly all, acute injury presumably having been repaired. The scatter of data is quite high for low divided doses, there being almost as many (omitted for simplicity) which show prolongation as shortening of life. The single dose curve rises more rapidly than linearly as LD_{50} is approached. The sources of the data are given in (16). LD_{50} is from 500 to 700 r for most of these strains. There is no established reason why data from different species should form a consistent pattern in this mode of plotting. They are less consistent than data on single strains.

Incomplete observations by HURSH and CASARETT (13) indicate that a given dose shortens life by about the same fraction of the normal expectancy in groups of rats exposed in early adult life or beyond middle age.

Direct measurements of residual injury as reduction in LD_{50} have been made no later than a few months after an initial dose. Such direct determinations when extended will be a more satisfactory test of the validity of the hypotheses depicted in Fig. 3 than the life-span data mentioned above.

Another factor to be discussed is whether the irreversible injury, or the constant α, is independent of dosage. It appears definitely to be larger with fast neutrons and alpha rays than with x- or gamma-rays (3). As measured by life-span shortening, it is also greater for single substantial doses of x- or gamma-rays than for divided doses even though all are delivered at the same dose rate in roentgens per minute.

Figure 4 shows the after effects of single and divided doses on life span in a number of strains of rats and mice. These data are plotted on the assumption that strains of different life spans and different LD_{50}'s will lose the same fraction of their life spans per unit dose measured in LD_{50}. Existing data

are not sufficiently accurate to decide whether this assumption is more correct than the one that the effects per roentgen are more similar in going from strain to strain or species to species.

It will be observed that according to Fig. 4 divided doses cause only about one-third the life-shortening per unit dose as that produced by single doses. Because some of the divided doses were given in increments as great as 120 r and because existing data are not sufficiently accurate to define small effects, it is probable that the curve for the smaller single doses coincides with that for multiple doses. It is certain, however, that substantial single doses such as 200 r (one-third LD_{50}) or more, have considerably more effect than the same dose in smaller increments.

The reason for this difference that immediately suggests itself, is that the irreversibility of the injury is some increasing function of its magnitude rather than the linear function assumed here. That this is not the correct explanation is indicated by the fact that repeated daily doses calculated to produce as much injury of the type defined here as a single substantial dose do not have the same effect on life span. There may be some unidentified dose dependent concomitant of injury which affects its reversibility. At this time, however, all that can be said is that α appears to be a constant independent of dose for doses of daily increments up to about 100 r but that it increases with dose with greater daily doses. That this larger effect of substantial single doses occurs at the time of irradiation and is not due to a dose dependent subsequent development is indicated by a single set of data (13). Such observations should be extended.

As predicted, the multiple dose curve of Fig. 4 is probably nearly linear. The single dose curve increases more rapidly than linearly if carried to higher doses than those depicted. This is to be expected because, according to the hypotheses, life shortening will be linear with dose only to the extent that the threshold curve of Fig. 3 is linear. As irreversible injury becomes substantial it will have more effect on life span per unit magnitude according to this curve.

Yockey (15) has postulated the identity of radiation damage with reduction of somatic genetic information, and has related the present formulation to the consequences of such damage in terms of information theory.

CONCLUSIONS

The hypotheses used appear to give a fairly accurate over-all description of radiation injury. The only one which is definitely known to be inaccurate is the last, which probably should be restated: Reparable and irreparable injury add in all proportions and death occurs when their sum attains a level which is some function, not fully determined, of the remaining life-expectancy.

Certain details, such as recovery rates, probably must be regarded as tissue- or region-specific rather than whole-body specific. This may also be true of irreversibility which has not been systematically studied in this regard. This latter problem is of particular interest with respect to human exposure, much of which, especially from internal emitters, is partial-body. However, even if each tissue, for complete description, requires a different set of constants A, α and β, this adds only complexity of detail and not of concept.

Irreversible radiation injury has the special property that it is closely related to premature aging abruptly laid down and probably persisting thereafter at a level constant or nearly so. This suggests the possibility that premature aging may be studied in young animals without waiting for them to die naturally. Of special interest is the possibility that irreversible injury may be prevented, at least in part, or altered once it has been laid down. This possibility should be studied in relation to exposure problems in man and also with respect to its bearing on natural aging. If, however, irreversible injury is wholly in the form of somatic mutations, as is often suggested, the possibility of altering it or its consequences would presumably be remote.

The acute injury described here in terms of radiation dose has antecedents in the form of disturbances of cellular structure and function from absorbed radiation and consequences in the form of the clinical syndrome of radiation sickness. Only the last stage has been at all well described in physiological terms, and the connections between the stages has not been elucidated at all. The ability to measure latent injury in terms of radiation dose should assist in deriving its description in biochemical or physiological terms. This is also true of irreversible injury.

Nearly all aspects of the long-term effects of radiation injury are markedly deficient in data, especially of those based on sufficient numbers of animals to be reasonably exact.

For this reason no formulation of the kinetics of the injury process can be adequately tested at present for its quantitative exactness. The virtue of a particular scheme is measurable rather in its ability to designate the phenomena involved, to make useful predictions and to serve as a basis for designing critical experiments.

REFERENCES

1. H. A. Blair: A formulation of the injury, life span, dose relations for ionizing radiations. I. Application to the mouse. *University of Rochester Report* UR-206 (1952).
2. H. A. Blair: A formulation of the injury, life span, dose relations for ionizing radiations. II. Application to the guinea pig, rat and dog. *University of Rochester Report* UR-207 (1952).
3. H. A. Blair: The shortening of life span by injected radium, polonium and plutonium. *University of Rochester Report* UR-274 (1953).
4. H. A. Blair: Recovery from radiation injury in mice and its effects on LD_{50} for durations of exposure up to several weeks. *University of Rochester Report* UR-312 (1954).
5. H. A. Blair: A formulation of the relation between radiation dose and shortening of life span. *Proc. Int. Conf. Peaceful Uses of Atomic Energy* **11**, 118–120, United Nations, New York (1956).
6. J. B. Hursh, and George W. Casarett: The lethal effect of acute x-irradiation on rats as a function of age. *Brit. J. Radiol.* **29**, 169–171 (1956).
7. D. Grahn: Personal communication.
8. C. W. Hagen Jr., and E. L. Simmons: Effects of total body irradiation on rats. Part 1, Lethal action of single, paired and periodic doses. *University of Chicago Metallurgical Laboratory Report* CH-3815 (1947).
9. A. L. Carsten, and T. R. Noonan: Determination of the recovery from lethal effects of lower body irradiation in rats. *University of Rochester Report* UR-455 (1956).
10. J. F. Thomson, and W. W. Tourtellotte: Effect of dose rate on LD_{50} of mice exposed to gamma radiation from cobalt 60 sources. *Amer. J. Roentgenol.* **69**, 826–829 (1953).

11. J. Storer: Personal communication.
12. J. P. Storaasli, S. Rosenberg, J. S. Krohmer, and H. L. Friedell: Effect of short interval fractionation on the lethal properties of total body radiation in rats. *Western Reserve University Atomic Energy Medical Research Project Report* NYO 4919 (1955).
13. J. B. Hursh, and G. W. Casarett: Unpublished observations.
14. R. Baxter: Unpublished observations.
15. H. P. Yockey: An application of information theory to the physics of tissue damage. *Radiat. Res.* **5**, 146-155 (1956).
16. H. A. Blair: Data pertaining to shortening of life span by ionizing radiation. *University of Rochester Report* UR 442 (1956).
17. P. S. Henshaw: Experimental roentgen injury. IV. Effects of repeated small doses of x-rays on blood picture, tissue morphology and life span in mice. *J. Nat. Cancer Inst.* **4**, 513–522 (1944).

SOME NOTES ON AGING

HARDIN B. JONES
Division of Medical Physics and Donner Laboratory
University of California, Berkeley, California

Abstract—Evidence of physiologic change with age uniformly points to a cumulative deterioration as age increases. Further degenerative change may occur proportionally to the amount of change already acquired. As age increases, incidence of degenerative disease and death increases exponentially. It is pointed out that, whatever aspect of body function is considered, e.g. functional members, metabolism, cellular activity, or blood flow characteristics, a relatively exponential increase in degeneration of body function occurs with increasing age of the individual. It is possible that each of these separately considered systems of aging is in partial equilibrium with the others, so that all general characteristics of change in functional vigor with time follow a similar course.

INCREMENTS of change in body structure and function occur as a phenomenon of aging. Usually, the term 'aging' is associated with deteriorative change, and as such is distinctively set off from those changes with age that are responsible for growth and development. However, even the period of development may be considered to have associated with every step some hazard that this step may not be achieved fully, thus adding an increment of imperfect function to the body. Such a deletion from full function, whether arising from genetic inheritance, developmental processes, or accidental mishap, may count just as much toward the accumulated deterioration we can manage to tolerate as does the deterioration of advanced age.

Experience of mishap accumulates throughout life. Some events, to be sure, have as little residual effect upon us as the whistle of the wind, but occasionally something of consequence occurs. As an example, it may be the crushing of a finger; although we usually recover, we can remember the event because of some persistent change—perhaps a scar, or a distortion of the nail, or even the loss of the finger.

Since, on the average, we live each day in a situation where there is some definite but slight chance that an event of misfortune may disturb us, then the longer we live under this average circumstance of risk, the more likely we are to find among us individuals showing physical impairment. Inspection of such a system leads to the probable conclusion that:

$$\text{Accumulated impairment} = \text{Mishap risk} \times \text{Time of exposure} \times \text{Fraction of function lost per mishap.}$$

But, the risk of occurrence of an unfavorable event is subject to increase as age increases, and the fraction of function lost per mishap may also increase as age increases. In this system, therefore, we can expect a relatively non-linear accumulation of average physical impairment as age increases; physical impairment may increase as some higher power of time lived than unity. This example

of a system contributing to aging reflects the exponential increase of morbidity and mortality that regularly is observed with increasing age.

There are other examples of impairment of body function that depend upon time lived and upon morbidity experience. A very general theory of impairment can be argued in which morbidity leads to morbidity, and mortality risk is some function of the integrated morbidity experience (1a). The following examples of functional disturbances may be cited to illustrate such relationship between morbidity and morbidity and between morbidity and mortality:

(a) The severity of toxic reaction usually increases more than proportionately to the poison dose.

(b) Radiation exposure induces ionization in tissues, and this morbidity in turn can induce morbidity proportional to the dosage. This holds both for acute effects and for life-span and carcinogenic changes.

(c) Risks of degenerative vascular disease are proportional to the extent of obesity (5).

(d) Risks of degenerative vascular disease are proportional to the disturbances of serum lipids in individuals followed over a segment of the adult life span.

(e) Death risks in diabetes throughout the past forty years have been undergoing a progressive reduction apparently proportional to the goodness of diabetic control.

(f) Dimming of primary senses (vision, touch, pain, and hearing) is associated with enhanced risks of trauma.

(g) Susceptibility to infectious disease is believed to be directly proportional to exposure intensity, and inversely proportional to defense mechanisms such as antibody levels and antibody generating capacity (1b); also, susceptibility to infectious disease can be quantitatively offset by administration of antibiotic agents.

(h) Proportional differences in death-rate risks among population samples throughout life span are related to sums of environmental and genetic factors.

Having noted examples of how morbidity and mortality risk can be dependent upon functional impairment, we can consider in greater detail evidence pointing to a widespread interdependence of physiologic systems. In vascular disease, occlusion may directly diminish blood flow in a small but critical segment of the body, as in coronary thrombosis. However, even though there is a measure of recovery from the acute episode, there may be a generalized insufficiency of circulatory function. Changes in blood flow caused by narrowing of the arterial channels may be expected to exact an effect upon function of the extremities, and DOBSON (2) has recently shown evidence for general dependence of the body's homeostatic mechanisms upon the proportional balance of regional blood flow. Thus, especially for the circulatory system, we can be certain that functional changes can influence the entire quality of body function.

A similar example of interdependence of disease is in the complications

of diabetes mellitus. This disease is not limited to the classic confines of its relationship to carbohydrate and intermediary metabolism: serious disturbances of lipid metabolism may also occur, linked with enhanced tendency for vascular changes; the term of pregnancy is frequently lengthened in diabetic mothers retinal changes may occur in diabetics, and the disease in general may be associated with somewhat early changes related to aging. There seems to be no reason to suspect that diabetes is a more complicated disease fundamentally than loss of islet-cell function or absence of insulin; but it does seem that the results of this functional deficiency can produce several different conditions that may even interact to compound the pathologic impact of the basic deficiency.

Another example of general disease being associated with a specific disease is observed in the follow-up of cancer patients. In cancer of the rectum, death from intercurrent disease may be just as likely as death from recurrence of the malignancy. There is also general evidence, from comparisons of mortality from disease in nineteen western countries, that high incidence of any one kind of disease is associated with high incidence of other diseases (1a). Some factors affecting adult health and life expectancy might be expected to be common to several kinds of overt disease; other factors influencing health may have a limited effect upon a single system. For example, in overweight individuals the increased risk of death is attributed to increased incidence of arterio-sclerosis and hypertensive disease, while the tendency toward cancer is not significantly changed from the average of the population. In radiation exposure, all major diseases may be enhanced. Leukemia, however, may be increased by a factor of 10, while other degenerative diseases are elevated less than twice. It is quite possible that some kinds of disease are less likely to occur following radiation exposure, even though the general trend is toward more severe and earlier degenerative disease following significant radiation exposure. Similarly, smoking generally enhances degenerative disease by a factor of 2 while lung cancer is increased tenfold. These observations point to the interrelationships in etiologic factors in disease, and the possibility that causative factors in development of degenerative disease may have interactions that accelerate the appearance and consequences of disease change.

Vascular Disease

Gofman and associates (3) have been able to show that the change in the wall of the artery in arteriosclerosis is essentially a linear thickening throughout aging. Thus, the shift toward occlusive change results from narrowing of a cylindrical tube by a progressive thickening of the mass, reducing the radius of the lumen. The function describing the reduction of blood flow in the artery involves the cross-sectional area of the artery, which is proportional to the square of the radius of the artery. Since blood flow in the artery is related to cross-sectional area, blood flow changes in arteriosclerosis are not proportional to time lived but rather vary as a power function of time. The fact that elasticity of the artery may fall off sharply as sclerotic thickening occurs probably accelerates the process. Thus, from several points of view, vascular change is not likely to produce a linear accumulation of disturbance with time lived, even though the basic feature of the disease is reasonably

established as a thickening of the artery wall proportional to lipoprotein elevation and duration of the condition.

Since vascular disease is a large component of degenerative disease and a contributing cause to other diseases, it is quite possible that exponentially-declining blood-flow capacity may in part determine the exponential pattern of increasing incidence of overt disease other than vascular disease.

Cancer and Aging

Throughout adult life, cancer incidence and cancer death rate are increasing exponentially; in most ways, this increase is remarkably similar to the above-described increase in heart disease tendency. ARMITAGE and DOLL (4) have ascribed this property to the fact that a succession of small changes necessarily precedes cancer. It is of interest to construct population samples of individuals known to have died of a given kind of cancer. When this is done, the increase in death rate in the cancer-destined population is remarkably like the increase in incidence of cancer in the population out of which it was taken (1a). Thus, we can be reasonably certain that the risk of cancer is increasing exponentially with age.

In contrast to the exponentially-increasing incidence of cancer with increasing age, individuals identified as having overt cancer have a constant death risk approximately independent of chronologic age. Therefore, it is a reasonable argument that changes characterizing the period prior to onset of cancer may be of many different kinds, each making cancer occurrence more likely, but the change representing incidence of cancer effects a single abrupt decrease in life expectancy.

It follows from this reasoning that many of the changes that accompany aging may be of consequence only as they allow a drastic and irreversible change into overt disease to take place. In vascular disease, the average degenerative change in the walls of the artery is of less consequence than the infarctions or vascular occlusive episodes that destroy peripheral tissue. Death may occur as a consequence of a random occlusive event, even though average changes in the arterial structure may be minimal.

Cellular Change and Aging

Cancer is usually considered to be an example of cellular change associated with aging, very possibly upon a basis of somatic mutational change. It should be noted that evidence for this is based upon an incidence of cancer exponentially increasing with age. While I, too, subscribe to this view, a similar phenomenon is seen in diabetes mellitus, a disease of deletion of function. It is quite possible that, in addition to changes in the quality of cells surviving with time, certain kinds of cells may survive aging with different likelihood. SHOCK (6) has evidence, for example, for a decline both in functional quality and numbers of cells in the kidney with age. It is reasonable to explore further the effects of declining numbers of cells with age. Instances, as in the case of disappearance of islet tissue in diabetes, may be observed in various tissues and are represented by epilation, appearance of channels in the fingernails, and disappearance of cells supplying sensory function of various kinds. These cells may disappear, but we do not know why.

In radiation effect, radiation exposure is related directly to enhancement of degenerative change, thus simulating the effect of aging. The similarity may be due in part to the random destruction of cells and partly to the alteration of function of cells. Within certain cells such as the marrow and the lymphatic tissue, or in embryologic development, radiation over a wide range of dose and for several species of mammals destroys approximately three cells out of every 1000 cells per roentgen of whole-body exposure (7c). At less than lethal exposures, this random destruction of cells proportionally to radiation exposure does not have a lasting effect upon the blood-forming tissues, since these cells rapidly regenerate. However, the average lethal dose of whole-body radiation exposure is estimated to involve a 50 per cent reduction in these cells. Somewhat the same changes occur in other body cells, the degree being dependent upon radiation sensitivity. (Some cells are known to be much more resistant to radiation than blood-forming cells.) The effects of radiation in diminishing the numbers of cells also seem to be about the same upon mammalian germinal cells as on blood-forming tissues; in both tissues, approximately two to three cells are affected per 1000 cells per roentgen. Thus it appears that each roentgen of exposure to tissues like the blood-forming system, the gonads, and the developing embryo may have about equal probability of either killing the cell directly or altering its chromosomal structure if it survives. Such changes may be suspected as having a role in inducing age change in the somatic tissues.

Leukemia induction by radiation, as evidenced from the analysis of Court-Brown and Doll and others (7a,b,c,d), is increased proportionally to radiation exposure. These changes are such that approximately 50 r of whole-body exposure produces a frequency of leukemia equal to its natural incidence and the effect is proportional to dose over a wide range. This is in remarkable agreement with genetic change in mammals; here, too, an exposure of 50 r produces approximately the same number of mutations as occur naturally in one generation.

REFERENCES

1a. H. B. Jones: A special consideration of the aging process, disease, and life expectancy. In: *Advances in Biological and Medical Physics*, **4**, 281–337, ed. by J. H. Lawerence and C. A. Tobias, Academic Press, Inc., New York (1956).

1b. G. A. Sacher: Reference to aging and antibody levels. Donner Laboratory Lecture, University of California, Feb., 1957.

2. E. L. Dobson: Homeostatic regulation of body fluid volumes: Role of a blood flow dependent control mechanism. *Fed. Proc.* **16,** 31 (1957).

3. J. W. Gofman: Serum lipoproteins and the evaluation of atherosclerosis. In: *Experimental methods for the evaluation of drugs in various diseases. Ann. N. Y. Acad. Sci.* **64,** Art. 4, 590–595 (1956).

4. P. Armitage and R. Doll: The age distribution of cancer and a multistage theory of carcinogenesis. *Brit. J. Cancer* **8,** 1–12 (1954).

5. L. I. Dublin and H. A. Marks: Mortality among insured overweights in recent years. 6th Annual Meeting, Association of Life Insurance Medical Directors. Metropolitan Life Insurance Company, October 1951.

6. N. Shock: Physiological aspects of aging. *Symposium on Aging*, Amer. Assoc. Advanc. Sci. Meeting, New York, Dec. 1956.

7a. W. M. Court-Brown and R. Doll: The incidence of leukaemia among the survivors of the atomic bomb explosion at Hiroshima and Nagasaki. In: *The Hazards to Man of Nuclear and Allied Radiations*, 84–87, Medical Research Council, London. (1956).

7b. H. B. JONES: A summary and evaluation of the problem with reference to humans of radioactive fallout from nuclear detonations. In: *Hearings before the Special Subcommittee on Radiation of the Joint Committee on Atomic Energy, Congress of the U.S., June 4–7, 1957*, 1116–1136 (also discussion, 1100–1108), Government Printing Office, Washington (1957).

7c. H. B. JONES: Estimation of effect of radiation upon human health and life span. *Proc. Health Physics Soc.* **1**, 114–126 (1956).

7d. E. B. LEWIS: Leukemia and ionizing radiation. *Science* **125**, 965–974 (1957).

CANCER AS A SPECIAL CASE OF A GENERAL DEGENERATIVE PROCESS*

HARRY AUERBACH
Division of Biological and Medical Research, Argonne National Laboratory, Lemont, Illinois

Abstract—Death rates or life table q_x values for populations throughout the world tend to exhibit a sixth power linear relationship with age when plotted on a log-log basis.

It is shown that the total deaths can be broadly separated into chronic degenerative causes and acute causes, with death from the degenerative causes increasing as the sixth power of age and death from the acute causes increasing in simple exponential fashion.

IN many cancer studies at this laboratory and elsewhere, attempts have been made to discover the underlying mechanism by which tumors come into existence. While a great many of these studies have been directed toward describing the process of carcinogenesis in biological terms, the statistical approach has also been productive. A recent study at Argonne National Laboratory, using analysis of vital statistics, indicates that cancer has some characteristics in common with the degenerative diseases.

The study was suggested by observations of others (1, 2, 3) that when the logarithm of death rate from cancer (either the total or that involving a specific site) was plotted against the logarithm of age at death, the result was usually a straight line. The slope of the line indicated a sixth power relationship, a fact that has been used to support several theories of carcinogenesis. Another interesting possibility—that the same linear relationship might be present in other causes of death—was recognized and investigated in the present study.

The question was first examined by analyzing the United States death rates for the years 1949–1951. Plots were made on the same log-log basis for several broad groups of causes of death. Five groups (circulatory system, malignant neoplasms, nervous system and sense organs, respiratory system, and genito-urinary system) showed a relationship of approximately the sixth power of age to a marked degree for age thirty and older, with departures from linearity being restricted to ages under thirty. The sum of these groups gave an almost perfect linear relationship from the age of thirty upwards (Fig. 1). These five groups represent the overwhelming majority of the chronic degenerative causes of death. The remaining three (infective and parasitic diseases, digestive system, and accidents) which did not show the linear relationship, represent the acute causes of death.

In order to find out whether the same situation obtained in other countries, a slightly different method had to be used. Life table data, which are available for most of the countries of the world, were used in the absence of reliable specific cause death rate statistics. The value used was q_x, the proportion of

* Work performed under the auspices of the U.S. Atomic Energy Commission.

persons alive at the beginning of the year of the specified age, who die during that year of all causes. The countries tested were United States, Canada, Israel (Jewish population), India, Union of South Africa (Asian population), Brazil, Japan, Portugal, Belgian Congo (African population), Costa Rica, El Salvador, Argentina, Ceylon, Finland, France and Norway. Similar sixth power relationships were exhibited by all. Deviations from linearity at

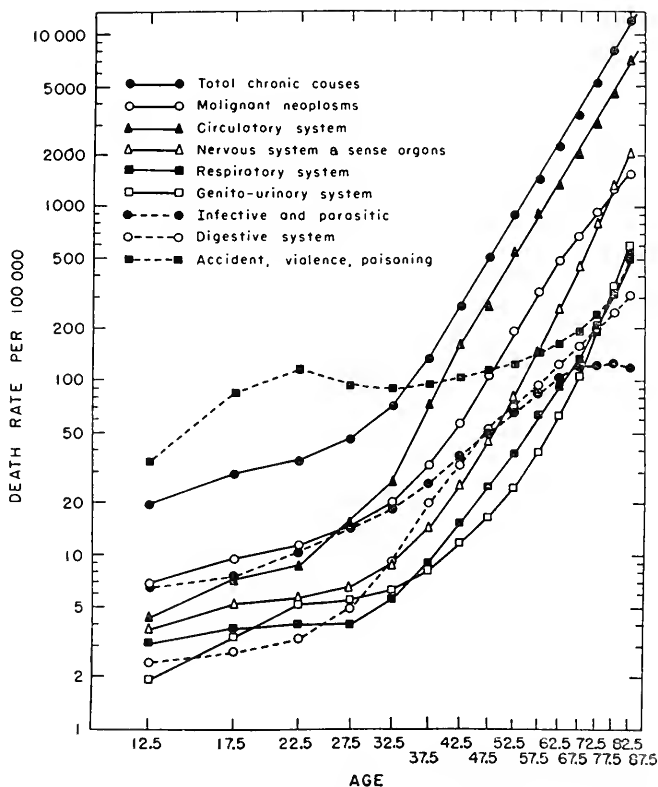


FIG. 1. Log-log plots showing the relationship of death rate to age in United States white males (1949–1951) for broad groups of causes of death. Solid lines, chronic causes; dotted lines, acute causes.

younger ages were always in the direction of the actual figures being higher than the extrapolated values.

