THE BOOK WAS DRENCHED

UNIVERSAL
LIBRARY

COLD SPRING HARBOR SYMPOSIA ON QUANTITATIVE BIOLOGY

VOLUME XIV

COLD SPRING HARBOR SYMPOSIA
ON QUANTITATIVE BIOLOGY

Founded in 1933

by

REGINALD G. HARRIS
Director of the Biological Laboratory
1924 to 1936

The Symposia were organized and managed by Dr. Harris until his death. Their continued usefulness is a tribute to the soundness of his vision.

The Symposium volumes
are published by the Long Island Biological Association
as a part of the work of the Biological Laboratory
Cold Spring Harbor, L.I., New York

COLD SPRING HARBOR SYMPOSIA ON QUANTITATIVE BIOLOGY

VOLUME XIV

Amino Acids and Proteins

THE BIOLOGICAL LABORATORY
COLD SPRING HARBOR, L.I., NEW YORK
1950

COMPOSED, PRINTED AND BOUND BY
GEORGE BANTA PUBLISHING COMPANY, MENASHA, WISCONSIN, U.S.A.

FOREWORD

When *Amino Acids and Proteins* was selected as the topic of this year's symposium, it was realized that there would be some duplication with the sixth symposium of our series, held in 1938, which dealt with *Protein Chemistry*. During the intervening decade, however, such rapid progress had been made in research with proteins that it seemed desirable to renew the discussion of this subject.

The program for this symposium was organized with the help of A. Mirsky, J. S. Fruton, and D. Shemin. The editorial work was done by Katherine Brehme Warren.

Meetings were held from June 8 to June 16, 1949. The registered attendance was 186. Expenses of the symposium, particularly those connected with foreign guests, were covered by a grant received from the Carnegie Corporation of New York.

M. DEMEREC

SOME SYMPOSIUM PARTICIPANTS

Top row: K. Linderstrom-Lang, W. H. Stein; I. Frantz, S. Moore; R. K. Cannan, H. B. Bull.
Second row: J. S. Fruton, P. C. Zamecnik, K. Linderstrom-Lang; I. Fankuchen; K. Linderstrom-Lang, M. L. Anson.
Third row: D. C. Hodgkin, A. Mazur, F. B. Seibert; E. L. Smith, J. S. Fruton; J. T. Randall, J. F. Danielli.
Bottom row: S. Simmonds, E. Fredericq; D. M. Wrinch, C. Fromageot; J. A. V. Butler, K. Pedersen.

SOME SYMPOSIUM PARTICIPANTS

Top row: A. Keston, S. Udenfriend, F. Sanger; D. C. Hodgkin, A. L. Patterson; R. Benesch, H. J. Teas, A. H. Doermann.
Second row: D. M. Wrinch, B. W. Low; E. Brand, J. G. Kirkwood; B. P. Kaufmann, P. A. Abelson.
Third row: M. Levy; A. Marshak, A. Rothen, I. Fankuchen; S. Emerson, R. Herriot.
Bottom row: M. L. Anson, S. Simmonds, J. A. V. Butler; M. Demerec; C. Fromageot, V. Menkin, J. S. Fruton.

LIST OF THOSE ATTENDING OR PARTICIPATING IN THE SYMPOSIUM

ABELSON, NEVA M., Children's Hospital, Philadelphia, Pennsylvania
ABELSON, PHILIP A., Carnegie Institution of Washington, Washington, D.C.
ADAMS, MARK, New York University College of Medicine, New York
ANSON, M. L., Continental Foods, Inc., Hoboken, New Jersey
BAER, HAROLD, College of Physicians and Surgeons, Columbia University, New York
BALLENTINE, ROBERT, Brookhaven National Laboratory, Upton, New York
BAPTIST, VICTOR H., Northwestern University, Chicago, Illinois
BARRY, GUY THOMAS, Rockefeller Institute for Medical Research, New York
BEISER, SAM M., College of Physicians and Surgeons, Columbia University, New York
BELLAMY, W. DEXTER, General Electric Research Laboratory, Schenectady, New York
BELSER, WILLIAM L., Carnegie Institution of Washington, Cold Spring Harbor, New York
BENACERRAF, BARNJ, College of Physicians and Surgeons, Columbia University, New York
BENDICH, AARON, Sloan-Kettering Institute for Cancer Research, New York
BENDITT, EARL P., University of Chicago Clinics, Chicago, Illinois
BENESCH, REINHOLD, Northwestern University, Chicago, Illinois
BENESCH, RUTH E., Northwestern University, Chicago, Illinois
BENTLEY, MARION BLANCHARD, New York University, New York
BERMAN, RUTH, Brooklyn, New York
BERTANI, G., Carnegie Institution of Washington, Cold Spring Harbor, New York
BLOCH, HUBERT, Public Health Research Institute, New York
BLOWNEY, SUSAN, Carnegie Institution of Washington, Cold Spring Harbor, New York
BORENFREUND, ELLEN, College of Physicians and Surgeons, Columbia University, New York
BRAND, ERWIN, College of Physicians and Surgeons, Columbia University, New York
BRAND, KATHLEEN, National Dairy Research Laboratory, Oakdale, New York
BRANDT, ARTHUR M., Sloan-Kettering Institute for Cancer Research, New York
BROWN, GEORGE B., Sloan-Kettering Institute for Cancer Research, New York
BRYSON, VERNON, Biological Laboratory, Cold Spring Harbor, New York
BULL, HENRY B., Northwestern University, Chicago, Illinois
BURKE, JOHN M., New York University College of Medicine, New York
BUSCHKE, WILLIAM, Manhattan Eye, Ear and Throat Hospital, New York
BUTLER, J. A. V., The Chester Beatty Research Institute, The Royal Cancer Hospital, London, England
CANN, JOHN R., California Institute of Technology, Pasadena, California
CANNAN, R. KEITH, New York University College of Medicine, New York
CARSTEN, MARY E., New York University, New York
CASPARI, ERNST, Carnegie Institution of Washington, Cold Spring Harbor, New York
CAVALIERI, L. F., Sloan-Kettering Institute for Cancer Research, New York
COGIN, GERALDINE, Brooklyn Polytechnic Institute, Brooklyn, New York
COLLINS, NANCY, New York University College of Medicine, New York
CRAIG, LYMAN C., Rockefeller Institute for Medical Research, New York
DALY, MARIE, Rockefeller Institute for Medical Research, New York
DANIELLI, J. F., Chester Beatty Research Institute, Royal Cancer Hospital, London, England
DARBY, HUGH H., Carnegie Institution of Washington, Washington, D.C.
DAVIS, BERNARD D., The Research Laboratory, U. S. Public Health Service, New York
DELAPORTE, BERTHE, Carnegie Institution of Washington, Cold Spring Harbor, New York

DEMEREC, M., Carnegie Institution of Washington, and Biological Laboratory, Cold Spring Harbor, New York
DE ROETTH, ANDREW, Columbia University, New York
DI CARLO, FREDERICK, Fleischmann Laboratories, New York
DISCHE, ZACHARIAS, College of Physicians and Surgeons, Columbia University, New York
DISSOSWAY, CAROLYN, Carnegie Institution of Washington, Cold Spring Harbor, New York
DOERMANN, A. H., Carnegie Institution of Washington, Cold Spring Harbor, New York
DOHERTY, DAVID G., Oak Ridge National Laboratory, Oak Ridge, Tennessee
EDELHOCH, HAROLD, Harvard Medical School, Boston, Mass.
EMERSON, MARY R., Pasadena, California
EMERSON, STERLING, California Institute of Technology, Pasadena, California
ERLANGER, BERNARD F., Columbia University, New York
FABER, MOGENS, College of Physicians and Surgeons, Columbia University, New York
FANKUCHEN, I., Polytechnic Institute of Brooklyn, Brooklyn, New York
FELIX, KURT, Frankfort, Germany
FOSTER, JOHN W., University of Maryland, College Park, Maryland
FRANK, SYLVIA, Columbia University, New York
FRANTZ, IVAN D., Massachusetts General Hospital, Boston, Massachusetts
FREDERICQ, EUGENE P., Duke Hospital, Durham, North Carolina
FREEMAN, BETTY B., Columbia University, New York
FRESCO, JACQUES R., New York University College of Medicine, New York
FREUND, E. H., National Dairy Research Laboratory, Oakdale, New York
FRICKE, HUGO, Cold Spring Harbor, New York
FRIEDMAN, SELMA, Columbia University, New York
FRIEDMAN, STANLEY, Johns Hopkins University, Baltimore, Maryland
FROMAGEOT, CLAUDE, Laboratoire de Chimie biologique de la Faculté des Sciences, Paris, France
FRUTON, JOSEPH S., Yale University, New Haven, Connecticut
FURCHGOTT, ROBERT, Cornell University Medical College, New York
GANZ, A. JEROME, Continental Foods, Hoboken, New Jersey
GAY, HELEN, Carnegie Institution of Washington, Cold Spring Harbor, New York
GOODGAL, SOL H., Johns Hopkins University, Baltimore, Maryland
GRANICK, S., Rockefeller Institute for Medical Research, New York
GREGORY, J. DELAFIELD, Rockefeller Institute for Medical Research, New York
HAUGAARD, G., National Dairy Laboratory, Oakdale, New York
HAUROWITZ, FELIX, Indiana University, Bloomington, Indiana
HAUSMANN, WERNER, Rockefeller Institute for Medical Research, New York
HAYASHI, TERU, Columbia University, New York
HERRIOTT, ROGER M., Johns Hopkins University, New York
HEYMAN, NATALIE, City College of New York, New York
HIRS, C. H. W., Rockefeller Institute for Medical Research, New York
HIRST, GEORGE K., Public Health Research Institute, New York
HODGKIN, DOROTHY CROWFOOT, Laboratory of Chemical Crystallography, University Museum, Oxford, England
HOTCHKISS, ROLLIN, Rockefeller Institute for Medical Research, New York
HSIE, JEN-YAH, Biological Laboratory, Cold Spring Harbor, New York
HUGHES, EDWARD W., California Institute of Technology, Pasadena, California
HUGHES, WALTER L., Harvard Medical School, Boston, Massachusetts
JACKSON, HORACE L., National Dairy Research Laboratory, Oakdale, New York
JOHNSON, ARNOLD H., National Dairy Research Laboratory, Oakdale, New York
JONES, CHASE B., Waltham, Massachusetts

KABAT, Elvin A., College of Physicians and Surgeons, Columbia University, New York
KAMEN, MARTIN D., Washington University Medical School, St. Louis, Missouri
KAUFMANN, B. P., Carnegie Institution of Washington, Cold Spring Harbor, New York
KAUZMANN, WALTER J., Princeton University, Princeton, New Jersey
KENT, PAUL W., Birmingham University, England
KESTON, ALBERT S., New York University-Bellevue Medical Center, New York
KIRKWOOD, JOHN G., California Institute of Technology, Pasadena, California
KLOTZ, IRVING M., Northwestern University, Evanston, Illinois
KRUEGER, R. C., Columbia University, New York
KUBU, EDWARD T., New York University College of Medicine, New York
LARK, GORDON, University of Chicago, Chicago, Illinois
LENNEY, JAMES, Fleischmann Laboratories, New York
LESH, J. B., Armour and Company, Chicago, Illinois
LEVINE, CELIA, Columbia University, New York
LEVY, MILTON, New York University-Bellevue Medical Center, New York
LINDERSTRØM-LANG, GERDA, Copenhagen, Denmark
LINDSTRØM-LANG, KAJ, Carlsberg Laboratory, Copenhagen, Denmark
LIPSHITZ, RAKOMA, Columbia University, New York
LIVELY, ETHELYN, Carnegie Institution of Washington, Cold Spring Harbor, New York
LOSPALLUTO, JOSEPH, New York University, New York
LOW, BARBARA W., Harvard Medical School, Boston, Massachusetts
LUCK, J. MURRAY, Stanford University, California
LUMRY, RUFUS, University of Utah, Salt Lake City, Utah
MANN, MARGARET, Biological Laboratory, Cold Spring Harbor, New York
MARKOWITZ, H., Columbia University, New York
MARSHAK, A., New York University College of Medicine, New York
MAURER, PAUL, Columbia University, New York
MAZUR, A., Cornell Medical College, New York
MCDONALD, M., Carnegie Institution of Washington, Cold Spring Harbor, New York
MCFARREN, EARL F., National Dairy Research Laboratory, Oakdale, New York
MENKIN, VALY, Temple University Medical School, Philadelphia, Pennsylvania
MILLER, LEON, University of Rochester, Rochester, New York
MILLS, J. A., National Dairy Research Laboratory, Oakdale, New York
MINDELL, FLORENCE, Hofstra College, Hempstead, New York
MOORE, STANFORD, Rockefeller Institute for Medical Research, New York
MUUS, JYTTE, Mount Holyoke College, South Hadley, Massachusetts
NELSON, THOMAS C., Columbia University, New York
NEURATH, HANS, Duke University, Durham, North Carolina
NORMAN, AMOS, Columbia University, New York
OKUDA, NAO, Carnegie Institution of Washington, Cold Spring Harbor, New York
PALADE, GEORGE, Rockefeller Institute for Medical Research, New York
PATTERSON, A. L., Bryn Mawr College, Bryn Mawr, Pennsylvania
PATTERSON, ELIZABETH K., Institute for Cancer Research, Philadelphia, Pennsylvania
PEDERSEN, KAI O., Institute of Physical Chemistry, University of Upsala, Upsala, Sweden
PENNOYER, JESSIE M., Carnegie Institution of Washington, Cold Spring Harbor, New York
PERLMANN, GERTRUDE E., Rockefeller Institute for Medical Research, New York
PETERMANN, MARY L., Sloan-Kettering Institute for Cancer Research, New York
PHILLIPS, CATHERINE, Columbia University, New York
PIROVANE, LOUISE, Temple University Medical School, Philadelphia, Pennsylvania
PRESSMAN, DAVID, Sloan-Kettering Institute for Cancer Research, New York

PRYTZ, Bo, Biological Laboratory, Cold Spring Harbor, New York
RANDALL, J. T., Kings College, London, England
REGNA, PETER, Chas. Pfizer and Co., Brooklyn, New York
ROLL, PAUL M., Sloan-Kettering Institute for Cancer Research, New York
ROTHEN, A., Rockefeller Institute for Medical Research, New York
ROTHER, MARIA-RENATE, College of Physicians and Surgeons, Columbia University, New York
SANGER, F., Biochemical Laboratory, Cambridge, England
SCHIFFMAN, GERALD, New York University College of Medicine, New York
SCHLESINGER, R. WALTER, Public Health Research Institute, New York
SCHULMAN, MARTIN P., University of Pennsylvania, Philadelphia, Pennsylvania
SEIBERT, FLORENCE, Henry Phipps Institute, Philadelphia, Pennsylvania
SHARPE, L. M., Brookhaven National Laboratory, Upton, New York
SHEMIN, DAVID, College of Physicians and Surgeons, Columbia University, New York
SIEGEL, ALBERT, California Institute of Technology, Pasadena, California
SIMMONDS, S., Yale University, New Haven, Connecticut
SLOBODIANSKY, EVELYN, New York University College of Medicine, New York
SMITH, DAVID J., Columbia University, New York
SMITH, EMIL L., University of Utah, Salt Lake City, Utah
SOBER, HERBERT A., National Cancer Institute, Bethesda, Maryland
SOLOMON, A. K., Harvard Medical School, Boston, Massachusetts
SPRING, HELEN, Carnegie Institution of Washington, Cold Spring Harbor, New York
STEIN, WILLIAM H., Rockefeller Institute for Medical Research, New York
STIMPSON, E. G., National Dairy Research Laboratory, Oakdale, New York
SULLIVAN, ROYAL A., National Dairy Research Laboratory, Oakdale, New York
SWANSTROM, MARYDA, New York University College of Medicine, New York
TANNENBAUM, STUART, Columbia University, New York
TAYLOR, JOHN FULLER, Washington University School of Medicine, St. Louis, Missouri
TEAS, HOWARD J., California Institute of Technology, Pasadena, California
TROLL, WALTER, New York University, New York
TUMERMAN, L., National Dairy Research Laboratory, Oakdale, New York
UDENFRIEND, SIDNEY, Washington University Medical School, St. Louis, Missouri
VAN VUNAKIS, HELEN, Columbia University, New York
WALLACE, BRUCE, Biological Laboratory, Cold Spring Harbor, New York
WAKSMAN, BYRON H., College of Physicians and Surgeons, Columbia University, New York
WARNER, ROBERT C., New York University College of Medicine, New York
WARREN, KATHERINE BREHME, Biological Laboratory, Cold Spring Harbor, New York
WEBER, IONE, New York University College of Medicine, New York
WEISBERG, S. M., National Dairy Research Laboratory, Oakdale, New York
WITKIN, EVELYN M., Carnegie Institution of Washington, Cold Spring Harbor, New York
WITTENBERG, JONATHAN, Columbia University, New York
WOOLLEY, D. W., Rockefeller Institute for Medical Research, New York
WRINCH, DOROTHY, Smith College, Northampton, Massachusetts
YAMINS, J. L., National Dairy Research Laboratory, Oakdale, New York
ZAMECNIK, PAUL C., Harvard University, Boston, Massachusetts
ZAMENHOF, STEPHEN, Columbia University, New York
ZIL, HELENA, New York City

CONTENTS

LIST OF PREVIOUS VOLUMES

Volume I (1933) Surface Phenomena, 239 pp.
Volume II (1934) Aspects of Growth, 284 pp.
Volume III (1935) Photochemical Reactions, 359 pp.
Volume IV (1936) Excitation Phenomena, 376 pp.
Volume V (1937) Internal Secretions, 433 pp.
Volume VI (1938) Protein Chemistry, 395 pp.
Volume VII (1939) Biological Oxidations, 463 pp.
Volume VIII (1940) Permeability and the Nature of Cell Membranes, 285 pp.
Volume IX (1941) Genes and Chromosomes: Structure and Organization, 315 pp.
Volume X (1942) The Relation of Hormones to Development, 167 pp.
Volume XI (1946) Heredity and Variation in Microorganisms, 314 pp.
Volume XII (1947) Nucleic Acids and Nucleoproteins, 279 pp.
Volume XIII (1948) Biological Applications of Tracer Elements, 222 pp.

HYDROLYSIS OF PROTEINS

HENRY B. BULL
Chemistry Department, Northwestern University Medical School, Chicago, Illinois

The hydrolysis of proteins has been of concern to chemists for a great many years. In spite of the large amount of work expended, important features of this reaction are still obscure; it is inherently a most complex sequence of reactions with which one has to deal.

So far, we in our laboratory have studied the kinetics of hydrolysis of egg albumin by hydrochloric acid and the molecular weight distribution of the peptides resulting therefrom, using a new spread film technique which we have been able to develop. We have also investigated to a limited extent the partial hydrolysis of egg albumin by barium hydroxide. We have been able to arrive at an approximate idea of the intrinsic relative rates of release of amino acids during the early stages of hydrolysis. For this purpose we have used filter paper chromatography. We have studied the kinetics of the hydrolysis of egg albumin by pepsin in some detail, and we are also prepared to make a preliminary report on the molecular weight distribution of the peptides resulting from the peptic hydrolysis of egg albumin.

The Peptide Bond

It is useful to review briefly some of the properties of the peptide bond. This bond as it exists in peptides has an abnormally short length for a carbon-nitrogen bond. Corey (1948) assigns it a length of 1.33 Å while the normal distance is about 1.47 A. This indicates that the peptide bond has considerable double bond character and is stabilized by resonance.

Massive groups attached to the α-carbons must impose considerable steric hindrance to the approach of a water molecule preparatory to hydrolysis of the peptide chain, and there are, no doubt, only certain configurations of the peptide chain which will permit such an approach.

From the dielectric constant studies of Marcy and Wyman (1941) and of Conner, Clarke and Smythe (1942), it can be concluded that while peptides show a marked degree of rigidity due to the high energy barriers for rotation about the various bonds in a peptide, a large measure of randomness in peptide configuration results.

Huffman (1942) has calculated the free energy change involved in the hydrolysis of some dipeptides and finds the energy change to be from about −3,000 to −4,000 calories per mole at 25°. This means, of course, that the equilibrium point is far in the direction of hydrolysis.

Haugaard and Roberts (1942) estimate that about 2,000 calories of heat are evolved per peptide bond when the peptide bonds in β-lactoglobulin are hydrolyzed by the action of pepsin.

It is to be expected that the two structural factors which would have the most important influence on the rate of hydrolysis of peptides would be 1) the nature of the amino acid residues adjacent to the peptide bond and 2) the length of the peptide chain.

Levene and co-workers (1932) investigated the rate of hydrolysis of a series of dipeptides by 0.5N NaOH at 25° and found the rate to be very much dependent on the size of the amino acid residues; the larger the residue, the slower the rate of hydrolysis. The size of the residue on the carbonyl side of the peptide bond was most important. This fact probably indicates that the water molecule attaches itself to the carbonyl group. More recently Synge (1945) has studied the rate of hydrolysis of a similar series of dipeptides in the presence of acid. His results parallel those of Levene closely. The influence of the size of the residue on the rate of hydrolysis was, however, less pronounced for acid than for alkali; alkali is apparently a more discriminating agent.

Kuhn, Molster and Freudenberg (1932) and later Freudenberg, Piazole and Knoevenagel (1938) studied the rate of hydrolysis of a series of glycine peptides of increasing length in the presence of one normal sodium hydroxide at 20°. They found that the rate of hydrolysis increases with the length of the peptide. A study of their data shows, however, that while the rate per peptide bond is substantially constant through the tetrapeptide, beyond this length the rate per bond progressively decreases.

Hydrolysis of Proteins by Acids and by Bases

In the hydrolysis of a protein either by acids or by bases, there is undoubtedly a large measure of randomness. It can be shown from statistical theory (Montroll and Simha, 1940) that under such circumstances there will be an accumulation of short chain peptides and the greater the degree of hydrolysis, the shorter will be the most favored length for the peptides. After hydrolysis of 10 to 15 percent of the peptide bonds, the peptides in greatest concentration will be di- and tripeptides.