On the basis of the demonstrated log-log sixth power linear plot of the chronic degenerative diseases, and the nonlinearity of the acute causes of death, an attempt was made to determine the relationship of acute causes to chronic degenerative causes in the total death rate. Life table values of q_x for United States white males for three periods, 1900–1902, 1929–1931, and 1949–1951, were plotted against age on the log-log basis. The usual departures from linearity at earlier ages were marked in the 1900–1902 period, less so in the 1929–1931 period, and still less in the 1949–1951 period, but all three curves tended to merge into a common straight sixth power line at the age of forty

and older (Fig. 2). This merging of the three plots at the age of forty and older demonstrates a well-known fact, that chronic degenerative diseases become most important as cause of death at the older ages; the values show essentially no change over the period from 1900–1951. On the other hand, at the earlier ages, the acute causes constitute almost 100 per cent of the total death rate, and chronic degenerative diseases represent only a minor fraction. Even though

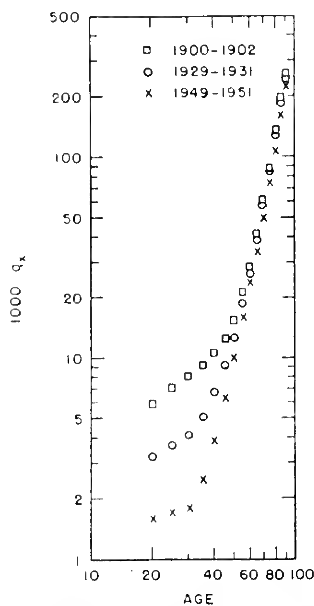


FIG. 2. Log-log plot of $1000 q_x$ against age for United States white males for 1900–1902, 1929–1931, and 1949–1951. Lines are omitted in order to show more clearly the good fit of the three sets of points on the straight portion of the curve.

it is apparent that acute causes have been decreasing steadily over the fifty-year period from 1900–1951, their presence is evident as departures from the straight line in the plots, with highest values at 1900–1902 and lowest at 1949–1951. The equation of the straight line for the forty years and older group was calculated on the assumption that the 1949–1951 values represented the least effect of acute causes of death on the over-all figures. This was regarded as giving q_x values for the degenerative causes of death which were then subtracted from the total q_x values for other countries. The resulting numbers were assumed to represent death rate attributable to acute causes of death.

The q_x's associated with these acute causes of death were found to plot as a simple exponential increase with age. The slopes were equal for the countries tested, but the intercepts were different and correlated in a general way with the levels of public health and medical care in the country concerned. Plots for four of the countries are shown in Fig. 3A. When the sum of the five major groups of degenerative diseases was similarly subtracted from the total causes of death in United States white males 1949–1951, leaving a residue

representing the acute causes of death, a similar exponential increase with age was apparent (Fig. 3B).

It therefore appears that the death process in man can be separated broadly

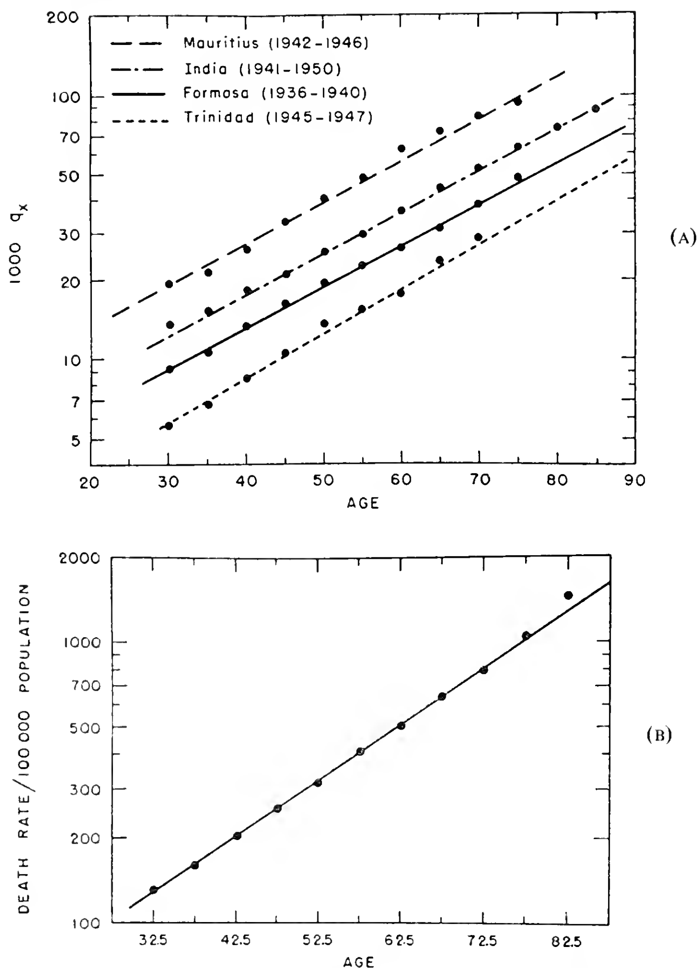


FIG. 3. A. Typical semi-log plots of acute portions of $1000\,q_x$ (total minus chronic), for four of the countries tested. B. Semi-log plot of death rate per 100,000 population from acute causes of death (total minus chronic), for United States white males, 1949–1951.

into chronic degenerative causes and acute causes, with death from the degenerative causes increasing as the sixth power of age and death from the acute causes increasing in simple exponential fashion.

Total and specific causes of death have previously been fitted by the Gompertzian or semi-log plot (4). However, it would appear from this work that the degenerative causes increase with age not at a constant rate of increase as predicted by the Gompertzian, and are therefore better fitted by the log-log

plot. On the other hand, the acute causes of death clearly are fitted by the Gompertzian function.

The presence of the sixth power relationship in a large number of different situations suggests a general underlying principle. Since we have no knowledge whatever of what this principle is in biological terms, we can only speculate that it could be a very general organizational scheme which provides about five redundant elements within each essential unit. An element might be a molecule, a cell or organelle (internal structural and functional unit of a cell), a group of cells, or a whole organ. The essential units might be separate or overlapping. Carcinogenesis may be a special case of this general degenerative process.

REFERENCES

1. J. C. FISHER and J. H. HOLLOMAN: A hypothesis for the origin of cancer foci. *Cancer* **4** 916–918 (1951).
2. C. O. NORDLING: A new theory of the cancer inducing mechanism. *Brit. J. Cancer* **7,** 68–72 (1953).
3. P. ARMITAGE and R. DOLL: The age distribution of cancer and a multi-stage theory of carcinogenesis. *Brit. J. Cancer* **8,** 1–12 (1955).
4. H. B. JONES: A special consideration of the aging process, disease and life expectancy. In: *Advances in Biological and Medical Physics* **4,** 281–337, ed. by J. H. Lawrence and C. A. Tobias, Academic Press, New York (1956).

DISCUSSION

QUASTLER: Two different functions have been proposed as representative of the course of the Gompertz function, $G(t)$, for the later part of the life span:

$$G_1(t) = a_1 + b_1 t$$

$$G_2(t) = a_2 + b_2 \ln t.$$

No author claims that either function is a perfect fit even for a limited interval. Still, it is worth showing that the difference between the two formulae is very small over a limited range. Let the center of this range be t^*; then

$$\Delta G_1 = \pm b_1 \Delta t$$

$$\Delta G_2 = b_2 \ln \left(\frac{t^* \pm \Delta t}{t^*}\right)$$

and for small values of $\Delta t/t^*$,

$$\Delta G_2 = \pm \frac{b_2}{t^*} \Delta t$$

It is said that the mortality rate (in the later part of the life span) doubles about every 8.5 years; hence $b_1 = 0.082$; and that it increases approximately with the 5.2th power of age, or $b_2 = 5.2$. These two values are compatible around $t^* = 63$ years, which characterizes the neighborhood in which both are claimed to be valid.

YOCKEY: If one plots survival data as AUERBACH does, one obtains curves which correspond to the Gompertz function for man and for many out bred wild-type organisms. For some in bred strains, particularly those which have a genetic defect, the survival curve may be of the form $\log l/l_0 = -\alpha\lambda^2$.

In Fig. 3 of my paper in Part V, I have plotted $\log l/l_0$ against the square of the age for several strains of mice. The dilute brown strain reported by MURRAY and HOFFMAN follows the

above equation quite closely for almost the entire life span, excluding only the first few months. On the other hand the DBF_1 hybrids (dilute brown female × C57 male) exhibit the Gompertz function as may be seen in Fig. 5 of that paper.

The dilute brown strain is characterized by a high rate of mammary cancer, while the hybrid has a low rate. The Marsh albino is another high-cancer-rate strain, which, although it does not have a survival curve of the form $\log l/l_0 = -\alpha\lambda^2$, does, when crossed with the C57, produce hybrids with a much longer life span. The survival curve is of the Gompertz type. Changes in the genetic characteristics associated with hybridization do not just change the constants of an equation of the Gompertz form, but rather the survivorship curve is of a different form.

FREE RADICALS AS A POSSIBLE CAUSE OF MUTATIONS AND CANCER*

WALTER GORDY
Department of Physics, Duke University, Durham, North Carolina

Abstract—The hypothesis set forth in this note is that free radicals produced outside the body may find their way into the body and produce mutations and/or cancer. The evidence for support of this hypothesis is the presence of radicals as detected by microwave paramagnetic resonance in several carcinogenic agents, and the fact that free radicals are now recognized by radiobiologists as being responsible for a large portion of mutagenic and carcinogenic effects of ionizing radiations.

FREE radicals may be loosely defined as molecular fragments which are characterized by a free valence or an unpaired electron. Because of their highly reactive nature they are not thought to exist in any significant quantity within the organic matter about us, although they are often postulated as important, transient intermediaries in organic and biochemical reactions. Within the past few years, however, microwave spectroscopists (1) have shown that free radicals can be readily detected in organic matter which has been subjected to some form of pre-treatment that can break chemical bonds. Such free radicals are produced in the combustion of organic matter—wood, paper, tobacco, coal, oil. They are produced in excessively cooked foods such as charred steak or scorched toast. They are produced in various forms of matter by ultraviolet light, by x-rays, or by atomic radiation.

The radicals are detected through their resonant absorption of microwave or radio-wave energy when they are placed in a magnetic field of the proper strength. This type of absorption spectrum is known as paramagnetic resonance or as electron spin resonance (2). Electrons in normal chemical bonds are paired in such a manner that their spins and magnetic moments cancel, and hence they exhibit no paramagnetic absorption. Paramagnetic resonance occurs only for the unpaired electrons of the disrupted chemical bond. It therefore provides a specific and powerful means of detecting and studying reactive free radicals within organic matter without interfering absorption or confusing signals from the normal stable molecules of the matter.

The surprising new evidence from paramagnetic resonance is not that free radicals can be easily produced but that they become trapped and stabilized and can be transported from place to place, even through the air within tiny particles of solid matter such as those in smoke. The nature of neither the radicals nor their cages is yet known definitely, although some radicals produced

* This research was supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command under contract No. AF18(600)-497. Reproduction in whole or in part is permitted for any purpose of the United States Government.

in amino acids and proteins by x-irradiation have been tentatively identified from the fine structure of their microwave resonance patterns (3). The information pertinent to the present discussion is that organic radicals produced by physical forces such as heat or irradiation outside the body can be taken into the body through the processes of eating, smoking, or normal breathing, or even by diffusion through the skin. Once inside the body, these radicals may themselves penetrate the cells or they may be converted to other radicals which do so. A radical containing an odd number of electrons must, in effect, meet and react with another radical before its free valence or uncancelled electronic moment is nullified. If it reacts with a normal organic molecule (which has an even number of electrons), another radical is produced. In fact, it is just this odd character which suggests that a lone radical might start a significant chain of events within a cell.

Many types of radicals which have been detected by microwave resonance are stabilized mainly within solid particles of matter. Normal chewing and mixing of food with saliva would tend to destroy them. This destruction may not always be complete, however. We have made tests which show that ordinary chewing of charred toast, beef, and other foods does not entirely kill the resonance signal of the radicals. Extremely small solid particles carrying radicals may diffuse into the tissues of the skin, stomach, or lungs where they would gradually dissolve and perhaps bring about damaging reactions as their radicals are released. Furthermore, these radicals are possibly stable in certain organic solvents which dissolve the solid cages and 'float' the individual radicals into the tissue. Such a solvent might assist in the production of cancer without being a primary cause of it. Strong resonances, like that shown in Fig. 1 for tobacco tar, are found for wood tar, coal tar, and other tars. H. Shields and the author have dissolved tars in organic solvents including benzene, acetone, and croton oil, and have found that the resonance of the tar radical remained strong. The role of agents such as croton oil, which are not themselves carcinogenic agents but which augment the effects of certain carcinogenic agents, may be that of facilitating the entrance of carcinogenic radicals into the body.

Radiobiology experiments (4) indicate that much of the effect of ionizing radiations on the cells themselves may be indirect; that is, irradiation produces a free radical in one part of the cell which later migrates to a more vital part of the cell where it may react to bring about a mutation. Alternately, the first radical formed may react to form a second radical, or a third, which finally causes the mutation. In particular, OH and OOH radicals have been postulated as important intermediaries in radiation damage. Of course a mutation might be brought about by a so-called direct hit, but indirect effects also appear to have significant consequences. We are proposing an extension of the indirect effects to include cases where the primary irradiation occurs entirely outside the injured body. In our laboratory, microwave evidence has been obtained to indicate that hydrocarbon radicals, R, produced by irradiation, are often converted to peroxide radicals, ROO, where they come in contact with oxygen. In the tissue such radicals might be further converted to the OH or OOH radicals, already under suspicion by radiobiologists.

The striking evidence which prompted this communication is the abundant paramagnetic resonance data for the existence of free radicals in many agents

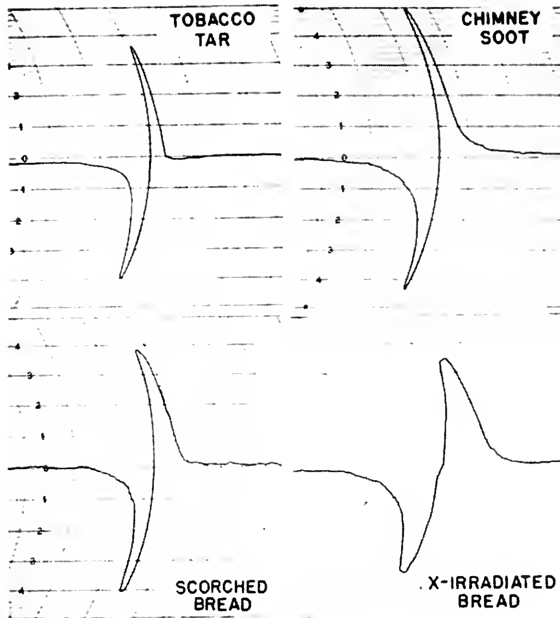


FIG. 1. Microwave electron spin resonances of radicals in some common substances. The tobacco tar was taken from an old pipe stem. Coal, wood, and other tars give similar resonances. The chimney soot was taken from the flue of an oil-burning furnace. Similar resonances were obtained for soot taken from the exhaust pipe of an automobile and from a wood-burning fireplace. Ordinary bread, unscorched and not irradiated, gave no detectable resonance in the same spectrometer.

to face p. 354

known or suspected to cause cancer. Among these are cigarette smoke, tobacco tars, various other tars, exhaust fumes from cars, smoke from home furnaces or industrial plants, and charred foods. I shall not attempt to cite the literature references for the various evidences that these are carcinogenic agents. It is well known that x-rays and other ionizing radiations can cause genetic mutations and are likewise carcinogenic agents. It is now well known from electron spin resonance that these ionizing radiations also produce radicals which in many biochemical solids (3) (including various proteins, carbohydrates, and fats) persist for long periods after the irradiation.

The carcinogenic effects of severe chemicals which produce burns of the flesh may possibly result from subsequent diffusion into the healthy cells of free radicals produced in the original, more violent chemical reaction causing the burn. It is known that a burn of the flesh from any source of heat has carcinogenic and mutagenic effects. Since we now know that the charring of any organic matter produces long-lived radicals, it seems probable that some of the carcinogenic and mutagenic effects may result from secondary activity of radicals produced by the burn. Of course chromosome linkages are broken as direct effects of the heat, but it seems probable that most of the cells exposed to the elevated temperatures in the burned area would be killed.

Certainly many known carcinogenic chemicals are not radicals, and I do not suggest that all cancer may be caused by radicals. However, many chemicals recognized as carcinogenic agents, not themselves radicals, may exert their carcinogenic activity indirectly through the production of radicals within the body. This would be analogous to the indirect effects of ionizing radiations already mentioned and might account for the seemingly parallel action of certain chemicals with ionizing radiations which has led to their being called radiomimetic chemicals (5). Many carcinogenic chemicals are large, aromatic, polycyclic hydrocarbons from which it would seem that free radicals might be easily produced.

The radicals are not convicted from 'guilt by association' with carcinogenic agents. Our proposal is not intended to be accepted *per se*, but is offered as a working hypothesis which can be put to rather objective test because of the powerful method of electron spin resonance now available for detection of radicals. That certain radicals are likely to be carcinogenic agents, or that some types can lead to genetic mutations, probably will not be questioned. Others, possibly some or all of those which are sufficiently stable in organic matter to be detected with paramagnetic resonance, may be perfectly harmless. I do not therefore recommend that we become suddenly alarmed about the radicals around us. I do think there is some justification for the careful study of these radicals which can be produced, transported, and taken into the body so easily. This study is made easier by the powerful new method of paramagnetic resonance for detection of such radicals.

If externally produced radicals are indeed dangerous, we can fortunately detect and avoid most of the ones we now are eating, breathing, or rubbing into our skins. INGRAM (1) has shown that the number of radicals produced by heating organic matter is a sensitive function of temperature. Tests in our laboratory on common foods such as meat and bread show that those cooked in a normal manner have no detectable resonances or only very weak resonances,

whereas burned food, scorched toast, charred steak, etc., have strong radical resonances. The temperature at which a cigarette is burned should have significant effect upon the number of radicals produced, although it may be impossible to produce smoke without producing radicals. If it proves harmful, we do not have to preserve our food by atomic irradiation.

Acknowledgement—Several of my students, HOWARD SHIELDS, HARVEY N. REXROAD, FRANK PATTEN, and GENE MCCORMICK, assisted with microwave magnetic resonance experiments in connection with the hypothesis proposed here. I am indebted to my wife for encouragement and for assistance with library reference work.

REFERENCES

1. See, for example, comments by:
   G. E. PAKE, J. E. INGRAM, J. CAMBRISSON, and J. UEBERSFELD, R. LIVINGSTON, W. GORDY *et al.*: *Disc. Faraday Soc.* **19**, 179, 184 (1955).
2. General treatments of the subject are given in:
   W. GORDY, W. V. SMITH, and R. F. TRAMBARULO: *Microwave Spectroscopy*, Chap. 5, J. Wiley and Sons, New York (1953).
   W. GORDY: Microwave and radio-frequency spectroscopy. In: *Techniques of Organic Chemistry*, Vol. 9, 71–185, ed. by A. Weissberger and W. West, Interscience, New York (1956). D. J. E. INGRAM: *Spectroscopy at Radio and Microwave Frequencies*, Butterworths, London (1956).
3. W. GORDY, W. B. ARD, and H. SHIELDS: Microwave spectroscopy of biological substances. I. Paramagnetic resonance in x-irradiated peptides; and II. Paramagnetic resonance in x-irradiated carboxylic and hydroxy acids. *Proc. Nat. Acad. Sci., Wash.* **41,** 983–1004 (1955).
4. E. C. EVANS: Effects of hydrogen peroxide produced in the medium by radiation on spermatozoa of *Arbocia punctulata*. *Biol. Bull.* **92**, 99–109 (1947).
   O. WYSS *et al.*: Role of peroxide in the biological effects of irradiated broth. *J. Bact.* **56**, 51–57 (1948).
   W. M. DALE: Some aspects of the biochemical effects of ionizing radiations. In: *Symposium on Radiobiology: The Basic Aspects of Radiation Effects on Living Systems*, 177–178, ed. by J. J. Nickson, J. Wiley, New York (1952).
   R. E. ZIRKLE: Speculations on cellular actions of radiations. In: *Symposium on Radiobiology: The Basic Aspects of Radiation Effects on Living Systems*, 333–356, ed. by J. J. Nickson, J. Wiley, New York (1952).
5. A. HADDOW: Comparative studies of the biological effects of ionizing radiation and of radiomimetic chemical agents. *Proc. Int. Conf. Peaceful Uses of Atomic Energy*, **11** 213–218, United Nations, New York (1956).

PART VI

INFORMATION NETWORKS

A PROBABILISTIC MODEL FOR MORPHOGENESIS*

MURRAY EDEN
Laboratory of Technical Development, National Heart Institute, National Institutes of Health, Bethesda 14, Maryland

Abstract—A program has been outlined for establishing relationships between the form of an organism and the minimal information content of the germ cell from which the organism was derived. A simple two-dimensional model has been chosen in order to explore the feasibility of such a program. A suitable information measure has been defined for this model and computations of information have been made for small aggregates of cells, as well as estimates for larger aggregates.

An arbitrary growth process has been formulated which is analogous to an assignment of virtually no information to the germ cell. The properties of this growth process have been studied and suggest that even such minimal information content in the germ cell is sufficient to specify the over-all form of the organism with high probability after a certain number of divisions have taken place. Possible ways of extending the model and increasing its embryological relevance have been suggested.

ONE of the problems that has been touched in this symposium is the particularly elusive subject that has been with biology since its beginnings as a science; that is, the general relationship between function and form and between growth and form. In the particular terms of discourse of this symposium the question might be phrased in this way: 'What is the minimal amount of information that is required in a fertilized egg, so that after a certain number of divisions and a certain length of time the egg will have developed into an organism that is recognizable as being a member of a certain species?'

A number of workers have estimated the information content of biological objects. DANCOFF and QUASTLER (1) computed values of information content relative to four different models; on the basis of atomic orientation, molecular structure, chromosome volume and a genotype catalogue. These authors were careful to specify the limitations of their computations. They write, 'We have arrived, by very tentative methods, to the result that the essential complexity of a single cell and of a whole man are both not more than 10^{12} nor less than 10^5 bits; this is an extremely coarse estimate, but is better than no estimate at all.' LINSCHITZ estimated the 'physical entropy' of a bacterial cell as being 10^{13} bits (2) and YOCKEY (3) has computed the information content of DNA based on its molecular size and on a postulated cryptographic relation between proteins and nucleic acids. Most of the values obtained have been very large, and one would presume that they are large enough to describe adequately the observable properties of a living organism. It may be that the growing organism requires a lot less information in the germ cell than is indicated by

* The work presented here was begun during the tenure of an United States Public Health Service Special Fellowship in the Department of Mathematics, Princeton University, Princeton, New Jersey.

estimates made on a molecular level. In the terminology of communication theory, the redundancy of the source may be extremely high. An examination of the literature indicates that there is some support for this view. Studies of properties of monozygotic twins have special relevance to this point. It may be noted in passing that the existence of twins or high multiplets derived from a single germ cell is in itself strong evidence of the presence of at least a small amount of redundancy in the germ cell (4). Monozygotic twins presumably arise from an identical genetic background and they develop into mature organisms that can be compared with regard to certain of their properties.

As long ago as 1876 GALTON (5) studied what he called 'The History of Twins as a Criterion of the Relative Power of Nature and Nurture'. NEWMAN, in a long series of publications begun in 1912 (6) has studied both human twins and armadillo quadruplets. The nine-banded armadillo is exceptional in that the female gives birth to monozygotic quadruplets. The scales or scutes on the back of an armadillo are regular and easily counted, even in the fetus. Newman prepared a fairly large statistical study on these quadruplets and he found a correlation coefficient for fifty-six sets of male quadruplets of 0.93 and for fifty-nine sets of females of 0.91. Still there was no identity in the scute counts.

Work of a similar character has been done by HANCOCK (7) on monozygotic calf twins, and by WENT (8) on genetically identical seedlings. The conclusion that may be reached on the basis of studies such as these is that even when embryonic growth starts from genetically isomorphic cells, by the time the organism has developed to maturity there is, it is true, a great similarity in the large, but at the cellular level there is very little similarity.

This would suggest that, aside from the genetic signals or instructions, there are certain statistical variables or environmental factors operating that permit the development of an organism to an ultimately recognizable form but require a good deal less information than would be required if every element in the structure of the organism had to be specified with microscopic exactitude.

An attempt to construct a theoretical model was made by TURING (9), who in 1952 posed the following problem. Given a group of identical cells arranged in some symmetrical configuration, e.g. a ring or a sphere; assume that each cell contains the same concentrations of certain chemical substrates and that the laws of diffusion and other classical physical laws hold. How can one devise a procedure whereby this homogeneous collection of cells could develop and differentiate so as to produce asymmetric or periodic forms?

Turing proposed accomplishing this in a way that does not do too much violence to biological understanding. He postulated certain hypothetical chemical reactions involving substrate, intermediates and enzymes, and built into this set of hypothetical reactions appropriate reaction rate constants so that the resulting reaction system would exhibit a special property: namely, that statistical fluctuations in the concentrations of chemical components in various cells would increase in amplitude so as to produce an instability and result in an asymmetric form or a form exhibiting periodicity. By use of a specific example he showed how a ring of cells might grow into something more or less petal-shaped with three or four lobes or petals. In another example he developed a mottled pattern on a two-dimensional surface.

The model described below is entirely mathematical; physical or chemical

phenomena are not considered. The principal concern of this model is the domain of forms an idealized organism can assume and the likelihood of an egg developing into such a form. The mathematics employed is elementary, and as in so much of combinatorial analysis, it is *ad hoc**. The attempt to establish a relationship between the form of an organism and the information content of the germ cell ancestor is treated here from a point of view that has some resemblance to that of statistical mechanics.

It is assumed here that there are only a finite number of different kinds of cells in any organism and a finite number of cells of each kind. We neglect the dynamic processes occurring continuously in an organism: changes within cells, the migration and movements of cells, the death of certain cells and the cleavage or maturation of others. If there is some well-defined way of describing the orientation of each cell in any organism relative to the other cells in that organism or relative to some arbitrary system of coordinates, then it is possible, in theory at least, to enumerate all the possible ways of arranging cells into different configurations. Some of these arrangements would be recognizable organisms, the overwhelming majority would not. In any case, these objects, both the recognizable and otherwise, are elements of the set of all possible configurations. This procedure might represent a means of defining a given species by certain restrictions on the possible orientations of cells and thus to identify the given species with a well defined subset of all possible configurations.

Most multicellular organisms can be said to arise from a single cell resulting from the fusion of two germ cells. It is true that there are certain biological objects, of which the slime-mold is a notable example, which take their form from the migration, coalescence and specialization of a number of free-living cells. However, such organisms are uncommon and will not be considered further.

This single germ cell divides into two cells and these cells will divide further and so on until maturity. Throughout the course of this branching process the growing organism will pass through a sequence of configurations, each of which is an element in the set of all possible configurations. If there is a relationship between successive configurations which is recursive, then a generating function can be constructed to describe the branching process. Generating functions are useful because they may afford a means of assigning a probability to each possible configuration. The actual model chosen for investigation has been simplified to the extent that its relation to biological reality is largely impressionistic. Its justification is heuristic, for the study of relatively simple systems may suggest methods of approaching the real systems which are so very much more complex.

The element of the model is called a *cell*. All cells are considered to be identical. We restrict ourselves to the consideration of arrangements of cells in two dimensions. The shape of the individual cell is unspecified (they may

* I should like to take this opportunity to express my debt to a number of mathematicians both at Princeton and at the Institute for Advanced Study with whom I have discussed this problem; and in particular to Professor VALENTINE BARGMANN, Professor WILLIAM FELLER, Dr HALE TROTTER and Dr NORMAN SHAPIRO for their stimulation and suggestions. Needless to say, the results and errors are my own.

be thought of as squares), but their positions are restricted to the points of a (two-dimensional) square lattice.

Any arrangement of cells on the lattice will be called a *configuration*, i.e. an arrangement of k cells will be called a k-configuration. If each cell in a configuration is adjacent to at least one other cell then such a configuration will be called *connected*. We will be interested only in connected configurations. The set of all possible (connected) k-configurations will be called the *k-array*. The number of distinct k-configurations, i.e. configurations that are not isomorphic under translations, reflections and rotations, will be called the *order* of the k-array, symbolized $N[k]$.

Each cell will have four *edges*, corresponding to the four nearest-neighbor lattice points. An edge will be called *open* if its corresponding lattice point is unoccupied by a cell, otherwise it is *covered*. Each cell also has four *corners* corresponding to points equidistant to four lattice points. A corner will be called an *inner corner* if it is at the center of a cluster of four cells.

The problem of enumerating all possible k-configurations is one that has, as yet, no easy solution. Similar combinatorial problems, arising in physics in what is called the order-disorder problem, have been considered by a large number of workers. Of particular relevance to the above problem is the work of VAN DER WAERDEN (10), KAC and WARD (11), and HIJMANS and DE BOER (12).

Certain bounds can be set for the order of the k-array. We can determine a lower bound for $N[k]$ by enumerating all members of a certain subset of $[k]$, i.e. the subset in which all save two cells have two edges covered. Two cells, i.e. the *ends*, have only one edge covered. It is even easier to enumerate a smaller subset of this 'two-ended' set. Consider an arbitrary lattice point as the origin of a random walk. Limit the choices for the first step and each succeeding step in this random walk to lattice points, either above or to the right. The kth cell will be added after $k - 1$ steps are taken. At each point there are exactly two possible choices, so that in all we have produced 2^{k-1} configurations. Since each configuration (except those that exhibit internal symmetry, in any case, a small fraction) occurs four times in 2^{k-1} configurations, the number of distinct configurations is 2^{k-3}. The restriction to two choice points is dictated by the necessity of avoiding cross-overs in the random walk. Obviously, each cross-over would have the effect of decreasing the number of occupied lattice points by one.

However, if the random walk is permitted three choice points, i.e. above, to the right and to the left, one can estimate the number of such walks of length, k, which contain no point adjacent to more than two occupied sites*. Such walks are isomorphic to the set of 'two-ended' k-configurations. This estimate was found to be very close to $(1 + \sqrt{2})^{k-3}$. In consequence the lower bound for the number of k-configurations may be raised to this value.

Upper bounds can also be computed using a somewhat different combinatorial technique. Consider any k-configuration. Arbitrarily choose one cell as the origin, and also arbitrarily choose one of the four possible orientations of the lattice. Identify this cell by 1 if it has a cell beneath, otherwise 0. Further, this cell may have a cell adjacent to it on the left; if so, assign a 1 to the next

* The mathematical details of the results presented in the text will be the subject of a separate publication.

digit in the identification; a cell above it, and a cell to the right. Thus, the first cell C_1 in a configuration is identified by four binary digits. Next, identify the adjacent cell which contributed the first '1' in the designation of the first cell, as the second cell C_2, the second '1', as the third cell, C_3, etc. Reorient the lattice so that the first cell is beneath the second. We construct the designation number of the second cell as we did for the first. However, this time there are only three binary digits required since the adjacency of C_2 to C_1 is already determined. Any of the cells adjacent to C_2 that have not yet been assigned a position in the order can be given one now in a perfectly well-defined way. It is obvious that this procedure can be continued until designation numbers have been obtained for each cell in the configuration. We thus have a well-defined word in $3k + 1$ binary digits and a possible 2^{3k+1} such words.

Since the initial cell and the orientation of the lattice were chosen arbitrarily, each district configuration (as usual, excepting those exhibiting some internal symmetry) will be given by $4k$ such words. Thus an upper bound for $N[k]$ is $2^{3k-1}/k$.

It is easily ascertained that a very large proportion of the 2^{3k+1} words do not represent k-configurations. These forbidden words arise for essentially the same reason that the unrestricted random walk on the square lattice fails to serve as an estimate of two-ended configurations. No simple relations have been found that will indicate which of the 2^{3k+1} words are permissible. However, one can generate a random sample of these words by a Monte Carlo procedure and arrive at a statistic that suggests that a satisfactory estimate of $N[k]$ is in the neighborhood of 2^{2k}.

Values of the bounds and the estimate mentioned above have been computed for certain values of k (Table I). This serves to give some idea of the

Table I. Estimate of Configurations for Large Arrays

$[k]$	Lower bound	Upper bound	Estimate
	$(1 + \sqrt{2})^k$	(2^{3k})	(2^{2k})
10	2.2×10^3	2.4×10^4	5.8×10^3
16	4.2×10^5	6.3×10^9	2.5×10^4
25	1.1×10^9	8×10^{17}	6×10^{12}
100	1.7×10^{18}	3×10^{85}	6×10^{57}
1000	1.7×10^{171}	3×10^{895}	6×10^{597}

magnitudes one might expect for configurations of large numbers of cells. So long as the number of cells is small, the distinct configurations can be exhibited with relative ease. This has been done up to $k = 8$ and the results are given in Table II.

In order to establish the assignment of a probability to each of these configurations, a simple and nearly featureless generating function was adopted. Starting with a single cell, equal probability is assigned to each of the four possible two-celled configurations. These are all isomorphic. This two-celled configuration has six open edges. Equal probabilities are assigned to each

edge. This time, of the six three-celled configurations obtained by adjoining a cell to an open edge, four are isomorphic to one of two three-celled configurations, and two to the other. Thus, the probability of the first three-celled configuration is 0.67 and the other 0.33. This procedure can be carried out indefinitely, in each case assigning equal weight to each open edge and adjoining a single cell at a time.

Table II. Number of Configurations in Each Array

k	$N[k]$	$\frac{N[k]}{N[k-1]}$
1	1	
2	1	1
3	2	2
4	5	2.5
5	12	2.4
6	35	2.9
7	108	3.1
8	367	3.4

While there is no biological organism that exhibits this pattern of growth, it has certain features in common with some tissue cultures, bacterial colonies or tumors, in that the cells are more or less undifferentiated. Growth in such

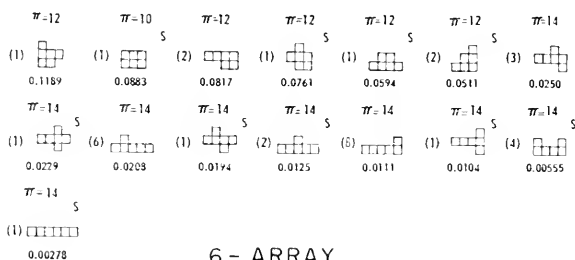


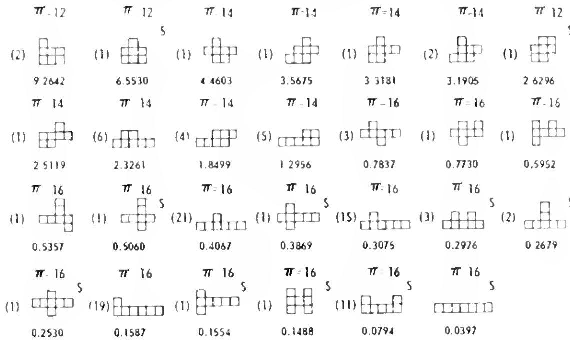
FIG. 1

biological objects has no preferential direction except that it is peripheral, a condition due most likely to the fact that diffusion of nutrient is too slow to permit any large number of cell divisions in the interior of the growth.

Exact computations have been carried out for the probability associated with each k-configuration up to $k = 8$. As before, computations for $k > 8$, while easily performed in principle, are prohibitively time-consuming. The configurations for $k = 6, 7, 8$ and their associated probabilities are assembled in Fig. 1, 2, 3.

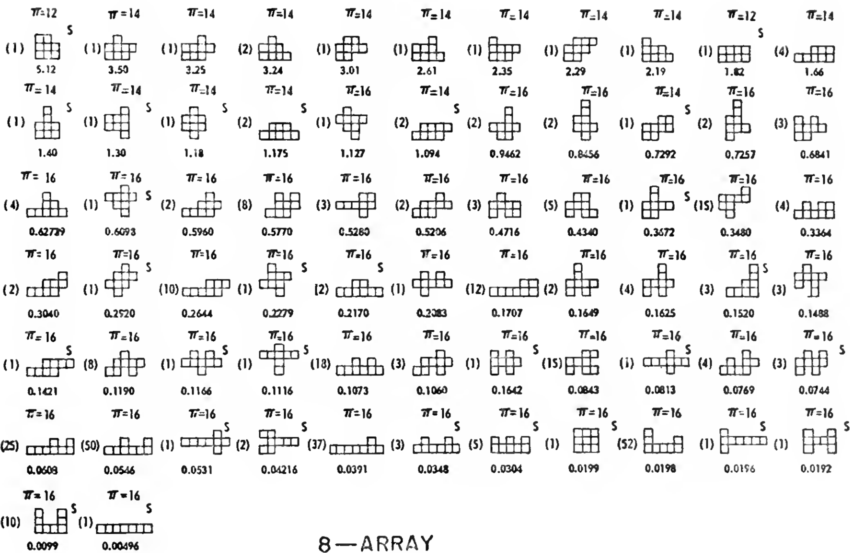
An unanticipated property of the particular generating function employed was revealed as a consequence of these exact computations. It will be observed in Fig. 1 to 3 that configurations have been grouped so that each different value of probability is recorded next to a single prototype configuration.

These configurations bearing the same probability, while they are not isomorphic in the sense mentioned earlier, have an important property in common. If each configuration is represented by a graph (13), identifying the cells as nodes



7 – ARRAY

FIG. 2



8 — ARRAY

FIG. 3

and the covered edges as branches between nodes*, it will be seen that all configurations represented by the same graph have the same probability. Certain other properties are also suggested by consideration of these small 'organisms'. The configurations with the largest number of inner corners† are

* This representation is analogous to the graph obtained by identifying countries on a map with nodes and common frontiers between countries with branches.

† We can use the perimeter, π, i.e. the number of open edges, instead of the inner corner, C, in describing the property in question since $\pi = 2(k + 1 - C)$.

most probable. It also appears that configurations with many short branches are more probable than those with a few long branches. Finally, it is also observed that as k increases, a decreasingly small fraction of the k-array carries the weight of probability. This is shown in Fig. 4 and in Table III. The k-configurations have been ranked in order of decreasing probability, that is,

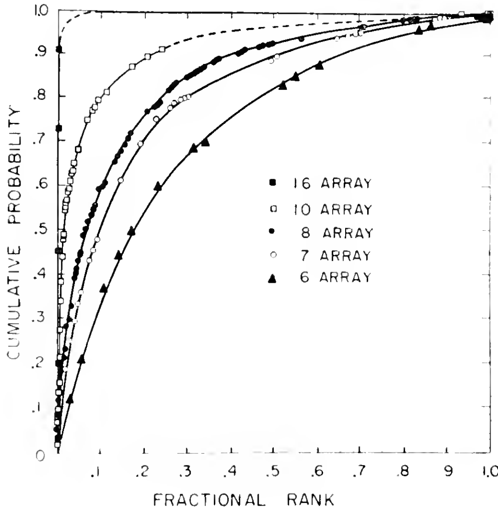


FIG. 4

Table III.

k	Probability of configuration of rank 1	Probability of least probable asymmetrical configuration $\left(\frac{Z^3}{k!}\right)$	Rank at $\Sigma p_i = 0.5$	Fractional rank at $\Sigma p_i = 0.5$
1	1.00	1.00	—	—
2	1.00	1.00	—	—
3	.67	.33	—	—
4	.33	.167	—	—
5	.40	.067	2	.167
6	.12	.011	6	.169
7	.093	.0016	12	.11
8	.051	.0002	24	.067
10	.021	2×10^{-6}	103	.020
16	.0015	3×10^{-15}	6400	.00002

the most probable configuration was designated 1, the next most probable 2, and so on, and the cumulative probability (as ordinate) was plotted against the rank divided by $N[k]$ (number of distinct k-configurations) as abscissa.

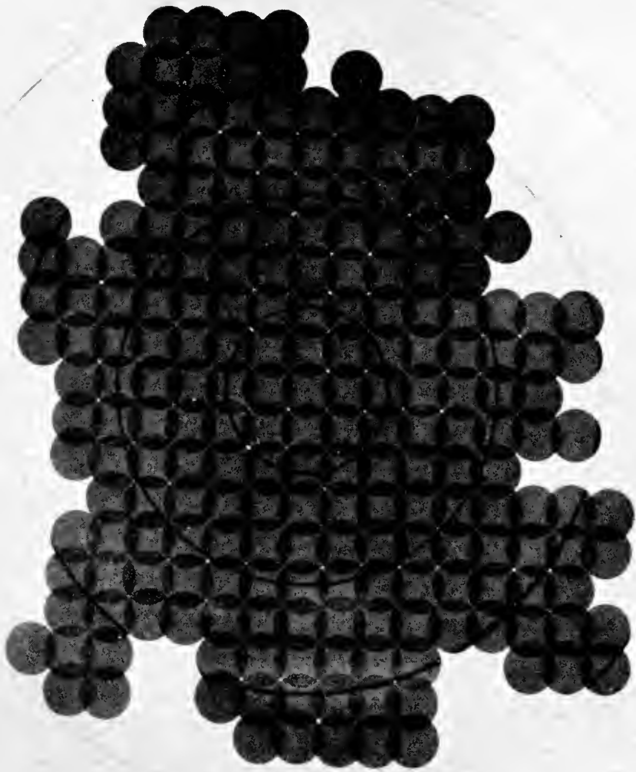


FIG. 5

o face p. 367

Since any direct extension of the model to larger values of k does not appear feasible, another procedure was adopted. Starting as before from a single cell, the edges were numbered, a random number table (14) was consulted to find a number equal to or less than 4, and then a cell was adjoined to the appropriately numbered edge. The open edges were renumbered, another number equal to or less than the number of open edges obtained from the table, and a new cell adjoined. In this way samples of 1000 10-configurations and 16-configurations were constructed. A few larger configurations were prepared by this procedure. One such containing 200 cells is shown in Fig. 5.

In the case of the sample of the 10-array, one configuration was obtained twenty-two times. Its probability was computed by the exact procedure described above and found to be 2.06 per cent. All the configurations containing four inner corners ($\pi = 16$) (maximal for $k = 10$) appeared more than ten times each. With very few exceptions, in order of occurrence, there followed the configurations with $\pi = 18, 20, 22$. There were eighty-three occurrences with $\pi = 24$ (no inner corners), but none of these was two-ended. Although it was impossible to enumerate all the configurations, by judicious use of the equality of probabilities found in configurations with the same graph, estimates were made of the numbers of configurations of each kind up to rank 1150. The data were plotted in Fig. 4. It can be seen that the portion of curve obtainable is very close to the ordinate axis.

A similar procedure was followed in the case of the sample of the 16-array. Here, estimates were considerably poorer, but the same general features were revealed (Table IV). The thirty-two possible configurations with $\pi = 18$

Table IV. Summary of 16-array Monte Carlo Sample

Perimeter	Configurations	Cumulative occurrence	Fractional rank	$E(p_i) \times 10^4$
16	1	0		
18	32	.028	1×10^{-6}	8.75
20	569	.202	2×10^{-5}	3.00
22	6250	.455	4×10^{-4}	.40
24	27,300	.728	1.2×10^{-3}	.16
26	148,500	.907	6.6×10^{-3}	.012
28	—	.967		
30	—	.995		
32	—	1.000		
34	—	1.000		

appeared twenty-eight times, or an expectation of occurrence of a particular configuration of 8.75×10^{-4}. (It is assumed that all configurations with identical values of π have approximately equal probabilities of occurrence.) It was estimated that the expectation of occurrence of a configuration of $\pi = 20$ was 3×10^{-4}; $\pi = 22$, 4×10^{-5}; and $\pi = 24$, 1.6×10^{-5}. In this sample of 1000 there were only five occurrences of configurations with $\pi = 32$, and no occurrences of $\pi = 34$, although a low estimate of the number of distinct

16-configurations with $\pi = 34$ would be 250,000. These estimations are plotted on the same figure as the computations for configurations of up to eight cells (Fig. 4). It can be seen that none of the estimates obtainable up to a cumulative probability of 0.907 can be distinguished from the ordinate axis.

The probability of the most probable k-configuration and of the least probable k-configuration are presented in Table III, for several values of k. It will be noted that the probability of the most probable configuration decreases slowly with increasing k. While there is no practical way to make exact computations of probability for large values of k, it may be conjectured that the probability of the first ranked configuration is proportional to $1/2^k$. On the other hand, the probability of the configurations of the lowest ranks falls very rapidly*. As with the estimates of the number of configurations, exact solutions for probability are readily obtained only for the two-ended configurations. As was noted earlier the number of such forms approximates $(1 + \sqrt{2})^k$ but the probability associated with each such form is $2^3/k!$

Information theory (15) suggests methods of defining appropriate measures for the distribution of probabilities as a function of k. If $N[k]$ is the number of distinct configurations containing k cells each, the maximal uncertainty for the k-array can be defined as $H_k^0 = -\lg N[k]$†. In a similar manner, an uncertainty can be defined for an arbitrary generating function, G_j, considered as an information source. $H(G_{j,k}) = -\sum_{i=1}^{N[k]} p_i \lg p_i$, in which p_i is the probability that the generating function G_j will terminate after $k - 1$ adjunctions in configuration ω_i. Further, a measure of relatedness (16) may be defined as $I(G_{j,k}) = [H_k^0 - H(G_{j,k})]$.

What does this mean in terms of information theory? Supposing we had a generating function or some procedure that produced every one of these unusual configurations with equal probability. Then the two numbers H_k^0 and $H(G_{j,k})$ would be identical. The uncertainty of such a generating function would be maximal. On the other hand, if the generating process were such as to specify, with probability 1, only one out of the total number of configurations, then the uncertainty of the generating process $H(G_{j,k})$ would be 0. As can be seen, $I(G_{j,k})$ for a given generating process carried out through k steps has been defined above simply as the difference of these two quantities. Very briefly then, this measure would suggest that if a knowledge of the generating process does not enable us to predict which of the possible configurations to expect after the process has gone along for k steps, then knowledge of the generating process provides no information. On the other hand, if one can devise a mathematical mechanism, that is, a generating process, that can specify the ultimate form of an organism exactly, then the generating process contains all the information it possibly can.

Applying this measure to the presently available data and the particular generating function introduced earlier, it is observed that $I(G_{j,k})$ increases with

* It will be noted that the probability of the most probable configuration exhibits a maximum at $k = 5$. This is an accident attributable to the fact that this particular 5-configuration is the only one containing a cluster and it is asymmetric. Such an accident is extremely unlikely to be found when k is large.

† The symbol 'lg' is used here to denote 'logarithm to the base 2'.

increasing k (Table V). Estimates have been made for $k = 10$ and $k = 16$ from the Monte Carlo samples. These estimates are certainly lower than the precise values since an estimate of p_i was not available for every configuration and the means for rather large groups of configurations were used instead. A functional

Table V. Entropy of k-*arrays*

k	$H(G_k)$	H	$I(G_k)$
1	0	0	0
2	0	0	0
3	.92	1.00	.08
4	2.19	2.32	.13
5	2.90	3.59	.69
6	4.54	5.13	.59
7	5.59	6.76	1.17
8	7.00	8.53	1.53
10	10.00	12.29	2.29
16	16.68	21.77	5.09

relationship between $I(G_{j,k})$ and k has so far not been found. One may conjecture that the relatedness increment $I(G_{j,k}) - I(G_{j,k-1})$ approaches 0.5 as k increases without limit. This may be interpreted to suggest that the rate of information accumulation in an organism constructed according to such a plan is half a bit per cell division.

Other measures have been suggested as being useful to our purposes. Following the terminology of McGill and Quastler (16), the *relative uncertainty* of the generating process is $D_{j,k} = \dfrac{H(G_{j,k})}{H_k^0}$ and the *redundancy* is $C_{j,k} = 1 - D_{j,k}$.

The redundancy evaluated from the results presented in Table V increases from a value of 0 for $k = 2$ to a value of 0.234. As with the measure $I(G_{j,k}) - I(G_{j,k-1})$, it seems plausible to expect that as k increases $C_{j,k}$ will converge to some value other than 0 or 1, but no procedure has as yet been found to test this conjecture and to determine the limit.

In a qualitative way, this increase in $I(G_{j,k})$ may be understood to mean that the featureless generating function considered above determines the configurations of large numbers of cells with a high degree of specificity. It is virtually a certainty that large configurations will be essentially circular in outline; that they will have a high density, i.e. they will contain very few 'holes' and short 'tentacles'. Thus if one considers the most probable outcomes of the generating procedure in the large, then these configurations appear to resemble one another very closely even though they exhibit no correspondence in detail.

It is true that the results obtained with such a simple model are far removed from the intricacy of development of living things. A few regularities in the most probable forms may be introduced by small modifications of the initial generating procedure. Objects that are ellipsoidal or cruciform or objects characterized by large numbers of branches have been developed by such modifications. However, it is unlikely that further complexity can be introduced

into the growth process without drastic modification of the generating function, in particular without consideration of the history of a particular growing configuration. The results of embryology suggest that the generating process must contain a set of instructions that will alter the pattern of growth on the condition that a given stage or over-all configuration shall have been reached, and that such a change in pattern of development may occur a large number of times during the process of maturation. It is certain that such a modified generating process will have a higher information content than the process considered in detail in this paper. It remains to be seen whether modifications of this character can be formulated and whether a mathematical treatment of the consequences is possible of achievement.

REFERENCES

1. S. M. DANCOFF, and H. QUASTLER: The information content and error rate of living things. In: *Information Theory in Biology*, ed. by H. Quastler, 263–273, University of Illinois Press, Urbana (1953).
2. H. LINSCHITZ: The information content of a bacterial cell. In: *Information Theory in Biology*, ed. by H. Quastler, 251–262, University of Illinois Press, Urbana (1953).
3. H. P. YOCKEY: An application of information theory to the physics of tissue damage. *Radiat. Res.* **5**, 146–155 (1956).
4. K. S. TWEEDELL: Identical twinning and the information content of zygotes. In: *Information Theory in Biology*, ed. by H. Quastler, 215–250, University of Illinois Press, Urbana (1953).
5. F. GALTON: The history of twins, as a criterion of the relative powers of nature and nurture. *J. Anthrop. Inst.* **5**, 391–404 (1876).
6. H. H. NEWMAN: *The Biology of Twins* (*Mammals*), University of Chicago Press, Chicago (1917).
7. J. HANCOCK: Monozygotic twins in cattle. *Advances in Genetics* **6**, 141–175 (1954).
8. F. W. WENT: The role of environment in plant growth. *Amer. Scient.* **44**, 378–398 (1956).
9. A. M. TURING: The chemical basis of morphogenesis. *Phil. Trans.* (*B*) **237**, 37–72 (1952).
10. B. L. VAN DER WAERDEN: Die lang reichweite der regelmässigen atomanordnung in mischkristallen. *Z. Physik* **118**, 473–488 (1941).
11. M. KAC and J. C. WARD: A combinatorial solution of the two-dimensional Ising model. Inst. for Advanced Study, Princeton, N.J. 1–15 (1954).
12. J. HIJMANS, and J. DE BOER: An approximation method for order–disorder problems. *Physica* **21**, 471–516 (1955).
13. D. KÖNIG: *Theorie der endlichen und unendlichen graphen*, Chelsea Publishing Co., New York (1950).
14. The Rand Corporation: *A Million Random Digits with* 100,000 *Normal Deviates*, The Free Press, Glencoe, Ill. (1955).
15. C. E. SHANNON, and W. WEAVER: *The Mathematical Theory of Communication*, University of Illinois Press, Urbana (1949).
16. W. MCGILL, and H. QUASTLER: Standardized nomenclature: An attempt. In: *Information Theory in Psychology*, ed. by H. Quastler, 83–92, The Free Press, Glencoe, Ill. (1955).

FUNCTIONAL GEOMETRY AND THE DETERMINATION OF PATTERN IN MOSAIC RECEPTORS

JOHN R. PLATT
University of Chicago, Chicago, Illinois

Abstract—Every visual pattern element—straight lines, curved lines, parallel lines, angles, periodicities—shows some *self-congruence* under translations or rotations. A random mosaic of detector cells, like the 10^8 cells of the human eye, can be used as a *null detector* to indicate this self-congruence during *scanning operations*. This *operational definition* of pattern is called *functional geometry*. It underlies the generation of precision optical and machine surfaces by the Whitworth, Rowland and Strong methods and theoretically can approach *infinite precision*, starting from rough materials. It converts a space pattern into time pattern repetitions whose accuracy is not limited by the mosaic structure. The spherical eyeball shape is generated by functional geometry, and its almost perfect rotation operations can establish among the retinal cells an *external Euclidean metric of perception-space* which is independent of the distortions of mapping on the retina or the cortex.