It is true, however, that neither acids nor bases lead to complete randomness of hydrolysis. For example, the approximate intrinsic relative rates of release of various amino acids from egg albumin by barium hydroxide and by sulfuric acid during the early part of the hydrolysis reaction have been determined. These determinations were made by filter paper chromatography technique (Bull, Hahn and

Baptist, 1949). The results of these studies are shown in Table 1.

A relative intrinsic rate of unity means that the amino acid is liberated at a rate proportional to its concentration in the protein while a rate greater than unity means that it is preferentially liberated;

TABLE 1. INTRINSIC RELATIVE RATES OF RELEASE OF AMINO ACIDS BY 5.0N H_2SO_4 AND BY 3.7N $Ba[OH]_2$ FROM EGG ALBUMIN AT 60°

Amino Acid	Intrinsic relative rate of release	
	H_2SO_4	$Ba[OH]_2$
Aspartic	1.56	.57
Ornithine	—	.26
Glutamic, lysine, serine (average)	.94	1.29
Threonine, arginine (average)	.51	4.07
Alanine, tyrosine (average)	1.36	.67
Valine, leucines (average)	.79	.55

the slower rate of release of valine and the leucines is notable. Conspicuous also is the differential effect of acid and of alkali.

A comparison of the relative rate of release of total free amino acids by alkali (1.45N NaOH at 66°, Warner, 1942) with the rate with acid (7.95N HCL at 60°, Bull and Hahn, 1948) for egg albumin reveals that the ratio of free amino acids released to total peptide bonds hydrolyzed is, during the first part of the reaction, about 3.5 times greater for alkali than for acid.

Over a substantial part of the hydrolysis reaction of egg albumin by 7.95N HCL there appears to be a nearly constant ratio between the number of peptide bonds hydrolyzed and the amount of peptide which cannot be heat coagulated at the isoelectric point of egg albumin. This ratio corresponds to about 55 peptide bonds per mole of egg albumin. This constant ratio means that the peptide bonds in protein hydrolyze at a much faster rate than do the peptide bonds in peptides produced from the protein hydrolysis. It may well be that the more coherent structure of the protein molecule is more conducive to the activation of peptide bonds. The hydrolysis of a peptide bond is an exothermic process, and it is possible that the heat evolved tends to activate neighboring peptide bonds in the protein molecule.

While, as noted above, upon acid hydrolysis of egg albumin the molecule splits on an average into 55 fragments, these fragments are not of equal size. Figure 1 shows the molecular weight distribution of the peptides resulting from the hydrolysis of egg albumin by 7.95N HCL at 30° after the isoelectric heat-coagulable has been removed (Bull and Hahn, 1948). This distribution was determined by the spread film technique.

As can be seen from Figure 1, there is no evidence for the accumulation in the hydrolysate of any considerable quantity of peptide of a given molecular weight; there is a spread of molecular weights. About 40 percent of the hydrolysate has a molecular weight below 1,000, and about 20 percent is between 1,000 and 10,000. There is little material whose molecular weight is between 10,000 and 30,000. A most curious feature is the relatively large amount of peptide of quite high molecular weight. The hydrolysis of egg albumin by acid departs very greatly from a random process; the greatest depar-

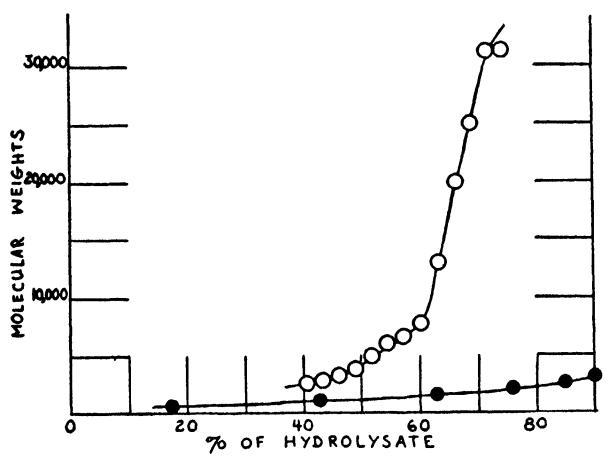


FIG. 1. Molecular weight distribution of peptides after removal of isoelectric heat coagulable material; two hour hydrolysate, 12.7 percent peptide bonds hydrolyzed at 30°. Also shown is the distribution of peptides on the basis of random hydrolysis for the same extent of hydrolysis (filled circles).

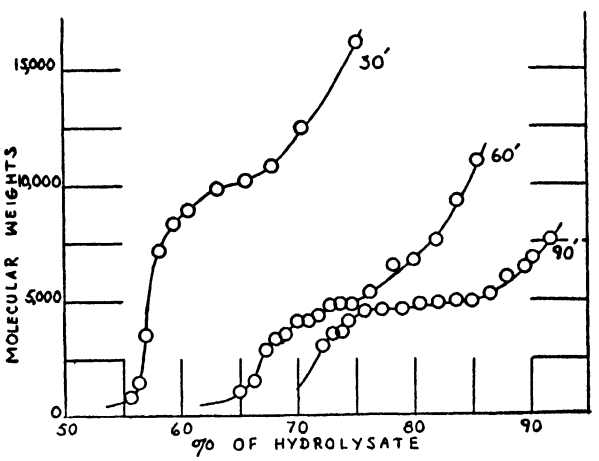


FIG. 2. Molecular weight distribution of peptides resulting from the action of 3.7N $Ba[OH]_2$ on egg albumin at 60° after removal of isoelectric heat-coagulable material for indicated times of hydrolysis.

ture is by way of this high molecular weight peptide fraction.

Figure 2 shows the molecular weight distribution of peptides resulting from the action of 3.7N $Ba[OH]_2$ at 60° on egg albumin.

It appears from Figure 2 that there is an accumulation of peptides in the alkaline hydrolysate whose molecular weight is about 5,000. The very high molecular weight peptides are missing from the alkaline hydrolysate.

Peptic Hydrolysis

We have concerned ourselves with the action of pepsin on egg albumin, and have studied certain kinetic aspects of this problem along with determinations of the molecular weight distribution of peptides resulting (Bull and Currie, 1949).

The extent of the action of pepsin on egg albumin was estimated from the amount of egg albumin which cannot be heat-coagulated at its isoelectric point. The maximum velocities (V) are expressed in moles of egg albumin solubilized per second per mole of pepsin. The molecular weight of egg albumin was taken as 45,000 and that of pepsin as 35,000. These maximum velocities were estimated by the appropriate Michaelis-Menten plots. The variation of the maximum velocities with the pH of the reaction mixture are shown in Figure 3.

According to the theory of absolute reaction rates, the rate of decomposition of the activated complex should be

$$\gamma C^* \frac{kT}{h}$$

Where γ is the transmission coefficient whose value approaches unity, C^* is the concentration of the activated complex, k is Boltzmann's constant, h is Plank's constant and T is the absolute temperature. The important question arises as to whether or not the Michaelis-Menten complex is identical with the activated complex. If this is true, then V should equal $\gamma kT/K_m h$ where K_m is the Michaelis-Menten dissociation constant. Setting γ equal to unity and substituting the value of K_m (3.4×10^{-4}) at pH 2.0 and at 30° C. and the values for the constants, we calculate that V should be 1.9×10^{17} moles per second per mole of pepsin. The experimental value for V is 0.077 moles per second per mole. Evidently, the Michaelis-Menten complex is not the activated complex; its concentration is very much greater than that of the activated complex. It is, therefore, necessary that the Michaelis-Menten complex pass over into an activated complex before the egg albumin can hydrolyze into peptides.

The Michaelis-Menten treatment also neglects the influence of the hydrogen ions. Since the reaction rate is so dependent on the concentration of hydrogen ions, its seems reasonable to assume that they enter directly into the catalytic reaction.

There are, no doubt, a variety of ways to formulate the above considerations; three simple possibilities are:

$$1)\ H^+ + E \rightleftarrows H^+E;\quad H^+E + S \rightleftarrows H^+ES \rightleftarrows [H^+ES]^* \rightarrow E + P$$
$$2)\ H^+ + S \rightleftarrows H^+S;\quad H^+S + E \rightleftarrows H^+ES \rightleftarrows [H^+ES]^* \rightarrow E + P$$
$$3)\ E + S \underset{K_1}{\rightleftarrows} ES;\quad ES + H^+ \underset{K_2}{\rightleftarrows} H^+ES \rightleftarrows [H^+ES]^* \rightarrow E + P \quad (K_3)$$

In the above E represents the enzyme, S the egg albumin, H^+ the hydrogen ions, P is the peptide and $[H^+ES]^*$ is the activated complex.

A consideration of the first and second formulation reveals that the maximum velocity of the reaction should be independent of the hydrogen ion concentration; this is contrary to experience so that both of these formulations must be rejected. The third formulation will be considered in detail.

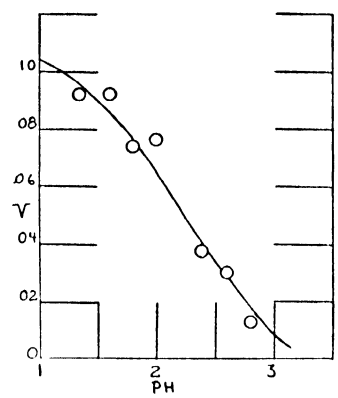


FIG. 3. Plot of experimental points for the maximum velocity of digestion of egg albumin against pH for 30°. The solid line has been calculated from Equation 5.

Now if we assume that the slowest reaction is the formation of the activated complex such that all steps leading up to this reaction are in substantial equilibrium, we can formulate the kinetics of the peptic hydrolysis quite simply. The velocity of the decomposition of the activated complex is

$$v = \gamma[H^+ES]^* \frac{kT}{h} \qquad (1)$$

Substituting the value of the equilibrium constants for the various steps in Equation 1, we have

$$v = \frac{\gamma \times H^+ \times E \times S}{K_1 K_2 K_3} \qquad (2)$$

We know that

$$E = E_0 - ES - H^+ES - [H^+ES]^* \qquad (3)$$

where E_0 is the total amount of enzyme added to the hydrolysate. Substituting Equation 3 in Equation 2 and rearranging, we find

$$1/v = \frac{K_1K_2K_3h}{\gamma kTSE_0} + \frac{K_2K_3h}{\gamma kTHE_0} + \frac{K_3h}{\gamma kTE_0} + \frac{h}{kTE_0} \qquad (4)$$

Evidently, when E_0 is unity and 1/v is plotted against 1/S, $K_1K_2K_3h/kT$ is equal to the slope of the line and is equal to K_m/V. The intercept on the Y-axis is

$$1/v = \frac{K_2K_3h}{\gamma kTH^+} + \frac{K_3h}{\gamma kT} + \frac{h}{\gamma kT} \qquad (5)$$

When 1/V is plotted against $1/H^+$, the slope of the line is $K_2K_3h/\gamma kT$ and the intercept is $K_3h/\gamma kT + h/\gamma kT$. It is thus possible to evaluate all the dissociation constants of the third formulation, and these are given in Table 2 for 30° and in Table 3 for 45°. The value of K_3 is based upon the assumption that γ is unity. Since we have the value of the dissociation constants at two different temperatures, it is possible to calculate the heats of dissociation as well as the entropy changes involved.

TABLE 2. DISSOCIATION CONSTANTS AND ENERGIES FOR STEPS INVOLVED IN THE DIGESTION OF EGG ALBUMIN BY PEPSIN AT 30° C

Constant	ΔF 30 Calories	ΔH Calories	ΔS E.U.
$K_1 = E \times S/ES = 7.4 \times 10^{-4}$	4,400	− 1,500	−19.2
$K_2 = H^+ \times ES/H^+ES = 7.2 \times 10^{-6}$	3,000	11,100	26.6
$K_3 = H^+ES/[H^+ES]^* = 5.7 \times 10^{13}$	−19,200	−31,400	−40.5

TABLE 3. DISSOCIATION CONSTANTS FOR STEPS INVOLVED IN THE DIGESTION OF EGG ALBUMIN BY PEPSIN AT 45°

$K_1 = E \times S/ES = 6.6 \times 10^{-4}$
$K_2 = H^+ \times ES/H^+ES = 1.7 \times 10^{-2}$
$K_3 = [H^+ES]/[H^+ES]^* = 4.9 \times 10^{12}$

As can be seen from Table 2, the heat of dissociation of the proton from the enzyme-substrate complex is about 11,000 calories and its pKa is about 2.14. The dissociation constant is about what is to be expected from the ionization of a carboxyl group, but the heat change involved is much too large for a single carboxyl group. It can be seen that the plot of the maximum velocity against pH is essentially a titration curve of the complex. It is possible that a number of carboxyl groups are involved and there may be multiple binding of protons in the formation of the activated complex; if there is considerable overlapping of the binding constants without electrostatic interference, the dissociation of the protons could still be approximately treated by the mass law on the assumption that only one proton was bound.

Figure 3 shows a comparison of the experimental values for the maximum velocity plotted against the pH of the hydrolysate. Also shown are the values calculated according to Equation 5 (solid line). The agreement is satisfactory.

The large heat needed for the creation of the activated complex is very nearly equal to the heat of activation for the denaturation of egg albumin which, at 30° C and in the acid region, is about 35,000 calories (Cubin, 1929). The creation of the activated complex quite possibly involves changes in the entire egg albumin molecule and probably is not confined to a single peptide bond.

As we have formulated the hydrolysis of egg albumin by pepsin, the collision rate between the pepsin and the egg albumin becomes secondary. It appears, however, that the Smoluchowski theory (1916) of the precipitation of colloidal solutions provides a more realistic approach to the calculation of the collision rate than does the gas collision formula.

According to Smoluchowski, the number of unit particles in a colloid at any time, t, is given by

$$n = \frac{n_0}{1 + 8\pi r D n_0 t} \qquad (6)$$

where r is the radius of the unit particle, D is the diffusion constant, n_0 is the number of unit particles at the beginning of the reaction. Evidently, the rate of disappearance of unit particles is twice as great as the collision rate and is at zero time equal to $-8\pi r D n_0^2$. Then the collision rate is $4\pi r D n_0^2$. Calculations show that the rate of collision of pepsin molecules with egg albumin molecules under the conditions which we have worked is about 1/200th of the rate as given by the gas collision theory.

Comparison of the number of peptide bonds hydrolyzed as determined by the Van Slyke amino nitrogen with the amount of protein which cannot

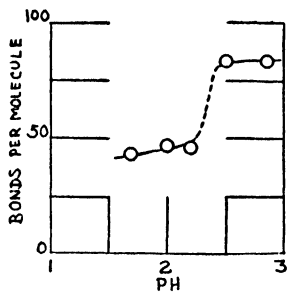


FIG. 4. Peptide bonds hydrolyzed per mole of egg albumin attacked by pepsin at 30° as a function of pH.

be heat-coagulated at the isoelectric point gives the average number of peptide bonds hydrolyzed per molecule of egg albumin attacked. We have determined this ratio as a function of pH, and these results are shown in Figure 4.

Including and above pH 2.5, the number of peptide bonds hydrolyzed per mole of egg albumin is about 84 while below pH 2.2 the ratio is about 44 bonds per egg albumin molecule. So far as can be judged, these ratios are independent of the reaction time; measurements were started as soon after the beginning of the reaction as feasible.

Shown in Figure 5 are the molecular weight distributions of the peptides resulting at pH 1.7 and pH 2.7.

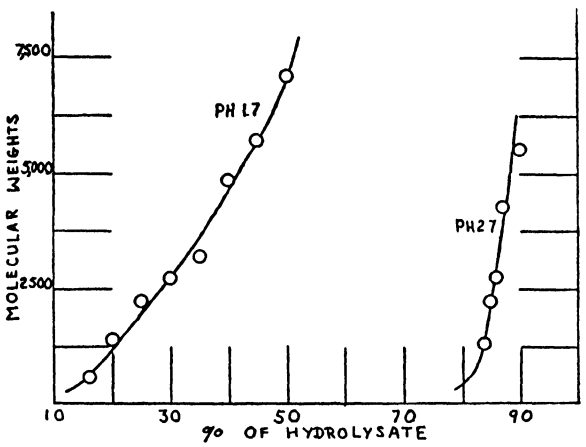


FIG. 5. Molecular weight distribution of peptides resulting from the action of pepsin on egg albumin at 30° after removal of isoelectric heat-coagulable material. Curve 1, pH 1.7. Curve 2, pH 2.7.

The reason for the lower ratio of bonds hydrolyzed per egg albumin molecules attacked, in the lower pH range, no doubt resides in the fact that there is a large quantity of high molecular weight material in the peptic hydrolysate at the lower pH values.

At pH values in excess of 2.5, the molar ratio between the bonds hydrolyzed and the molecules attacked is, as we have seen, about 84. This means that on the average, the egg albumin molecule is split into tetra- and pentapeptides, and since there are few or no free amino acids released and practically all the peptides have a molecular weight below 1,000, the tetra- and pentapeptides must be in abundance. It must also be noted that this extensive bond breakage per molecule occurs during the beginning of the hydrolysis reaction where there can hardly be any question of peptide rearrangements and of co-substrate formation. Quite evidently, the Bergmann (1942) specificity requirements for pepsin are, as far as the hydrolysis of egg albumin is concerned, meaningless. Pepsin simply hydrolyzes every fourth or fifth bond in the egg albumin molecule.

Our research on the hydrolysis of proteins has been supported by Corn Products Refining Company, Swift and Company, The Abbott Laboratories, and the Division of Grants of the U. S. Public Health Service.

REFERENCES

BERGMANN, M., 1942, A classification of proteolytic enzymes. Advances in Enzymology *2:* 49-68.

BULL, H. B., and HAHN, J. W., 1948, The acid hydrolysis of egg albumin. I. Kinetic studies. J. Amer. Chem. Soc. *70:* 2128-2131.

1948, The acid hydrolysis of egg albumin. II. Molecular weight distribution of peptides. J. Amer. Chem. Soc. *70:* 2132-2135.

BULL, H. B., HAHN, J. W., and BAPTIST, V. H., 1949, Filter paper chromatography. J. Amer. Chem. Soc. *71:* 550-553.

BULL, H. B., and CURRIE, B. T., 1949, Peptic hydrolysis of egg albumin. I. Kinetic studies. J. Amer. Chem. Soc. *71:* 2758-2760.

CONNER, W. P., CLARKE, R. P., and SMYTHE, C. P., 1942, The dielectric investigation of polypeptides. I. The dielectric increments of amino acids and polypeptides. J. Amer. Chem. Soc. *64:* 1379-1387.

COREY, R. B., 1948, X-ray studies of amino acids and peptides. Advances in Protein Chemistry *4:* 385-406.

CUBIN, H. K , 1929, The denaturation of proteins. Part V. Denaturation by acid. Biochem. J. *23:* 25-30.

FREUDENBERG, K., PIAZOLE, G., and KNOEVENAGEL, C., 1938, Polymere anhydride des glycins und der asparaginsäure. Ann. Chem. *537:* 197-204.

HAUGAARD, G., and ROBERTS, R. M., 1942, Heats of organic reactions. XIV. The digestion of beta-lactoglobulin by pepsin. J. Amer. Chem. Soc. *64:* 2664-2671.

HUFFMAN, H. M., 1942, The heats of combustion and free energies of some compounds containing the peptide bond. J. Phys. Chem. *46:* 885-891.

KUHN, W., MOLSTER, C. C., and FREUDENBERG, K., 1932, Zur Kinetik des Abbaus Polymer-homologer Ketten: Die Geschwindigkeit der Hydrolyse einiger Aminosaure-Derivate. Ber. dtsch. Chem. Ges. *65:* 1179-1183.

LEVENE, P. A., STEIGER, R. E., and ROTHERS, A., 1932, The relation of chemical structure to the rate of hydrolysis of peptides. VII. Hydrolysis of dipeptides by alkali. J. biol. Chem. *97:* 717-726.

MARCY, H. O., and WYMAN, J. JR., 1941, Anomalous dispersion of dipolar ions. J. Amer. Chem. Soc. *63:* 3388-3397.

MONTROLL, E. W., and SIMHA, R., 1940, Theory of depolymerization of long chain molecules. J. Chem. Phys. *8:* 721-727.

SMOLUCHOWSKI, M. U., 1916, Diffusion in infinite space. Phys. Z. *27:* 557-583.

SYNGE, R. L. M., 1945, The kinetics of low temperature acid hydrolysis of gramicidin and of some related dipeptides. Biochem. J. *39:* 351-355.

WARNER, R. C., 1942, The alkaline hydrolysis of egg albumin. J. Biol. Chem. *142:* 741-756.

DISCUSSION

ANSON: There has been much evidence, both from the studies of acid hydrolysis by Bull and of enzymatic hydrolysis by Tiselius and others, that

the first pieces of proteins produced by hydrolysis are rapidly hydrolyzed further. In the usual course of hydrolysis one thus obtains a mixture of intact protein and small pieces. I have, however, obtained some suggestive, but quite inconclusive evidence that, despite the usual explosive character of hydrolysis, it is possible to isolate big pieces of proteins from the first products of enzymatic digestion.

If a protein is digested by a proteinase to the extent of only a few percent and the undigested protein is precipitated by 0.2N trichloracetic acid, half or more of the digestion products not precipitated by 0.2N trichloracetic acid can be precipitated by 1N trichloracetic acid. As digestion proceeds further, the fraction of the products of digestion precipitated by 1N but not by 0.2N trichloracetic acid becomes less and less (Anson, 1940, J. gen. Physiol. *23*:695; see page 699). It remains to be found out by direct experiment whether or not the fraction of slightly digested protein precipitated by 1N trichloracetic acid does in fact contain the much desired big pieces of proteins. This could be done by measurements of end groups and by the usual molecular weight measurements.

BULL: In the case of acid hydrolysis as well as of peptic hydrolysis a fair fraction of the hydrolysate has, during the early stages of hydrolysis, a molecular weight in excess of 30,000 as revealed by film balance studies. This high molecular weight material cannot, however, be heat coagulated. This is no doubt the material which Dr. Anson has encountered in his work. The nature of this large peptide fragment is obscure. Alkalin hydrolysis does not give rise to this high molecular weight peptide.