A variety of second-stage and higher-stage neuroanatomical structures would have to be grown for tracking and detecting pattern repetitions. These would almost certainly include *delay lines* and *null-transmitter cells* to transmit only the identical parts of multiple input patterns.

Such *pattern-perception* in the mature network is equivalent to determination of the initially unknown space relationships or *addresses* of the random detector cells. A *non-addressed mosaic* requires much less initial assembly information than a *pre-addressed mosaic*, but requires a long *learning* and *growth* time for *address-determination* after operation begins. It has other quasi-human characteristics, since to determine addresses it *consumes information* in *abstracting properties*, draws *analogies*, shows *closure*, makes *symbols*, *learns from experience*, incorporates *functional memories* in the network structure, and apparently might even need to *sleep*. But the self-congruences of functional geometry would impose certain paradoxical and Kantian restrictions on the learning process, such that only certain congruent types of experience can be learned at all, and only certain congruent types of address-connections can be formed, regardless of what the experiences are.

THIS paper revolves around the problem of visual pattern perception by the human eye and brain. It is an attempt to generalize the problem; to restate it in a language suitable for electrical networks; and to see what basic physical principles might be involved, what detailed neural relationships might be required, and how these principles and relationships restrict and determine the general properties of such networks.

The eye has millions of simultaneously active photodetectors. The theory of connections in such a system is still in a primitive state. It is therefore necessary to begin by introducing and explaining a number of new terms which will be needed in the analysis.

I. MOSAIC RECEPTORS

Single-element and Multiple-element Receptor Systems

A feedback mechanism or a neural network or a social organization is a decision network connecting sensory-receptor inputs with motor-effector outputs. The system may have *single-element receptors* or *multiple-element receptors.* An example of a single-element receptor is a phototube. Another is a proprioceptive muscle spindle cell. In the simplest case each of these might actuate a single-channel feedback loop or reflex arc leading to a one-coordinate output function of time. There may be non-linear circuit elements in the loop that pulse or chop or clip or average or stabilize the input or otherwise transform it. Nevertheless, each feedback signal from a single-element receptor remains a one-dimensional time signal except as it may be trivially or artificially split into several components.

Multiple-element receptors consist of many functionally similar single-element receptors acting simultaneously. If each of these has its own private reflex arc, independent of the others, to its private motor output, the system is merely an additive system of single-element receptors. But to avoid conflict in the motor responses, it is desirable to reduce their independence. In this case, the simultaneous inputs can be combined in a *decision network* which selects a single complex response from the output field, with suppression of conflicting alternatives. Some of the physical and mathematical relationships in such a network were discussed earlier (1).

The receptor organ of such a system becomes a *mosaic receptor* with a pattern and hierarchy of connections to the decision network. It is an advantage if the network is concentrated into a compact central switchboard where extensive interconnections can be made quickly and cheaply.

Examples of mosaic receptors are the 10^8-element retina of the human eye, the basilar membrane of the ear, and the olfactory membrane. The retina wiii be treated as the prototype of such systems. Mechanical mosaic receptors have also been constructed such as the 10^2-element assembly of sensory pins in the reading head of a punch-card sorter or reader. A social mosaic receptor would be the 10^2 traveling salesmen sent out by a business organization. The relatively low complexity of these man-made systems means that they are inferior to their biological counterparts by more orders of magnitude than almost any other man-made devices.

It is true that some artificial receptor systems are more elaborate than the two mentioned. A television camera iconoscope tube with its 10^6 separate resolvable spots is an example. But at present, the iconoscope signals are scanned and sent in sequence into a single output channel, undergoing only the most rudimentary inter-comparisons or decisions, such as stabilization, contrast or color balance. Likewise the 10^9 grains of a photographic emulsion, although they form a very fine-grained system, do not feed into any decision network until they are transduced onto the human retina.

Pre-addressed and Non-addressed Mosaics

An *address*, in computer nomenclature, designates a point in the network at which a signal may be located. But in a mosaic receptor, the address of an

input element is only partly specified by its *network-address*. It is incomplete unless the location in space, or *space-address*, is also given, at least relative to the other elements, since both address components effect the kinds and combinations of messages sent through the network.

Mosaic receptors may be *pre-addressed* or *non-addressed*. In a pre-addressed system, each receptor element has a specified space address and network address. It is completely connected in a unique and permanent way to the decision net before the net begins to operate. In a non-addressed system, the space address of an element, or its network address, or both, may need to be determined after operation begins.

This may be the main difference between the insect eye and the human eye. The insect eye consists of a close-packed array of uniform receptor elements. Because of their uniformity, they lie in long parallel lines. Absolute genetic determination of the connections from each element to its neighbors and to the decision net might be easy: a pre-addressed system.

Straight lines in the field of view that fire all elements on one of the principal lines of such an array should be easy to distinguish from curved lines, if such a distinction were biologically useful. But straight lines in any other general direction would be hard to distinguish from curved, without very elaborate inter-connections in the network; and therein might lie the limitations of a pre-addressed system.

The human retina escapes this impasse. It appears to make no such distinction between straight lines in different directions. And indeed under a microscope the cones in our foveas appear to be close-packed but sufficiently non-uniform that no straight line arrangements are more than a few cones long (2).

Assembly Information

This useful randomness seems inevitable from assembly considerations. A non-random biomechanical assembly of 10^8 elements distributed over several square centimeters of the retina with individual tolerances of better than 1 micron would be almost inconceivable. Even if this could be achieved, the complexity of a non-random wiring diagram for any system of 10^8 input elements, geometrically regular or irregular, would be almost impossible for the chromosomes to specify, as Pitts has emphasized (3).

And so the randomness, if it has solved one dilemma, has evidently created another. The addresses of the retinal elements are now uncertain. All straight lines have been made equal by a device which appears to make it impossible for the eye to identify straight lines at all!

On the other hand, if this new problem could be solved—and the present paper aims to show that it can—non-addressed systems would evidently have one tremendous advantage over comparable pre-addressed systems: their economy of assembly information. In pre-addressed systems, if the interconnections among m elements are to be specified in advance, the assembly information must increase with a power of m at least as large as two and perhaps much larger.

This elaboration of initial design specification and mechanical assembly detail is what makes our artificial electronic networks slow and expensive to

manufacture. Sooner or later the increase with increasing m will limit the size of the pre-addressed systems we can construct, no matter how much the assembly process is speeded up.

But for a non-addressed system, even with 10^8 or 10^9 elements, a very few specifications of the general assembly or growth patterns may suffice (3). The construction is cheaper, whether measured in assembly information, in time or money. Obviously there is a price. It is that the addresses of all retinal elements must now be *learned—after* operations begin. The construction is speeded up; the attainment of full operating efficiency is delayed until address-determination is completed. But the non-addressed system constructed with a given amount of assembly information can eventually become far more complex and 'intelligent' than its pre-addressed counterpart.

This initial incompetence may be why, in evolution, the non-addressed organisms only become prominent when parental care appears in family systems like those of birds and mammals. The long learning time for large m might be connected with the long childhood of the more intelligent species.

Actually there may be no sharp boundary in biology between the pre-addressed and the non-addressed. On evolutionary grounds alone, a vitally necessary fraction of the human brain must certainly be pre-addressed. The autonomic nervous system may be largely so constructed. Reflex actions and probably color vision seem to have this character. The non-addressed sections of our networks, although perhaps responsible for our most characteristically human activities, may be a late and still secondary addition to a large pre-addressed core—as Dr. Sacher stressed in his comments on this paper.

It is often asserted that nerves and synaptic connections do not grow. This might be true for the pre-addressed sections; but it should be false for the non-addressed sections. Address-learning in a network necessarily means creation or change of connections. Change of neural connections means growth or atrophy or both. If new synaptic connections do not grow, they must at least be selectively and permanently activated or deactivated during the address-determining process.

The Pattern Question

Whatever the economy of assembly, the question remains: Can randomly arranged elements be used to make discriminations of straight lines or of any other types of *pattern elements?*

There is evidently an intimate connection between the *perception of pattern* and the *determination of the addresses* of the retinal elements. To make the question more precise, let us number the elements $123 \cdots j \cdots$ in as nearly the same way as possible in all retinas, and set up coordinate axes as nearly alike as possible. The randomness means that element j will have different address coordinates, x_j, y_j in every retina. Or better, we might specify addresses by relationships rather than coordinates, giving them forms such as 'Element j is collinear between elements g and p'. This address might be right in one retina, wrong in another.

Such an uncertainty of internal pattern has to be resolved *within* the network. The question is then: Can the coordinates $x_j y_j$ be determined, or can the straight-line or other geometrical spatial relations of element j to many other

elements $\cdots g \cdots p \cdots$ be determined, within the receptor network and solely by its normal functional operations? And if so, how?

The present paper aims to show that at least one simple method exists for this *functional determination* of addresses. It can be called the method of *functional geometry*. It seems feasible for use at least in an artificial mosaic system. It may or may not be the method used by the eye or by any other biological system, although many of the results here strongly suggest that it is. In any case, its existence removes a principal conceptual difficulty of non-addressed mosaic receptors. And the examination of one particular method can help sharpen up our experimental inquiries as to what methods of address-determination and pattern-perception actually are used in biological systems.

II. FUNCTIONAL GEOMETRY

There is a class of geometrical operations that is of great importance in the highest precision machine work and in anatomy, especially in the joints of vertebrates. The operations are related to group theory but, as we shall see, they might form the axiomatic basis of a separate systematic mathematical discipline. If this discipline were ever created, an appropriate name for it would be functional geometry.

Generation of Perfect Surfaces

An illustrative operation of this class is that by which an optician or an amateur telescope maker grinds and polishes a spherical lens or mirror surface (4). A rough blank of glass is placed against another rough blank of glass or metal or pitch, with grinding or polishing powder between them. The blanks are pressed and rubbed together by hand or by a rather crude and loose grinding machine, as shown schematically in Fig. 1A. The operation continues with

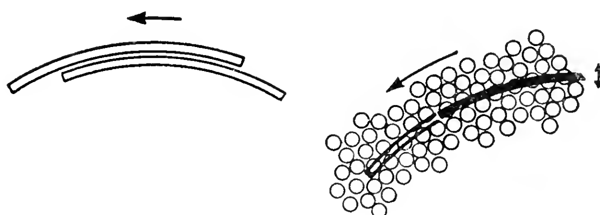


FIG. 1. Self-congruence of a sphere or a circular arc under random translation.

successively finer grades of powder. Finally each of the surfaces approaches a perfectly spherical shape to a precision which may be one-tenth of a wavelength of light, or better if desired.

Theoretically, if edge effects are neglected, *the method can approach infinite precision.* Its practical precision is limited only by the patience of the optician and the accuracy of available testing methods. The accuracy of approximation to a perfect sphere can be many orders of magnitude higher than the accuracy

of the initial blanks or the accuracy of construction or operation of the grinding machine.

Usually the optician also wants a particular curvature, convex or concave, but this is a separate question which need not concern us here. The curvature determination is not automatic and it is the automatic approach to perfection by these methods which is the point of interest.

In order to produce a spherical surface, the motions of the grinding machine must be (a) relative translation of the blanks in both coordinates along their surfaces, and (b) relative rotation of the surfaces. Each motion must be randomly independent of the others, relatively unconstrained by the machine. This is why the grinding machine must be loosely coupled. A grinding machine that couples the motions in any regular way or whose translations have some arbitrary fixed relation to the axis of rotation would 'over-determine' the system and damage the rate of approach to a spherical surface or the attainable precision. The surfaces are self-centering, determining their own centers more and more precisely as the polishing proceeds.

The reason these particular motions generate a spherical surface is that this is the *only* surface that satisfies the following *functional definition:* A spherical surface is one of two surfaces that is everywhere in contact regardless of relative translation or rotation against each other.

For one surface alone, this could be made a statement of *displacement congruence:* A *spherical surface* is self-congruent for all translations or rotations in the surface. A complete *sphere* is self-congruent for all rotations in the surface; that is, about any axis normal to the surface. (Three degrees of freedom. Any two rotational degrees of freedom imply the third.)

The functional geometry of such definitions is conceptually more fundamental than either Euclidean or analytic geometry. To say with Euclid that 'a spherical surface is a surface in which every point is at the same distance from a fixed point', is to require points, fixity and measures of distance. To say that 'the equation of a sphere is $x^2 + y^2 + z^2 = R^2$' is to require also a coordinate system. But functional geometry generates perfect surfaces by only using two of the most primitive notions: identity (congruence) and displacement.

The motions involved in these definitions are those of the continuous translation and rotation groups of group theory. The definitions can therefore be generalized to surfaces representing other group operations, including discrete groups:

> Real surfaces approaching indefinitely close to a mathematically perfect form can be generated by mechanical processes that enforce displacement self-congruence under a particular set of group operations. The set determines the shape of the surface. The surface is self-centering and defines its own special centers and axes in space more and more precisely as the operation proceeds.

In practice, what development of these other operations has been done has come from the makers of precision screws and ruling-engines, especially WHITWORTH, ROWLAND (5) and STRONG (6). STRONG emphasized the opposition between these 'inherently precise' methods (self-congruent surfaces) and the traditional 19th-century semi-precision methods of 'kinematic design' which he had described earlier (4), and the superiority of the self-congruent method.

'The construction methods of greatest precision are all primitive methods' (6). The following are some examples.

A *surface of revolution* is self-congruent for rotation about its axis. (One degree of freedom: Strong method for thrust bearings.)

A *screw* is self-congruent for simultaneous rotation about its axis and translation along it. (One degree of freedom: Rowland method of lapping.)

A *cylinder* is self-congruent for all rotations about its axis and translations along it. (Two degrees of freedom: Strong prescription for lapping a cylinder). A *cylindrical surface* section is self-congruent for pure translations in the surface with no component of rotation about a line normal to the surface.

A *gear* of n identical teeth, $360°/n$ apart in angle, is self-congruent for any of n different angular displacements about its axis. (One continuous degree of freedom plus one discrete. In the Strong method, the gear is polished within a kind of open-ended squirrel cage of n lapping bars or pawls that slide between the teeth. The cage is rotated by one bar after every stroke, and any initial irregularity in either the gear or the cage is polished away.) The group operations are those of the discrete group, C_n. Functional geometry can therefore generate perfect right angles or other angles.

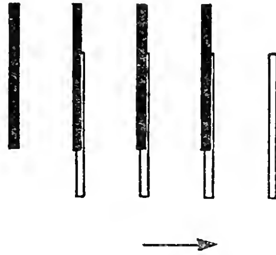


FIG. 2. Self-congruence in translational periodicity.

By analogy with the screw and the gear, a *cylindrical surface with straight parallel equally-spaced identical grooves* (possibly helical) is self-congruent for continuous translation in one direction in the surface and discrete translations in the other, as indicated in Fig. 2. (In principle, the precision ruling of surfaces might be accomplished in this way.) Perfect translational periodicities in two or three dimensions might be generated in succession.

There are more sophisticated possibilities on moving beyond ordinary group theory: A *plane* is one of *three* surfaces of which any pair can be placed in contact everywhere regardless of relative translations or rotations against each other. (In making optical flats by the Whitworth method of lapping, three flats are generated simultaneously by being polished against each other, with frequent interchange of pairs to prevent development of concave or convex surfaces).

Restated in terms of displacement congruences: *A plane* is self-congruent for all translations and rotations in the surface and for two-fold rotations about an axis in the surface. Note that three-fold rotations about such an axis would be impossible. This exemplifies a fundamental physical restriction on possible generating processes, of a kind we will encounter shortly in the biological cases.

The sophistication of this definition of a plane is that it is antecedent to the definition of a straight line in this geometry and requires no definitions of lines or axes or coordinate systems or rectilinear translations.

Other surfaces can be generated by grinding and lapping operations that maintain only a line of contact between two self-congruent surfaces, such as two surfaces of revolution rotating about skew-perpendicular axes.

Biological Examples

Any two physiological surfaces that are pressed and rubbed together continuously must exhibit displacement congruences approximating mathematical perfection.

The familiar chicken drumstick has at its lower end a perfect surface of revolution sweeping through an angle of about 270°. (One degree of freedom. It may be slightly helical, since the revolution is not complete.) The helical grooves on the narwhal tusk may be generated by displacement congruences as it grows from its socket. Ball-and-socket joints are likewise familiar in anatomy, with accurately spherical surfaces. (Three degrees of freedom.)

The eyeball-and-socket is perhaps the most perfect instance of this type. The spheres must be very precise if there are not to be considerable changes of pressure during normal rotations. The oculomotor musculature provides all three rotations, about the X-axis (vertical axis), the Y-axis (transverse horizontal), and the Z-axis (longitudinal horizontal). Functional geometry provides a precise self-centering specification of the center of the spheres and therefore of the reference point about which all the operations of the three-dimensional continuous rotation group can be carried out.

What is more important, these motions provide the necessary displacements by which *the displacement congruences of any pattern in the external field may be detected by the retina.*

To anticipate the results of the next section, if an arc in the external field produces an excitation pattern on the retina (Fig. 1B), the pattern can remain unchanged during a displacement of the fixation point along the arc if, and only if, the arc as seen from the eyeball is either a straight line or the arc of a perfect circle, with constant curvature. This is the two-dimensional analogue of the functional definition of a perfect sphere given above.

Likewise a set of lines in the field is parallel and equidistant if and only if the excitation pattern can remain unchanged as the fixation point moves from one line to the next or moves along the lines (Fig. 2). This is the analogue of the functional definition of a surface with parallel equally-spaced grooves.

These are indeed the kinds of pattern judgment that the human eye makes most precisely. Our peculiar sensitivity to changes of curvature and to non-parallelism and non-periodicity is well known in model-making and in pattern tracing and analysis.

An extreme case is the curvature-continuity judgment involved in 'vernier acuity'. If two ends of a line join imperfectly in the middle, the eye can still perceive the break when the lateral displacement is as small as 2 seconds of arc—1/30th the diameter of a retinal cone (7). Regardless of what neural connections might be needed to make such a discrimination, it is obvious that the judgment must depend on a physical operation of inherently high precision, inherently

unlimited by the coarseness of the mosaic structure and the randomness and uncertainty of cone locations.

Functional geometry offers such a method, since it can generate indefinitely high precision out of arbitrarily coarse materials crudely manipulated. With it, the practical limitation in precision could be a very refined signal-to-noise limitation, that is, an intensity-judgment-time limitation, as we shall see, and not a coarse mosaic structure limitation. For the internal as well as the external eye, it would be characteristic of biological systems to make use of such a method, conceptually simple, operationally precise, making only minimum demands on the accuracy of assembly, and capable of being driven to higher and higher precision as needed under the pressure of natural selection.

III. DETERMINATION OF ADDRESSES

A. Scanning in Vision

DITCHBURN and co-workers (8–10), and RIGGS and co-workers (11–12), have shown that vision disappears unless the field is continuously scanned by the eye. The scanning is normally provided by 'physiological nystagmus', or 'fixation tremor.' When a subject is fixating as steadily as possible, the following eye movements are present:

'(i) a tremor of amplitude of the order of 15 sec arc and frequency ranging from 30 to 80 c.p.s.

'(ii) a series of 'flicks' of up to 20 min arc occurring at irregular intervals ranging from 0.03 sec to 5.0 sec.

'(iii) slow drifts in the intervals between flicks.' (DITCHBURN (9)).

The movements are involuntary. They continue undiminished even when an image has been stabilized on the retina, so that they do not seem to have quantitative feedback character, at least for fixation of a point source; but the flicks do tend to produce recentering after the image begins to drift off the fovea.

The frequency, amplitude and sequence of the movements as presently known would be consistent with assigning the tremor jerks to the successive single neural spike inputs in the normal trains of spike pulses to the ocular muscles; with assigning drift to the unbalance between these jerks in opposed muscles at slightly different spike frequencies; and with assigning flick to a final sudden burst of spikes to the less active muscle which redresses the unbalance and recenters the system. There is no vision during the flick movement. This demonstrates an intimate oculomotor interaction with the retinal output, complementary to the interaction which will be postulated later.

When these movements are stopped by optically stabilizing the retinal image, vision is lost within a second or two. It can be restored by flicker modulation of the light intensity or by reintroduction of some image movement.

The necessity for scanning in maintaining vision might have been anticipated. It is probably no surprise to a biologist to find such a mechanism used to counteract the effects of adaptation and fatigue in receptors that need to be continuously sensitive; nor to a biochemist to find that the electrochemical shock waves corresponding to nerve impulses are reduced in frequency as the photochemical steady state is approached; nor to a physicist to find that a.c. operation of a phototube is the best way to avoid d.c. drifting. Many retinal

potential and neural spike results obtained with steady-state illumination may have to be reexamined for their relevance to the process of vision.

Insects and amphibia and other lower orders seem to hold their heads and eyes rigid for long periods. The need for scanning suggests this might permit selective detection of moving objects in the field. (The insect eye perception problem treated earlier was an artificial way of pointing up the general pattern problem, and not an attempt to describe the real workings of the insect eye.) At some point up the scale, a scanning tremor in the eye might have been an evolutionary predecessor of wide-angle motion.

(It is not only vision that requires 'scanning'. A variation of input stimulus is needed to maintain sensitivity in touch and in smell. This strongly suggests that a search be made for a similar mechanism in hearing, by which the 'sound image' might be scanned up and down the basilar membrane to provide continual change of stimulation, to prevent local fatigue, and to sharpen tonal discrimination.)

B. Determination of Addresses

In any field of study, it is always a hopeful sign to find two or more unexplainable effects and not just one; for this opens up the possibility that the two will explain each other. In the retina, we are confronted first with the pattern-perception address-determination problem and then with the strange importance of scanning. Putting these together, it appears that scanning might be a particularly straightforward method for functional determination of addresses in a non-addressed mosaic receptor. And this is functional geometry. Several theorems suggest themselves.

The Fundamental Operations

1. *Sequence of Elements*—During random scanning over visual fields containing some structure such as sharp discontinuities or boundaries, if retinal elements *i*, *j*, *k* are triggered in similar patterns in succession far more often in the time-sequences *ijk* or *kji* than in the sequences *jik*, *jki*, *ikj*, or *kij*, then:

(1a) there are some boundaries in the external field that are relatively stable during the scanning motion;

(1b) *j* lies on the image of a point in the field between the corresponding points for *i* and *k*; and

(1c) the eye movement for one of the sequences *ijk* is opposite to that for the other *kji*.

2. *Collinearity*—During random scanning over visual fields containing sharp boundaries, if all the elements in a certain large set *fgh* · · · *k* are excited simultaneously in the same way (d.c.; or a.c., as by tremor across a boundary) and if this excitation continues unchanged throughout a short drift movement, then:

(2a) there is a linear boundary in the field;

(2b) elements *fgh* · · · *k* lie on the image of that boundary; and

(2c) the drift movement is parallel to that boundary.

The photodetector inputs produced by tremor movement could provide a *gradient discrimination* across the boundary which, when combined with drift, as suggested in Fig. 1 for a curved line, could give an especially delicate determination of addresses (3). Thus, for cells distributed roughly along the image, one traverse might produce firing in a reproducible sequence *fkghjfigf* · · ·, the

sequence depending on some function of the relative transverse cell displacements, sensitivities and repetition rates, and on the image boundary gradient. 'Unchanged excitation' would mean successive recurrences of this same sequence, and changes in the sequence could correspond to changes in the image amounting to only a small fraction of a cell diameter. The *gradient-discriminating power* would then be limited essentially by signal-to-noise considerations rather than by mosaic structure and it might be far higher than the static mosaic resolving power, as numerous authors have suggested. *The transmitted self-congruence signal, whatever it is, need contain no trace of the static mosaic structural irregularities. It is also independent of differences in the sensitivity of different receptor cells and could remain unchanged even if a few of them should fail completely* (*closure*). These would be important biological advantages for the self-congruence method of address-determination.

2′. *Parallelism*—If the elements $fgh \cdots k$ of Operation 2 also are grouped into r subsets each of whose excitations can be duplicated for r different transverse displacements, with a different set of displacements for each subset of elements, then:

(2a′) there are r parallel linear boundaries in the field;
(2b′) each subset lies on the image of one of these boundaries; and
(2c′) the first drift movement is parallel to the boundaries, while the discrete transverse displacements are not.

It is typical of functional geometry that it *simultaneously* limits (2a) the type of external pattern that can be interpreted (2b), the type of internal relationship that can be organized, and (2c) the operational motions that can produce a coincidence of the two. This situation for pattern structure is no different from that for the eyeball, where functional geometry simultaneously limits the shape of the external socket, the shape of the ball, and the possible movements and musculatures. We shall see over and over that the functional geometry, if it is the address-determining method, is neither experience nor structure but stands outside them both, imposing an inescapably limited selection of forms on the only experiences we can perceive and the only structures we can create.

Point (2c), the establishment of retinal relationships and proprioceptive oculomotor signals relative to each other, as suggested by HELMHOLTZ, is not the least important aspect of address determination, now that proprioceptive muscle spindles are known to be present (13, 14).

Note that it is the boundaries in the external *field that are linear, and* not *their retinal images, when self-congruence is the method of discrimination. Likewise the projections on the cerebral cortex can have any kind of twist, distortion or discontinuity—which they have—without destroying a* functional *definition of collinearity and parallelism.*

The ambiguous word 'linear' is used in theorems (2) and (2′) so as to postpone for a moment the question of how well these procedures will distinguish a perfectly straight line from a perfect circular arc of very slight curvature. But aside from that question, a mosaic detector is seen to be in principle far more accurate than a single-element detector. With the latter, straight lines could be discriminated by tracing them out, perhaps using small hunting movements superimposed on a long sweep, but the accuracy is limited by the accuracy of

the analog position-sensing circuits. With mosaic detectors, the sensitivity to imperfections in a line pattern is not affected by analog errors.