BUTLER: Is Dr. Bull sure that all the heat precipitated material is undigested protein? Peptic digests of proteins easily deposit, at some pH's, insoluble materials (plasteins). Does anything analogous occur under his conditions?

BULL: The question of whether or not some of the peptide is brought down during the heat coagulation of the intact egg albumin is dealt with in my reply to the discussion of Dr. Neurath, where it is shown by recovery experiments that such precipitation, under the conditions of our experiments, is small.

If the filtrate resulting from the heat coagulation of a peptic hydrolysate be concentrated to about a third of its volume in a sausage casing by evaporation at room temperature, a slight precipitate results.

BUTLER: With regard to the specificity of the action of the enzyme, we have found with chymotrypsin that there is an initial rapid specific action followed by a slow, less specific proteolysis. Has Dr. Bull found any indication that the bonds broken in the initial stages may be different from those broken toward the end of the reaction?

BULL: Such information as we have does not indicate that pepsin behaves like chymotrypsin in regard to bond specificity. The ratio of bonds hydrolyzed to egg albumin attacked is, as nearly as we can judge, constant over a broad range of the reaction extending to a point as early in the reaction as it is practical to make measurements.

HAUROWITZ: Could the larger split products of peptic hydrolysis not arise by disaggregation of ovalbumin? Are they not sub-units of a larger polymeric molecule?

BULL: Egg albumin, unlike some proteins, does not easily dissociate into sub-units. I am inclined to believe that the fragments which are produced by hydrolysis arise principally if not exclusively from the hydrolysis of peptide bonds.

HAUROWITZ: Can secondary closure of peptide bonds not interfere with the results? Such a formation of peptide bonds seems inprobable in non-enzymatic hydrolysis. However, the formation of covalent bonds between haem and globin has been observed, when haemoglobin is either boiled with NaOH or exposed to the action of trypsin.

BULL: The resynthesis of peptide bonds is, of course, always a possibility and nothing which we have done throws any light on this problem. The hydrolysis studies which we have made have been concerned with the early stages of hydrolysis where the probability of such resynthesis is least.

HAUROWITZ: Are there any indications of a core in the ovalbumin molecule? If there were a difference between superficial and interior parts of the molecule, they should be revealed by the paper chromatography of the hydrolysates.

BULL: We have no evidence for the existence of a core in the egg albumin molecule.

NEURATH: I am under the impression that Dr. Bull's most interesting results are predicated on the accuracy of the method which he has used in differentiating between unhydrolyzed protein and split products. The method of isoelectric heat coagulation may not be entirely valid in this respect not only because of the coprecipitation of split products with the heat coagulum, as already mentioned in this discussion, but also because it is known from Kleczkowski's work that the heat coagulation of a protein may be inhibited by other proteins which do not coagulate themselves under the same conditions. Thus, protein fragments of different isoelectric points, formed during the initial phase of hydrolysis, may conceivably mask the heat coagulation of unchanged egg albumin.

BULL: We hydrolyzed egg albumin by pepsin and prepared solutions of peptide free from egg albumin, the protein being removed by isoelectric heat coagulation and filtration. We then added this peptide to egg albumin solutions without the addition of pepsin. Aliquots of these solutions were removed and heat-coagulated at the isoelectric point of egg albumin in the usual manner. These solutions were

made up to volume and filtered. Total nitrogens were run on aliquots of the filtrates and the amount of peptide in the filtrates calculated. The recoveries of the peptides are shown in the table below:

GRAMS OF PEPTIDE RECOVERED FROM EGG ALBUMIN SOLUTIONS AFTER ISOELECTRIC HEAT COAGULATION

Grams protein per 100 cc.	Grams peptide added per 100 cc.	Grams peptide recovered per 100 cc.
6.10	0.159	0.159
1.89	0.049	0.046
1.02	0.026	0.025
0.62	0.016	0.017
4.46	1.62	1.63
3.96	1.62	1.64
2.63	1.62	1.67
0.98	1.62	1.63

Evidently, the method of isoelectric heat coagulation is capable of yielding good recoveries of added peptide. It is, accordingly, probable that no large error is involved in following the course of the digestion of egg albumin by pepsin by means of this method.

NEURATH: I should like to ask Dr. Bull also whether the hydrolysis of amide bonds in the protein occurs and if so, whether he has applied the proper corrections in his calculations.

BULL: We have been unable to detect any production of ammonia during the hydrolysis of egg albumin by pepsin. In the case of the acid hydrolysis of egg albumin, the ammonia has been determined and corrections applied.

NEURATH: Since I shall not be in attendance of the Symposium when other papers on proteolytic enzymes will be presented, I should like to offer now a few comments on the significance of the calculations based on the determination of the temperature dependence of hydrolysis rates. Such measurements have been carried out in our laboratory in great detail, primarily of the chymotryptic hydrolysis of specific peptide and ester substrates but also of the hydrolysis of specific peptides and esters by carboxypeptidase and trypsin. These have been calculated for the rate-determining step of the activation of the Michaelis-Menten complex, and corrected to a standard state defined by the moles of substrate activated per mole of enzyme-substrate complex, per second, at the pH optimum of each enzyme at 25°. The calculated heats of activation are uniformly lower than those given here by Dr. Bull, *i.e.*, ranging from about 10 to 16 kcals. per mole, as compared to Bull's value of 31 kcals. The changes in entropy of activation are likewise lower than those reported here for the peptic hydrolysis of egg albumin, *i.e.*, about −10 E.U., as compared to Dr. Bull's value of −40 E.U. Our values are comparable to those previously reported by Dr. Butler.

I can conceive of two explanations for the considerably higher values presented by Dr. Bull: First, Bull's values are only slightly higher than those found in the literature for the acid- or base catalyzed hydrolysis of specific peptides and thus may be indicative of the acid hydrolysis of egg albumin, though I am aware that the standard states for these catalytic reactions and for enzymatic catalysis need not necessarily be comparable. Nevertheless, one should recognize the fact that the measurements of the peptic hydrolysis of egg albumin were performed in solutions of high acidity and that the enzymatic contribution to the energetic constants can only be evaluated by comparison with analogous measurements in the absence of the enzyme.

The other possible explanation is that the energetic constants given by Dr. Bull refer not to protein hydrolysis but to the peptic denaturation of egg albumin. This suggestion is being offered because of the resemblance of the present energetic constants to those found for protein denaturation in general and also because of the widely-held notion that the first step in the enzymatic hydrolysis of proteins is denaturation. Since the measured rates are those of the slowest step, all we would have to assume is that in the present case, and perhaps in every such case, the rate of denaturation is slower than the rate of proteolysis. Whichever explanation is more nearly correct, I believe that calculations of energetic changes accompanying the enzymatic hydrolysis of peptide or ester substrates, which will be published in full elsewhere, represent more nearly the true state of affairs.

BULL: The value of −31 kcals per mole for the heat and −40 E. U. per mole apply to the reverse reaction; to compare these values with those obtained by the usual formulation, the signs should be reversed. As stated, these large values probably indicate that the entire egg albumin molecule is activated before it can undergo hydrolysis. Undoubtedly, this activation can be regarded as a form of denaturation, although how profitable such a point of view may turn out to be is questionable.

Dr. Neurath's experiments on the influence of temperature on the rate of hydrolysis of specific peptides by chymotrypsin were conducted at the pH optimum of the enzyme at 25° C. The heat of ionization of the enzyme-substrate complex has thus, by necessity, been neglected. This neglect has introduced an element of ambiguity into his calculations of the energies of activation. It is, therefore, difficult to compare our values in which the heat of ionization has been taken into account with those of Neurath and of Butler.

Under the conditions which we have worked, the amount of hydrolysis of egg albumin in the absence of pepsin (acid hydrolysis) is so small that it is undetectable.

NEURATH: Finally, I should like to ask Dr. Bull whether he would care to comment on the effect of enzyme concentration on hydrolysis rates, with particular consideration of the competitive influence of reaction products on the hydrolysis of remaining protein substrate.

BULL: We have made no detailed investigation of the influence of varying enzyme concentration on the rate of digestion. It has been found that the addition of peptide obtained from a peptic hydrolysate to a fresh enzyme-substrate system greatly inhibited the digestion of egg albumin by pepsin.

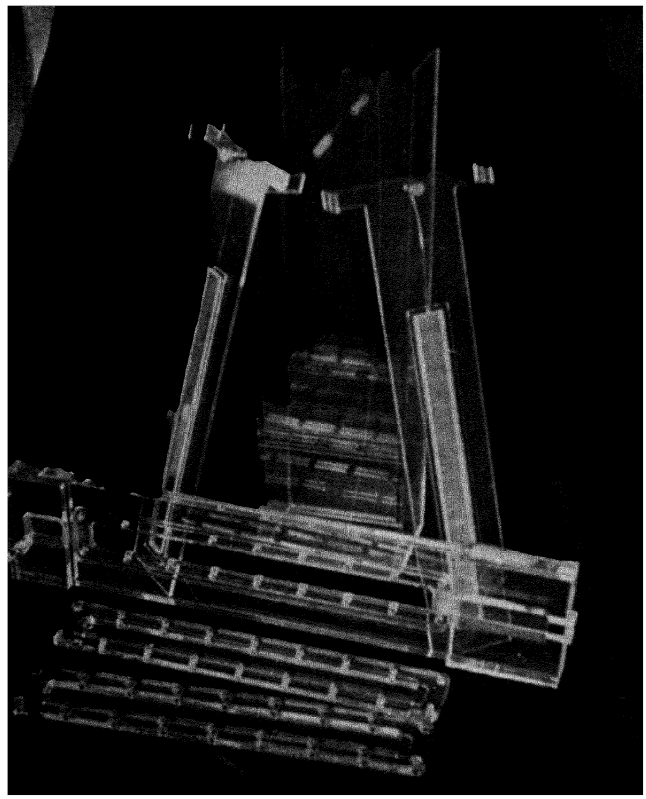


FIG. 1. Electrophoresis-convection apparatus.

THE FRACTIONATION OF PROTEINS BY ELECTROPHORESIS-CONVECTION

JOHN R. CANN AND JOHN G. KIRKWOOD
Gates and Crellin Laboratories of Chemistry, California Institute of Technology, Pasadena, California
(Contribution No. 1292)

A method of fractionation of proteins in solution utilizing a combination of electrophoretic and convective transport of the components was first suggested by Kirkwood in 1941 and investigated experimentally by Nielsen and Kirkwood (1946) several years later. More recently an electrophoresis-convection apparatus of improved design has been described and successfully used to fractionate ovalbumin, horse diphtheria antitoxin pseudoglobulin, the bovine serum proteins and bovine γ-globulin (Cann, Kirkwood, Brown and Plescia, 1949; Cann, Brown and Kirkwood, 1949).

The method is based upon the same principles as those of the Clusius column, except that horizontal electrophoretic transport instead of transport by thermal diffusion is superimposed on convective transport in a vertical convection channel. The fractionation scheme may be briefly described as follows. Two reservoirs connected by a vertical channel, of width sufficiently small to ensure laminar flow, contain a solution of the protein to be fractionated. Upon application of a horizontal electric field, differential transport of the mobile components across the channel takes place, producing a horizontal density gradient depending upon the composition gradients. Under the action of gravity, the density gradient induces convective circulation in the channel with a velocity distribution qualitatively similar to that of the Clusius column. The result of the superposition of the horizontal electrophoretic transport and vertical convective transport is movement of the mobile components from the top reservoir to the bottom reservoir at rates depending on their mobilities, with a relative enrichment of the top reservoir with respect to the slow components and of the bottom reservoir with respect to the fast components. In order to avoid contamination of the solution by electrolysis products, the walls of the convection channel are constructed of semi-permeable membranes, separated from the electrodes by buffer solution. The electric field across the channel is maintained by the electric current carried by the ions of the buffer electrolyte, to which the membranes are permeable. The exterior buffer solution is replenished by a circulation system at rate sufficient to prevent electrolysis products from reaching the membranes.

The purpose of this paper is to point out the essential features of electrophoresis-convection and to illustrate the utility of the method in protein fractionation.

THE ELECTROPHORESIS-CONVECTION APPARATUS

The electrophoresis-convection apparatus may be considered as being composed of five principal parts: (1) the fractionation cell consisting of a cell block, two face plates, and semi-permeable membranes; (2) the box housing the fractionation cell and electrodes; (3) the electrode assemblies; (4) the buffer circulating system; and (5) the power pack. A photograph of an unassembled fractionation unit is presented in Figure 1. On assembly the cell consists essentially of a narrow vertical channel connecting upper and lower reservoirs. That portion of the channel effective in fractionation is formed by the rectangular space between two sheets of semi-permeable membrane. The cell, containing a buffered solution of the protein to be fractionated, is immersed in buffer solution to within about an inch above the bottom of the upper reservoir. The electrode assemblies are placed in the box on opposite sides of the cell, and an electric field applied across the channel of the cell. During operation electrolysis occurs which tends to change the pH of the buffer solution. To counteract this, buffer is circulated vertically around the cell. The temperature of the system is regulated in a constant temperature cold room operating at 4° C.

The material used in the construction of the fractionation cell and its housing must be electrically non-conducting, exhibit no swelling when in contact with aqueous solutions, possess dimensional stability, and be readily machined with high precision. Lucite appeared to be the material most suitable for this purpose and was chosen as the construction material.

The cell block consists of an upper and lower reservoir between which is a vertical rectangular slot. The reservoirs are connected to the central slot by means of narrow vertical channels passing through the body of the block. A recess is milled around the edge of the rectangular slot on both faces of the cell block. The cell block is supplied with two small legs. The reservoirs are supplied with valves for sampling; and the top reservoir is open to the atmosphere. The capacities of the upper and lower reservoirs of the apparatus used in these studies are 100 and 50 ml., respectively. The top and bottom reservoir capacities of the apparatus pictured in Figure 1 are 15 and 10 ml., respectively.

The face plates are frames which fit into the recesses around the periphery of the central slot of the cell block. Around the inner edge of the frame is a

shoulder constructed to fit into the central slot. The face plates clamp the sheets of semi-permeable membrane into place against the bottom of the recess. Since the membranes form the face-walls of the effective portion of the channel, the height of the shoulders control the channel wall separation. In the apparatus described here the wall separation is 0.037". The face plates are bolted to the cell block with machine screws. If metal screws are used it is necessary to insulate electrically the heads and tips of the screws. The use of plastic screws would, of course, eliminate the necessity for electrical insulation. Sheets of semi-permeable membrane are made from cellulose sausage casing.

The inside dimensions of the box which houses the fractionation cell and electrode assemblies are such that a snug fit is obtained between the sides of the box and the edges of the cell block. This minimizes loss of electric field by leakage around the cell. The box is supplied with buffer inlet and outlet tubes.

Each electrode assembly consists of 2 mil platinum foil mounted in a lucite frame. The dimensions of the platinum foil correspond to those of the inner periphery of the face plates. The dimensions of the frames are such that the strips of platinum and the effective channel of the cell are aligned when the cell and electrode assemblies are housed in the box. Since platinum occludes large quantities of electrolytic hydrogen it is advisable to alternate the polarity of the electrodes from run to run.

Circulation of buffer around the cell is by gravity flow. The buffer flows from an aspirator bottle into the box housing the cell and electrodes. After circulating vertically around the cell the buffer is discharged into a flask. A centrifugal pump periodically pumps the circulated buffer back into the aspirator bottle. It is desirable to operate under conditions of field strength and current density which do not lead to electrolytic decomposition of the buffer anions. Under these conditions the original pH of the buffer solution is restored on mixing, and the buffer may be recycled in the circulating system without replenishment.

Operation of the Electrophoresis-Convection Apparatus

Depending upon the electrophoretic properties of the particular protein system under investigation, fractionation can be accomplished by three modes of operation. In the first method the operating pH is such that all the components of the heterogeneous protein are either on the alkaline or acid side of their isoelectric points. Under these conditions the components are differentially transported out of the upper and into the lower reservoir, the fractionation depending upon the difference in mobilities of the constituent proteins. The greatest separation of components is obtained by so choosing the operating time that half of the major component is transported out of the upper reservoir. This method is suitable for the separation of components differing but slightly in their isoelectric points.

The separation of a protein mixture possessing discrete mobility and isoelectric point spectra, *e.g.*, serum, into its constituent proteins is accomplished by the isoelectric procedure. In this method of operation one of the constituents of the heterogeneous protein is immobilized by operating at its isoelectric point. The mobile components are transported out of the upper and into the lower reservoir, leaving the immobilized component in the upper reservoir. Complete exhaust of the mobile components from the top reservoir, although closely approach under ideal conditions is sometimes inhibited by various disturbing factors. The most important disturbing factors appear to be osmotic transport of solvent from the exterior buffer solution into the cell and the establishment of a stationary state before exhaust when the mobilities of some of the components are of opposite sign at the operating pH. Optimum operating conditions must therefore be determined by pilot fractionations. When applicable this procedure is far more efficient than the first method of operation.

The successive separation of the components of a protein mixture can be accomplished by the isoelectric procedure as follows. The component with the most alkaline or acid isoelectric point is first separated from the other in several successive stages by operating at its isoelectric point. The composite of the top cuts of these stages is further processed to purify the desired component. The bottom cut of the last of these stages is a concentrate of the mobile components. The process is repeated until the mixture has been resolved.

In the case of a protein which migrates as a single boundary in an electric field but possesses a specified mobility distribution as revealed by reversible electrophoretic boundary spreading, *e.g.* γ-globulin, fractionation is accomplished by means of a modified isoelectric procedure. In this procedure the fractionation is carried out at a pH displaced by an arbitrary amount from the mean isoelectric point of the heterogeneous protein. Transport in the apparatus leads to a redistribution of the protein ions such that the fractions withdrawn from the top and bottom reservoir possess mobility distributions differing from that of the original protein. Fractions possessing different mean mobilities and isoelectric points are obtained by proper choice of the operating pH's. Transport proceeds to a stationary state in which the top fraction is isoelectric at the operating pH.

The application of these methods of fractionation to particular heterogeneous proteins will be discussed below.

Fractionation of Ovalbumin

Within the pH range of 5 to 10, crystalline ovalbumin is resolved electrophoretically into two components A_1 and A_2, A_2 being the slower moving

component (Longsworth, Cannan and MacInnes, 1940). The relative concentration of A_2 in freshly prepared material is reported to vary from 15 to 25 percent. A third component, A_3, has also been found to be present in small amounts (MacPherson, Moore and Longsworth, 1944; Alberty, Anderson, and Williams, 1948). A_3 is less mobile than A_1 and A_2 at pH 6.8. At its average isoelectric point pH 4.58 and ionic strength 0.1 (Longsworth, 1941), ovalbumin is not resolved electrophoretically into its components (Alberty, Anderson and Williams, 1948).

Since the components of ovalbumin possess approximately the same isoelectric point and since the differences in the mobilities of the components are relatively small, the partial fractionation of this material affords a good example of the application of the first method of operation described above to a naturally occurring protein. Although component A_1 converts irreversibly into A_2, the rate of conversion is slow enough to be negligible in the present investigation (MacPherson, Moore and Longsworth, 1944).

Fractionation experiments were carried out in phosphate buffer, pH 6.8 and ionic strength 0.05. At this pH and ionic strength the mobility of component A_1 is -7.08×10^{-5} $cm^2sec^{-1}volt^{-1}$, A_2, -5.93×10^{-5}, and A_3, -5.2×10^{-5}. In these experiments the three components were differentially transported out of the top and into the bottom reservoir at rates depending on their mobilities, with a relative enrichment of the top reservoir with respect to components A_2 and A_3 and of the bottom reservoir with respect to component A_1. The results of a series of single stage fractionations are presented in Table 1. The tabulated relative concentrations of A_2 were obtained by extrapolation of the apparent electrophoretic distributions to zero protein concentration at a constant ionic strength of 0.1 and pH 6.8. The contribution of component A_3 to the electrophoretic pattern is included with that of component A_2. The efficiency of separation is expressed in terms of the top separation factor defined by the relation

$$f_T = \frac{x_i}{x_i^0} \cdot \frac{1 - x_i^0}{1 - x_i} \qquad (1)$$

$$x_i = \frac{C_i}{\Sigma_j C_j}$$

In this case x^0_i and x_i are the initial and final ratios of component A_2 concentration to total protein concentration in the upper reservoir, respectively.

The field strength employed in these experiments was varied from 0.4 to 3.1 volts/cm. The duration of the runs were such as to transport about half of component A_1 out of the top reservoir. It will be noted that under the condition of half transport the top separation factor is independent of the field strength. It was also found that at constant field strength and time, the fraction of protein transported out of the top reservoir and the separation factor are independent of the initial protein concentration within the range 1 to 3 percent. These observations and the magnitude of the top separation factor are in agreement with the predictions of a mathematical theory of transport in the electrophoresis-convection channel.

TABLE 1. FRACTIONATION OF OVALBUMIN BY ELECTROPHORESIS-CONVECTION AT pH 6.8 AND IONIC STRENGTH 0.05

Run	E,[a] volt/cm.	t,[b] hr.	Total protein concentration % by wt.			Percent component A_2			
			Initial	Top	Bottom	Initial	Top	Bottom	f_T
1	0.4	2	1.2	0.8	1.6	30	34	28	1.20
2	1.2	1	1.1	0.4	1.4	30	35	29	1.26
3	1.2	1	1.6	0.9	2.1	30	34	26	1.20
4	1.2	1	3.3	1.6	3.8	32	38	31	1.30
5	1.6	¾	1.3	0.7	1.7	33	37	29	1.19
6	3.1	½	1.0	0.5	1.4	30	34	27	1.20

[a] Nominal field strength.
[b] Duration of run.

The average separation factor of 1.23 obtained in this series of runs indicated the possibility of accomplishing considerable enrichment of the slow moving components by successive batch fractionations, the top cut of each stage of fractionation serving as the starting material for the succeeding stage. Thus, starting with a solution of ovalbumin containing 20 percent A_2, after ten stages of fractionation the relative concentration of A_2 in the final top cut would be 67 percent. Such a three-stage batch fractionation was carried out at a field strength of 1.2 volts/cm and an operating time of 1 hr. The initial total protein concentration was 4.8 percent, and the initial relative concentration of A_2 was 22 percent. The solution withdrawn from the top reservoir after the third stage of fractionation analyzed 0.9 percent total protein and 31 percent A_2. This represents an over-all top separation factor of 1.59.

These experiments demonstrate that the separation of a mixture of proteins possessing nearly the

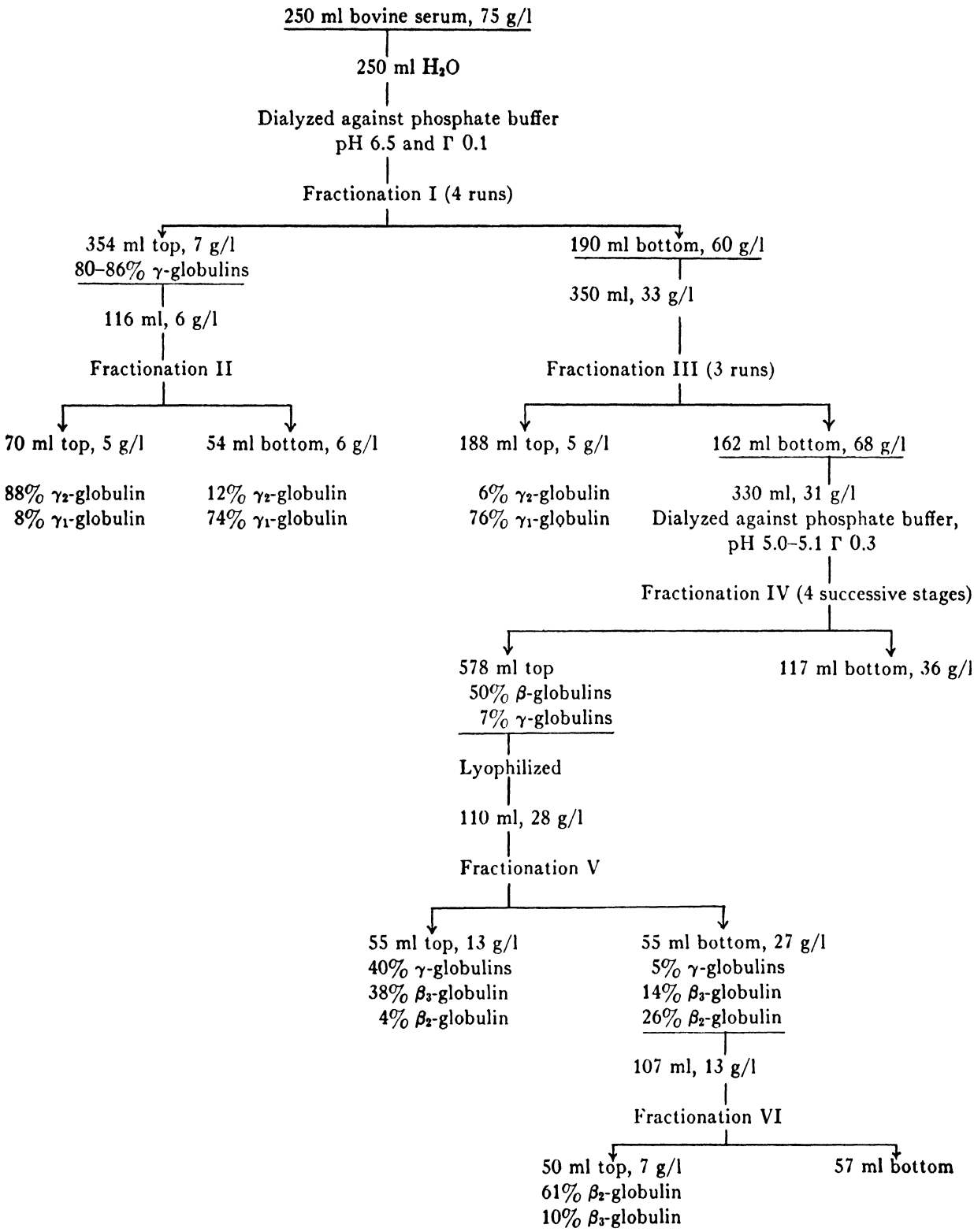


FIG. 2. Schematic fractionation of bovine serum.

same isoelectric point but different mobilities at pH's removed from the isoelectric point can be accomplished by electrophoresis-convection. However, it should be emphasized that when applicable the isoelectric procedure is more efficient by far than the procedure used to separate partially the components of ovalbumin.

Fractionation of Bovine Serum

Bovine serum, which is a mixture of at least seven proteins with appreciable differences in isoelectric points, is well suited for a test of the isoelectric fractionation procedure. The electrophoretic pattern of the serum used in this study is presented in Figure 3a. The mobilities and relative concentrations of the electrophoretic components of the serum are presented in Table 2. (In addition to the β_1- and β_2-globulin, a third β-globulin was present in the serum. This component, designated as β_3-globulin, had a mobility of -2.44×10^{-5}. Because of the poor resolution of this component in the serum, its contribution to the electrophoretic pattern has been included with that of β_2-globulin. However, in the case of the β-globulin fractions obtained in Fractionations V and VI, the β_3-globulin is sufficiently resolved electrophoretically to justify distinguishing between β_2- and β_3-globulin.) All electrophoretic data presented in this section were obtained in barbital buffer, pH 8.7 and ionic strength 0.1.

Table 2. The Mobilities ($-\mu \times 10^5$) and Relative Concentrations of the Components of Bovine Serum

	Albumin	α_1	α_2	β_1	β_2	γ_1	γ_2
Mobility	6.52	5.76	4.62	3.86	3.08	1.97	1.28
Relative concentration	43	3	13	4	22	8	7

Separation of the γ-globulin fractions from the serum was carried out in phosphate buffer, pH 6.5 and ionic strength 0.1. The isoelectric points of the γ-globulins appeared to be close to this pH. The other components of the serum migrated towards the anode on electrophoresis at pH 6.5. To separate the β-globulins runs were carried out at pH 5.0-5.1 and ionic strength 0.3. The scheme used in the fractionation is shown diagrammatically in Figure 2. The electrophoretic patterns of Figures 3, 4, and 5 follow the course of fractionation. The γ-globulins were separated from the serum and purified in Fractionations I, II, and III. β-globulin fractions were obtained in Fractionations IV, V, and VI.

The separation of crude γ-globulin fractions was accomplished in Fractionation I. Since the γ-globulins were practically immobilized by operating near

Table 3

A. Conditions for Separation of γ-Globulin Fractions

Fractionation	Run	E, volt/cm.	t, hr.	V_t, ml.	$c°$, g/100 ml.	V_0, ml.	f_T
I	1	3.1	14	110	3.7	7	23
	2	6.1	14	115	3.7	45	24
	3	3.1	24½	117	3.7	12	35
	4	3.1	27	155	3.7	9	35
II		3.1	24	116	0.6	8	4.2
III	Composite of 3 runs	3.1	15	350	3.3	—	46

B. Distribution of Electrophoretic Components into Fractions

Fractionation	Run	Fraction	Albumin	α_1	α_2	β_1	β_2[a]	γ_1	γ	γ_2
			percent of fraction							
Serum			43	3	13	4	22	8		7
I	1	top	1	—	3	—	16	67		13
	2	top	2	—	—	1[b]	16		81	
	3	top	1	—	—	—	13		86	
	4	top	1	—	—	—	13	63		23
	Composite of 1, 2, 3, 4	bottom	49	3	15	3	21		9	
II		top	—	—	—	—	4	8		88
		bottom	4	—	—	—	10	74		12
III	Composite of 3 runs	top	7	1	2	3	5	76		6
		bottom	53	3	16	2	20		6	

[a] Includes β_3-globulin.
[b] Includes trace of α_2-globulin.

their isoelectric points, only the albumin, α- and β-globulins were transported out of the upper and into the lower reservoir. Four runs were made in order to study the influence of field strength, time of operation, and volume of solution being fractionated upon the efficiency of separation. In all cases the initial concentration of the protein solution was

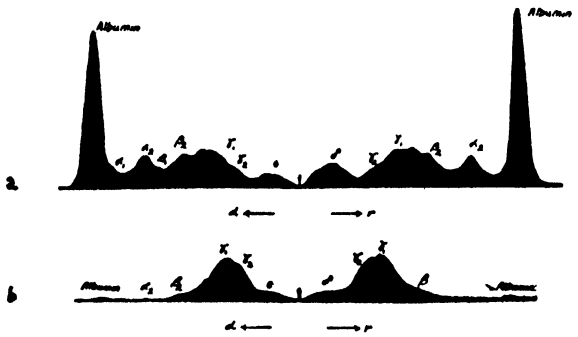


FIG. 3. Electrophoretic patterns obtained in barbital buffer, pH 8.7 and ionic strength 0.1; a, bovine serum; b, top fraction of Fractionation I, Run 1.

3.7 g protein/100 ml solution. (Osmosis, increasing the volume of protein solution, occurred during these runs. The volume of osmoid depended upon the conditions of fractionation. It was found that osmosis increases with increasing field strength and increasing protein concentration and decreases with increasing ionic strength. It is desirable to minimize osmosis, since the passage of water through the membranes into the channel interferes with the convection current, and decreases the efficiency of separation.) The fractionation data are presented in Table 3, where

c^o = initial concentration of protein solution in g per 100 ml solution.

V_i = initial volume of protein solution in ml.

V_0 = volume of osmoid in ml.

The other symbols have the same meaning as in Table 1. The efficiency of separation is expressed in terms of the top separation factor defined by eqn. (1). In the case x^o_i and x_i are the initial and final ratios of the concentration of the immobilized component to total protein concentration in the top reservoir.

The initial separations yielded γ-globulin fractions, the compositions of which compare favorably with that of the crude γ-globulin fraction obtained by Hess and Deutsch (1948) using ethanol precipitation. Thus the top fraction of Run 4 had the composition 86 percent γ-globulin, 13 percent β-globulins and 1 percent albumin, while the γ-globulin separated in crude form by ethanol precipitation was of composition 85 percent γ-globulin and 15 percent β-globulins. Figure 3b represents the electrophoretic pattern of a representative crude γ-globulin fraction. Partial separation of γ_1- and γ_2-globulin was also accomplished in these initial fractionations. The relative concentrations of γ_1- and γ_2-globulin in the serum were 8 percent and 7 percent respectively. The top fraction of Run 1 analyzed 67 percent γ_1-globulin and 13 percent γ_2-globulin, and the top fraction of Run 4, 63 percent γ_1-globulin and 23 percent γ_2-globulin. Electrophoretic resolution of the γ-globulin peak was not obtained on analysis of the solutions from the top reservoir of Runs 2 and 3.

Removal of albumin, α-globulins and β-globulins from a composite of the initial γ-globulin fractions obtained in Runs 1, 2, and 3 was accomplished in Fractionation II. The electrophoretic patterns of the resulting top and bottom cuts are shown in Figures 4a and 4b, respectively. Analysis of the solution in the top reservoir yielded 96 percent γ-globulins and 4 percent β-globulins. Considerable separation of γ_1- and γ_2-globulin was also accomplished in this experiment. Referring to Table 3, it will be noted that the relative concentrations of γ_1- and γ_2-globulin in the upper reservoir were 8 percent and 88 percent, respectively, while in the lower reservoir they were 74 percent and 12 percent, respectively.

Although the calculation of a separation factor for this fractionation is complicated by the transport of γ_1-globulin, it appears that the efficiency of fractionation was markedly lower than those obtained in Fractionation I. The only difference between Run 1 of Fractionation I and Fractionation II was that the concentration of the protein solution in the former run was 3.7 g/100 ml as compared to 0.6 g/100 ml in the latter run. The influence of concentration upon operating efficiency was confirmed by the results of two runs carried out for the purpose of separating crude γ-globulin fractions from serum. Except for concentration,

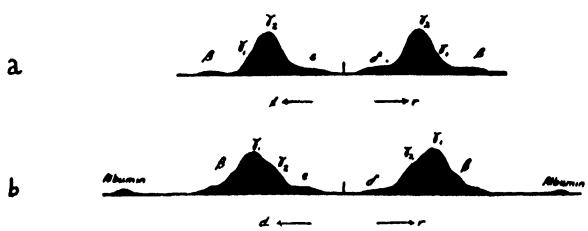


FIG. 4. Electrophoretic patterns obtained in barbital buffer, ph 8.7 and ionic strength 0.1; a, top fraction of Fractionation II; b, bottom fraction of Fractionation II.

the operating conditions of these runs were identical with those used in Run 1. It was found that the separation factor was 23 for an initial concentration of 2.9 g/100 ml and only 7 for an initial concentration of 0.8 g/100 ml.

The material taken from the bottom reservoir of Fractionation I was further processed in Fractiona-

tion III to remove γ-globulins. The data are presented in Table 3. Analysis of the solution withdrawn from the top reservoir yielded 76 percent γ_1-globulin and 6 percent γ_2-globulin. The solution taken from the bottom reservoir contained only 6 percent γ-globulins. Thus, the concentration of γ-globulins in the serum was reduced by about 60 percent in two successive stages of fractionation. Further fractionation would have eventually removed the lower reservoir, leaving the β_2-globulin in the top reservoir. Under the same conditions of field strength and duration of electrolysis the operating efficiency in this instance was less than in the case of the separation of the γ-globulins. There are two reasons for this. First, the mobilities of the migrating components were less than in the case of the γ-globulin fractionations; and the rate of horizontal electrophoretic transport was smaller. Also, at the

TABLE 4

A. CONDITIONS FOR SEPARATION OF β-GLOBULIN FRACTIONS

Fractionation	Stage	E, volt/cm.	t, hrs.	$c°$, g/100 ml.	f_T
IV	a	2.3	29	3.1	4.3
	b	3.5	31	2.9	7.5
	c	3.5	29	2.1	—
	d	3.5	27	—	—
V		2.3	30	2.8	—
VI		2.3	30	1.3	3.7

B. DISTRIBUTION OF ELECTROPHORETIC COMPONENTS INTO FRACTIONS

Fractionation	Stage	Fraction	Albumin	α_1	α_2	β_1	β_2	β_3	γ
			percent of fraction						
Serum			43	3	13	4	22[a]	—	15
IV	a	top	16	8	9	3	52[a]	—	12
		bottom	61	3	17	1	14[a]	—	4
	b	top	20	7	13	1	55[a]	—	4
		bottom	65	3	17	4	10[a]	—	1
	composite c and d	top	28	9	20	4	35[a]	—	4
	d	bottom	70	4	17	3	6[a]	—	—
V		top	6	—	12[b]	—	4	38	40
		bottom	29	5	18	3	26	14	5
VI		top	10	—	14[b]	—	61	10	5
		bottom	39	5	22	1	24	4	5

[a] Includes β_3-globulin.
[b] Includes α_1 and β_1-globulin.

all of the γ-globulins, but for the purpose of this investigation that was not warranted.

Subsequent to the separation of γ-globulin fractions, crude β-globulin fractions were separated from the serum by fractionation at pH 5.0-5.1 and ionic strength 0.3. Two further stages of fractionation served to purify these crude fractions. The pertinent data are presented in Table 4. The top separation factor is defined by a relation analogous to eqn. (1).

At pH 5.0-5.1 the β_2-globulin appeared to be close to its isoelectric point. As a result the mobile serum proteins were transported out of the upper and into operating pH the residual γ-globulins in the serum were opposite in sign to those of albumin, α-globulins and β_1-globulin. The horizontal electrophoretic transport of a component counter to that of other mobile components decreases the horizontal density gradient across the channel, thereby decreasing the convective velocity. Indeed, a steady state should eventually be established after which no further change in the composition of the solutions in the reservoirs will occur.

The separation of crude β-globulin fractions was accomplished by four successive stages of fractionation, Stage a, b, c and d of Fractionation IV. The

material from the bottom reservoir of Fractionation III served as the starting material for Stage a; the bottom fraction of Stage a was refractionated in Stage b, etc. Each stage represents a composite of two to three runs, the starting material and operating conditions being the same in each case. The electrophoretic pattern of a representative crude β-globulin fraction is shown in Figure 5a. The purest of these crude fractions had the composition 20 percent albumin, 20 percent α-globulins, 1 percent β_1-globulins, 55 percent β_2-globulin, and 4 percent γ-globulin. Electrophoretic analysis of the material

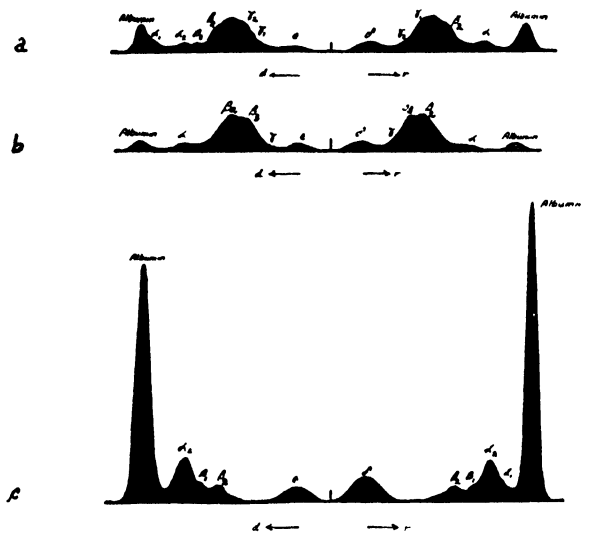


FIG. 5. Electrophoretic patterns obtained in barbital buffer, ph 8.7 and ionic strength 0.1; a, top fraction of Fractionation IV; b, top fraction of Fractionation VI; c, bottom fraction of Fractionation IV, Stage d.

obtained fom the bottom reservoir of Stage d yielded no γ-globulin and only 6 percent β_2-globulin, Figure 5c. Thus, four successive stages of fractionation resulted in a 70 percent reduction of the relative concentration of β_2-globulin.

After adjusting the concentration by lyophilizing, the composite of the crude β-globulin fractions served as starting material for Fractionation V. The results of this fractionation are quite interesting. Analysis of the material in the top reservoir yielded 42 percent β_2- and β_3-globulin and 40 percent γ-globulins. The bottom reservoir contained 40 percent β_2- and β_3-globulin and 5 percent γ-globulins. Apparently the γ-globulins, which had mobilities opposite in sign to the other mobile components, migrated into the upward convection current and were concentrated in the top reservoir. Considerable separation of β_2- and β_3-globulin was also accomplished. The top reservoir contained 4 percent β_2- and 38 percent β_3-globulin, and the bottom reservoir 26 percent β_2- and 14 percent β_3-globulin.

Finally, the solution taken from the bottom reservoir in Fractionation V was diluted and reprocessed in Fractionation VI. Analysis showed that the resulting top fraction contained 71 percent β-globulins. The electrophoretic pattern of this material is shown in Figure 5b. Further purification of the β-globulins was not warranted for the purpose of this investigation.

A theoretical analysis of transport in the electrophoresis-convection column allows the calculation of time of exhaust, θ, of a protein of mobility u, diffusion constant D, and initial concentration c° from a top reservoir of volume V. In a solvent of density ρ_0, viscosity coefficient η_0 and at a field strength E, one obtains eqn. (2) where b is the channel width, l the channel length, K an apparatus

$$\bar{\theta} = \frac{10^{-4}KVD}{hblu^2E^2} \qquad h = \left(\frac{2\eta_0 Dl}{\alpha\rho_0 g c^0}\right)^{1/4} \tag{2}$$

constant of the order of magnitude of one hour, g the acceleration of gravity, and $\alpha\rho_0$ the density increment produced by one gram of protein per 100 ml of solution. The fraction transported in time t is a function of t/θ. Since this equation is valid for low field strengths, estimations of exhaust times at high field strengths are correct only in order of magnitude.

From eqn. (2), it is seen that in an isoelectric fractionation of specified duration, the top separation factor for the immobilized component should increase with increasing field strength and mobility of the mobile component and remain relatively insensitive to the initial protein concentration. Although eqn. (2) is qualitatively confirmed by a number of experimental fractionations reported here, certain disturbing factors play a role which makes it undesirable to work at very high field strengths or at very low protein concentrations.

There appear to be three principal disturbing influences. The first is osmotic transport of solvent into the channel from the external buffer solution. This effect is reduced by decreasing the field strength and increasing the ionic strength. The second is the establishment of a stationary state before complete exhaust of the mobile components from the top reservoir. This effect occurs only when some of the mobile components have mobilities of opposite sign. When operating conditions cannot be chosen to avoid mobilities of opposite sign, some sacrifice in separation efficiency must be accepted. The third disturbing effect is the destruction of laminar flow in the channel, which can be inhibited by increasing the viscosity of the solution or by decreasing the field strength and the channel wall separation. Although a thorough investigation has not been made, it is surmised that decreased fractionation efficiency observed at low protein concen-

trations with correspondingly low viscosities, may be due to this effect.

In general, it is necessary to determine optimum operating conditions, minimizing the effect of the disturbing factors by pilot fractionations. However, except in unusual cases, separation factors of a satisfactory magnitude are attainable under conditions far removed from the optimum.

The Fractionation of Bovine γ-Globulin

The results obtained in the fractionation of bovine serum discussed above suggest that electrophoresis-convection should supplement the ethanol fractionation of biological tissues and fluids as carried out by Cohn, *et al.* (Cohn, Luetscher, Oncley, Armstrong, and Davis, 1940; Cohn, Strong, Hughes, Milford, Ashworth, Melin, and Taylor 1946). It is apparent that sub-fractionation of the plasma protein fractions obtained by alcohol precipitation should be an important application of the method of electrophoresis-convection. Its utility has been demonstrated by a sub-fractionation of bovine γ-globulin prepared by ethanol precipitation, Fraction II of bovine plasma. γ-Globulin was chosen because of its known heterogeneity and its immunological importance.

In the case of a protein which migrates as a single boundary in an electric field but has a specified mobility and isoelectric point distribution as revealed by reversible electrophoretic boundary spreading, fractionation is accomplished by the modified isolectric procedure. In this procedure the pH for fractionation is so chosen as to be several tenths to one pH unit removed from the average isoelectric point of the protein. Transport in the apparatus leads to a redistribution of the protein ions such that the fractions withdrawn from the top and bottom reservoir possess mobility and isoelectric point distributions differing from those of the original protein. Using this procedure γ-globulin has been separated into eight fractions of different mean mobilities and isoelectric points.