In these theorems, a discriminated boundary will be perfectly straight if the longitudinal Z-axis of the eyeball passes through it and if there is no Z rotation during the scanning motion. Eye movements about the Z axis during fixation seem not to have been measured as yet, but it seems unlikely that physiological tremor about X and Y would be unaccompanied by tremor about Z. It will be convenient here to consider the analogue of these theorems for pure rotation about Z, with no X and Y components, and to come back later to the question of how the present theorems will be changed, and what new theorems will be valid if all three rotations are present.

3. *Circularity*—During pure rotational scanning about the Z-axis over visual fields containing sharp boundaries, if all the elements in a certain large set $fgh \cdots k$ are stimulated in the same way and if the stimulation continues unchanged throughout this kind of scan, then:

(3a) there is a boundary in the field which is a circle or circular arc as seen from the eye;
(3b) elements $fgh \cdots k$ lie on the image of that boundary; and
(3c) the-Z axis of the rotation passes through the center of the circle.

3′. *Concentricity*—If the elements $fgh \cdots k$ of Operation 3 are grouped into r subsets whose elements can be re-excited in the same local patterns by discrete sets of X, Y rotations (in analogy with the Operation of 2′) then:

(3a′) there are r concentric circular boundaries in the field; and so on.

Concentricity is also one of our very delicate discriminations, as is shown by many gunsight designs.

4. *Translational Comparison*—During a random scanning drift movement (in X and Y alone) over visual fields containing sharp boundaries, if a certain time pattern of excitation of elements $bcd \cdots$ is repeated after a certain fixed time-delay (that is, a certain displacement) by elements $fgh \cdots$ in a one-to-one correspondence with the $bcd \cdots$ excitation pattern, then:

(4a) there is a stable pattern, fixed or undergoing translation, in the external field; and
(4b) there is a constant translational separation in the field between points whose images fall on elements b and f, c and g, d and h; and so on.

4′. *Translational Periodicity*—If the elements $bcd \cdots fgh \cdots$ of Operation 4 can be divided into r subsets, where r is greater than 2, each of whose excitations can be duplicated for r different displacements, with the excitations of several subsets simultaneously duplicated for certain displacements, then

(4a′) there is a stable translationally periodic pattern in the field, with r repetitions; and so on.

Translational comparison is of course the theoretical procedure for establishing a metric in a space of unknown geometry. The precision of translational inter-comparisons between the patterns in the two eyes is the basis of depth perception. Under the most favorable conditions it approaches the same high angular precision that was found in vernier acuity. This fact alone seems to require a physical operation that can create a high-precision *translational* metric

for the retinal elements: an operation for each eye that can translate elements with fixed relations from one part of the field to another—that is, scan—during the intercomparison process.

A figure with bilateral symmetry, whenever its median line is defined, has translational periodicity for lateral scanning. This might be the basis for whatever accuracy the human eye has in judging such symmetry.

5. *Angular Comparison*—During pure rotational scanning about the Z-axis in fields containing sharp boundaries, if a certain time pattern of excitation of elements $bcd \cdots$ is repeated after a certain fixed time delay by elements $fgh \cdots$, then:

(5a) there is a stable pattern in the field, fixed or rotating about the Z-axis;

(5b) there is a constant angular separation, with respect to the Z-axis, between elements b and f, c and g, d and h; and so on.

5. *Angular Periodicity*—If a relation like 4′ is satisfied for pure rotational Z displacements, then:

(5a′) there is a stable angularly periodic pattern in the field, with r repetitions; and so on.

Because of the limited range of Z-rotation in the human eye (about 20°), our perception of angular periodicities may lose precision rapidly for larger angles. Some of this acuity may be recovered by treating the angular periodicity as a bilateral symmetry, converting the judgment into one of lateral translational periodicity.

The metric of the 'space' of the addresses established by these operations is that of the rotation-space of the eyeball and not that of the retina or cortex surface.

All these operations have been internal operations, specified so as to depend only on the internal properties of the decision net and its scanning system, and to be as independent as possible of the object and properties of the external field except for the minimum requirement that there is some variety of structure and that at least sometimes its changes and motions are slow compared to those of the eyeball.

Displacements and motions in the external field could lead to another similar list of theorems which would establish similarly an *external metric* and an expected external behavior, whose familiar translational and other constancies—comparable to the congruences produced by the internal operations—we might finally interpret as invariant 'objects' in the field, uniform motions, and so on. The external metric may or may not be consistent with the eyeball-rotational metric. Probably the external metric is the primitive one, with the scanning metric providing a sophisticated refinement. Inconsistency between the two may be the source of many optical illusions.

But since theorems (1) to (5) suffice to establish in several different ways that consistent address determination within the network is at least physically possible it seems more important to go on now to see how it would be anatomically possible.

C. *Possible Types of Neural Connections Required*

Proprioceptive Coordinate Specification

What anatomical connections are needed for proprioceptive sensing and control? The requirements and some possible ways of solving them, at least

for an artificial system with some quasi-neuronal properties, will be seen if we consider the problem of scanning along a curved line, with Z-rotation of the retina to follow the curve, as in Fig. 1.

If the differential muscle stress or strain or rate of change of either (whichever is the principal sensed variable) about the Y-axis has a fixed ratio r to that about the X-axis, the eye will sweep up along a line at an angle *arctan* r to the horizon. If the differential neural spike frequencies f_x and f_y from the two muscle pairs give a quasi-logarithmic representation of the muscle action, a fixed ratio r corresponds to a fixed frequency difference, $f_x - f_y$, which we can call F. A subtractive-frequency mixer tube, and perhaps a similar *subtractive mixer neuron*, could be devised which would combine two synaptic inputs so that an output pulse is produced only when the input pulses are simultaneous, as suggested in Fig. 3. With suitable cell sensitivity and time constant, this output is the beat

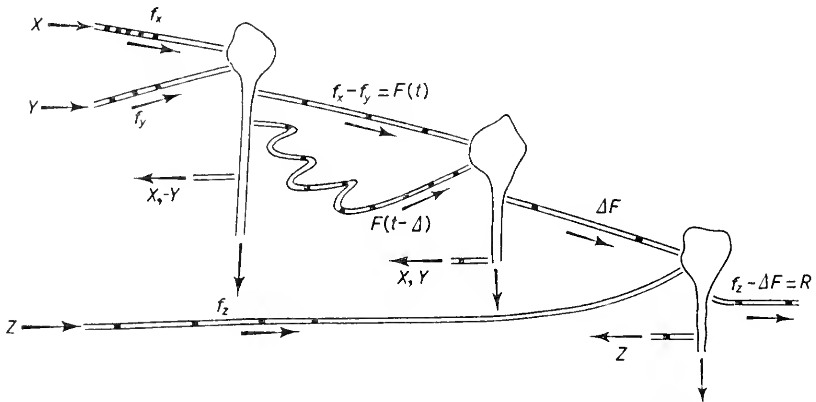


FIG. 3. Possible proprioceptive analog connections for scanning along a uniform curve.

frequency, F, the difference of the two input frequencies (1). A neuron used in this way might be called a *difference cell*. (Determination of the sign of the difference might require another cell.) If the output frequency F is fed back with proper sign to the X and Y muscles, a constant direction of motion can be stabilized.

The rate of change of direction could be sensed by introducing a *time delay* and comparing $r(t)$ with $r(t - \Delta)$ or $F(t)$ with $F(t - \Delta)$ by a second subtractive neuron which generates the frequency ΔF—a *time-differential cell*. With a lower sensitivity and a shorter time constant, the same type of cell could be made to fire only for a certain constant pulse interval in the inputs, equal to the time delay. This would be a constant-frequency detector, and could be called a *null cell*.

A constant Z-rotation of the retina to follow the change of X, Y direction in scanning a curve could be detected by a third subtractive neuron which generates the frequency difference $f_z - F$. Call this frequency R. It can again be held at a constant value by suitable feed back into the Z motion, as shown at the right side of Fig. 3. Such a tracking motion permits generalization of theorem (3) and (3′):

3″. *Decentered Circularity*—During combined translational and rotational scanning, if the conditions of theorem (3) or theorem (3′) are satisfied, then (3a) and (3b) or (3a′) and (3b′) are valid; but:

(3c″) the Z-axis of the rotation does not pass through the center of curvature of the boundary or boundaries.

The oculomotor motions that could be directed by this kind of analog control correspond to the crude motions of the grinding machine in generating spherical surfaces. They do not need to be exact. They need only to be capable of making retinal displacements across the field sufficiently good that during the tremor movements the retinal excitation will have an adequate chance to signal if it is congruent with the original pattern. This signal could also interact with the oculomotor system to make the scanning more stable and accurate. In practice, visual acuity for moving patterns is considerably reduced (15), presumably because of the increased tracking errors and decreased chance of a congruence signal.

We can now see the effects of combined motions on theorems (2) and (3). Evidently Operation (3), the examination of circles, gives a functional self-centering specification of the axis of rotation, even if it is off the Z-axis, whenever congruence is maintained. Any tremor about other axes simply provides useful scanning motions.

But Operations (2) and (2′), the examination of collinearity and parallelism, cannot discriminate perfect straight lines from perfect circular arcs of large radius except by invoking the accuracy of the sensing and analogue control, a discrimination of much lower accuracy than the mosaic self-congruence discriminations. It seems that our perception of such differences is in fact small unless there are known straight lines nearby permitting a self-congruence test for parallelism. There is a familiar illusion in which a comparison straight line appears curved in the opposite direction from a curved line. This shows an uncertainty in the analogue system, which tends to scan along the bisector so as to give the figure bilateral symmetry.

The Z rotations of the eyeball during scanning of curved patterns evidently deserve examination. In the classical Zöllner illusion (parallel lines appear to be non-parallel when crossed by oblique converging lines) there might also be a Z-rotation of the retinal coordinate system, perhaps in the sense of stabilizing the local foveal pattern and the local bilateral symmetry axis as the fixation point oscillates from one of the parallel lines to the other.

With further combinations of difference cells and time-differentiation cells and feedbacks, probably the tracing out of any pattern by scanning could be converted at a high enough stage into a constant output from some subtractive neuron. If adjustments of the scanning rates at various points in the pattern are also introduced, probably changes of size and distortions of shape could even be accommodated in a constant output at a still higher stage neuron. We can dimly visualize how this might proceed by stages to a neuron capable of producing a fixed output, or better, a total motion of advance or retreat, whenever so specific an object as a particular person is scanned, regardless of aspect and light.

Even if in later life the proprioceptive tracing of patterns by scanning becomes subordinate to direct mosaic pattern-perception, the long stages of

finger tracing of block letters and large patterns by children and newly-sighted adults suggests its early importance. Studies of the developmental pathology of pattern-perception with partial oculomotor paralysis might be instructive.

Null Detectors and Delay Lines

What kinds of neural connections might be needed in the mosaic to determine addresses as these operations are performed?

The discriminating self-congruence information is always of the form 'constant repetition of the same pattern' or 'repetition after a time delay'. What is needed from the photodetectors is the information 'absence of change'. They are being used as *null detectors*. Unstable detector elements in the laboratory are often used in the same way whenever the utmost accuracy of measurement and simplicity of interpretation is wanted. The rather complex relationships that can be established when using *mosaic* receptors as null detectors seem not to have been explored before.

To signal 'no change' we could use a null cell of the type already described. But another good way to examine stability of pattern would be to have a cell with two input channels of different lengths, like two of the channels in Fig. 4,

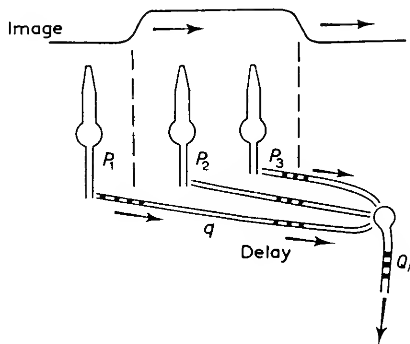


FIG. 4. Cone addresses from delayed coincident pulses to a null-transmitting velocity-detector cell.

or of different diameters and travel times, where the cell sensitivity is such as to require simultaneous spikes from both channels in order to produce an outgoing spike in its axon. (The word 'channels' is used to avoid the experimentally unsettled question as to whether these could be all-or-none dendrites of the cell, if such exist, or two excitatory synapses with different delays in the axons, or collateral processes from the cells of the preceding stage.)

Such input channels would be *delay lines* like those used in nuclear physics to distinguish certain events and particles, to eliminate random spurious counts, and to measure velocities. They might be used for all these purposes here. The axon output of such a cell, as shown in Fig. 4, combines three important properties. It is (a) a null indicator, firing only for those patterns that are identical in the input channels. It is (b) an indicator of a particular delay time or difference of times in the channels. And it is (c) a pattern transmitter, since the pattern is not lost. Let us call this a *null-transmitter cell*, and if the

delay is different from zero, a *delay cell.* Such cells would have several possible special applications.

In the auditory system, a delay cell with binaural channels would be a useful *direction-indicator*.

In the retina, a second-stage delay cell like that in Fig. 4 would indicate a particular image velocity-component in the plane of the paper, and could be called a *velocity detector cell.* One with several inappropriate delays might be sensitive to almost any motion or flicker. This may be one of the functions of the widely branching horizontal cells that are so numerous at the periphery of the retina, since this is a region particularly sensitive to motion.

Velocity detector cells would give useful correction signals to the oculomotor system.

Under Operation 1, it is easy to see how delay cells with input channels from retinal elements i, j, k might be preserved in the organism if their time lags are in either the spatial sequence ijk or the sequence kji, but might atrophy from disuse or at least rearrange their channels if their time lags are in any other sequence. And the ijk delay cells would be a different group from the kji cells. Such a principle of natural selection and differentiation might be applicable to all types of second-stage and higher-stage cells.

It may be profitable to examine these or other kinds of time-delay connections in trying to make a model of color-vision, since it now appears that this may involve a comparison of signals from cones at different times as the photochemical substance in each one goes through some time sequence of spectral transformations under illumination and perhaps under scanning.

To summarize these exploratory notions, it appears that the types of neural connections that would be useful for address determination, at least in an artificial system, would include: difference cells; time-differential cells; null cells; null-transmitter cells; delay lines and delay cells; and velocity-component cells.

The outputs of such second and third-stage cells apparently can signal all the self-congruences required in the basic operations of functional geometry. All the geometrical patterns defined by local group theory congruences can be signaled without using the retinal elements in any way except as null detectors.

If natural selection favors those cells, together with their oculomotor connections, which signal repeated self-congruences under scanning; then in the mature organism each retinal cell will feed into many second-stage cells, each of which expresses a useful functional relationship between that retinal cell and some others. The address of the cell has indeed been *determined.* The mature network-address becomes an expression of the space-address.

Evidently the functional geometry of scanning a pattern is a way of converting its space congruences into identical time patterns. It therefore could be said, as some have said (16), that in the mature organism the appearance of a certain time pattern at a certain point in the network creates an 'expectation' of its repetition at an adjacent point, and stimulates oculomotor and other movements normally appropriate for the accurate fulfillment of this expectation.

The accuracy of address-determination in these operations depends on the temporal accuracy of the delay lines and not on the spatial structure, which can be largely eliminated from the congruence signals, and therefore from the perceived patterns.

In the adult, the need for continuous redetermination of addresses becomes smaller. As Dr. W. A. Arnold of the Oak Ridge National Laboratory pointed out when this paper was first given, the words on a printed page that is illuminated for only 10^{-4} sec, too short for any scanning, can be read (by anyone who can read) quite normally over an area comparable with the foveal diameter. The addresses have already been established and need little if any reconfirmation from the operations that generated them.

How complete and rigid they are might be discovered if we knew the limits of developmental distortion of the retina after, say, the age of 3. Or if adults could work with optical systems giving subtle distortions of a few minutes of arc in the shape and topology of patterns in the foveal region, to see the effect of the loss of addresses on line, circle and pattern perception, on reading and the identification of persons, and how soon—and how far back in the network—a new set of addresses would be learned.

IV. NECESSARY PROPERTIES OF NON-ADDRESSED SYSTEMS

Certain properties would seem to be necessary characteristics of at least the early stages of all non-addressed and address-determining networks. What is interesting is that many of these seem to be familiar aspects of higher human behavior, restated in receptor-network terms. We may properly inquire how far our more complex activities can be subsumed under a generalized functional geometry, and how far our more complex experiences are organized by means of generalized displacement congruences.

A. Operational Characteristics

The null-transmitter delay cell which it seems necessary to invoke for any network making time-delay comparisons of patterns would be a suitable prototype for much of our higher neural organization. In the mature organism, after the structure and connections of such a cell have been stabilized—that is, after the addresses of its input channels have been determined—the cell will always collate two or more input patterns in a standard way to produce a simplified vital output. Let us focus attention, first, on the nature of the outputs, then on the inputs, and finally on the process as a whole.

Abstraction of Invariant Pattern Properties

The output of each such cell signals to the higher stages the presence of some particular kind of simple or complex pattern in the lower stages. This implies in turn a pattern in the first-stage images, representing a pattern in the *external field*. The process is *abstraction*. Pattern is another name for congruences or *invariances* in this field.

At a given instant, some fifth or tenth stage cell may be signaling 'That is the letter R'. Simultaneously some other set of elements and delay lines is abstracting from the *same* retinal elements the information 'It is in my wife's handwriting'. A third neuron connected to these elements says, 'It is black'; a fourth, 'It is large'; and so on, for details and context and background and all the other components. Perhaps some still higher neuron also signals the unification of these separate pattern properties and others into a word,

glanced past in a tenth of a second. These signals have a one-to-one correspondence with Platonic *properties* of 'R-ness', 'Blackness', 'Largeness', and so on.

Of course, a pre-addressed network might also give the same kind of information that any of these neurons gives. Or it might give information trivial or inscrutable for us, such as 'Slope of sharpest edge, 103° 8′'; or 'Five corners, two arcs concave to the left'.

In fact, *any* neuronal output in any connected network could be thought of as indicating *some* kind of pattern invariance. In this sense there are *no* addresses to learn! But most of these invariances in an arbitrary synthetic network would be worthless for biological survival. A major evolutionary problem for non-addressed systems must have been the facilitation of principles of connection leading to the appearance of cells, like the delay cell, capable of perceiving *useful* invariances, color, velocity, topology, and so on.

Every internal invariance or imposed relationship of signals in time and space gives rise to essential redundancies that can be eliminated from the higher-order signals with no loss of external invariance information. There is a reduction by a factor of about 10^2 between the 10^8 elements of the retina and the 10^6 elements of the optic nerve. Possibly this represents the elimination of some of the redundant scanning constancies of types such as those described in Section III that are implicit in the oculomotor operations and in the constraints of the kinematic rotational metric. These regularities would then acquire an inescapable *a priori* character so far as the higher operations of the network are concerned.

The external field may also contain many redundancies that are not important in a given situation. For many purposes it suffices to know that the animal is a wet, friendly dog, and we do not need the concurrent retinal information that he is opaque and continuous and in contact with the sidewalk.

We have not considered here how such an 'attention' to certain patterns and suppression of others might take place. However, the facilitation of signals through one neuron by means of a change of its sensitivity produced by feedback from the 'expectations' of a higher-order neuron (representing an earlier wet-dog experience pattern) may not be different in principle from the facilitation of oculomotor tracking movements by feedback from the 'expectations' of second-stage or third-stage retinal velocity-detector cells. It may be helpful in many problems to think of attention and expectation as generalized tracking devices.

The elimination, first, of field information that does not fit into the familiar useful second-stage patterns or categories, then of the redundant internal patterns, and finally of the temporarily unimportant and unattended-to external patterns, shows qualitatively how and why information is *consumed* in the course of abstracting invariances from a mosaic receptor (1). It is not lost or damaged by transmission in the sense usually considered in single-channel communication theory. Instead, it is used up, somewhat in the way that energy is used up in doing mechanical work. More mosaic input information is consumed in abstracting out a higher-level decision or invariance than in a lower-level one. It is consumed in the sense that the detailed input information is irrecoverable, non-reconstructable, from the output. Only by the almost

impossible process of applying m independent abstracted relations simultaneously could the m independent inputs be inferred. But this consumption or 'loss' of information is not biological loss but a gain, since it represents the selection of the biologically relevant item from the confusing irrelevant flux.

After a lifetime of suppression of the less valuable pattern and field details, adults finally attend only to the later neuron outputs or abstractions and seem to lose the eidetic ability to bring forth the exact early-stage patterns of instantaneous retinal excitation, except as they can reconstruct them approximately from their appropriate or inappropriate collection of output neurons. This may show the cessation of early-stage rearrangements, which finally become completely pre-addressed as far as new experiences are concerned.

Analogy Perception

We may look not only at what has been abstracted—the outputs—but at what has been compared—the inputs. The elementary process in address-determination was the *comparison* of excitation patterns at two different times. A network whose neurons can signal identities or similarities of pattern—displacement congruences—is an *analogy-perceiving* network. Much, if not all, of what we call intelligence may be the ability to perceive successive analogies at higher and higher levels of abstraction, a multiple repetition of a single basic neural process of organization.

Artificial pre-addressed systems are not generally able to perceive any analogies except between those sets of inputs that they are wired up to treat as equivalent. The value of mechanical mosaic detectors such as the punch-card reading-head lies in the fact that they are wired up to perceive obscure informational analogies and not any of the space or time pattern analogies of the kind that we perform easily in retinal abstraction.

Thoughts and Symbols

Pattern and analogy perception resemble some important aspects of the higher-order process we call thought. We might make a limited definition of a thought as the realization of previously unperceived pattern-relationships. A thought could then be represented by an operator equation,

$$P_2{}^m q P_1{}^n = Q_P{}^{m+1},$$

where $P_1{}^n$ is a pattern of the nth stage of abstraction; $P_2{}^m$ is one of the mth stage; q is the time delay or other transformation operator which relates pattern P_2 to P_1. And Q_P is the $(n+1)$st or $(m+1)$st (whichever is higher) stage of relationship; it is the pattern of the P's, a pattern of patterns. It is primarily a realization signal—a displacement-congruence signal—but it may also contain some or all of the common elements of the P patterns. We might distinguish if necessary between the possibility of the thought and the continued existence of the thought-relationship; and between the *insight*, or first assertion of the thought, and the repeated use of the established thought.

If $P_2{}'^m$ has no perceivable relationship to $P_1{}^n$, there is no thought,

$$P_2{}'^m q P_1{}^n = 0 \qquad \text{for all } q.$$

The next stage of thought might be to perceive a relationship between two different Q's,

$$Q_2^{m+1} r Q_1^{n+1} = R Q^{m+2}$$

and so on to stages of any order.

This operator-form of the equation seems to be the simplest familiar form that expresses all the required relationships. It suggests that Q can be regarded as a characteristic value or output value of the relationship operator when the latter connects the state-functions P. Or Q is a *symbol* of the relationship between P_1 and P_2. The equations suggest the possible usefulness of a formal *calculus of abstraction.*

In the first equation, if either P_1 or P_2 or q is changed, Q is different and generally vanishes. A little reflection will show that it is typical of thought, as it is of such operator equations, that there is only a restricted class of pairs of P's that have *any* relation to each other. The q's are sharply limited at the same time as the P's; catalogues of the possible q's have been made by various philosophers. And there is only a restricted class of realization signals, Q, in any case. If the equation is to have a non-vanishing value, it imposes simultaneous restrictions on all four variables. Some P's may never show any congruence. Some q's may operate forever in vain. And the Q's may take one or two values so repeatedly that they become independent of what particular P's are present.

The significant thing here is that these same equations would also describe many of the processes and structures in an address-determining network. So, in the functional determination process, q could be the functional geometry displacement operation, P_1 and P_2 the geometrical surfaces or the geometrical patterns of excited cells before and after the operation, and Q the signal of self-congruence.

In the velocity-component cell of Fig. 4, q could be the delay operation, P_1 and P_2 the pulse patterns in two of the input channels, and Q the coincidence signal output. In the neuroanatomical structure of the same hypothetical cell, q represents the delay line or lines, P_1 and P_2 the input cells of the previous stage, and Q the cell itself or its output axon or other output processes.

At a higher stage of the network, we might imagine that each of these structural elements in a particular neuron may also be connected to a verbal motor stage, with the relationships among these structural connections again represented by the same equations; and probably this structural parallel would be repeated in the language structure among the words themselves.

An address-determining system therefore seems to be necessarily a symbol-creating system. Whether regarded from the process aspect or the signal aspect or the structure aspect, a relationship or pattern of patterns in each case becomes represented by a symbol.

The close parallel between neuronal transmission and logical operations has been discussed for many years and is of great importance in the computational and logical performance of decision nets. But the present equations and hypothetical models suggest that analogy-perception, 'closure', 'insight' and other apparent 'extrapolations' from the known—in short, thinking—

may also be a necessary normal and even rather simple aspect of the neural connection process in any system that can determine its addresses.

B. Growth Characteristics

Learning

We saw earlier that a non-addressed system must be initially incompetent and needs a long learning time. This learning or address-determination requires inputs containing pattern regularities—that is, *experience*. For the retina, the experience may be generated by the external environment alone, or from this environment as scanned by the eyeball; in either case it is external to the retina.

In either case it generates spaces and metrics independent of the retina. Scanning is probably the visual counterpart of exploratory oral and manual manipulation which defines the 'spaces' of taste and touch. Probably the 'externality' of the visual metric, plus the simplicity and universal identity of the scanning operations of all eyeball-spheres about their centers, help to account both for the Kantian *a priori* character, and for the public and universal character, of visual space. This is contrasted with the situation, for example, in vocal or tone-quality space, which depends on the complex interaction of hidden muscular movements, and is perhaps the most incommunicable of our public spaces.

The network can learn only those types of regularities it has experienced. Two networks should develop somewhat different pattern perceptions if their environmental regularities or scanning schemes are systematically altered. A non-addressed network which is forced to operate for a long time in a structureless environment, like a blindfolded and insulated animal or human in the RIESEN and HEBB experiments, should and does have seriously defective pattern-perception and response. One can see how the formation of simple and accurate early-stage addresses in a network would be very important in facilitating fast accurate pattern-perception at later stages.