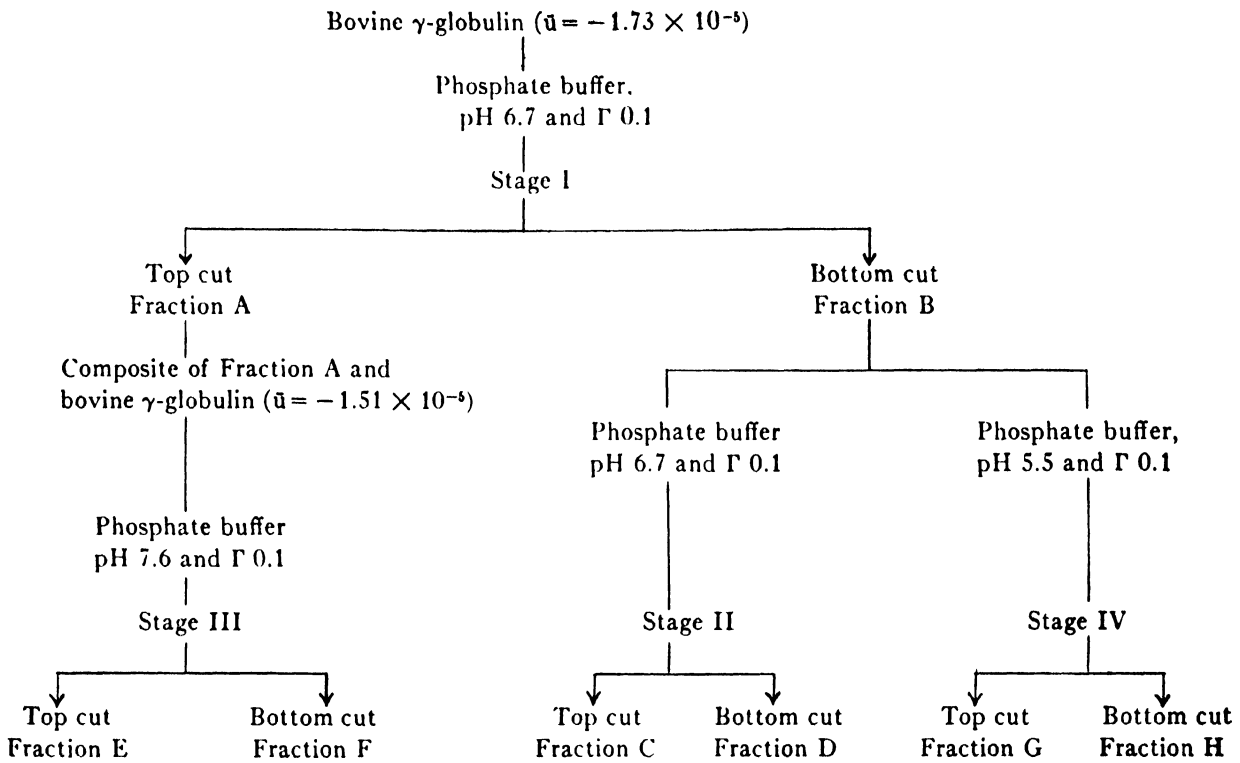


FIG. 6. Schematic fractionation of bovine γ-globulin.

Two different samples of bovine γ-globulin were used in this investigation. The one sample had a mobility of -1.73×10^{-5} in barbital buffer pH 8.7 and ionic strength 0.1. The other had a mobility of -1.51×10^{-5}. The isoelectric points of these two samples in cacodylate buffer (0.08N NaCl −0.02N Na cacodylate) were 6.5 and 6.75, respectively. The γ-globulin migrated as a single boundary during electrophoresis. The fractionation of γ-globulin was carried through four stages. The scheme used in fractionation is shown diagrammatically in Figure 6. Since the starting material for

Stage III is a composite of Fraction A and unfractionated γ-globulin, it does not represent a true stage in the fractionation scheme. However, for convenience we will refer to it as such.

The first stage of fractionation was carried out on 100-120 ml of a 2-3 percent γ-globulin solution equilibrated against phosphate buffer, pH 6.7 and ionic strength 0.1. γ-Globulin with a mobility of -1.73×10^{-5} served as the starting material in this stage. The operating pH was several tenths of a pH unit removed from the average isoelectric point of the γ-globulin. Eight runs were made in order to study the influence of field strength and duration of electrolysis upon the efficiency of fractionation. (Osmotic transport of solvent into the channel from the exterior buffer solution increased the volume of protein solution during these runs. The rate of influx of solvent was found to increase with the field strength.) The results of a representative series of runs are presented in Table 5, where

$\bar{u}$ = mean mobility ($cm^2\ sec^{-1}\ volt^{-1}$) at pH 8.7 in barbital buffer, ionic strength 0.1.

β = standard deviation of the mobility distribution ($cm^2\ sec^{-1}\ volt^{-1}$).

The standard deviations of the mobility distributions were determined by electrophoretic spreading experiments at the mean isoelectric points of the proteins in cacodylate buffer (0.08N NaCl, −0.02N Na cacodylate). β is referred to as a heterogeneity constant (see appendix). The other symbols have the same meaning as in previous tables.

TABLE 5. FIRST STAGE OF FRACTIONATION OF BOVINE γ-GLOBULIN

A. EXPERIMENTAL CONDITIONS

Run	E, volt/cm.	t, hr.	c^0, g/100 ml.
1	3.1	24	3.0
2	2.6	47	3.2
3	1.6	42	2.4
4	1.6	77	2.4
5	1.0	118	2.8

B. PROPERTIES OF γ-GLOBULIN FRACTIONS

Run	Fraction	Yield g. protein	$-10^5 \times \bar{u}$	$10^5 \times \beta$
1	top	—	1.52	0.64
	bottom	—	2.10	—
2	top	1.4	1.36	0.62
	bottom	1.6	2.07	—
3	top	1.1	1.33	0.67
	bottom	1.4	2.12	—
4	top	1.4	1.33	0.66
	bottom	1.4	2.02	—
5	top	1.3	1.38	—
	bottom	1.6	1.98	—

With the exception of Run 1 these initial separations all yielded the same top fraction, designated as Fraction A. Fraction A had a mobility of -1.35×10^{-5} and an isoelectric point of 7.03 in cacodylate buffer, about 0.5 pH unit greater than the mean isoelectric point of γ-globulin. The slope of the mobility—pH curve in the neighborhood of the isoelectric point was found to be -0.74×10^{-5}. The heterogeneity constant determined for this fraction is 0.65×10^{-5}.

The top fraction resulting from Run 1 had a mobility of -1.52×10^{-5}, which is significantly greater than that of Fraction A, and the same heterogeneity constant as Fraction A. It appears that the optimum conditions for fractionation are 2.6-1.0 volts/cm at operating times of 48-118 hrs., respectively. A number of routine fractionations of γ-globulin, employing a field strength of 1.6 volts/cm and an operating time of 48 hrs., have been carried out in this laboratory. The results obtained in these runs have shown that fractionations accomplished by electrophoresis-convection are very reproducible.

The bottom fractions, Fraction B, resulting from the first stage of fractionation had mobilities ranging from -1.98 to -2.12×10^{-5}.

The second stage of fractionation was carried out on 120 ml of about a 2 percent solution of a composite of Fractions B resulting from Runs 1, 2, and 4. The protein solution was equilibrated against phosphate buffer, pH 6.7 and ionic strength 0.1. In this experiment osmotic transfer of solvent into the channel from the external buffer solution was practically negligible. The data are presented in Table 6, where

I.P. = isoelectric point in cacodylate buffer (0.08N NaCl −0.02N Na cacolylate).

$\Delta\bar{u}/\Delta$pH = slope of the mobility—pH curve in the neighborhood of the isoelectric point.

The resulting top fraction, Fraction C had a mobility of -1.63×10^{-5} and an isoelectric point of 6.47, which is very close to the mean isoelectric point of γ-globulin itself. The bottom fraction, Fraction D, had a mobility of -2.20×10^{-5} and an isoelectric point of 6.01, about 0.5 pH unit lower than the mean isoelectric point of γ-globulin. Both fractions were found to have the same heterogeneity constant as Fraction A.

Fraction A represented 45 percent of the original γ-globulin, Fraction C 19 percent and Fraction D 36 percent. These fractions possessed Gaussian mobility distributions. The mobility distribution of the original γ-globulin was not Gaussian, of course. However, as a first approximation the mobility dis-

tribution can be adequately represented as a Gaussian function whose standard deviation is taken as that of the actual mobility distribution. (A complete representation of the mobility distribution entails, of course, the determination of the third and higher moments of the refractive-index gradient curves, which give the deviations from a Gaussian distribution, eqn. (3) of the appendix.) The standard deviation, β, of the actual mobility distribution was found to be 0.67×10^{-5}. The mobility distribution of each fraction at pH 6.5 has been normalized to an area corresponding to its weight fraction of the γ-globulin and plotted in Figure 7. The solid curve in Figure 7 represents the sum of the three distributions. In Figure 8 this composite curve is compared with the Gaussian probability function which represents to a first approximation the mobility distribution of γ-globulin. The mobility distribution of γ-globulin is normalized to unit area. It will be noted that the agreement between the two curves is good.

On the basis of electrophoretic characterization the difference between Fraction C and γ-globulin are relatively small. The mobility of Fraction C was found to be -1.63×10^{-5} and that of γ-globulin -1.73×10^{-5}. The two proteins have approximately the same isoelectric points. The chief difference between the two proteins is that Fraction C has a Gaussian and γ-globulin a non-Gaussian distribution of mobilities. Fraction C appears to be a center cut of γ-globulin resulting from the separation of two fractions, A and D, possessing mobility distributions the centroidal axes of which are situated on either side of the centroidal axis of the mobility distribution of γ-globulin.

A composite of Fraction A and γ-globulin (-1.51×10^{-5}) served as the starting material for Stage III. It will be noted that the mobility of this γ-globulin is the same as that of the top cut of Run 1, Stage I. This stage of fractionation was carried out in phosphate buffer, pH 7.6. This pH is about 0.6 pH unit on the alkaline side of the mean isoelectric point of the starting material. The fractionation was carried out in duplicate. The pertinent data are presented in Table 6.

TABLE 6. FRACTIONATION OF BOVINE γ-GLOBULIN

A. EXPERIMENTAL CONDITIONS

Stage	Run	pH	E, volt/cm.	t, hr.	c^0, g/100 ml.
II		6.70	1.6	48	~2
III	1	7.60	1.7	52½	2.8
	2	7.58	1.7	53½	2.4
IV		5.48	1.6	51½	2.3

B. PROPERTIES OF γ-GLOBULIN FRACTIONS

Stage	Run	Fraction	Yield g. protein	$-10^5 \times a$	I.P.	$10^5 \times \beta$	$-10^5 \times \frac{\Delta a}{\Delta pH}$
II		top (C)	1.0	1.63	6.47	0.67	1.3
		bottom (D)	1.7	2.20	6.01	0.65	0.78
III	1	top (E)	1.1	1.24	7.31	—	0.35
		bottom (F)	2.0	1.78	6.51	0.59	0.89
	2	top (E)	1.2	1.25	—	0.55	—
		bottom (F)	1.5	1.69	—	—	—
IV		top (G)	0.6	2.25	5.74	0.63[a]	0.64
		bottom (H)	2.3	1.81	6.41	0.77[a]	0.81

[a] Non-Gaussian mobility distribution.

The electrophoretic properties of the resulting top cut, Fraction E, are quite striking. The mobility of Fraction E was -1.25×10^{-5}. The mean isoelectric point was found to be 7.31, about 0.8 pH unit greater than that of the original γ-globulin (-1.73×10^{-5}). Furthermore, the heterogeneity constant was significantly lower than the values found for the other top fractions; and $\Delta\bar{u}/\Delta pH$ was considerably lower than the values for the other fractions. The bottom cut, Fraction F, had a mobility of -1.74×10^{-5} and an isoelectric point of 6.51.

The top and bottom fractions of Run 2 were re-

combined in proportion to their relative concentrations in the starting material. The mobility and heterogeneity constant of this composite were -1.45×10^{-5} and 0.59×10^{-5}, respectively, which are to be compared with the calculated values of -1.54×10^{-5} and 0.67×10^{-5}.

A composite of Fraction B and γ-globulin (-1.51×10^{-5}) appeared to be roughly comparable to the γ-globulin of mobility -1.73×10^{-5}. The composite was fractionated under the same conditions as in Stage I. The resulting bottom fraction had a mobility of -1.91×10^{-5}, an isoelectric point of 6.29, and a heterogeneity constant of 0.75×10^{-5}. This material possessed a Gaussian distribution of mobilities at its mean isoelectric point, although at pH 8.7 the electrophoretic pattern was non-Gaussian and skewed. Differences in the dependence of mobility upon pH for the various components of the fraction undoubtedly are responsible for this departure of the mobility distribution from a Gaussian. This fraction, which appeared to be the same as Fraction B, served as the starting material for Stage IV.

The operating pH in Stage IV was 5.5 or 0.8 pH units on the acid side of the mean isoelectric point of Fraction B. The material withdrawn from the top reservoir, Fraction G, had about the same mobility as Fraction D. However, its isoelectric point was about 0.3 pH unit lower than that of Fraction D, and its mobility distribution was non-Gaussian. The bottom cut, Fraction H, which had a mobility of -1.81×10^{-5} and an isoelectric point of 6.41, possessed a skewed non-Gaussian distribution of mobility with a rather large standard deviation. The calculated mobility and heterogeneity constant for a composite of the two fractions, in proportion to their relative concentrations in the starting material, are -1.91×10^{-5} and 0.76×10^{-5}, respectively.

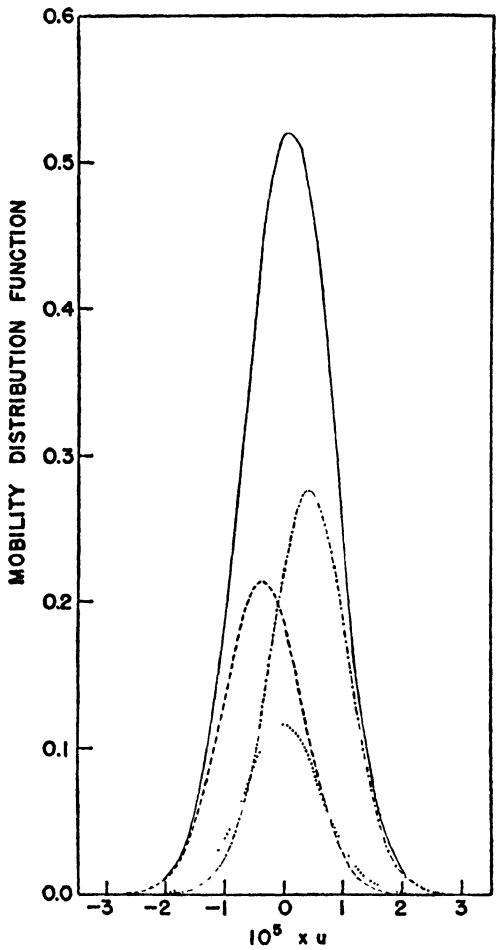


FIG. 7. Mobility distribution functions at pH 6.5; - · -, Fraction A; . . . , Fraction C; - - -, Fraction D; ——, sum of the mobility distributions of Fractions A, C and D.

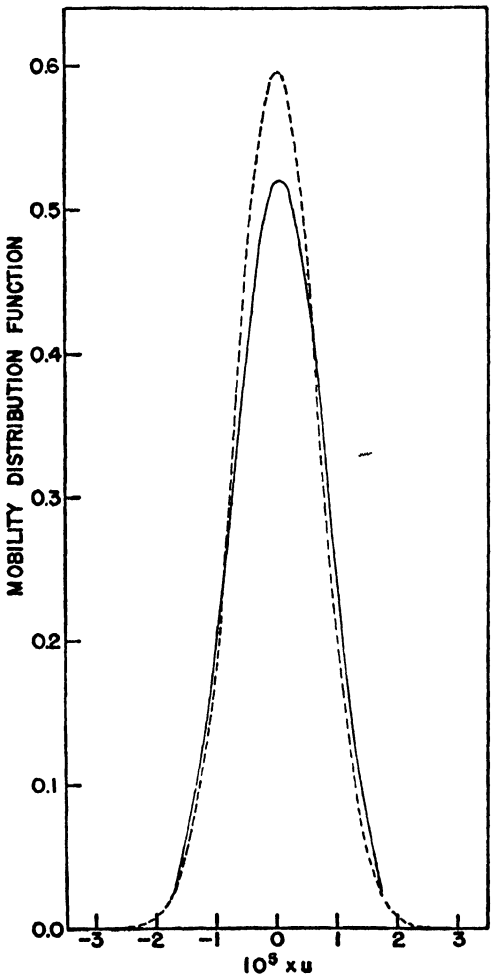


FIG. 8. Mobility distribution functions at pH 6.5; ——, sum of the mobility distributions of Fractions A, C and D; - - -, bovine γ-globulin ($\bar{u} = -1.73 \times 10^{-5}$).

These calculated values are in excellent agreement with the experimental values quoted above for Fraction B.

To summarize, γ-globulin has been separated into eight fractions which constitute mean mobility and mean isoelectric point spectra which range from -1.25×10^{-5} to -2.25×10^{-5} and from 7.31 to 5.74, respectively. The various fractions can be distinguished from one another by at least two electrophoretic properties. Thus, Fractions C, F, and H have about the same mean isoelectric points but different mean mobilities at pH 8.7, and different mobility distributions in the neighborhood of their respective mean isoelectric point. The considerable

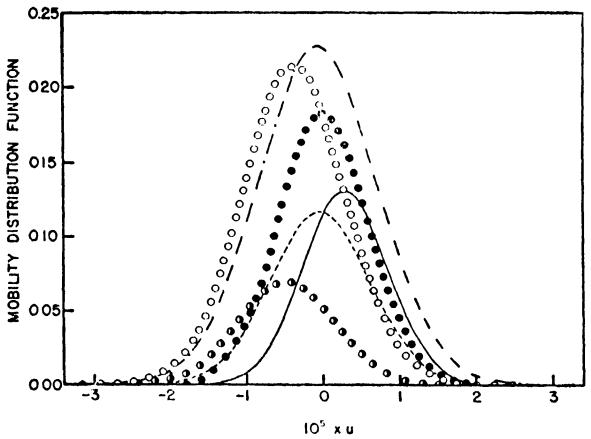


FIG. 9. Mobility distribution functions at pH 6.5 of the six unique fractions of bovine γ-globulin; – –, Fraction C, ○, Fraction D; —, Fraction E; ●, Fraction F; ◑, Fraction G; – · –, Fraction H.

variation in $\Delta\bar{u}/\Delta pH$ among the fractions is probably a reflection of chemical and structural differences. The mobility distribution of each of the six unique fractions has been normalized to an area corresponding to its weight fraction of γ-globulin (-1.73×10^{-5}) and plotted in Figure 9. The relative concentrations of Fractions E and F were calculated on the assumption that the material fractionated in Stage III was Fraction A and not a composite of Fraction A and γ-globulin (-1.51×10^{-5}). Except for Fractions G and H the boundary spreading technique failed to reveal any perceptible departure of the mobility distributions from a Gaussian. However, as a first approximation the mobility distribution of Fractions G and H can be represented as a Gaussian probability function, whose standard deviation is taken as that of the actual mobility distribution (see appendix).

The theory of transport in an electrophoresis-convection channel predicts that the fractionation of a protein possessing a Gaussian mobility distribution, with specified first and second moments, will result in top fractions which also possess Gaussian mobility distributions. Furthermore, the second moments of the mobility distributions of the top fractions will be the same as that of the original protein. The first moments will, of course, be different than that of the original protein. The theory also predicts that transport in the apparatus will be stopped by attainment of a stationary state when the first moment of the mobility distribution of the material in the top reservoir vanishes. Both of these predictions have been approximately realized in the experiments reported in this paper. The mobility distribution of the material fractionated in Stage IV was non-Gaussian and skewed at pH's removed from the mean isoelectric point. As a result, the top fraction obtained in this stage exhibited a non-Gaussian mobility distribution at its mean isoelectric point.

APPENDIX

The electrophoretic inhomogeneity of γ-globulin and its fractions is evidenced by reversible boundary spreading. The rate of spreading of an electrically homogeneous protein boundary under conditions such that convection and anomalous electrical effects are avoided should be no greater than that due to diffusion alone. However, in the case of an inhomogeneous protein the refractive-index gradient curve is spread simultaneously by diffusion and by the differences in the mobilities of the constituent protein ions.

It is a result of the general theory of reversible electrophoretic boundary spreading (Brown and Cann, 1949) that the mobility distribution of a heterogeneous protein, q(u), can be expressed in terms of the moments of the gradient curve taken about the centroidal axis by means of an infinite series, eqn. (3). D is the diffusion constant, and E the electric field strength.

$$q(u) = \frac{1}{\sqrt{2\pi\beta}} e^{-u^2/2\beta^2} \left\{1 + \sum_{j=3}^{\infty} \frac{C_j}{j!} (-i)^j \alpha^j H_j(iu\sigma/Et_E\beta^2\alpha)\right\}$$
$$\beta^2 = (\sigma^2 - \sigma_0^2 - 2Dt_E)/E^2t_E^2 \qquad (3)$$
$$\alpha = \sqrt{1 - 2(\sigma/\beta E t_E)^2}$$

σ_0^2 and σ^2 are the second moments of the gradient curves at the moment of application and at time t_E after application of the electric field. β is the standard deviation of the mobility distribution, referred to as a heterogeneity constant. H_j is the j-th Hermite polynomial. The coefficients C_j are related to

the higher moments of the gradient curves, *e.g.*,

$$C_3 = \overline{\chi^3}/\sigma^3, \quad C_4 = \overline{\chi^4}/\sigma^4 - 3.$$

If the gradient curves are Gaussian in form, the mobility distribution is Gaussian. Departures from a Gaussian distribution of mobilities are given by the third and higher moments of the gradient curves. β may be calculated from eqn. (4). D* is the "apparent diffusion constant" calculated from the second moments of the gradient curves during electrophoresis.

$$D^* = \frac{\sigma^2 - \sigma_0^2}{2t_E} = D + (E^2\beta^2/2)t_E \qquad (4)$$

A plot of D* *vs.* t_E is a straight line which extrapolates back to the normal diffusion constant at zero time. β may be calculated from the slope $E^2\beta^2/2$.

Summary

The utility of the method of electrophoresis-convection is well illustrated by the representative fractionations described in this paper. The high efficiencies of separation and the large quantities of material fractionated in a single run, without sacrifice of purity of the fractions, coupled with the ease of manipulations and economy of time promise to make electrophoresis-convection a valuable method for the fractionation of naturally occurring protein mixtures and for the subfractionation of protein fractions obtained by ethanol precipitation.

References

ALBERTY, R. A., ANDERSON, E. A., and WILLIAMS, J. W., 1948, Homogeneity and the electrophoretic behavior of some proteins. J. phys. colloid Chem. *52:* 217-230.

BROWN, R. A., and CANN, J. R., 1949, Extension of the theory of reversible electrophoretic boundary spreading of proteins. J. phys. colloid Chem.: in press.

CANN, J. R., BROWN, R. A., and KIRKWOOD, J. G., 1949, The fractionation of bovine serum proteins by electrophoresis-convection. J. Amer. chem Soc. *71:* 1609-1614.

1949, The fractionation of γ-globulin by electrophoresis-convention. J. Amer. chem. Soc.: *71:* 2687-2691.

CANN, J. R., KIRKWOOD, J. G., BROWN, R. A., and PLESCIA, O. J., 1949, The fractionation of proteins by electrophoresis-convection. An improved apparatus and its use in the fractionation of diphtheria antitoxin. J. Amer. chem. Soc. *71:* 1603-1608.

COHN, E. J., LUETSCHER, J. A., JR., ONCLEY, J. L., ARMSTRONG, S. H., JR., and DAVIS, B. D., 1940, Preparation and properties of serum and plasma proteins. III. Size and charge of proteins separating upon equilibration across membranes with ethanol-water mixtures of controlled pH, ionic strength and temperature. J. Amer. chem. Soc. *62:* 3396-3400.

COHN, E. J., STRONG, L. E., HUGHES, W. L., MULFORD, D. J., ASHWORTH, J. N., MELIN, M., and TAYLOR, H. L., 1946, Preparation and properties of serum and plasma proteins. IV. A system for the separation into fractions of the proteins and lipoprotein components of biological tissues and fluids. J. Amer. chem. Soc. *68:* 459-475.

HESS, E. L., and DEUTSCH, H. F., 1948, Biophysical studies of blood plasma proteins. VIII. Separation and properties of the γ-globulins of the sera of normal cows. J. Amer. chem. Soc. *70:* 84-88.

KIRKWOOD, J. G., 1941, A suggestion for a new method of fractionation of proteins by electophoresis-convection. J. chem. Phys. *9:* 878-879.

LONGSWORTH, L. G., 1941, The influence of pH on the mobility and diffusion of ovalbumin. Ann. N. Y. Acad. Sci. *41:* 267-285.

LONGSWORTH, L. G., CANNAN, R. K., and MACINNES, D. A., 1940, An electrophoretic study of the proteins of egg white. J. Amer. chem. Soc., *62:* 2580-2590.

MACPHERSON, C. F. C., MOORE, D. H., and LONGSWORTH, L. G., 1944, Changes in the electrophoretic composition of crystalline hen's egg albumin with age. J. biol. Chem. *156:* 381-382.

NIELSEN, L. E., and KIRKWOOD, J. G., 1946, The fractionation of proteins by electrophoresis-convection. J. Amer. chem. Soc. *68:* 181-185.

Discussion

ANSON: The success of Dr. Kirkwood's elegant combination of electrophoresis and convection ought to remind us that there are many combinations of methods for the fractionation of proteins which remain to be investigated adequately. I should like to call attention to the neglected ion-accelerated diffusion, which is a combination of diffusion and a sort of electrophoresis in which the electromotive force is set up in the solution by the difference in the rates of diffusion of small and large ions.

Ordinary diffusion is not very good for separating different proteins. First, diffusion of the large protein molecules is slow. Secondly, the rate of diffusion is not very sensitive to the size of the protein molecule, since it depends on the cube root of the radius.

In the ordinary diffusion of proteins, charge effects are eliminated by having salt present or by having the protein isoelectric. In ion-accelerated diffusion, charge effects are produced by the addition of acid or base to salt-free protein. Thus if salt-free protein is allowed to diffuse from hydrochloric acid solution into water, the protein diffuses at a rate between the slow rate of diffusion of protein molecules in salt solution and the much faster rate of diffusion of the small chloride ions. The small chloride ions tend to diffuse ahead of the large protein molecules into the water and they drag the oppositely charged protein ions along with them.

The rate of ion-accelerated diffusion of a protein is not only greater than the rate of diffusion in salt solution but is much more specific for the protein. Whereas the rate of ordinary diffusion depends only on the size and shape of the protein molecule, the rate of ion-accelerated diffusion, like the rate of electrophoresis, depends on the charge of the protein at the pH used.

KIRKWOOD: Dr. Anson's suggestion is a very interesting one which merits investigation.

BUTLER: I have been experimenting with electrophoresis down closed tubes packed with porous

materials such as glass wool and asbestos fiber. Good separations of both proteins and amino acids can be made in this way when the mobilities differ by 20 percent or more. How would this compare with the efficiency of Dr. Kirkwood's apparatus?

KIRKWOOD: We are able to carry out fractionations with adequate efficiency with mobility margins of less than 20 percent.

NEURATH: I believe you said that the electrophoretic pattern of the *gamma*$_2$ globulin fraction had the shape of a Gaussian distribution curve and that the pattern obtained from an isolated cut again resembled a Gaussian distribution curve. I am wondering whether these curves were really Gaussian when plotted in normal coordinates and whether the identical distribution of the curves of the original fraction and of its sub-fraction does not follow necessarily from the assumption that the curves are Gaussian which they may not be? I should also like to ask Dr. Kirkwood whether he has any physical interpretation of this finding.

KIRKWOOD: The theory of the fractionation unit predicts that the top cut of a Gaussian distribution in mobility will also be Gaussian. Our gamma globulin fractions appeared in most cases to be nearly Gaussian. However, it is difficult to estimate higher moments than the second with much precision in a boundary spreading experiment.

STUDIES ON POLYPEPTIDES AND AMINO ACIDS BY COUNTERCURRENT DISTRIBUTION

LYMAN C. CRAIG, J. DELAFIELD GREGORY AND GUY T. BARRY
The Rockefeller Institute for Medical Research, New York

Because of the importance of amino acids, polypeptides and proteins an attempt to apply nearly every method thus far devised by chemists for fractionation, characterization and proof of purity of organic substances has at some time been made. It is thus only logical to attempt application of the method and viewpoint of "Countercurrent Distribution." Perhaps some definition of countercurrent distribution should be given for the benefit of those who are not familiar with the method (Craig and Post, 1949) and the published work which has dealt mainly with antimalarials, penicillins, fatty acids, antibiotics, pyrimidines, streptomycins etc. The method depends on a stepwise fractionation process involving the distribution of solute between two immiscible phases in a series of contacting units. It is performed in such a manner that for a single solute the fraction present in any unit, at any stage in the process, will exactly correspond to a given term of a particular binomial expansion. It is strictly an equilibrium process since before they are separated, the immiscible phases are equilibrated for a length of time well in excess of that experimentally required to reach equilibrium (Barry, Sato and Craig, 1948). For solutes with constant partition ratios at different concentrations the fraction present in a given cell is fixed by the value of the partition ratio, the relative volumes of the phases and the number of transfers. All of these requisites can be measured directly and no assumptions are required.

Countercurrent distribution thus coincides closely to the ideal process which forms the yardstick usually used for evaluating the efficiency of truly continuous column processes. Since efficiency calculations of chromatography and partition chromatography on several occasions (Martin and Synge, 1941; Moore and Stein, 1948) have indicated a relatively short column often to have in excess of 1000 theoretical plates, it would, at first thought, not appear rewarding to try to develop the strictly stepwise procedure for actual use. The relation of the two procedures, however, is by no means as simple as these calculations would indicate.

Although the transition from algebra to calculus theoretically is analogous to the transition from the stepwise to the truly continuous process, many pertinent experimental pitfalls may easily be overlooked when such an analogy is carried too far. In fact it is even difficult to set up experiments which will be strictly comparative in every respect for the continuous and discontinuous process. Furthermore, the value of a separating tool depends not only on high resolving power under a particularly favorable set of circumstances but also on other factors such as capacity, reproducibility with different mixtures or range of adherence to the ideal, ease of analysis and ease of isolation if the fractions obtained are in solution, general applicability etc. On many of these points the discontinuous process rates rather well. Its chief disadvantage would appear to be the labor involved in equilibrating and in making the necessary transfers when high numbers of transfers are required. This is purely a mechanical problem and is one which has concerned our group for the past few years.

Several apparatuses for accomplishing multiple extractions have been devised and studied extensively. One of these is the steel machine (Craig and Post, 1949) which has permitted up to 53 extractions to be made with one operation and up to 300 transfers to be made with profit. A more recent device operates on a different principle. It is made entirely from glass tubing with individual interlocking units. A single unit is shown in Figure 1, upper left.

The two phases containing the solute are placed in chamber A. The mixture is equilibrated by rocking back and forth at an angle of approximately 35°. The phases are allowed to separate and the apparatus is then tilted backward to an angle slightly more than 90° (Fig. 1 upper center). The upper phase from A at this position decants through B into C. The lower phase is of such volume that it all remains in A. On righting the unit again to the position of Figure 1 upper right, the contents of C flow out through D into the next unit since the units interlock as shown. The lower picture shows 108 of these units in a single train. Such a train permits 108 nearly quantitative extractions to be made in a single operation and up to 1000 transfers (approximately 102,000 extractions) to be made with profit.

We are now building an improved design based on this principle which has several times the number of tubes and also are attaching motors, time clocks etc. so that the fractionation will be fully automatic. The design lends itself admirably to such mechanization. If it should prove entirely successful the number of "plates" which can be brought to bear on a problem will be limited only by the laboratory space available, the cost of the glass units, solvents, etc.

However, as previously stated, high numbers of

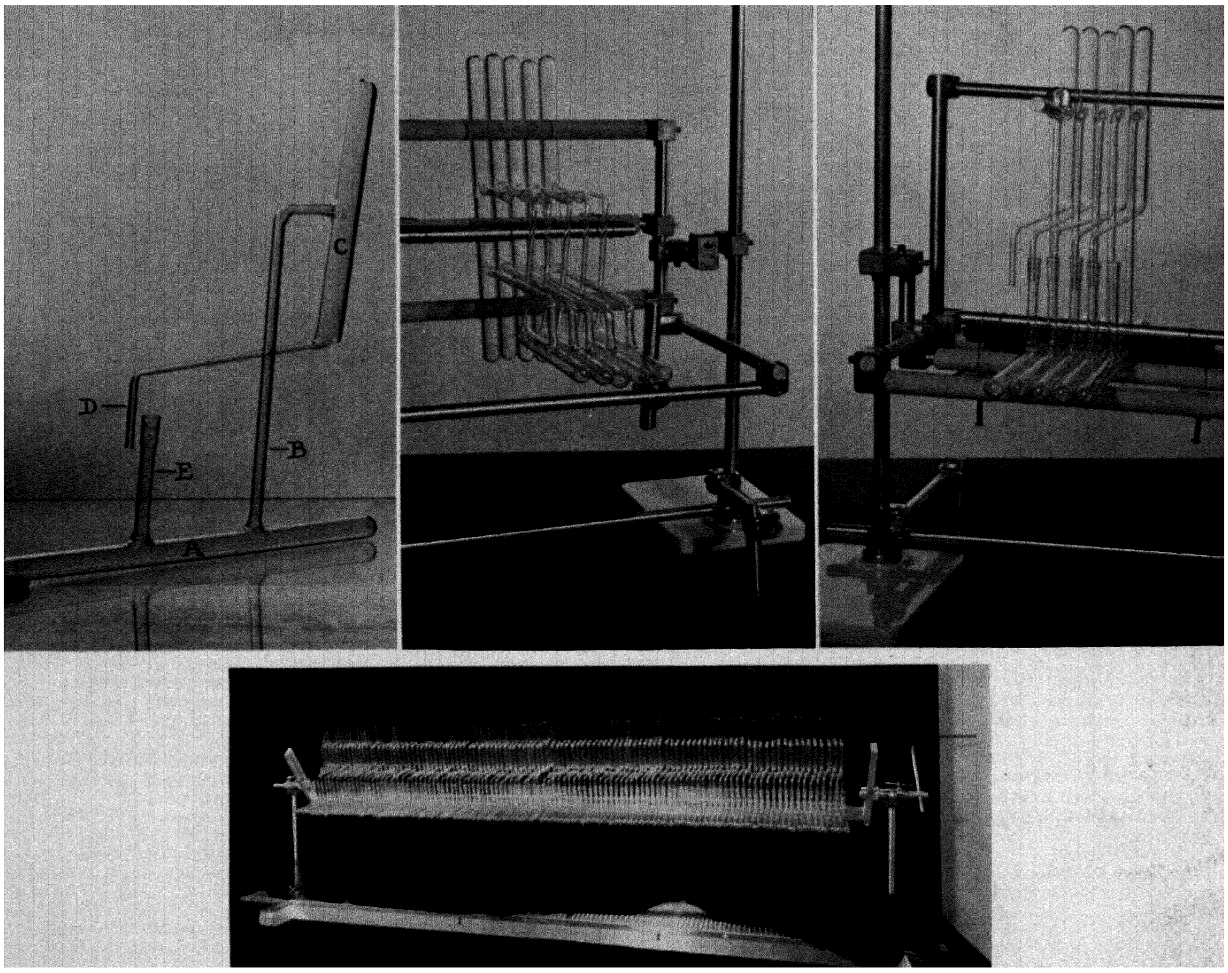


FIG. 1. Glass countercurrent distribution apparatus.

"plates" do not insure the required separating power. Systems must be found which will conform well enough to the ideal so that the high numbers of transfers applied can be used with advantage. For example, the amino acids are of "Zwitterion" character. They do not tend to associate strongly and are the most ideal substances we have yet found from the standpoint of a constant partition ratio at varying concentrations. As will be shown later, we can therefore calculate with assurance that on the basis of the systems we have thus far studied, the satisfactory separation of the amino acids from protein hydrolysates can be a matter of the intelligent use of a distribution apparatus containing several hundred tubes.

To our dismay, however, we have found that the higher polypeptides do not behave so ideally from the standpoint of constant partition ratios. Nonetheless, in spite of this, fractionation has been achieved with each of the four classes we have thus far studied, namely, gramicidin, tyrocidine, gramicidin-S and bacitracin.

In any structural study the first concern of the organic chemist has classically been that of obtaining a pure sample of substantial amount which could be proven to be pure by accepted methods at hand. However, it is not easy to obtain really decisive evidence that a polypeptide preparation of

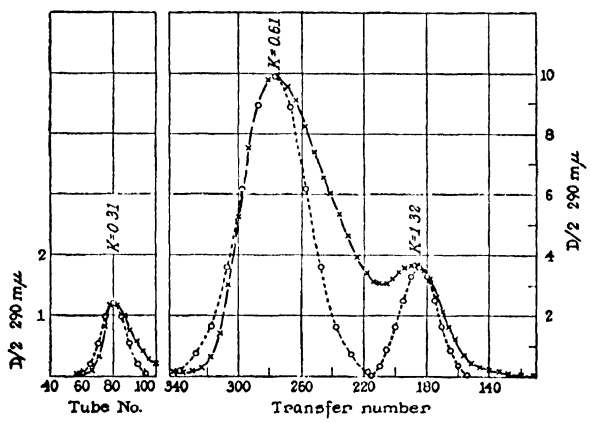


FIG. 2. Distribution pattern for a sample of crystalline gramicidin. Ext., x——x; Calc'd, o——o.

molecular weight in the range of 1000 to 10,000 is indeed a single substance. Even when heterogeneity has been indicated by some physical procedure, the exact amount of impurity usually remains in doubt and the problem has not been solved until some method of removing the impurity has been found. This basic problem of purity has been our chief objective in the polypeptide work thus far. It has mainly involved the search for suitable systems.

Crystalline gramicidin (Hotchkiss, 1941) gave the pattern shown in Figure 2, with the system methanol-water-benzene-chloroform (23:7:15:15 volumes). Three definite gramicidins are separated and at least one more is indicated. We have isolated the three in crystalline form and have found them to differ in their amino acid content both quantitatively and qualitatively. Thus A, the centrally located one in the pattern, has the amino acids reported by Hotchkiss (1941) and Gordon, Martin

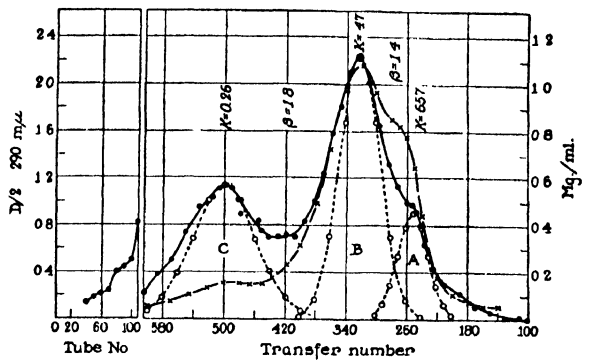


FIG. 3. Distribution pattern for a sample of crystalline tyrocidine Exp. wt., ●——●; Ext. 280 Mu., x——x; Calc'd, o——o.

and Synge (1943) while B, the smaller peak on the left of the pattern, contains these amino acids and phenylalanine in addition. C, the peak on the right, contains tyrosine instead of phenylalanine. Since gramicidin has been thought to be a single substance (Gordon-Martin-Synge, 1943; Synge and Tiselius, 1947), the need for extreme caution regarding this point is emphasized. It is logical to assume that the situation is much worse with proteins because of their greater complexity.

When ideal solutes, *i.e.*, those with constant partition ratios, are distributed, the experimental distribution will agree with a distribution calculated by the binomial expansion. If the partition ratio changes with concentration, a skewed curve is obtained. When the shift is too severe excessive tailing results as in chromatography under similar circumstances and a poorer separation usually results. With the gramicidins marked skewness did not occur at 100 transfers (Gregory and Craig, 1948) but was much more apparent in the pattern of Figure 2 which involved several hundred transfers. It is only logical to expect this difficulty (or advantage in some cases) when a high degree of separating power is reached since a correspondingly higher degree of conformance to ideality is also required.

Other systems studied have not shown as much specificity for the gramicidin mixture as the one used for Figure 2 has shown. For instance, a system containing heptane, cyclohexane, dioxane and water gave a pattern which showed separation of gramicidin C from the others but in which the A and B bands were completely merged. Unfortunately, there appears to be no rule for choosing the solvents or the proportion of solvents which will be "selective"

or specific for separating a given mixture. Trial runs must be made.

When crystalline tyrocidine hydrochloride (sample supplied by Dr. Dubos) was studied, it was found necessary to use an entirely different system in order to get the proper range of partition ratio. Moreover, tyrocidine is a basic peptide and if a neutral system is used, certain tubes will probably become strongly basic due to the stripping action of the countercurrent principle in removing chloride. Instability is likely to result. Accordingly, a system

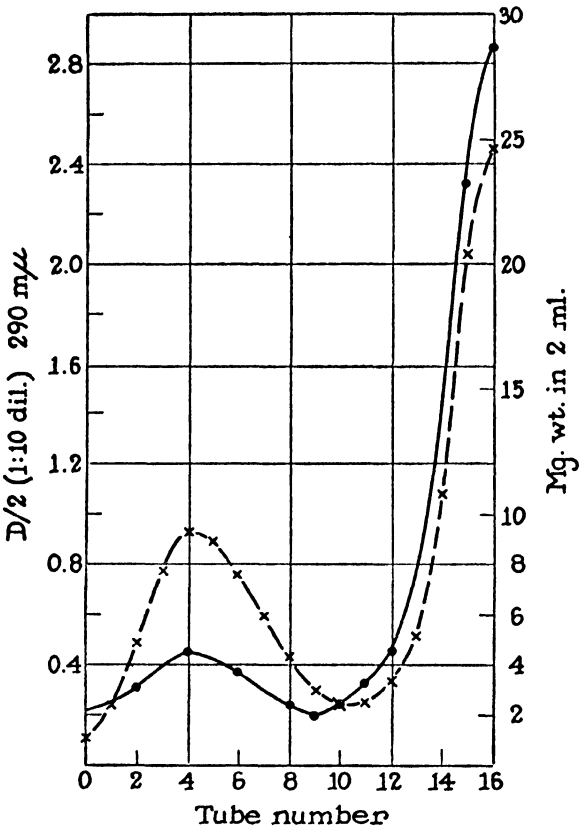


FIG. 4. Large scale 16 transfer distribution pattern of a sample of tyrothrycin. Wt., ●——●; Ext. x——x.

containing hydrochloric acid has been employed. A system containing methanol, chloroform and 0.1N hydrochloric acid (2:2:1 volume ratio) gave the pattern shown in Figure 3 with 2 gms. of material. In this pattern the concentration was determined by both weight and extinction at 290 Mμ.

A minimum of five components is indicated. The relative proportions of the three major components can be correctly estimated by the areas of the theoretical distributions A, B and C. All three of these components differ in their extinction coefficient at 290 Mμ. Since absorption at this wave length is mainly due to tryptophane it may be inferred that the peptides differ in their tryptophane content. This was confirmed by hydrolysis and paper chromatography. The tryptophane spot was very strong for C but was weak for A. Paper chromatography did not reveal qualitative differences and did not suggest quantitative differences for the other amino acid components which were those reported for tyrocidine (Gordon, Martin and Synge, 1943).

The various tyrocidine components were isolated in crystalline form. A complete description of their properties will be given elsewhere. Other systems for distributing tyrocidine have been investigated but they have not given further information in regard to the composition of the mixture.