This picture of non-addressed learning exemplifies HEBB's conclusion (17) that many adult pattern-perceptions having introspectively the most instinctive and self-evident or necessary character are in fact perceptions that had to be learned at some very early age. It is early experience that selects the address-connections that are to be permanent; it is the permanent address-connections that create expectations and pattern-organizations in later experience.

Nevertheless, there is a double paradox in the present picture. (a) There are possible external input experiences that *cannot* determine address connections. And (b) There are internal pattern-perceptions, just as there are eyeball-shapes, which have grown or have been learned and yet have *not* been determined by the particular experiences or motions that contributed to the learning process.

Both these conclusions would follow from the operator equations of the last section that were supposed to represent network structure. The first point is obvious and almost trivial. External field patterns and their images are usable only if they fall within the limitations set by the network growth mechanisms and assembly principles. Motions of too high a velocity, patterns of far too coarse or far too fine a structure, fields of diffuse clouds with no

sharp boundaries, dazzle patterns with the proper kaleidoscopic confusions of rapidly disappearing and reappearing spots, the diabolical fields of DITCHBURN that refuse to be scanned—in fact, any patterns that deny the analogies that the network is prepared to detect—are probably all equivalent to a structureless environment in their failure to produce organization behind the retina.

The second point becomes obvious when we consider functional geometry and the profound limitations it imposes. Just as the lens-grinding machine with sufficient freedom of motion necessarily produces spherical surfaces no matter what it starts with, so the scanning eyeball necessarily generates 'external' space relations or addresses corresponding to the continuous three-dimensional rotation group, no matter what is the structure of the external field. Likewise at the retinal level, any scanning retina necessarily acquires a unique perception for continuous lines of constant curvature and for parallel lines and lines or points periodically spaced, which it can never accord to patterns violating these displacement-congruences.

These natural congruence relations may play the same organizing and aesthetic role in vision that octaves and simple frequency relations play in hearing.

It is peculiar to functional geometry and it is extremely important for philosophy that these necessary relations are neither given to the visual system by any particular external field or experience, nor are they logically implicit in the structure of the network, even when we include the analogy-detecting structure. They are Q's that, like spheres, turn up invariably, no matter what the P's or q's. They have rather the character of geometrical preconditions simultaneously imposed on both the external field and the network organization if *any* learning is to be possible. And they are not imposed *by* the scanning operation, even though it does mediate between the field and the network—any more than the spherical shape is imposed on the lens *by* the loose random grinding machine, or on the eyeball *by* the muscles. They are more like *a priori* requirements, mathematical absolutes, that determine the only kinds of experience that can be organized and the only forms that learning can take, if learning is to be done at all.

Functional geometry thus may be the origin of the Kantian epistemological limitations on thinking, as represented by 'the synthetic *a priori* categories of the apperceptive dialectic.' PITTS has described this as 'perhaps the most fundamental problem of neurophysiology and psychology' (3). It appears that much, if not all, of Kant's theory of knowledge can be translated word for word into the language of inputs, structures and geometries in an address-determining network.

Functional Storage

A neuron that has been selected by experience, so that its output signals an experienced pattern, constitutes a *storage* of the experience—a memory. The storage is not 'dead storage' but a functional link which permanently changes the operation of the larger net and which remains part of it. The address-determining connections therefore constitute a *functional storage* of experience. At any instant, the net *is* the memory; the memory *is* the net.

This kind of storage differs in an essential way from that of a pre-addressed

network, where the inputs do not permanently modify the connections. The latter must have a separate storage unit capable of modification and isolated from the main network except for controlled temporary interactions. So our electronic computers have their fast-access and slow-access storage units. Our social decision networks have their files and libraries more or less insulated from the functioning decision-personnel. This necessary but misleading subdivision of our artificial systems into operating units and storage units may be the reason why so many investigators in the past have searched—unsuccessfully—for a special memory organ in the human brain. Functional storage may be a typical property of biological systems, a further manifestation of their usual simplicity and efficiency.

As an example, consider the genetic material of the cell, which at the present time is supposed to consist of a few species-specific macromolecules, such as DNA or RNA. In a newly-formed cell, such a molecule has two functions (although they might not be separate functions): to initiate the steps up the ladder of chemical syntheses of specific cell materials; and to duplicate itself. But this is functional storage: the chemical structure and reactions are the expression of the heredity; the heredity *is* the chemical structure.

Likewise in the production of antibodies by antigens, the chemical record of the first antigenic experience is preserved in the antibodies (or in the chemical information in the antibody-producing cells), ready to find instant chemical expression when a second essentially identical experience occurs. The record *is* the specific chemical protection; the protection *is* the record.

On a grosser scale, evolution is functional storage. The coming of the cold is shown in our fur and feathers and families. The record of the ancient temperatures and salinities may be in our blood and tears.

The speed and efficiency of social decision networks might be increased if they could incorporate this lesson, and replace some of their file cabinets by continuously repeated appropriate functional modification in the decision channels.

Time Constants

Address-determination must go on at a certain regulated rate. This is probably faster for early-stage neurons and slower for later ones, but the order of magnitude should be well-defined for a given network.

In the adult human brain, the indications are that roughly 50 milliseconds elapse between distinguishable perceptions or decisions—one 'moment', in the STROUD terminology. This is of the order of fifty of the millisecond repetition intervals or synapse intervals of an individual cell, which seems to be a reasonable relationship (1). Knowing this time constant, we can make some numerical estimates of brain rates and capacities. These estimates are naive and probably false in detail, but they are explicit and rather instructive.

Thus suppose that there is one new perception every moment and that it may be preserved in a memory, represented by a single changed neural connection. The now-classical experiment of micro-electrode stimulation during brain surgery shows at least that if certain points are stimulated, a complete, detailed and specific memory is indeed evoked. Combining this with the working hypothesis suggested by QUASTLER and others (18), that the

waking brain appears to be processing input information at a constant rate, it would appear that a human brain may be making changed neural connections at rates up to 10^6 per day. The necessary sequential spatial order in these connections might be the origin of our sense of temporal order in our memories.

It may be no accident that this rate adds up to the order of 10^{10} to 10^{11} neurons per lifetime, comparable to the total number of neurons estimated to be contained in the adult brain; although of course a major fraction of these may be pre-addressed, unchanging after birth. (This number has also been computed as the minimum number of neurons required in a fully-developed decision-net serving 10^8 to 10^9 input elements (1). But there is no necessary conflict between these two points of view, since it is a familiar property of biological systems that they represent simultaneous optimization of different considerations—as in the two-point resolution of the eye, which is simultaneously limited by diffraction, by aberrations, and by the mosaic cell size.) By this reckoning, less than one neural junction in a thousand would be changed per week, which might account for the difficulty of detection of histological changes.

With such a specific moment-by-moment localization of new connections, the increasing loss of memory in older persons might be the result of cumulative damage to the neurons, such as radiation damage or microhemorrhages; or it might be due to a kind of saturation of the address-determining connections, so that either no new relationships are perceived in the continuing flux of inputs, or else those that are perceived are no longer able to modify the network.

These numerical estimates are not unreasonable; and even if the one-moment one-neuron assumption were dropped, it would not be surprising from the general dimensional considerations in the physics of the problem to find that that assumption would give correct order-of-magnitude relations between the time-constant, the lifetime and the number of neurons and its rate of change. Such a situation is common in order-of-magnitude calculations.

But this estimate of the rates is defended only so that it can be attacked on other grounds: for it leads to another important biological dilemma, and one that might have an interesting resolution. For it must be remembered that the brain is not merely an electrical network; it is also a biological network —living, breathing, and growing. And a neural connection time of 50 milliseconds is *orders of magnitude* too short for the usual cell growth time or atrophy time. While electrochemical channels or barriers might be formed or sudden changes of shape might take place in milliseconds, these can only occur for cells that are already present.

A few years ago it was supposed that a way out of this dilemma would be to let the new perception or thought be initially established as a closed self-maintaining loop of neural electrical excitation, which could persist long enough afterward for the cell growth and structural change to take place. But a succession of apparently negative experiments seems to have caused this notion to be largely abandoned.

There is an alternative. It is to let the neural growth take place, not after, but *before* the chemical and electrical connection to the network, as the electrician carries his coils of wire to the site before he hooks them up. The order-of-magnitude gap between the time constants can be got over by supposing

that the slow growth of new cells or random *potential* connections occurs in *parallel*, thousands or millions of cells at a time, while the fast new decisions or perceptions or insights occur *sequentially*, hooking up one cell at a time or a small group. Such a sequence might resemble the activity-stimulation-proliferation-organization sequence in other tissue. And while this specific suggestion may again be wrong, its accuracy is less important than its general bearing on the time-constant problem, which suggests that epochs of growth may need to be separated from epochs of decision in a biological address-determining network.

This possibility seems to deserve experimental inquiry. Perhaps our limited time span of intellectual attention, and the 'subconscious' solution of problems, and the role of sleep, especially in the infant, in preparing new cells to be ready for new (waking) connections or learning or decisions, should be re-examined from this point of view.

C. *Artificial Non-addressed Systems*

The truck driver is trained in childhood to perceive and respond appropriately to cars, stop-lights and pedestrians of whatever kind. In this pattern and analogy-perception he excels any arrangement of photocells yet created.

A pre-addressed decision-net might be able to operate with his small highway tolerances and high speeds if it had his 10^8-element resolving power and wide field of view. But it would not be safe in the unpredictabilities of the open road. For this job, a non-addressed mosaic is needed, capable of learning new patterns. Otherwise the appearance of a new type of car or a new type of hazard on the road will cause the machine to be sent back to the factory for a complete rewiring of the circuits to establish the new invariances and their analogies with the old cars and the old hazards.

It is important that the new hazard be recognized by *analogy* and not by trial and error. Direct highway experience would eliminate quickly a number of types of 'learning' computers that have recently been devised, in which the internal strategies are altered according to *experienced* successes or failures, but in which there is no pattern-extrapolation or 'insight'.

The possible construction of artificial non-addressed 10^4- to 10^9-element systems with complete decision nets and with 10^4- to 10^9-element outputs may deserve consideration. Primitive pattern-perceiving networks might be useful for narrowing the band-width of communication channels, if not for crude vehicle guidance. They might be useful internal elements in high-speed analogue and digital computers, where their stupidity could be partly compensated by the speed of operation. There they might simplify the presently elaborate programming operations; and could speed up computations requiring many simultaneous substages of qualitative judgment or identification under distortions or transformations, where the total judgment is more elaborate than can be quickly represented by the coincidence of two digital words.

The complete theory of artificial non-addressed systems with their many quasi-human characteristics will be fascinating. Evidently in many respects it may be simpler and more physical than present theories of digital computers and single-channel systems. It would include questions of optimization of different aspects of mosaic detection, such as rates and cell sizes, the proper

balance among second-stage detectors of different kinds, and of foveal versus peripheral vision, the role and mechanism of fixation and attention, and the whole output-selection problem which has been ignored here. Theoretical consideration might lead to principles of neural connection, unused in biological systems, which would produce entirely different kinds of 'intelligence' in the organization of the input fields.

Actual construction of at least the first stages of a non-addressed system might even be relatively easy. Since the receptor elements do not need to be wired individually, they can be laid down en masse, like the 10^9 crystals of a photographic emulsion. The first-stage neuron layer, second-stage layer, and so on, could be laid down similarly. The crystals could not be compact in shape, like those of the emulsion, but would have to be interbranching needles. But the first successful device might be many orders of magnitude more complex than anything now made.

To create such a device would require a number of really penetrating chemical or electrical inventions, but perhaps not a prohibitive number. Oculomotor outputs for scanning and tracking might have solutions close to the present standard single-element solutions. The main problem would be to guarantee that the neuron connections will tend to grow in such directions as to support any congruences in the chemical or electrical time-patterns, and will tend to be dissolved otherwise.

With elements having 10^{-8} second time-constants (comparable to transistors) the potential learning speed of such a device would be 10^5 times faster than that of a human being (2 hours = 20 years). Such speeds could not be fully realized because the initial address-determination will be limited by scanning speeds and motor-output speeds and by the chemical speeds of deposition of successive layers. But these potential speeds and these limitations are comparable to those of a digital computer; the latter being similarly held back by the slowness of programming and input and by the slow storage access speeds.

A pattern-perceiving device so much faster than a human being and with a full range of inputs and outputs would pose grave problems of education, manipulation and control, problems different from those of a digital computer and more difficult; but the rewards would be correspondingly greater if these problems could be solved.

Acknowledgement—This study was stimulated by a number of conversations with Professor Richard L. Meier, now of the University of Michigan.

REFERENCES

1. J. R. Platt: Amplification aspects of biological response and mental activity. *Amer. Scient.* **44**, 180–197 (1956).
2. S. L. Polyak: *The Retina*, Fig. 40, University of Chicago Press, Chicago, Ill. (1941).
3. W. C. Pitts: *Proc. Western Joint Computer Conference at Los Angeles*, 1–3 *March* 1955, 108–110, Inst. Radio Engng., New York (1955).
4. J. Strong: *Procedures in Experimental Physics*, Prentice-Hall, New York (1938).
5. H. A. Rowland: The Screw. *Encyclopedia Britannica*, 11th ed. (1913).
6. J. Strong: New Johns Hopkins ruling engine. *J. Opt. Soc. Amer.* **41**, 3–15 (1951).

7. W. H. MARSHALL and S. A. TALBOT: Visual mechanisms. In: *Biological Symposia*, ed. by H. Klüver, **7**, 117–164, Jaques Cattell, Lancaster, Pa. (1942).
8. R. W. DITCHBURN: Eye-movement in relation to retinal action. *Optica Acta* **1**, 171–176 (1955).
9. R. W. DITCHBURN and D. H. FENDER: The stabilized retinal image. *Optica Acta* **2**, 128–133 (1955).
10. R. W. DITCHBURN: Stabilized interference fringes on the retina. *Nature, Lond.* **177**, 434 (1956).
11. L. A. RIGGS, F. RATLIFF, J. C. CORNSWEET, and T. N. CORNSWEET: The disappearance of steadily fixated visual test object. *J. Opt. Soc. Amer.* **43**, 495–501 (1953).
12. L. A. RIGGS, J. C. ARMINGTON, and F. RATLIFF: Motions of the retinal image during fixation. *J. Opt. Soc. Amer.* **44**, 315–321 (1954).
13. S. COOPER and P. M. DANIEL: Muscle spindles in human extrinsic eye muscles. *Brain* **72**, 1–24 (1949).
14. *a.* S. SUNDERLAND: A preliminary note on the presence of neuromuscular spindles in the extrinsic ocular muscles in man. *Anat. Rec.* **103**, 561–2 (1949).
    b. N. C. R. MERRILEES, S. SUNDERLAND and W. MAYHOW: Neuromuscular spindles in the extraocular muscles in man. *Anat. Rec.* **108**, 23–30 (1950).
15. E. J. LUDVIGH: Symposium on Physiol. Psychology, Pensacola, Mar. 10–11 *Office of Naval Research Symposium Report ACR*-1, p. 125 (1955).
16. D. M. MACKAY: Towards an information-plan model of human behaviour. *Brit. J. Psychol.* **47**, 30–43 (1956).
17. D. O. HEBB: *The Organization of Behaviour*, J. Wiley and Sons, New York (1949).
18. H. QUASTLER, *et al.*: Human Performance in Information Transmission, I–V, *Tech. Report 62, 68, 69, 70, Control Systems Laboratory*, Univ. of Ill., Urbana, Ill.

PART VII

THE STATUS OF INFORMATION THEORY IN BIOLOGY*

A Round Table Discussion

H. Auerbach, C. Ehret, S. Freed, L. S. Frishkopf, E. E. Jacobs†, B. N. Jaroslow, P. D. Klein, A. L. Koch, H. R. Mahler, H. Quastler, W. S. Yamamoto‡

Edited by Henry Quastler

Information theory is very strong on the negative side, i.e. in demonstrating what cannot be done; on the positive side its application to the study of living things has not produced many results so far; it has not yet led to the discovery of new facts, nor has its application to known facts been tested in critical experiments. To date, a definitive and valid judgment of the value of information theory in biology is not possible.

The first attempts to apply information theory to biological studies have been met with varying degrees of enthusiasm, ranging from outright rejection to statements like: 'Information theory furnishes a person with a sort of thread which would allow him to sense out a continuum in the order of the universe'; 'A means of relating the existence of life to the non-existence of life'; 'A quest for regularities in irregular phenomena'. This is an extremely vast span of reactions to a proposition of admittedly limited scope. Many of the reactions refer not to information theory as such but more generally to interdisciplinary endeavours, and to system sciences, both of which are characteristically represented by information theory.

Interdisciplinary meetings are always, or almost always, an exhilarating experience to all. They allow some sub-groups of scientists of a number of breeds to communicate with each other in a way that is in general impossible with the rest of the breed to which the particular scientist belongs. To put it another way, interdisciplinary meetings factor out scientists in a different way than occurs normally, and allow them fruitfully so to aggregate. Information theory, with its 'interdisciplinary' generalization of the entropy concept, provides a common meeting group for many disciplines; what is more, it has in many actual instances provided strong rapport between representatives of widely separated disciplines. The value of the communications aspects

* On the evening following the Conference, eleven participants gathered for an informal session to discuss how they felt about the proceedings they had witnessed. The informal debate which ensued was transcribed and re-arranged into a coherent account. In doing this the editor tried to be objective.

† Dept. of Physiological Chemistry, University of California.

‡ School of Medicine, University of Pennsylvania.

of information theory was demonstrated by the apparent harmony that prevailed at this three-day meeting despite the varied professional disciplines represented by the papers and the participants. But the devil's advocate will argue that of potential meeting grounds there are many, and the fact that information theory actually has provided one does not necessarily imply that it was an essential ingredient of success. We must not follow the lure of each and every interdisciplinary beacon; most actual results are still obtained within the safe limits of the established disciplines.

More important is the problem of system sciences in general—that is, of all sciences that deal with the whole rather than with parts, with general principles rather than detailed specifications, with patterns rather than specific mechanisms. This is a many-faceted problem. There are, first of all, the situations of the 'forest-and-the-trees' type; for instance, in physics: a complete description of every particle in a gas would contain implicitly all thermodynamic parameters—but in this form the information would be useless. Or, to use an example in biology: if we knew the chemical constitution of all substances in all cells, together with all details of distribution, chemical kinetics, in brief, if we had reached the biochemical millenium—then we still would not necessarily know which of all these details are significant on the next higher level of organization, although presumably this information must be implicitly contained in the known details. Are we simply up against a psychological limitation? We seem able to think only of so much detail within any single train of thought. Faced with amounts of detail considerably beyond our mental capacity we begin to select: in the course of such selection important features are eliminated almost as readily as unimportant ones. Knowledge is not usable for human minds unless it is organized in blocks with not too much detail in each.

There is more behind the desire to look at the whole rather than the parts than just an awareness of psychological limitations. There are relations within a whole which cannot be expressed in terms of parts alone. This is the fact expressed in the proposition 'the whole is equal to more than the sum of its parts'. The very mixed reactions which this proposition elicits are probably due to a failure to state exactly in what way the whole is more than the sum of its parts. There should be little disagreement if the 'more' refers to propositions concerning the whole which are qualitatively different from any proposition which can be made about any of its parts. Information theory provides a very convenient formalism to state this situation: the *total* information content of the whole is exactly equal to the sum of the information contents of the parts—where the description of each part includes all possibilities of connection with other parts; the mutual dependency of parts being organized into a whole causes a mutual reduction of uncertainty; therefore, the amount of *non-redundant* information associated with the whole is *less* than the sum of the information contents of the parts; the difference is exactly the information content of the constraints, or all those propositions which apply only to the whole and not to its parts in isolation. This formulation may be of some help in making clear a puzzling aspect of the whole-parts relation.

The preference for dealing directly with wholes rather than with parts is greatly supported by contemplation of the way living things are organized. The most striking feature is the existence of the organizational pattern with

several distinct levels of organization. Some levels are more sharply defined than others, and the hierarchy of levels is not always unambiguous. Still, they are pervading enough that the intelligibility and validity of any statement in biology depends on proper agreement with organizational hierarchy. Now, one of the outstanding features in biological organization is that quite obviously only a small amount of the features obtaining at a given level has observable effects on the next higher level. Hence one of the most urgent problems, on any level, is that of determining what details are involved in the communication to the next higher level—but this is precisely one form of the problem of the 'whole and its parts'. Thus, in studying the whole rather than its parts we seem to act as organelles, cells, organs do.

So we have good reasons to believe in the importance of the systems approach. Still, it remains no more than a belief—and there exists an equally strong belief that only intense preservation of details will yield major biological breakthroughs and that it would be a 'young miracle' if really important contributions would come to biology without intensive examination of details. So we have extreme misgivings either way—and those misgivings seem to be destined to be with us forever. There exists no rigid calculus telling which formalism must be used on what data to achieve a major discovery.

The present conference was arranged to explore the applications of information theory to the study of living things. This is a new field, and one cannot say, at this time, which approach is going to be most successful. Accordingly, the scope of the program was extremely wide. It was natural to question how much the various papers had contributed to furthering the purpose of the meeting. There was general agreement that some papers had contributed verymuch, some a moderate amount, some little or nothing; there was however, notable disagreement about which papers belong in which category.

There exist a few cases where information theory was used in dealing with problems which could have been solved in other ways; and there are very many cases where problems have been solved by various methods which could have been, possibly, solved more easily by using information theory. The coin problem that RAPOPORT talked about falls in this category; so do the cryptographic studies of GAMOW and YČAS. Information theory is so general that its domain of applicability is very broad; one cannot name one situation in which by the use of information theory one cannot get some understanding on what is happening on an abstract basis. But one is always beset by the niggling doubt that the application may not be proper. One can in many situations obtain results which seem to clarify understanding or increase the sharpness of a description to an extent which was not possible prior to the use of information theoretic methods. On the other hand, such results seem often suspended in mid air, away from the results of conventional disciplines. The important question then, at this nascent stage of affairs, is something which is repellent to the scientific mind, the assessment of the 'worthwhileness' of the answers information theory seems to give. It seems plausible to assume that information theory should be useful where communication is critical, where messages are to be transmitted in the presence of noise, and where one might assume that some optimization is approximated; biologists are inclined to invoke the Darwinian mechanism of random trials with perpetuation

of successful attempts; one is reluctant to admit any basic element of purpose—yet it might be better to bring it to the surface for a dispassionate inquiry.

The question arose whether it was preferable to use information theory only in a semi-quantitative fashion, to account for general trends in observed data, or buttressed by measurements. The advantage of working with actual numerical estimates is obvious, but against it is the irreducibly relative nature of information measures. No unanimity existed concerning this question. There is general agreement that data properly usable are scarce, that there is a slight risk involved in using data from the literature which are inadequate for this purpose, and that the procurement of more pertinent and better data will yield material which would hardly have been produced otherwise. Future meetings might be designed to give stimulus and continuity to production of data which are more cogent and amenable to information theory. There was general agreement that further meetings should and will be arranged—and that information theory is here to stay in biology.

AUTHOR INDEX

Abderhalden, E. 81
Abelson, P. H. 83
Acher, R. 80
Ada, G. L. 90
Afzelius, B. A. 222
Akamatsu, N. 148
Alcock, R. S. 125
Allen, A. O. 270
Allmann, D. W. 61, 124, 126
Alper, T. 299
Anderson, E. 219, 222
Anderson, E. H. 300
Andjus, R. K. 178
Anfinsen, C. B. 80, 109, 113, 114, 116, 125, 126, 278, 289
Ard, W. B. 242, 252, 253, 256, 278, 354, 355
Armington, J. C. 379
Armitage, P. 344, 347
Aschaffenburg, R. 73
Ashby, R. W. 188
Askonas, B. A. 125
Astbury, W. T. 219
Astrachan, L. 139
Atwood, K. C. 309
Aub, J. C. 148
Auerbach, H. 347, 351, 399
Augenstine, L. 51, 61, 103, 106, 109, 114, 116, 120, 193, 194, 202, 218, 239, 287, 288, 289, 292

Bachofer, C. 290
Bahn, A. 81
Baker, W. K. 300
Baldwin, W. F. 306
Balis, M. E. 138
Balser, M. 34
Bang, F. B. 222
Bar-Hillel, Y. 194
Barltrop, J. A. 283
Barnafi, L. 80, 87
Barron, E. S. G. 258, 283
Barry, J. M. 72
Batty, I. 84
Baxter, Robert 337
Beams, H. W. 219, 222
Beerstecher, E., Jr. 207
Behrens, O. K. 79, 87
Bell, P. H. 75
Bellamy, W. D. 298
Bender, M. 86
Benedict, W. H. 294, 308, 310
Benfrey, B. J. 87
Benzer, S. 76
Berg, P. 131
Berman, M. 181, 182, 183
Bettelheim, F. R. 72
Bier, M. 113
Bigelow, J. N. 226
Binet, L. 322
Birch, L. C. 303
Black, F. L. 90
Blair, E. A. 154
Blair, H. A. 188, 294, 331, 333, 337
Blank, A. A. 41, 42
Bleaney, B. 242
Block, R. J. 90, 93
Bollman, J. L. 152
Boltman, E. T. 83, 138
Boltzmann 4, 325
Bond, V. P. 307, 308
Born, M. 265
Borysko, E. 222
Bourliere, F. 322, 328
Bowers, K. D. 242
Boyd, G. A. 294
Boyd, W. C. 212
Brachet, J. 88
Brambell, F. W. R. 84
Branson, H. 101, 105, 106, 108, 109, 110 117, 193, 197, 198
Brawerman, G. 88
Brillouin, L. 21, 198
Britten, R. J. 83
Brody, S. 327
Bromer, W. W. 79, 87
Brown, D. M. 86
Brues, Austin M. 294, 309
Bruice, T. C. 207
Bucher, N. L. R. 148
Buettner-Janusch, V. 86
Bulbenko, A. 61, 124, 126
Burdette, W. J. 53
Burger, M. 322
Burnett, G. 72
Burton, M. 289
Butler, J. A. V. 262, 269, 272

Calderon, H. 222, 223
Calkins, G. N. 222, 223
Calvin, M. 283
Cambrisson, J. 353
Campbell, P. N. 125
Cannon, W. B. 318, 322

Carlson, J. G. 55
Carnap, R. 194
Carsten, A. L. 335
Carver, E. 106
Casarett, G. 294, 333, 334, 336, 337, 338
Cavallini, D. 86, 94
Chance, B. 180
Chandrasekhar, S. 321
Chantreene, H. 125
Chapman, W. H. 307, 308
Chargaff, E. 63, 83, 90, 91, 93, 94
Chase, H. H. 39
Chatton, E. 219, 221
Cheng, H. G. 127, 131
Christenberry, K. W. 294, 308, 310
Christensen, B. G. 143, 148
Clark, A. M. 306
Clarke, J. M. 307
Clayton, R. S. 294, 309
Cloudman, A. M. 294, 309
Cohen, B. L. 73
Cohen, M. R. 222
Cohn, M. 125
Cole, D. 80
Cole, L. J. 308
Collinson, E. 285
Combrisson, J. 241
Condliffe, P. 113
Condon, E. U. 248, 253, 255
Consden, R. 81
Cooper, G. 114, 117, 120
Cooper, S. 381
Corey, R. B. 108, 109, 110, 111, 117, 254
Cornsweet, J. C. 379
Cornsweet, T. N. 379
Cote, Louis J. 95
Cotter, G. J. 84
Court-Brown, W. M. 345
Cowie, D. B. 83, 106, 113, 131
Cramer, H. 324
Crane, H. 117
Crick, F. H. C. 51, 52, 53, 63, 65, 72, 90, 91, 92, 95, 117, 141, 298
Cronbach, L. J. 192
Cronkite, E. P. 307, 308
Crossman, E. B. 306

Dainton, F. W. 285
Dale, W. M. 258, 290, 300
Dalhamn, T. 219, 222
Danchakoff, V. 126
Dancoff, S. M. 53, 57, 111, 190, 193, 202, 218, 298, 359
Daniel, P. M. 381
Danielli, J. F. 197, 221
Davey, W. P. 293
Davie, E. W. 81
Davies, J. 290
Davis, B. D. 73
Dawson, R. M. C. 126
DeBoer, J. 362
DeBusk, A. G. 88
DeFinetti, B. 231
DeFremery, D. 90
DeGroot, S. R. 198
Deich, T. L. 73, 75, 77
DeMarko, L. 86, 94
DeMoss, J. A. 131
DeRobertis, E. 219, 222
Deutsch, T. 73
Devine, R. L. 219, 222
Dewey, D. L. 71
Diandier, K. M. 132
Dimond, A. E. 278
Dirr, K. 79
Ditchburn, R. W. 379
Dixon, J. J. 75, 79
Dobson, E. L. 342
Dolansky, L. 24
Dolansky, M. 24
Doll, R. 344, 345, 347
Donohue, J. 142,
Dorner, R. W. 90
Dounce, A. L. 71
Doyle, B. 271, 274, 288, 289
Dreyer, W. J. 79
Dublin, L. I. 342
Dulbecco, R. 305
Dunnebacke, T. H. 222, 223
Duryee, W. R. 222, 223
Du Vigneaud, V. 74, 80, 81, 87

Eakin, R. E. 207
Ebert, J. D. 125, 126, 129, 132
Edds, M. V., Jr. 39
Eden, M. A. 188, 359
Edsall, J. T. 267
Ehrenberg, L. 258
Ehret, C. 191, 218, 219, 222, 223, 224, *224–5*, 399
Eldjarn, L. 258, 259
Elias, P. 34
Ellis, M. E. 308
Elson, D. 83, 90, 91, 93, 94
Ephrussi, B. 125
Epstein, P. 324
Erickson, J. 114, 117, 120
Erlanger, J. 154
Evans, E. A., Jr. 138
Evans, E. C. 354
Evans, J. V. 73
Exner, F. M. 300
Eyring, H. 117, 120, 267, 287, 290

Fahmy, M. J. 309
Fahmy, O. G. 309

Falcone, G. 71
Fano, R. M. 11
Farmer, C. J. 130
Fatt, P. 159, 160, 161
Fawcett, D. W. 219, 222
Feinstein, R. N. 84
Felix, K. 74, 79
Feller, W. 213, 361
Felton, —. 44
Fender, D. H. 379
Fenton, P. F. 39
Fine, P. C. 198
Fink, R. M. 294
Finkelstein, P. 121, 283
Fisher, J. C. 347
Fisher, R. A. 4, 325
Flavin, M. 72, 80, 125, 126
Fleischer, S. 127, 131
Fling, M. 73
Flower, S. S. 328
Fox, S. W. 87, 101
Fraenkel-Conrat, H. 70, 73, 76, 77, 80, 81, 117
Francis, M. D. 125, 131
Franck, J. 239, 255, 257, 262, 263, 264, 267, 273, 287, 289, 291
Freed, S. 171, 173
Frenkel, A. 294
Fricke, H. 271, 272, 283
Friedberg, F. 86
Friedberg, W. 126, 127, 131
Friedell, H. L. 335
Friedewald, W. F. 300
Frishkopf, L. S. 61, 153, 161, 164, 399
Fritz, F. 44
Fromageot, C. 80, 86
Fuerst, C. A. 280
Furth, J. 294, 308, 310

Gale, E. F. 131
Galton, F. 360
Gamow, G. 8, 51, 52, 61, 63, 71, 78, 83, 91, 92, 95, 101, 107, 189, 194, 401
Garen, A. 222, 223
Gates, F. L. 299
Gay, H. 219, 222, 223
Gergeley, J. 113
Geschwind, I. I. 75, 79, 80, 87
Gey, G. O. 148
Gierer, A. 70
Giese, A. C. 306
Giles, N. H., Jr. 55
Gladner, J. 113
Glinos, A. D. 61, 148
Godin, C. 125
Gofman, J. W. 343
Golay, J. E. 34
Goldblith, S. A. 280
Goldstein, L. 88
Gordon, A. H. 81
Gordon, W. G. 86
Gordy, W. 239, 241, 242, 251, 252, *252–3*, 253, 254, 256, 263, 278, 289, 292, 353, 354, 355
Gowen, J. W. 54, 55, 293
Grahn, D. 293, 334
Gray, L. H. 262, 300
Green, D. E. 219
Green, D. W. 73
Green, F. C. 79
Greenstein, J. 114, 117, 120
Grier, G. W., Jr. 44
Griffith, J. S. 91, 92
Grindlay, J. H. 152
Gros, F. 88
Gross, J. 131, 221
Grubb, H. M. 291
Guild, W. R. 269, 274, 278, 280
Gundalch, G. 79
Guthneck, B. T. 94
Guyer, M. G. 214

Haddow, A. 355
Hagen, W., Jr. 335
Hagihara, B. 84
Hald, A. 156
Ham, A. W. 150
Hamilton, K. A. 294, 309
Hampton, M. 84
Hancock, J. 360
Handa, D. T. 101, 144
Harkness, R. D. 151
Harrington, W. F. 113, 116
Harris, A. 125, 131
Harris, D. G. 174
Harris, H. 73
Harris, J. 113
Harris, J. I. 74, 75, 79, 80, 87
Hart, E. J. 270
Hartley, R. V. L. 4
Haurowitz, F. 113, 115, 126, 127, 131
Hayes, P. M. 283
Hearon, J. Z. 143
Heath, D. C. 306
Hebb, D. O. 392
Heck, W. W. 72
Heimberg, M. 125
Hemmings, W. A. 84
Henshaw, P. S. 293, 333
Herr, E. B., Jr. 306
Herriott, R. M. 125
Hershey, A. D. 144
Hetzer, H. O. 54
Hijmans, J. 362
Hill, R. 113, 116
Hins, C. H. W. 81, 86, 87

Hinshelwood, C. N. 143
Hirohata, R. 79
Hoagland, M. D. 131
Hodge, A. J. 219, 222
Hodgkin, A. L. 202, 203
Hoffman, J. G. 307, 308, 351
Hogness, D. S. 125
Hollaender, A. 54, 298, 299, 300
Holloman, J. H. 347
Holmes, B. 295
Homeyer, P. G. 87
Honnen, L. 79
Horio, T. 84
Horowitz, N. H. 73
Hough, A. 86
Howatson, A. F. 150
Howe, E. E. 71
Huettner, A. F. 219, 222
Huggins, C. 283
Hughes, W. L. 219
Huisman, T. H. J. 73
Hunt, C. C. 159, 160, 161
Hursh, J. B. 333, 334, 336, 337, 338
Hutchinson, F. 269, 274, 278, 280, 290
Hutchison, C. A. 241
Hvidt, Aa. 113, 116

Ingram, D. J. E. 242, 353, 355
Ingram, J. E. 353
Ingram, V. M. 73, 80
Itano, H. A. 73

Jacobs, E. E. 399
Jacobsen, C. 109
Jacobsen, E. 148
Jacobson, B. S. 301, 306
Jager, B. U. 86
Jaroslow, B. N. 211, 399
Jarrett, H. S. 252
Jensen, E. V. 283
John, E. R. 234
Johnson, M. P. 131
Jolles, G. 86
Jones, H. B. 188, 329, 341, 342, 343, 344, 345, 350
Jonix, J. H. P. 73
Jordan, D. O. 141
Jutisz, M. 81

Kac, M. 362
Kahn, J. R. 75, 80
Kalman, S. 113
Kalnitsky, G. 113, 116
Karreman, G. 191
Katz, B. 159, 160, 161
Katz, J. J. 178, 284
Kaufman, B. P. 55
Kauzman, W. 109, 268
Kay, L. M. 79, 81
Keller, E. B. 131
Kelner, A. 298
Kendell, M. G. 325
Kendrew, J. C. 74
Kennedy, E. P. 72, 116
Keynes, R. D. 202, 203
Kharasch, N. 207
Kihlman, B. 146
Kimball, A. W. 55, 305, 308
Kimmel, J. R. 86
King, J. W. B. 73
Kitai, R. 71, 80, 87
Klein, D. 193, 204, 399
Klemmer, E. 24
Knight, C. A. 90
Koch, A. L. 61, 101, 136, 137, 138, 139, 140, 141, 144, 145, 146, 239, 283, 284, 290, 399
Koechlin, B. A. 86
Komintz, D. R. 86
Konig, D. 365
Korik, S. B. 125
Koros, Z. 73
Koshland, D. 116
Krohmer, J. S. 335
Kroner, T. D. 79
Kunitz, M. 113
Kutsky, R. J. 133

Lagerstedt, S. 150
Laki, K. 86
Lambert, W. V. 54
Lamont, W. A. 139, 141, 144, 146, 284
Lampen, J. O. 139
Landmann, W. A. 271
Landsteiner, K. 215
Lange, K. 86
Lansing, A. I. 309, 310, 313
Lark, C. T. 138
Lark, K. 88
Laurila, U. R. 80
Lawler, H. C. 74, 81, 87
Lea, D. E. 55, 262, 270, 272, 289, 298, 301
Leak, J. C. 72
Lee, R. H. 307, 308
LeGette, J. 79
Leone, C. 271
Lepper, R. 219
Leslie, P. H. 304
Levinthal, C. 219
Levy, H. R. 101, 144
Lewis, E. B. 345
Li, C. H. 75, 79, 80, 87, 113
Lietze, A. 127, 131
Linderstrom-Lang, K. 109, 113, 116
Lindgren, F. T. 131
Linschitz, H. 201, 359
Littlefield, J. W. 131
Livingston, R. 353

Lloyd, D. P. C. 157, 158
Locker, R. H. 79, 81
Lockingen, L. S. 88
Loftfield, R. B. 125, 131
Loiseleur, M. J. 178
Lorand, L. 75, 86
Lorentz, E. 293
Lovelock, J. E. 178
Low, B. W. 253
Lucas, F. 81, 87
Luck, C. F. 252
Ludvigh, E. J. 385
Lumry, R. 117, 120, 267, 287, 290
Luria, S. E. 299, 300, 305
Luzzati, D. 138
Lwoff, A. 221, 222, 223

Maalow, O. 88
Maas, W. K. 73
Mackay, D. M. 387
Magasanik, B. 88
Magdoff, B. S. 90, 93, 95
Mahler, H. R. 61, 124, 126, 180, 399
Marack, J. C. 74
Margenau, H. 182, 183
Markham, R. 90
Marks, H. A. 342
Marsh, W. H. 75, 80
Marshall, W. H. 378
Maylow, W. 381
Mazia, D. 71
Merrilees, N. C. R. 381
Metcald, R. G. 294
Metz, C. B. 219
Michel, R. 72
Middlebrook, W. R. 75
Milhalyi, E. 113
Miller, D. G. 173
Miller, E. C. 94
Miller, J. A. 94
Miller, J. G. 234
Miller, W. C., Jr. 280
Minick, O. T. 219, 222
Minto, W. L. 294
Mirsky, A. E. 267, 287, 288,
Molnar, S. 234
Monier, R. 81
Monne, L. 218
Monod, J. 125
Montagna, W. 39
Montie, D. B. 90, 93
Moore, D. 222, 223
Moore, S. 81, 86, 87
Morgan, C. 222, 223
Morgenstern, O. 38
Morowitz, H. J. 101, 107, 108, 193, 239, 276, 277, 281
Mortimer, R. K. 301, 302, 310, 311, 312
Moscona, A. 222
Moshman, J. 294, 308, 310
Muller, H. J. 55
Murai, K. 79
Murphy, G. M. 182, 183
Murphy, J. B. 126
Murphy, J. S. 222, 223
Murray, Wm. S. 307, 308, 351
Muus, J. 86
Myerson, S. 291

McConnell, W. B. 73
McCormick, G. 242, 252, *252–3*
McCulloch, W. S. 234
McFadden, M. L. 86
McGarr, J. J. 79
McGill, Wm. J. 38, 161, 163, 164, 369
McIntyre, A. K. 157, 158
McLaren, A. D. 121, 274
McLean, J. R. 131
McMorris, R. O. 201

Nagel, E. 222
Naughton, M. A. 74
Neumann, J. von 38
Neurath, H. 79, 81, 113, 114, 117, 120
Newman, H. H. 360
Newton, G. 131
Nickerson, W. J. 71
Niu, C. I. 76, 80, 81
Noonan, Thomas, R. 335
Nord, F. 113
Nordling, C. O. 347
North, A. C. T. 73
Northrop, J. H. 125, 222, 223
Novelli, G. D. 131
Novick, A. 136, 137, 141, 145, 146
Nowell, P. C. 308
Nozaki, M. 84
Nybom, N. 301, 306

Oakley, C. L. 84
Oberle, E. M. 280
Ohno, K. 80, 82
Okunuki, K. 84
Oliff, W. D. 303
Orgel, L. E. 91, 92
Orlans, E. S. 74
Ottesen, M. 80, 113, 116
Ouchterlony, O. 214
Oudin, J. 211, 214
Owen, J. 242
Owen, M. E. 302, 310, 311, 312
Ozawa, H. 74, 76, 77, 79, 87

Pake, G. E. 252, 353
Palade, G. E. 219, 222
Paleus, S. 74, 79

Pappas, G. D. 219, 222
Pardee, A. B. 88, 144
Parish, H. D. 86
Parker, J. R. 328
Parker, S. L. 303, 304, 307
Parkes, A. S. 178
Parrish, R. G. 74
Passman, J. M. 80
Pauling, L. 108, 109, 110, 111, 117, 254, 267, 287, 288
Pearl, R. 54, 303, 304, 307
Peart, W. S. 75
Pecher, C. 154, 155, 156, 157, 159, 160, 161
Perkinson, J. D., Jr. 126
Perlmann, G. 113, 116
Perry, B. T. 90
Pihl, A. 258, 259
Pitelka, D. R. 219
Pitts, W. 234
Pitts, W. C. 373, 374, 380, 393
Platt, J. R. 107, 123, 291, 294, 371, 372, 389, 394, 395
Platzman, R. 239, 255, 262, 263, 264, 265, 266, 267, 268, 270, 273, 287, 289, 291, 292
Plaut, W. 88
Pollaczek, H. 79
Pollard, E. C. 262, 269, 272, 274, 278, 280, 287, 288
Polson, A. 83
Polyak, S. L. 373
Pomper, S. 309
Pon, N. 113
Pontecorvo, G. 76
Popenoe, E. A. 74, 81, 87
Porter, K. R. 219, 222
Porter, R. R. 79
Posternak, T. 79
Potts, B. P. 222, *224–5*
Powers, E. L. 219, 222, 223, 224
Preer, J. R. 226
Prestidge, L. S. 88
Price, J. M. 94
Proctor, B. E. 280
Prosser, C. L. 318
Purvis, J. L. 87
Putnam, F. 114, 117, 120, 138

Quastler, H. 3, 38, 39, 41, 53, 58, 111, 135, 187, 190, 191, 193, 194, 197, 202, 207, 209, 211, 216, 218, 221, 226, 281, 298, 351, 359, 369, 394, 399

Raacke, I. 113
Raacke, I. D. 75, 79
Rabinowitch, —. 255, 257
Rahn, Otto 54, 298, 309
Rall, W. 159, 160
Ramachandran, L. K. 73
Ramasarma, G. B. 90
Ramler, W. J. 270
Ramsey, N. F. 247, 248
Ranson, R. M. 304
Rapaport, A. 196, 230, 401
Rashevsky, N. 193, 222, 223, 225
Ratliff, F. 379
Ray, R. 114, 287, 289
Rebhun, L. I. 219, 222, 223
Redfield, R. 289
Reed, I. S. 34
Rees, E. D. 178
Ressler, C. 80, 87
Rexroad, H. N. 242, 252, 356, *253–3*
Rhondin, J. 219, 222
Rich, A. 71, 72, 78, 83, 93, 107
Richards, F. 113
Ridenour, L. N. 6
Riggs, L. A. 379
Rimoldi, H. J. A. 234
Robbins, K. C. 74
Roberts, E. 90
Roberts, R. 106, 113
Roberts, R. B. 83
Roche, J. 72
Rocklin, S. R. 270
Rogers, W. 113, 116
Romanoff, A. J. 130
Romanoff, A. L. 130, 131
Ronzoni, E. 86
Roos, P. 80, 87
Roper, J. H. 76
Rose, H. 222, 223
Rose, I. A. 139
Rosenberg, S. 335
Rosenblith, W. A. 61, 153, 157
Rossi-Fanelli, A. 86, 94
Rost, G. 94
Roth, L. E. 219, 222, 223, *224–5*
Rowland, H. A. 376
Rupe, C. O. 130
Russ, S. 293, 294
Russell, W. L. 55, 294, 301, 312
Rylander, P. N. 291
Ryle, A. P. 71, 80, 87, 283, 284, 286

Sacher, G. A. 293, 294, 301, 317, 322, 323, 326, 327, 329
Sachs, H. 131
Sanger, F. 71, 74, 80, 87, 283, 284, 286
Satake, K. 74, 76, 77, 79, 87
Savage, L. J. 231
Sayre, E. V. 173
Sbara, A. J. 300
Schellman, J. 113, 116
Schoenfeld, R. 181, 182, 183
Schott, R. G. 54
Schram, G. 70

Schroeder, W. A. 79, 81
Schroedinger, E. 117, 189
Schuler, R. H. 270
Schwartz, D. 71
Schweigert, B. S. 94, 139
Scott, G. M. 293, 294
Scott, J. F. 140
Seegers, W. H. 86
Seguela, J. 219
Sekuzu, I. 84
Semmett, W. F., 86
Seno, N. 79
Setlow, R. B. 269, 271, 272, 274, 278, 280, 288, 289
Shannon, C. E. 4, 21, 44, 53, 187, 234, 298, 368
Shapiro, B. 258, 259
Shaw, J. T. B. 81, 87
Shields, H. 242, 251, 252, *252–3*, 253, 254, 256, 278, 354, 355
Shields, J. 74
Shimura, K. 79
Shive, W. 207
Shock, Nathan 344
Shortley, G. H. 248, 253
Shumway, N. P. 75, 80
Siekevitz, P. 131
Silberstein, H. E. 294
Silverman, R. A. 34
Simmons, Eric L. 335
Simon, H. A. 92, 95, 123
Simpson, M. V. 125, 131
Sinex, F. M. 72
Singer, S. J. 178
Sinn, L. G. 79, 87
Skanidze, I. K. 132
Sleggs, L. T. 75, 80
Sloan, G. J. 252
Smart, H. F. 178
Smellie, R. M. S. 88
Smith, A. U. 178
Smith, E. 113, 116
Smith, E. L. 86
Smith, H. 71
Smith, J. D. 90
Smith, L. F. 71, 80, 87
Smith, M. J. 307
Smith, S. G. 81, 87
Smith, W. V. 242, 353
Soreni, E. T. 73, 75, 77
Spiegelman, S. 125
Sri Ram, J. 113
Stacy, R. W. 201
Stadler, J. 55, 293
Stapleton, G. E. 298, 300
Staub, A. 79, 87
Stearn, A. 114, 117
Stearn, A. E. 268, 287, 288
Stein, W. H. 81, 86, 87
Steinberg, D. 109, 113, 125, 126
Stent, G. 280
Stepka, B. 301
Stetten, M. R. 72
Stevens, K. W. H. 242
Stockell, A. 86
Storaasli, J. P. 335
Storer, John 335
Strange, R. E. 71
Strong, J. 375, 376, 377
Sunderland, S. 381
Susman, N. 148
Suter, G. M. 294
Swanson, C, P, 55
Swick R. W. 101, 144
Swift, H. 219, 222, 223
Symonds, P. 86
Szafarz, D. 88
Szent-Györgyi, A. 113
Szilard, L. 4, 136, 137, 141, 145, 196
Szorenyi, E. T. 73

Tabroff, W. 79
Tahmisian, T. N. 219, 222
Talbot, S. A. 378
Tapley, D. F. 283
Tartar, V. 223, 225
Taylor, J. H. 219
Teller, E. 120
Terminiello, L. 113
Thompson, A. R. 80
Thompson, E. O. P. 74, 79, 80, 81, 86
Thompson, W. D'Arcy 191, 219, 225
Thomson, J. F. 335
Tobias, C. A. 300, 301, 305
Tolman, R. C. 198
Tourtellotte, W. W. 335
Townsend, J. 252
Trambarulo, R. 242, 353
Trippett, S. 80, 87
Tristram, G. 289
Trucco, E. 191
Tumanishvili, S. 132
Tuppy, H. 74, 79
Turba, F. 79
Turing, A. M. 37, 360
Turner, J. E. 127, 131
Tweedell, K. S. 360

Uebersfeld, J. 241, 353
Uhlenbek, G. E. 320
Umbarger, H. F. 144
Upton, A. C. 294, 308, 310
Uretz, R. B. 299, 300, 305, 312

Van Der Schaaf, P. C. 73
Van der Waals 109

Van der Waerden, B. L. 362
Van Roggen, A. 252
Van Roggen, L. 252
Van Slyke, D. D. 72
Vaughan, M. 109, 113, 125, 126
Velick, J. F. 86
Velick, S. F. 86, 125
Volkin, E. 139

Waelsch, H. 131
Walker, E. P. 328
Wallace, B. 309, 312
Walter, H. 124, 126, 127, 131
Walton, B. P. 131
Wang, M. C. 320
Ward, J. C. 362
Warren, F. L. 73
Wassermann, F. 221
Watson, J. D. 51, 52, 53, 63, 65, 90, 117, 141, 298, 299
Watson, M. L. 219, 222, 223
Weaver, W. 4, 53, 187, 298, 368
Webster, G. C. 131
Weinstein, H. J. 222
Weiss, P. 125, 126
Weissmann, S. I. 252
Weisz, P. B. 39, 225
Wenneker, A. S. 148
Went, F. W. 360
Wertz, J. E. 242, 252
Westfall, J. A. 219
White, H. E. 248, 253
White, W. F. 74, 75
Wicks, L. F. 86
Wiener, N. 4, 21, 187
Williams, D. T. 201
Williams, G. R. 180
Williams, R. B. 126
Williams, R. J. 55, 207
Williamson, M. B. 80
Wilson, P. W. 144
Windsor, E. 71
Winnick, T. 125, 131
Winzler, R. J. 208
Wohlfarth-Bottermann, K. E. 219, 222
Wolf, G. 72
Wolken, J. J. 222
Wollenberg, A. 80
Wood, J. L. 126
Wood, T. H. 300, 305, 306
Woods, P. S. 219
Worden, R. E. 201
Work, E. 71
Work, T. S. 125
Wyatt, G. R. 63
Wyss, O. 354

Yalow, R. 290
Yamamoto, W. S. 399
Yamashita, J. 84
Yates, R. A. 144
Yčas, M. 8, 51, 52, 55, 61, 63, 68, 69, 70, 71, 78, 83, 87, 88, 92, 101, 102, 107, 108, 189, 194, 401
Yearkes, A. W. 328
Yearkes, R. M. 328
Yockey, H. P. 50, 54, 55, 188, 297, 298, 299, 308, 327, 338, 351, 359
Yuasa, Akira 219

Zamecnik, P. C. 131
Zavoisky, E. 241
Zcheile, F. P. 174
Zelle, M. R. 54, 298
Ziegler, J. 86
Zimmer, K. G. 258
Zinder, N. D. 222, 223
Zirkle, R. E. 270, 300, 301, 305
Zwartouw, H. T. 71