A preliminary attempt has been made to use these methods for fractionating the parent mixture, tyrothrycin, from which gramicidin and tyrocidine were crystallized. A 16 stage distribution starting with 60 gm. of tyrothrycin gave the pattern shown in

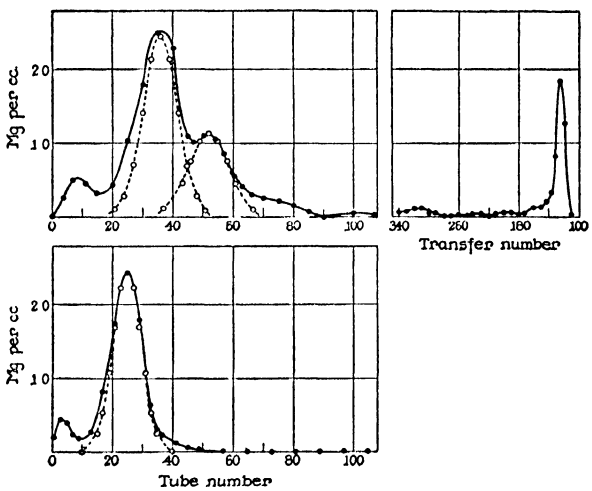


FIG. 5. Distribution patterns of samples of gramicidin-S. Wt., ●——●; Calc'd, o——o.

Figure 4. We are indebted to the Wallerstein Co. for generously supplying this sample and also for samples of gramicidin. The two groups gramicidin and tyrocidine are clearly separated. More illuminating, however, were 100 transfer distributions of material occurring in the tubes (11-13) of the minimum of Figure 4. These revealed the presence of at least three more polypeptides not present in either gramicidin or tyrocidine. Altogether at least eleven different polypeptides have been separated from tyrothrycin.

Following the work with tyrothrycin it appeared interesting to investigate gramicidin-S since this substance is a basic polypeptide of the tyrocidine class (Synge, 1945). It was thought that the latter might have components which would prove to be identical with some of those from tyrothrycin. This has not proven to be the case although a spectrum of

polypeptides did appear in the distribution with crude gramicidin-S. The sample of this material was supplied by Sharp and Dohme. The system used for fractionating tyrocidine gave the upper pattern shown in Figure 5. The component on the right did not prove to be of polypeptide character but at least four distinct polypeptides are shown in the remainder of the pattern. One of these agreed in properties with a sample purified by recrystallization and supplied by Professor Synge. A distribution of the latter is shown in the lower pattern. Upon hydrolysis and paper chromatography the different components all showed small differences in their amino acid content. The position of the different bands and therefore the partition ratio is much different from any of the components of tyrocidine.

The antibiotic, bacitracin, discovered by Johnson, Anker, and Meleney (1945) in a certain *B. subtilis* culture has been an interesting polypeptide type to study. It is most stable in solution at an acid pH. A representative distribution is shown in Figure 6. We are indebted to the Commercial Solvents Co. for liberal samples of this antibiotic. One of the most satisfactory systems found for distributing this polypeptide, and used in the above pattern, was made by equilibrating equal volumes of 2-butanol and 3 percent acetic acid.

Material recovered by concentration and freeze drying from the smaller peak on the right showed low antibiotic activity. It also had a lower nitrogen content than material recovered from the main peak although it had essentially the same amino acid spectrum on hydrolysis followed by paper chromatography. The material recovered from the main peak had the highest activity we have yet been able to reach. Material from the region just to the left of the main peak (tubes 80 to 100 and on into transfer number 370) had a lower activity but showed an additional amino acid on hydrolysis and paper chromatography. The small peak on the far left was also less active than the main peak. These polypeptides contain ten or eleven amino acids as will be shown later on.

It is not our intention in the present treatment to deal extensively either with technique or possible structure. Rather the results with these polypeptides have been chosen from the data we have accumulated in order to show probable limitations in separating, isolating and proving purity and perhaps certain possible means of minimizing the difficulties. The significant points of the work thus far appear to be the following:

1. Unquestionably a considerable amount of clear-cut resolution of mixtures of closely related polypeptides can be accomplished under the mildest of conditions by means of the distribution technique and a measure of confidence therefore can be cautiously placed in the method of countercurrent distribution as a criterion of purity. Separations have been accomplished with mixtures which have not thus far yielded to satisfactory resolution by other techniques.

2. The resolution can be made with sufficient sample so that a more or less prolonged structural study on purified material will be possible.

3. Certain information has been derived in regard to the deviations from ideality and approaches for minimizing these difficulties have been suggested.

The latter of these points merits further discussion since it would seem that the problem of deviation now becomes the main limiting factor in the separation of nearly all types of complex substances of higher molecular weight, provided they have sufficient stability to withstand the conditions of the

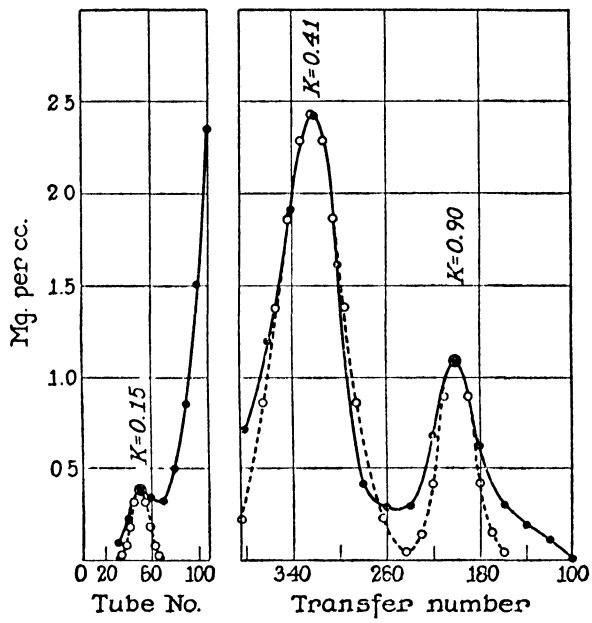


FIG. 6. Distribution pattern of a sample of bacitracin. Wt., •——•; Calc'd, o——o.

fractionation. Fractionation can be regarded as inherently a dilution process and there is no way to completely circumvent the steadily decreasing concentration level throughout a series. Of course, it is possible to choose a system in which there will be less deviation but here favorable separation factors are often sacrificed.

When, however, a given deviation must be faced it can be minimized to a considerable extent. If the sample is all placed in a single tube initially a rapid change in concentration results during the earlier transfers. Thus, at a K of 0.41 and an initial concentration of 85 mg./cc. in the upper phase, the maximum concentration emerging from 100 tubes would be of the order of 4 mg./cc. and the minimum concentration of interest would be of the order of 0.04 mg./cc. Now, if the sample were scattered initially in the first 15 tubes equally as was done in the run of Figure 6 (3 gm. sample) less shift in

concentration would result since the initial concentration is only 5.6 mg./cc. instead of 85 mg./cc. and yet the same total weight of sample is taken. A narrower range of concentration shift is thus involved and less skewing results. Scattering the sample initially in a series of tubes has been discussed elsewhere (Gregory and Craig, in press). It only produces a small loss in separating power.

Where overlapping due to skewing results, a shift to a different concentration level may succeed. Extraction offers sufficient range to make this possible. The basic cause of the shift in partition ratios with change of concentration is analogous to the cause of the formation of azeotropic mixtures in fractional distillation. In the latter case it is well known that choice of a different pressure may resolve the azeotrope.

The first point above also merits a little further discussion, perhaps qualification, particularly since the known classes of polypeptides, the ergot alkaloids, the penicillins and others as well as those studied here, appear to occur in nature as groups of closely related members rather than as individuals. Whether or not any apparently homogeneous band isolated from a distribution indeed represents a single chemical species cannot be determined with complete assurance at the present time. A small difference in structure might not cause a shift in partition ratio sufficient to be detected. We can only use the best criterion available as a working hypothesis until something better is demonstrated, realizing that the best may be inadequate.

Redistribution of a single band in a different system is often of considerable value. We have done this with a number of the bands in the previous charts. The results have not materially changed the conclusions thus far presented and accordingly will not now be treated further. With the polypeptides, redistribution has not always given a decisive result since one does not have a choice of widely divergent systems and redistribution in a less specific system loses part of its value. A tendency toward instability and also deviations from the ideal make comparison with a calculated distribution more difficult. Rather than spend time redistributing in the same solvent it would appear more profitable to have a sufficient number of units in the train so that the bands of interest are completely resolved without further work. It is anticipated that the apparatus now in construction in our laboratory will greatly help in this problem.

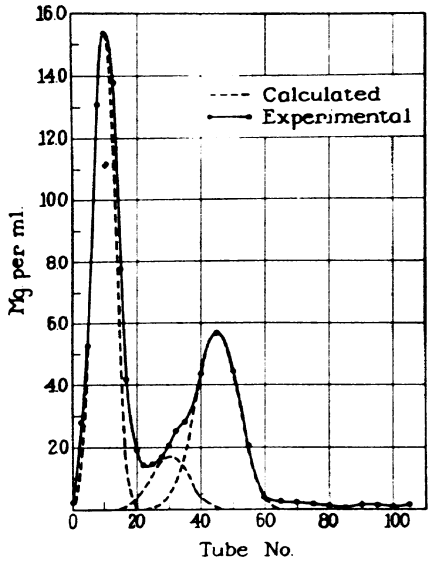
One approach to the problem of purity is linked to the problem of the study of structure and involves the quantitative determination of the fragments resulting on degradation. It would seem that the most promising quantitative approach at the present time with the polypeptides and where the fragments are known amino acids, is chromatography, particularly with the starch column or with some other adsorbent and with the technique of Moore and Stein (1948, 1949).

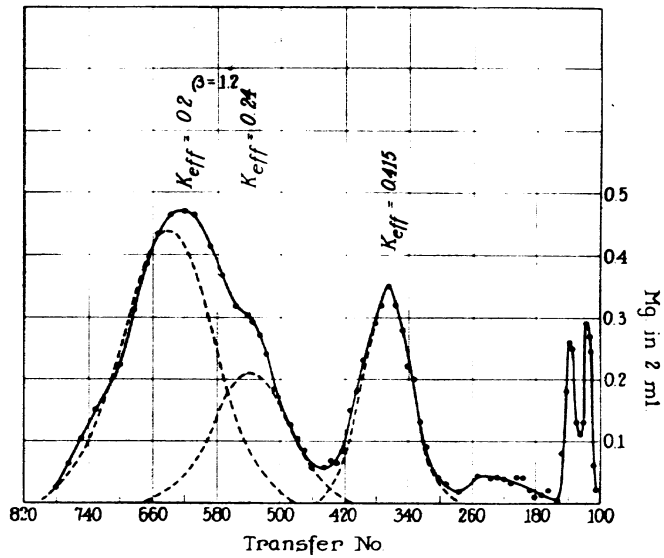
The great interest and the development of such a type of column was stimulated by the original ingenious idea of Martin and Synge (1941). They adsorbed one phase of a liquid-liquid two phase system on an adsorbent so that the adsorbed phase could be held stationary and thus present mechanical conditions for an ideal countercurrent liquid-liquid extraction process. A good correlation was obtained between the rate of travel of a number of bands of the aliphatic amino acids and their individuality determined liquid-liquid partition ratios. They therefore assumed the process to be a liquid-liquid extraction process with the solid adsorbent only providing mechanical support. If this view is correct, the process would differ from countercurrent distribution only in that the latter is strictly discontinuous.

Our distribution studies have served to throw doubt on the original view of Martin and Synge and have led us to almost the opposite view, namely, that liquid-liquid extraction plays little if any role in such separation processes. There appears to be only chance correlation in the aliphatic amino acids beween the rates of travel of individual bands in the distribution apparatus and on paper chromatographs or on starch. Studies on rates of interchange of solutes between the two phases in liquid-liquid systems also are against the original view as well as certain other theoretical considerations. In one respect this is unfortunate for us since information obtained quickly from paper chromatograms is of no help in selecting systems for distribution. On the other hand, since the processes are so basically different there is a greater chance that one method will succeed where the other fails.

Irrespective of the real basis for the particular type of chromatography called "Partition Chromatography," countercurrent distribution certainly, in its present stage of development, does not have the resolving power of certain types of chromatography for the amino acids. On the other hand, its resolving power does not diminish when attempts to fractionate larger amounts of material are made and, as mentioned earlier, resolving power with ideal solutes can be mainly a matter of labor. Where isolation is desirable, it would therefore appear logical to use chromatography as an analytical tool to control the separations made with the distribution technique. Paper chromatography has proven ideal for this purpose.

Thus far in distribution a single system has proven suitable for separation and isolation of not more than six to eight amino acids in a single run. The longer train now under construction will naturally improve this situation and for this reason we have reserved extensive study until the proper equipment becomes available.





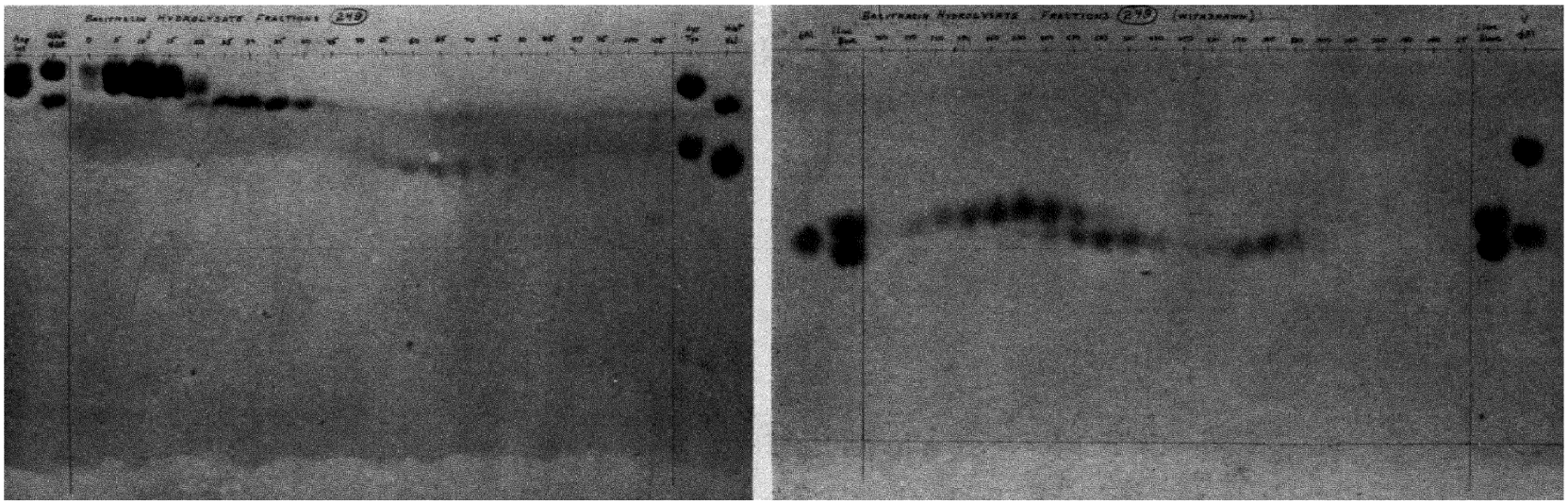


FIG. 7. Distribution pattern of a hydrolysate of bacitracin.
Wt., ●——●; Calc'd, - - - - -

In our present work with the hydrolysis products of bacitracin, ten spots on the paper chromatograph have been obtained. Eight of these have been isolated in crystalline form by countercurrent distribution. Seven of these are in analytically pure form and the eighth appears to be an unknown amino acid which can be traced to an impurity. One of the distributions is shown in Figure 7. Here 2 gm. of the crude hydrochlorides resulting on evaporation of the hydrolysate was used directly. The system was made by equilibrating 2-butanol with an aqueous solution containing 30 percent ammonium acetate and 7 percent excess ammonia. This system permitted the analysis to be made by weight.

Samples from appropriate tubes were spotted on a broad paper chromatogram so as to identify and give the position of the bands. The position of the paper spot was controlled by a known amino acid and the number taken from the fraction in turn related this spot to the band on the distribution pattern.

The first two bands did not show spots. Moreover, a previous run with somewhat milder conditions of hydrolysis gave much larger bands in these positions. In this case crystalline material was isolated from the first band and on further hydrolysis it yielded phenylalanine and isoleucine. Analytical data were all in good agreement for a dipeptide even though no spot appeared with ninhydrin.

The second peak on further hydrolysis showed a spot corresponding to ornithine in addition to ones for phenylalanine and isoleucine though the intact peptide showed no spot on paper. The next large peak proved to be optically inactive phenylalanine. Next appeared two overlapping peaks of leucine and isoleucine. Theoretical curves were constructed by locating the peaks on the basis of the center of the band of spots from the paper chromatograph. This permitted the intelligent selection of tubes for isolation. Leucine proved to be levorotatory but partially racemic. Isoleucine proved to be the partially racemic dextro form.

Weak spots were obtained showing a band at tubes 60 to 80. From these tubes some 20 to 25 mg. of crystalline material was isolated but its identity is unknown as yet. The band ending at tube 60 is ammonium chloride. A deviation occurs on the left side of this band and the paper spots indicate this region to contain histidine. This was easily isolated in analytically pure form as the hydrochloride by direct crystallization. It proved to be the levo isomer.

The high band on the left was a mixture which probably would separate with higher numbers of transfers. Other systems, however, are more efficient from the standpoint of labor.

By changing to other systems cystine, cysteine and glutamic and asparatic acids have been isolated in analytically pure form but ornithine (only indicated by a spot on the filter paper) has not as yet been fully resolved from the lysine.

It is thus apparent from our preliminary attempts to isolate the amino acids from the hydrolysis of a complex polypeptide that we probably will be able to isolate the majority of the amino acids in sufficient amount for final identification or even structural study if the amino acid should prove to be an unknown one. We have found it possible to set up systems which will permit the isolation of tryptophane, methionine, threonine, alpha-aminobutyric acid, proline alanine and glycine from synthetic mixtures even with our present equipment although a considerable amount of labor is involved. It would not appear worthwhile to present this work at the present time since it is only of preliminary character but to reserve its presentation until we have been able to apply a train several times the length of our present one and thus simplify the over-all process and discussion. As the matter now stands, too many different systems are involved.

One final point on the hydrolysis of bacitracin appears worth mentioning. Our first attempt at hydrolysis resulted in the isolation of peptides which would have been missed had we not used weight analysis in obtaining the distribution curves. The peptides appear to contain d-amino acids and are difficult to hydrolyze. Christensen (1943) found this to be true with gramicidin and isolated valylvaline, a result which we also verified with the above technique. The valylvaline isolated gave a very poor color yield with ninhydrin.

In considering partial hydrolysates it has been anticipated that the peptides in a given system would tend to have partition ratios intermediate between those of the free amino acids contained in the peptide. This has not proven to be the case in our work thus far, particularly when the system contains a considerable concentration of salt. If the salting out effect continues to be true, the future of countercurrent distribution in the determination of the structure of the more complex polypeptides is much more promising than it appeared when the work was begun. A study is therefore under way to investigate the effect of salts on the selectivity of certain systems. For example, in the system 2-butanol/30 percent ammonium acetate-7 percent ammonia, the partition ratios of tryptophane and valylvaline were 0.75 and 0.93 respectively. They thus appeared as an overlapping band from the hydrolysis of gramicidin with the peptide slightly in advance. When this overlapping band was redistributed in tertiary amyl alcohol/water the partition ratios were 0.58 and 0.18 respectively with the peptide now trailing. This permitted complete separation.

As has been mentioned previously (Consden, Gordon and Martin, 1947) it is unlikely that any single separating tool will prove adequate for the

problem of partial hydrolysates of proteins and higher polypeptides. Countercurrent distribution would appear to be an additional tool to supplement those already available.

Finally, we would like to point out that our success with the polypeptides thus far studied does not necessarily mean that we can fractionate and purify intact proteins. We have hopes for some of the more stable ones which have appreciable alcohol solubility and expect to take up this problem in the near future.

References

BARRY, G. T., SATO, Y., and CRAIG, L. C., 1948, Distribution studies. X. Attainment of equilibrium. J. biol. Chem. *174:*209-215.

CHRISTENSEN, H. N., 1943, The isolation of valylvaline from gramicidin hydrolysates. J. biol. Chem. *151:*319-324.

CONSDEN, R., GORDON, A. H., and MARTIN, A. J. P., 1947, The identification of lower peptides in complex mixtures. Biochem. J. *41:* 590-596.

CRAIG, L. C., and POST, O., 1949, Apparatus for countercurrent distribution. Anal. Chem. *21:* 500-504.

GORDON, A. H., MARTIN, A. J. P., and SYNGE, R. L. M., 1943, Compositon of gramicidin. Biochem. J. *37:* 86-92.

1943, The amino acid composition of tyrocidine. Biochem. J. *37:* 313-318.

GREGORY, J. D., and CRAIG, L. C., 1948, Countercurrent distribution of gramicidin. J. biol. Chem. *172:* 839-840.

1949, The analytical specificity of countercurrent distribution. Ann. N. Y. Acad. Sci., in press.

HOTCHKISS, R. D., 1941, The chemical nature of gramicidin and tyrocidine. J. biol. Chem. *141:* 171-185.

JOHNSON, B. R., ANKER, H., and MELENEY, F. L., 1945, Bacitracin: A new antibiotic produced by a member of the *B. subtilis* group. Science *102:* 376-377.

MARTIN, A. J. P., and SYNGE, R. L. M., 1941, A new form of chromatogram employing two liquid phases. Biochem. J. *35:* 1358-1368.

MOORE, S., and STEIN, W. H., 1948, Partition chromatography of amino acids on starch. Ann. N.Y. Acad. Sci. xlix: 265-278.

1949, Chromatography of amino acids on starch columns. Solvent mixtures for the fractionation of protein hydrolysates. J. biol. Chem. *178:* 53-77.

STEIN, W. H., and MOORE, S., 1948, Chromatography of amino acids on starch columns. Separation of phenylalanine, leucine, isoleucine, methionine, tyrosine and valine. J. biol. Chem. *176:* 337-364.

SYNGE, R. L. M , 1945, Gramicidin-S: Over-all chemical characteristics and amino-acid composition. Biochem. J. *39:* 363-367.