SUBJECT INDEX

Acetabularia 88
Acetyl DL-valine *252–3* (Fig. 7)
Acroblast 219, 222
ACTH 69, 75, 76, 104, 107
Actin 79, 86
Additivity, in information functions 21
Address 191, 371
—, determination of 379, 380, 387, 394
Address learning 374
Adenine 63, 90, 136, 137, 138, 141, 145
— deoxyriboside 138, 282, 294
Adenosine deaminase 146, 286
Adenylic acid 70, 90
Adrenocorticotropin 68, 79
Aging 188, 293, 297, 308, 317, 326, 339, 341, 344, 347
—, premature 293
Alanine 70, 90
Albumin 104, 114, 128, *252–3* (Fig. 10)
—, egg *See* Ovalbumin
— peptides 128
Alcohol dehydrogenase 86
Aldolase 86, 288
Allele 73, 311
Allelic states 190
Allelomorph 301, 302
Alphabet 8, 15, 18, 211
—, binary 8, 20
Ambiguity 31
Amino acid residues 50
— — sequence 105, 135, 194
— — —, phylogenetically common ancestral 107
Amino acids 18, 65, 71, 76, 84, 105, 128, 242, 252, 253, 263, 273
— —, X-irradiated 252
Amino-adipic acid, α 71
Amoeba 88
Amylase 114
Analogy perception 390
Anaphe maloneyi 87
Antibodies 192, 342, 394
Antibody formation 211
Antigens 192, 211, 214, 394
Antimetabolites 193
—, artificial 206
—, natural 206
Antimutagens 145
Antiserum 214
Apo-enzyme 204
Arginine 70, 90, 109
Armadillo, nine banded 360
Arrangement, linear 18
Ascorbic acid 208
Asparagine 70
Aspartic acid 70, 90, 109
Auger cascades 270
Averaging, effect of 23
Axon 386

Bacillus subtilis 277
Bacterium 136, 178, 215, 277, 299, 301
Binary digits *See* Bits
Binary symbol 9
Biological material, inactivation of 262, 276 *et seq.*, 287
Bioplast 218
Biosynthesis 117, 145
—, of protein 124
—, purine 138
Biotin 208
Biotopology 218
bis-2, 4-dinitrophenyl-L-cystine 284
bis-DNP cystine 284
— — —, gamma-ray irradiation of 285
Bits 21
Blepharoplast 219
Bohr magneton 247
Bonds, interhelical 110
—, rupture by ionization 269
Born energy of polarization 265 *et seq.*
Bovine albumin *252–3* (Fig. 10)

CO_2, level in blood, as O_2 regulator 7
—, — — —, as symbolization 7
Caffeine 136, 137
Cage effect 252, 255, 257, 273, 353
Calandra oryzae 303, 306
Cancer 342, 343, 344, 347, 353, 355
Cariboxypeptidase 79, 86
Carcinogenic agents 354, 355
Carcinogens 294, 309, 347, 351
Carotene 173, 208
Casein, α 79, 86
Cat, cochlear nucleus 157
Catalase 84
Category 6, 10, 12, 40
Cerebral cortex 381
Channel 30
— capacity 32, 194
— — theorem 31, 34, 39, 298
Channel communication 31
—, noise-free 31

Chemical potential 199
Chick 208
Chick embryo *See* Embryos
Chlamydomonas 301, 306
Chloracetyl-DL-alanine *252–3* (Fig. 5)
Chloramphenicol 88
Chlorophyll 174
Choline 208
Chromidion 218
Chromosome 63, 65, 219, 224, 312, 359, 373
— aberrations 55, 298, 301
Chymotrypsin 113
Chymotrypsinogen 79, 87, 104
Cilium 219
Citruline 71
Clupein 79
Coacervation 219
Cocarboxylase 204
Cochlear nucleus 157
Code 30, 34, 219, 298
—, antigen–antibody 211
—, binary 9
—, DNA-protein 52, 67, 194
—, error correcting 32, 281
—, Fano 11, 15, 16, 18
—, Morse 8, 10, 41, 193
—, over-lapping 67, 92
— groups, combination of 10
Coding
— problem, protein text 87
Co-enzymes 192, 204
Co-factors 204
Collagen 72, 79
Collinearity 380
Communication 25
— channel 31
— engineering 187
— systems 30, 35
Conditioned reflexes 7
Congolese moth *See Anaphe maloneyi*
Congruence, displacement 376
—, self 376
Conjugation 309, 310
Conservation principle, one-sided 189
—, two-sided 189
Constraint, measures of 27, 188
Constraints 37, 105
Continuity, in information functions 21
Control, in the animal and in the machine 187
Correlation intersymbol 83
Cortex surface 383
Creatine-phosphate-ATP system 225
Cross reaction 212
Cryptography 70
Cybernetics 5
Cysteamine 259
Cysteine 70, 101, 129, (Figs. 9, 10), 253, 255, 258
— residues 109
Cystine 71, 90, 101, 251, *252-3* (Figs. 4, 9), 253, 255, 256, 258, 259, 271, 283, 284, 288, 290, 292
Cystinyl residue 71
Cytidine deaminase 286
Cytidylic acid 70, 90
Cytochrome 74, 79
Cytoplasm 63, 65, 89, 150, 218
Cytosine 63, 142

Dahlbominus fuscipennis 306
Dancoff's principle 50, 53, 111, 190, 192, 298
Darwinian machines 188
Death, definition of 54
Decision theory 5
Decoding 6, 31
Degenerative diseases, chronic 348
Denaturation 283, 292
—, of proteins 117, 267, 268, 271, 274, 286, 287
Deoxyribonucleic acid *See* DNA
Diabetes mellitus 342, 344
Diamino pimelic acid 71
Dictyosome 219, 222
Dielectric absorption 264
— properties 263
Differentiation, capacity for 125
Dipeptides, combinations of 78, 82
Diphenyl picryl hydrazyl (DPPH) 249
Diploid survivorship 298, 301
Direct action 253, 258, 272, 276, 300, 301, 354
Disaccharide 212
Displacement congruence 376
Disulfide bond 116, 239, 255, 257, 258, 259, 271, 278, 283, 286, 287 *et seq.*, 292
Disulfide interchanges, radiation-induced 283 et seq.
DNA (deoxyribonucleic acid) 51, 52, 63, 70, 76, 88, 101, 117, 138, 245, 269, 272, 282, 298, 359, 394
— inactivation of by radiation 269, 272
— synthesis, steady-state flux of 142
— Watson and Crick model of 51, 55, 63, 65, 90, 142, 189
Drosophila 53, 303, 304, 312
— *melanogaster* 293, 304, 307, 337
— *subobscura* 307

Earthworm *See Lumbricoides terrestris*
E. coli *See Escherichia coli*
Edestin 104
Egg albumin *See* Ovalbumin
Electrochemical potential 200, 261, 292

Electron-spin resonance 241–261, 292, 353
— hole (electron vacancy) 257, 258, 272, 292
— hole trap 257, 258, 272, 278
— trap 256, 257, 259, 265, 266
Electrophysiology 197
Embryonic protein synthesis 130
Embryos 124
Emulsin 114
Encoding 6, 12, 16, 31
Endoplasm 221, 222
Ensemble 20, 191, 307
—, Gibbsian 277
—, of genetic messages 53
—, of organisms 51, 52, 297
—, probability measures of 193
Entropy 21, 44, 52, 119, 171, 198, 269, 287, 298, 317, 324, 359
—, irreversible production of 198
— surface, equilibrium position on 200
Enzymatic inhibition 139
Enzymes 63, 111, 121, 125, 140, 192, 197, 242, 263, 276, 278
Enzyme-substrate complex 171, 180, 206
Equilibrium position, on the entropy surface 200
Equivocation 31, 52, 55, 302
—, in the germ line 298, 309
Error detection and detection 32
Escherichia coli 73, 83, 92, 101, 106, 137, 138, 215
— —, B/r strain of 300
— —, purine metabolism in 138, 144
Estrogens 208
Euchlanis triquetra 309, 310
Europium chloride 173
Events, groups of 12
—, rare 42
—, real 7, 8
—, symbolic 7
Evolution 112
Excited molecules 273
Experimental design theory 5
Eye, human 373
—, insect 373
—, spot 222

Fano code *See* Code, Fano
Fatty acids 242
F-center 241
Feather quill, irradiated Fig. 11
Feedback 188, 223, 372, 385, 389
—, negative 148
Fibrinogen 75, 87
Fission, binary 223
Fixation tremor 379
Flagellum 219
Fluctuations 160, 200, 317, 320, 324, 328
Fluorescence spectrum of solutions 173
"Fluxes" 199
Folic acid 208
Force of mortality 298, 299
"Forces" 199
Formate 138
Franck–Condon principle 255, 257
Free energy 117
— radicals 242, 243, 246, 252 *et seq.*, 258, 259, 283, 353
Frog 208 *See Rana pipiens*
—, sciatic nerve 154, 155
Fruit fly *See Drosophila melanogaster*
Functional geometry 375

g-Factor 243
Game theory 38
γ-Globulin 79, 86, 212
Gastrocnemius nerve 158
— —, in decapitated cats 157
Gaunine 63
Gene 73, 88, 224
— code 310
— mutations 298, 301
General systems theory 5
Genetic information 51, 70, 190, 306, 338
— —, storage and transfer of 298
Genetic locus 190
— noise 52, 55, 189, 298, 309, 313
Genome 53, 54, 55, 310
Germ cell 359
— line, equivocation 298, 309
Giant starfish 215
Gliadin 73, 75, 104
Globulin 79, 129, 212
Glucagon 79, 87
Glutamine 70
Glutamic acid 70, 90, 109, 131
Glutathione 253, 255, 258, 259
Glyceraldehyde dehydrogenase 86
Glycine 70, 90, 138, 242
Glyco-protein 212
Glycyl DL-valine *252–3* (Fig. 6)
Glycyl-glycine *252–3* (Fig. 10)
Glycyl glycine, deuterium substitution 354
Glycyl glycyl glycine *252–3* (Fig. 5)
Golgi apparatus 222
Gompertz function 323, 350, 351
Grain beetle *See Calandra oryzae*
Gramicidin, action spectra for 288
Grasshopper *See Melanoplus differentialis*
Green algae 306
Growth hormone 104
Guanine 63, 136, 137, 138, 141
— deoxyriboside 138
Guanosine 138
Guanylic acid 70, 90
Guinea pig 208

Habrobracon juglandis 306, 307
Haploid 300, 306
— survivorship 298
Heart, chicken 129
Helix 65, 110, 117, 142
—, alpha 254
—, double 141, 190
Hemocyanins 212
Hemoglobin 73, 75, 79, 85, 86, 90, 94, 104, 114, 268, 272
Hepatectomy 148
Heterophile reaction 212
Histidine 70, 90
Hit theory *See* Target theory
Holo-enzyme 204
Homeostasis 318, 326, 342
Homocysteine 71
Hormones 191, 204, 242
Horse mussel 215
Humoral system, of communication 148
Hydrogen bond 109, 117, 266, 287 *et seq.*
Hydroxyglutamic acid 71
Hydroxylysine 70, 72
Hydroxyproline 70, 72
Hyperfine structure 250
Hypertensive peptide 75, 80
Hypoxanthine 137

Immunopolysaccharides 88
Inactivation of proteins, by ionizing radiation 262, 287
Independence 21
Indirect action 258, 300, 301, 354
Individuality 54
Influenza A virus 90
Information 230, 317, 324
—, chemical 394
—, conservation of 189
—, decay of 55
—, destruction of by ionizing radiation 239
—, Fisher 325
—, genetic 51, 70, 190, 306, 338, 359
—, measure of 18, 20, 25, 40, 169, 188
—, rate of production 198, 199
—, redundant 33, 189, 292
Information, replication 87
—, storage 61, 87, 393
—, transfer 61, 87
—, transmission of 112, 211, 298
— and entropy 21, 44, 171
— and intelligence 6
— assembly 373
— content 20, 21, 39, 41, 103, 116, 212, 218, 235, 276, 309, 359, 400
— —, carried by protein sequence 103, 113
— —, configurational 105, 114
— —, destruction of 297
— —, of molecules 193
Information content, of tracer data 181
— —, per printed letter 44, 281
— —, structural 111
— —, total 105
— —, total carried by proteins 113
— continuum 6
— functions 26, 31, 43
— measures, Shannon–Wiener 21
— measures 103
— —, determination of 169
— —, functions of ensembles 192
— —, of constraint between letters 281
— —, relativity of 40, 191, 193, 235
— —, the search for invariants 190
— —, topological representation of molecules 193
— representation 6, 7, 8, 9
— theory 5, 7, 50, 230, 276, 399
— —, domain of in biology 187
— transfer 65, 125
— transmission, by biological co-factors 204
— —, from genetic to somatic material 189
— —, process of 112, 189
— —, rate of 44
Inheritance, cytoplasmic 88
Inhibition, enzymatic 139
Inhibitor 208
Injury, irreparable 331
—, irreversible 336
—, radiation 271, 331
—, reparable 331
Inosine arsenolysis, inhibition of 140
— hydrolase 140
Input information 32
Insulin 71, 74, 104, 107, 114, 125, *252–3* (Fig. 12), 256, 268, 288, 290, 343
— (cattle) 80
— 'A' chain 66, 69, 80, 87
— 'B' chain 69, 80, 87
Intelligence and information 6, 7, 20
Internal emitters 338
Interphase 224
Intersymbol influence 56, 67, 71, 78, 105, 115
Invertase 115
Ionization 243, 255, 264 *et seq.*, 278
Ionization, in proteins 266
Ionizing radiation 252 *et seq.*, 262 *et seq.*, 273, 283, 331, 333, 355
— —, latent injury from 271, 272, 331
Isoleucine 70, 90
Isopropylbenzene 291

Karyoplasm 221
Keratin protein, β 253
Kinetodesma 224
Kinetosome 224

Kinety 224

Lactalbumin, α 86
Lactogenic hormone 104
Lactoglobulin, β 80, 86, 104
Lacto peroxidase 114
Language 8, 44, 104, 111, 193
L-cystine 284
Late effects, from radiation damage 301
LD_{50} 332, 333
Learning 392
LET 269, 270, 279 *et seq.*
Lethal bound, limit 298, 318, 326
Leucine 70, 90
Leucosin 115
Leukemia 343, 345
Life expectancy 333, 334, 344
— span 312, 331, 334, 342
Limitations, of information theory 192
Linear programming theory 5
Lipo-protein 212
Liver, chicken 129
— regeneration 148
Logical machine 37
Longevity 331
Lumbricoides terrestris 215
Lysine 70, 72, 90, 109
Lysozyme, papaya lysozyme 80, 86, 107, 113

Macrostate 118
Malignant neoplasms *See* Cancer
Markoffian, fluctuation process 320
Mass action, law of 151
Maximum likelihood estimate 213
Maxwell's demon 120, 196
Melanophore 80, 87
Melanoplus differentialis 215
Membrane phenomena 197
Membranes, exclusive 198
—, indifferent 197
—, responsive 198
Memory 393, 394
Message 30, 33
—, genetical 51
— entropy 52, 54, 298
— sets 191
Metabolism, purine 138
Methionine 70, 90
Methylcholanthrene 294, 309
Methyl xanthines 136
Micelle 109, 218
Micro-organisms 101, 131, 208
Microsome 66, 132
Microstate 118
Microtus agrestis 303
Microwave spectroscopy 242, 353
Mitochondrion 132, 219

Molecular energy 118
— dissociation process 264
Molecules, irreversibly inactivated 120
Monosaccharide 212
Mono-DNP cystine 284
Mono-2, 4-dinitrophenyl-L-cystine 284
Morbidity 342
Morphogenesis 218, 359
Morse code *See* Code, Morse
Mortality 317, 319, 342
Mosaics, non-addressed 372
—, pre-addressed 372
Moth, Congolese (*Anaphe moloneyi*) 87
Motoneurons 157
Mouse 293, 304, 309, 333, 337, 351
—, LAF_1 294, 307
—, R.F. 294, 307
Mutagenesis 136
—, chemical 145
Mutagenic agents, chemical 136, 145, 309
— —, DNA 137
— —, RNA 137
Mutagenic effects 355
Mutagenicity 137
Mutants, biochemical 300
Mutation rate, spontaneous 58, 136, 141, 145
Mutations 65, 71, 74, 84, 137, 298, 353, 355
—, point 55, 56, 146
—, recessive lethal 301
Myeloma globulin 86
Myoglobin 74, 80, 86, 104
Myosin 104, 113

Nebenkern 219
Negentropy, rate of production of 199
Neoplasms, malignant *See* Cancer
Nerve fiber 153
— membrane 201
— optic 389
— protein 86
Nervous system 347
Neural connections 383
Neural network 372
— systems 153
Neural thresholds 153
Neurospora 73
Neuron (*See also* Motoneurons) 384, 389, 393, 395
Nicotinamide 208
Nissl substance 222
Noise 31, 35, 51, 53, 165, 189, 281, 319, 379
—, chemical 172
—, genetic 52, 189, 298, 309, 313
—, measures of 188
—, random Gaussian 320
—, "white" 55, 320

Noise-and-redundancy-theorem 188, 189, 298
Non-addressed systems 373, 388, 396
Norleucine 71
Nuclear hyperfine structure 244
Nucleic acid metabolism, enzymes of 140
— acid residues 191
— acids (*See also* DNA, RNA) 18, 138, 144, 209, 242
Nucleolus 219
Nucleotide pairs 51, 52
Nucleotides 8, 18, 65, 70, 76, 91, 298
Nucleus 219

Odd-coin problem 233, 401
OH radicals 258, 259, 284
Operations analysis 5
Order 268, 276, 362
Orderliness, living things feed on 117, 189
—, problem of destruction of 188
Organ, control 191
Organ, target 191
Organelle decision trees 221
Organelles 218
Organization 36, 39, 218, 262
Organ-specific transfer 129
Origin of life 278
Ornithine 71
Output information 32
Ovalbumin (egg albumin) 80, 114, 128, 212, 268, 272, 289
—, peptides 128
Ovum, fertilized 51, 63
Oxytocine 66, 69, 80, 87

p-Amino benzoic acid 208
Pantothenic acid 208
Papain 80, 86, 113
Parallelism 381
Paramagnetic resonance (*See* Electron-spin resonance)
Paramecium 220–225, 306
Paschen–Back effect 246
Pathology of aging 293
Pathology of radiation damage 293
Pattern perception 371, 374, 388
Pauli principle 243
Pellicle unit 224
Peniculus 225
Pepsin 80, 104, 113, 116, 268
Pepsinogen 87
Peptide bonds 263
Peptides 215, 242, 252
—, X-irradiated 253
Peroxidase (milk) 268
Phenylalanine 70, 90
Philodina citrina 309, 310
Phosphoglucomutase 116
Phosphorylase 86, 144
Phosphorylated derivatives 138
Phosphoserine 70, 72
Physiologic variables 319, 322
Plasmagenes 219
Plasmapheresis 148
Ploidy 312
Point mutation 55, 56, 146
Poisson distribution 83, 107
Polarization, electronic 266
—, orientation 267
—, secondary-bond 266
—, vibrational 266
Polypeptides 273
Polypeptide structures, minimum entropy 117
Polystyrene, X-irradiated 256
Pooling, effect of 24
Praseodymium chloride 172
Pre-addressed 373
Preservation, of cells, tissues at low temperatures 178
Probabilities, conditional 28
—, unequal 13
Probability density 247
— distribution 231, 298
— —, joint 320
Problem-solving process 234
Prolactin 80
Proline 70, 72, 90
Protamine 74
Protein 18, 202, 204, 209, 212, 242, 263, 273, 283, 287
— and peptide chains 87
—, composition of 90, 101
— precursors, incorporation into chick embryos 128
Proteins, denaturation of, reversible 117, 267, 288
—, genetic determination of 52, 73
—, homologous 74
—, inactivation of 262 *et seq.*, 287
—, ionized 257
—, radiation damage in 253, 262 *et seq.*, 287
—, terminal residues of 77
Protein specificity 18, 50
— structure 103, 111, 267
— — and language 105, 112
— synthesis 65, 125, 135, 144, 298
— —, mechanism for 65, 125
Prothrombin 86
Proto-RNA 89
Pseudo-holo-enzyme 204
Psychology 191, 231
Punctuation mark 8, 69, 91

Purines 63 *et seq.*, 139
—, biosynthesis 138
—, derivatives, structure 136
—, incorporation and mutagenicity 137
—, metabolism 138
—, phosphorylases 140
Purine-purine transglycosidase 140
Purine-pyrimidine transglycosidase 140
Pyridoxine 208
Pyrimidine phosphorylase 140
Pyrimidines 65

Quandrulus 222, 225
Quantum theory 196, 231, 242, 247
q_x 347

Radiation (*See* Ionizing radiation, Ultraviolet radiation)
— damage 241, 252 *et seq.*, 262 *et seq.*, 276 *et seq.*, 283 *et seq.*, 287 *et seq.*, 308, 395
— —, protection from 273
— exposure 342
— hazards to man 297, 301, 338
— injury 253, 331
— —, irreversible 273, 339
Radiation sensitivity 269
Radicals, free 259
—, lifetime of 243, 252 *et seq.*
Radiobiology 190, 239, 252 *et seq.*, 262 *et seq.*, 276 *et seq.*, 283 *et seq.*, 287 *et seq.*, 292 *et seq.*, 293, 297
Radiomimetic chemicals 355
Radiomimetics 309
Random networks, theory of 188
— variables 230
Rat 148, 208, 293, 335, 337
Rate processes 327
Rattus natalensis 303
Reaction kinetics 56
Receptors 372
Recessive lethal mutations 301, 312
Recovery 308, 322, 335
—, from radiation damage 332
Redundance 3, 33, 111, 116, 369
—, in the germ cell 359, 360
—, measures of 188
Redundant information 3, 33, 189
Rennin 115
Replication, capacity for 125
Representability condition 31
Representation, binary 8, 17, 20
— theorem 17, 188
Residues, correlations between adjacent 78
Reticulocytes 94
Retina 383
Reversibility 121
Riboflavin 204, 208
Ribonuclease 69, 81, 86, 87, 104, 110, 116, 125, 285, 288
Ribonucleic acid *See* RNA
Ribonucleoprotein 150
—, associated basophilic bodies 150
RNA 8, 52, 66, 70, 88, 90, 92, 101, 131, 138, 278, 394
Rotifer 178, 309, 313
Ruling engines 376

Salivary amylase 86
Salmine 81, 104
Salmonella gallinarum, *typhimurium* 54
Salt bridges 109, 267
Sample space 231, 235
Sarcosome 222
Sciatic nerve, of frog 154, 155
Screws, precision 376
Sea pen 215
Selection 322
— rules 191
Semantic information 194
Sequence 51, 107, 113, 135, 191, 194
Serine 70, 90, 138
—, protein-bound 72
Serum, normal blood 148
— albumin 74, 81, 86, 104, 128, 150, 289, 290
— —, chicken 127
Silk, irradiated 353
—, X-irradiated *252–3* (Fig. 8)
— fibroin (*Bombyx*) 81, 104
Site, sense, non-sense 91
Solanin 115
Somatic line 309
Somatic mutation 344
Soret band 174
Southern bean mosaic virus 90, 93
Specificity *See* Protein specificity
—, antigenic 211
—, code of 194
Spectroscopic splitting factor, *g* 242
Spin-orbit coupling 249
Spleen homogenate 308
Spores 276
S-S bond 116
Staphylococcus aureus 131
Stentor 223
Stereocilium 222
Stimulus, threshold of 154
— threshold variations, possible sources of 159
Stochastic process 321
S^{35} 126
Sulphur linkages, in proteins 109, 116, 239, 257, 278, 283, 287, 292
Survival curves, exponential 300
— —, sigmoid 300

Survivorship curves 51, 297, 317
Symbol, 8, 30, 390
Symbolic events *See* Events, symbolic
— representation 7
Symbolization 7
Synapse 157
Synaptic delay 234
System analysis 37
— parameters 182
— theory 5
Systems 36 *See also* Communication systems
—, multipart 38
—, two-part 28, 30

Target 276, 281
— theory 276, 298, 300, 301, 305
— volume 276, 289
Tautomeric form, of purines 141
Temperatures, chemistry and biochemistry at low 171
Template 67, 71
Theobromine 136, 137
Theophylline 137
Thermal killing 297, 305, 306
Thermodynamics 52, 192, 196
—, equation of state in 307
Thermodynamics, irreversible 194, 198
—, second law of 117, 120, 196
Thiamine 208
Threonine 70, 90
Threshold, fluctuating 153, 161
Thymine 63, 141
Thyroglobulin 72
Thyroxine 72, 207, 208
Tissue homogenates, incorporation of activity from 130
T-Measure 27, 29
Tobacco mosaic virus 70, 76, 81, 90, 93, 278
Toluene 291
Tomato bushy stunt virus 90
Torulopsis utilis 126, 131
Transducer 31
Transforming principle 276
Transphosphorylase 86
Tribolium confusum 293
Trichocyst 220, 221
Triticum durum, vulgare 73
Tropomyosin 81, 86, 104
Tropylium ion 291
Trypsin 114, 268, 288
Trypsinogen 81, 87
Tryptophan 70, 90
Tumors 309
Turnip yellow virus 90, 93
Twins, monozygotic 360
Tyrosine 70, 90, 109
Tyrosine-O-sulfate 72

Ultraviolet radiation 252, 263, 273, 274, 287, 290, 297, 299, 306, 312, 353
— action spectra 274, 288
Uncertainty 19, 27, 32, 41, 230, 235, 368
—, conditional 28
—, —, average 30
—, functions 207
—, joint 27
—, mappings in generalized space 184
—, measure of 19, 21
—, relative 369
—, unit of 185
Unitization 29, 218
Unsaturated fatty acids 208
Uridylic acid 70, 90
Urosil 66

Valine 70, 90
Van der Waals forces 109
Variables 25
Variance 322
Vascular disease 342, 343
Vasopressin 66, 69, 74, 81, 87
Vibriolysin 115
Virus 223, 276, 299, 301
Vitamins 204, 242
—, B 207
—, D 208
Vole 304

Wool 81
Word 8, 107, 212, 363

Xanthine 136, 137
X-rays 243, 252, 290, 293, 300, 305, 337, 353, 355

Yeast 215, 300, 301, 310
—, diploid 301, 305
Yeast invertase 268
Yule distributions 123

Zein 104
Zöllner illusion 385