SYNGE, R. L. M., and TISELIUS, A., 1947, Adsorption experiments with gramicidin and related substances. Acta. chem. Scand. *1:* 749-762.

Discussion

HOTCHKISS: Anyone who has attempted to fractionate by ordinary methods some of the natural products which Dr. Craig and his associates have successfully separated is in the best position to appreciate the power of the distribution methods. We shall probably never be able to know, definitely, how pure our preparations of gramicidin and tyrocidine of nine years ago were, since the culture and mode of processing the crude material have inevitably been undergoing modification during the intervening time. However, it is clear that our earlier criteria of purity were inadequate to assure chemical individuality of these peptides. And it is doubtful whether we should have been able to recognize it at that time, if we had achieved partial separation of any component gramicidins which we may have had. It may, therefore, be important to stress that countercurrent distribution has not only provided a technique for separating such closely related substances, but has also furnished a series of excellent additional criteria of purity not previously available, namely the distribution coefficients.

In connection with the comments of other discussants, I would like to say that the question whether there exist families of closely related peptides or proteins is not conclusively illuminated by the case of the gramicidins.

One possibility suggested by Dr. Sanger, that the different members arise from genetically different strains, does not seem probable. Usual precautions have been taken at various times to use single-colony or "pure" strains of the parent organism, although it is not possible to say what is the state of the strains presently being used. Furthermore, a recent report describes thirty-four new penicillins produced by the same mold culture when different organic acid derivatives were provided as precursors. Malogen, nitro, ether, thioether, heterocyclic groups, etc. were introduced in this manner. It may be concluded from this example alone that biological synthetic systems do not always show absolute specificities, since they respond in some instances to a change in the available precursors. It seems equally possible that contingencies in the amino acid metabolism at different stages of growth of a bacterial culture could give rise to such differences as for example those apparently small ones that exist between A, B, and C in the gramicidin "family."

In connection with another possibility raised by Drs. Stein, Woolley, Smith, that chemical alterations in an original gramicidin molecule are produced biochemically during growth or chemically during isolation, we can only point out that this might require the opening and re-closing of a cyclopeptide ring. The usual preparation of gramicidin involves an autolysis, weak acidification, solvent extraction and recrystallization from hot organic solvents. Anhydridization of a dipeptide to form a six-atom ring under such conditions might occur, but a fifty- to one hundred-atom ring would not be more likely to form than any equivalent condensation between open chain molecules. Since in separate experiments gramicidin can be prepared with avoidance of heating, of acid, or of any one of the solvents, there is no sign that the general gramicidin

structure is a chemical artifact. Nevertheless, some chemical or physical association allows it to be more water-soluble in the crude culture medium than it is found to be after extraction.

Dr. Fruton has suggested that the antibacterial activity of some of these peptides seem unexplained by their structures as outlined so far, mentioning that attempts to synthesize peptides related to gramicidin-S have yielded no active materials. To this I should like to remark that it can be stated with some assurance that gramicidin-S and the tyrocidines (basic peptides), but not the true gramicidins A, B, and C (neutral peptides), have a cytolytic action upon bacteria and other cells, releasing soluble metabolites from the cell at just those concentrations in which they are bactericidal. This action they share with surface active agents of diverse structure such as lauryl esters of amino acids, or basic detergents containing no amino acid, such as fatty amines, high alkyl quarternary ammonium soaps, and even the anionic detergents such as the high alkyl sulfates and phenols. On the other hand some thirty non-surface active antiseptics tested did not have this action. The tyrocidines and gramicidin-S might appear then as examples of biological agents of which intimate knowledge of the chemical structure is of no particular use in understanding or duplicating the biological activity.

STUDIES ON THE CYTOCHEMISTRY OF PROTEINS

J. F. DANIELLI

The Chester Beatty Research Institute, Royal Cancer Hospital, London, England

I. Introduction

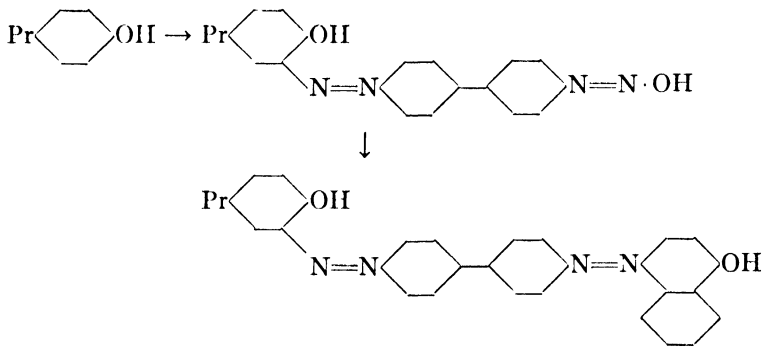
Cytochemistry can, in principle, provide information concerning a) the distributions of proteins in cells, b) the sites of synthesis of proteins, c) the sites of degradation of proteins and d) the physiological significance or action of proteins. To provide this information a great variety of methods is necessary. At present there are few methods available, and of these few, most are either known to be inadequate or not yet proved to be sound. Among the reasons for this unsatisfactory state of affairs is the fact that many of the methods are those of the biochemist, intended originally for use on the test-tube scale, applied virtually unchanged to the cytological level. In this paper several new methods are described which have been designed specifically for cytochemical purposes. Among the points to which particular attention was paid in the designing of these methods were 1) the need for avoiding reactions which involve diffusible reaction products, and 2) the need for having several independent methods which should give the same result. It must be emphasized that this is a preliminary report, and very few studies have been made by the methods described here. The necessity for new methods is at present paramount. The most profitable procedure for the next few years will be to investigate the procedures for making these new methods quantitative, to ascertain the limitations of these procedures, and only after that to change the emphasis to their detailed exploitation.

When attempting to design new methods in this field there are at least three different approaches. The first of these is to identify a protein by its activity: for example, an enzyme may be localised if it will catalyse a reaction, the products of which are insoluble. Secondly, it would be of value if the total number of —CO·NH— groups could be estimated in a particular part of a cell. It is possible that the biuret reaction could be modified to provide this information, or that observations of absorption in the infrared region will prove adequate. The third method is to devise reactions specific for the reactive groups in protein side-chains and end-groupings: this last approach is the one which was used in the present studies.

It is difficult to obtain reagents which will react exclusively with one group in a protein. It has therefore been necessary to use reagents of two types, 1) non-chromogenic blocking reagents, and 2) chromogenic reagents. The procedure is then to choose a reagent which will be chromogenic for the group which is to be studied. This reagent will usually also combine with other groups. Consequently use of the chromogenic reagent must be preceded by use of non-chromogenic reagents which block the groups which are competent to react with the chromogenic reagent but which it is not desired to study.

II. Use of Diazonium Hydroxides

The main tissue components which will react with diazonium hydroxides are histidine, tryptophane and tyrosine. So far the most valuable reagents of this type have been tetrazotised benzidine and dianisidine. (See Plate I, Fig. 1.) The procedure is to allow the tissue sections to react with a considerable excess of the diazonium hydroxide. The colour so developed is often sufficiently intense, but where this is not so, intensification may be obtained by first washing the section free of excess diazonium hydroxide, and then exposing it to a solution of a phenol or aromatic amine, when further coupling occurs (Danielli, 1947). For example,



In the above and subsequent structural formulae, Pr represents the protein molecule.

Among the blocking agents which may be used

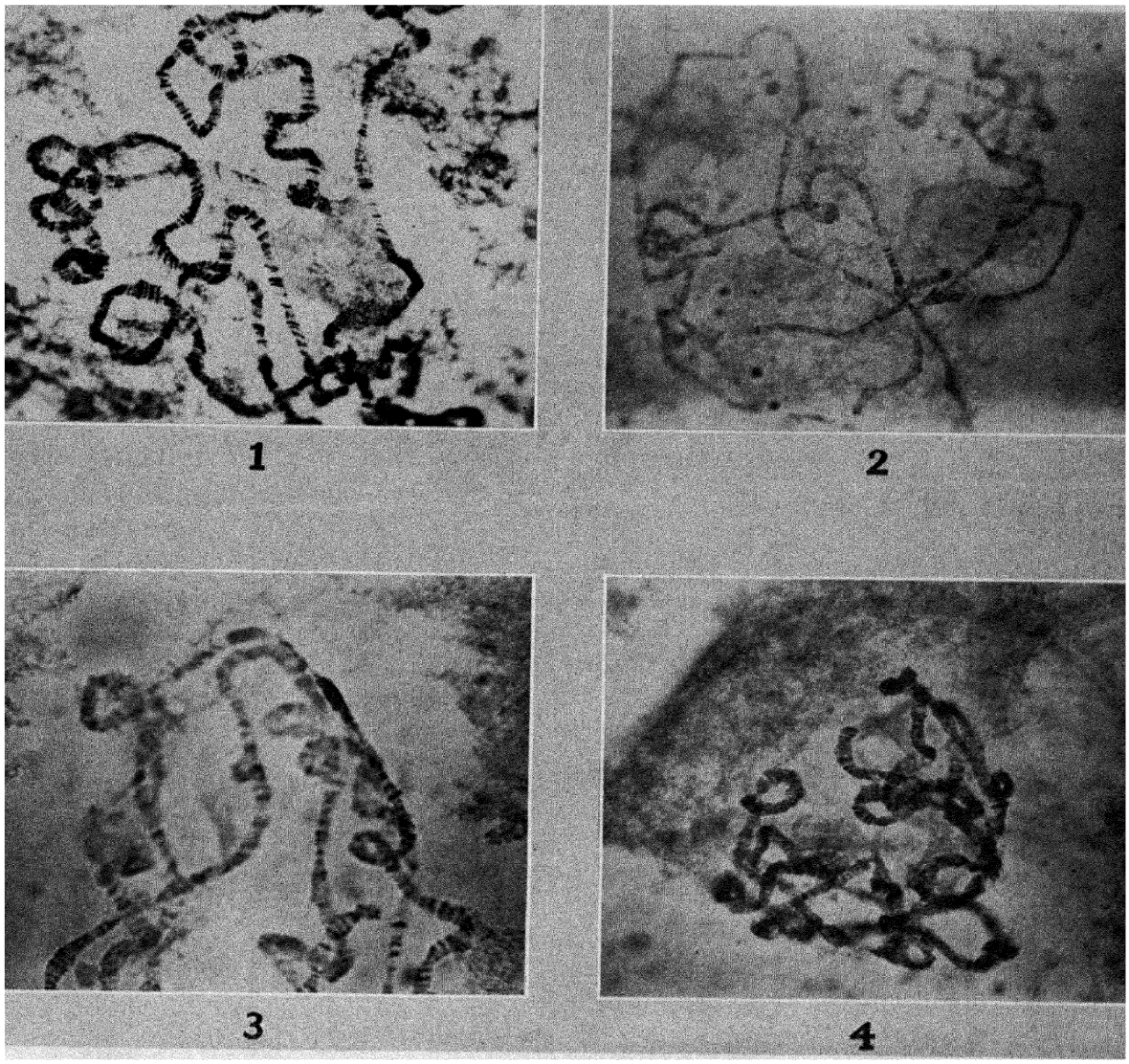


PLATE I

FIG. 1. Drosophila chromosomes: fixative 45 percent acetic. Sum of histidine, tryptophane and tyrosine, etc., shown by treatment with tetrazodised benzidine followed by β naphthol.

FIG. 2. As Fig. 1, but benzoylated in 10 percent benzoyl chloride, 90 percent pyridine for 12 hours to destroy histidine, tryptophane and tyrosine. In the photographing of Figs. 1 and 2, exactly the same illumination and exposure were used. But to obtain an effective print it was necessary to overdevelop Fig. 2, *i.e.* the diminution in absorption on benzoylation is even greater than indicated by the plates.

FIG. 3. Drosophila chromosomes: fixative 45 percent acetic. Sum of NH_2 and SH shown by treatment with dinitrofluorobenzene.

FIG. 4. As Fig. 3. Treated with nitrous acid, followed by p. nitrophenyl isocyanate, then reduced, diazotised, and linked to H acid. Tyrosine and other phenols demonstrated.

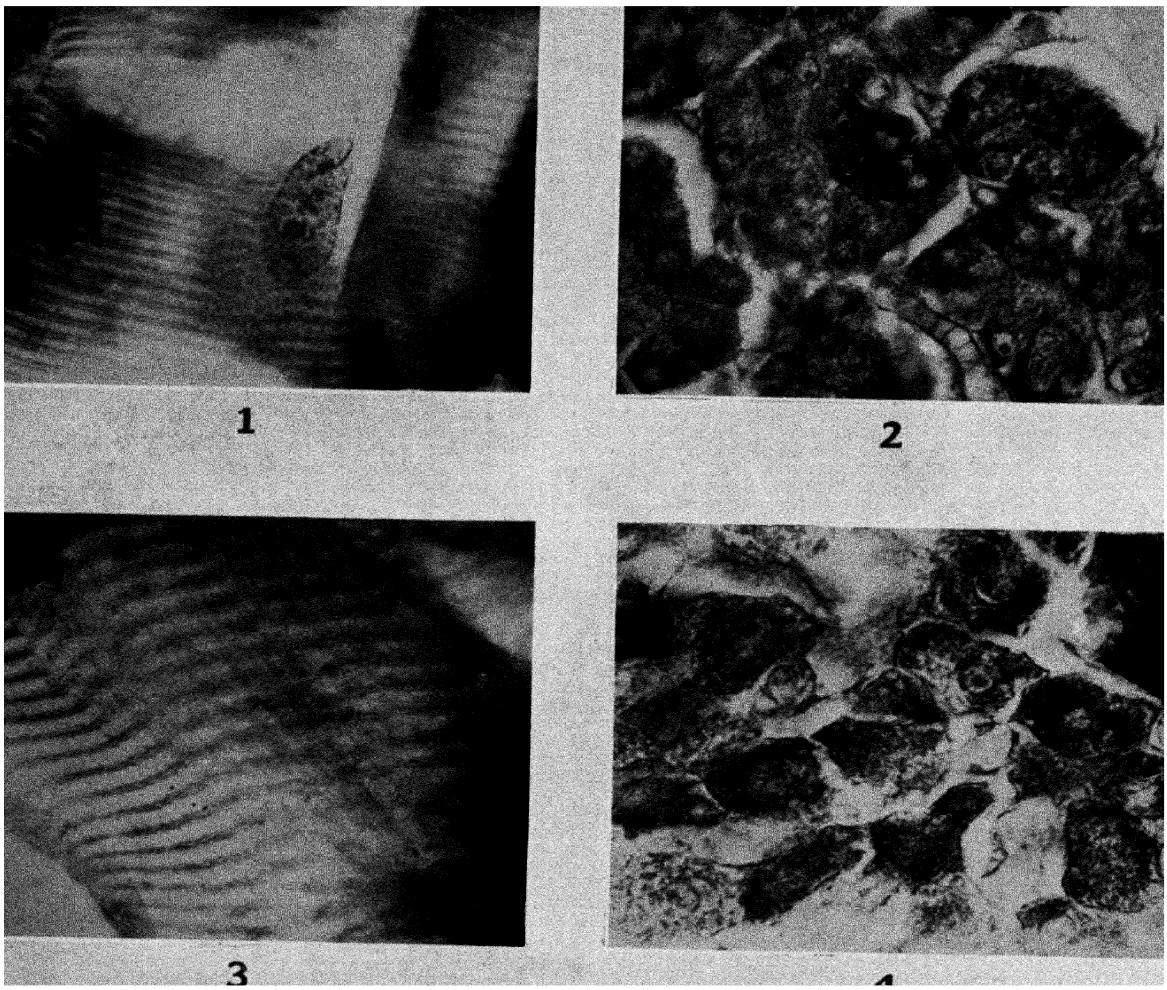


PLATE II

FIG. 1. Squash of newt larva tail. Treated with dinitrofluorobenzene, then reduced, diazotised and linked to H acid.

FIG. 2. Rat pancreas. Treated with dinitrofluorobenzene, then reduced, diazotised and linked to H acid.

FIG. 3. Squash of newt larva tail. Treated with p. nitrophenyl isocyanate, then reduced, diazotised and linked to H acid.

FIG. 4. Rat liver, showing sites of combination with $NH_2 <\!\!=\!\!> N(CH_2CH_2Cl)_2$, diazotised and linked to H acid.

before reaction with a diazonium hydroxide are 2:4 dinitrofluorobenzene for tyrosine and other phenols, performic acid for tryptophane, and benzoyl chloride which destroys the reactivity of histidine, tryptophane and tyrosine (Table 1). (See Plate I, Fig. 2.) There are many other blocking agents which can be used. Some of these will be referred to later.

TABLE 1. THE ABILITY OF CERTAIN AMINO ACID RESIDUES TO REACT WITH DIAZONIUM HYDROXIDES AFTER TREATMENT WITH VARIOUS REAGENTS

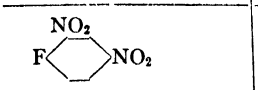
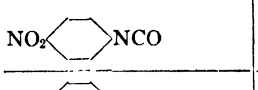
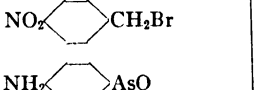
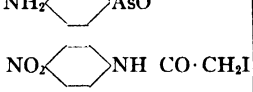
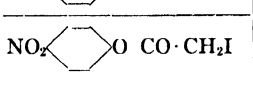
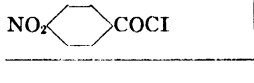
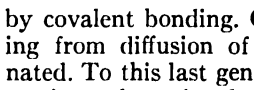
Reagent	None	2:4 dinitro fluorobenzene	Performic acid	Benzoyl chloride
histidine	+	+	+	−
tryptophane	+	+	−	−
tyrosine	+	−	+	−

III. USE OF NITRO-COMPOUNDS

There are many reagents for organic compounds which contain nitro groups. Some of these compounds yield coloured products after reaction. Thus according to Sanger, whilst tyrosine combines with 2:4 dinitrofluorobenzene to give a colourless product, amino and SH groups give yellow compounds. (Plate I, Fig. 3.) Consequently, after blocking either SH or NH_2 with another reagent, the cytochemical distribution of the remaining group may be determined. However, the colour intensity is usually too low to be satisfactory.

Very great increases in colour intensity may be obtained by the following procedure (Plate I, Fig. 4). After treating a section with a reactive nitro compound, the excess of the reagent is washed away, after which the nitro group is reduced to NH_2 with a reducing agent such as stannous chloride. The section is again thoroughly washed, after which the aromatic NH_2 group is diazotised by exposure to a nitrous acid bath at 0° C. After washing away the excess of nitrous acid, the section is placed in a bath containing an aromatic amine or a phenol, when coupling occurs and an azo dye is formed at the site of the group which is being studied. For example,

TABLE 2. SOME NITRO REAGENTS AND GROUPS FOR THE LOCALISATION OF WHICH THESE REAGENTS ARE PARTICULARLY USEFUL

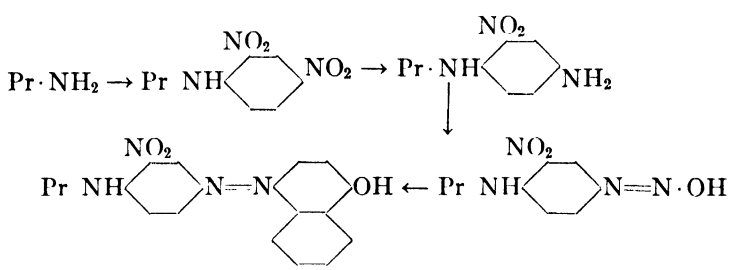
	Tyrosine	SH	NH_2	CO_2H	CHOH
	+	+	+		
	+	+	+		
				+	
		+			
		+	+		
		+			
					+

by covalent bonding. Consequently artefacts resulting from diffusion of reaction products are eliminated. To this last general statement there is one exception—there is always the possibility that the protein molecules themselves may diffuse in the sections during the procedure. Methods will be described later for investigating this point.

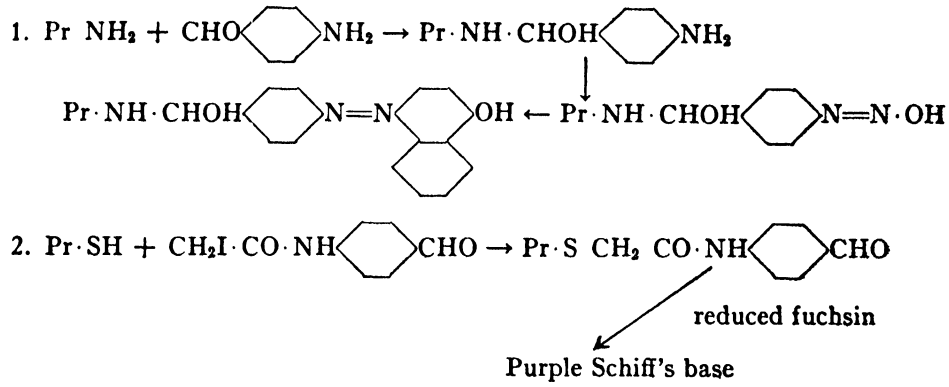
In Table 2 is given a list of some of the more readily available nitro reagents, with an indication of those protein groups for the localisation of which a given reagent is likely to be particularly suitable.

IV. ALDEHYDE REAGENTS

A third family of reagents which holds much promise includes those containing an aldehyde group and another reactive group. In some cases the aldehyde group will react with a group in a protein, such as NH_2, and the second reactive group of the aldehyde, say NH_2 or NO_2, may be subjected to the procedures described in the previous section. In other cases the second reactive group is allowed to react with a group in the specimen; then after appropriate washing the specimen is exposed to reduced fuchsin, which reacts with the aldehyde to give a purple colour. Examples of the reactions are:



As is the case with the diazonium hydroxide method discussed in the previous section, the dye is linked to the group which is under investigation



A number of reagents of this type are now being investigated. Table 3 shows some reagents which are being studied in connection with the cytochemistry of $-NH_2$ and $-SH$ groups.

V. Blocking Agents

An essential feature of the type of method which has been described in this paper is the use of blocking agents, which are used prior to use of the chromogenic reagent, to eliminate some of the types of group with which the chromogenic reagent is competent to react. Table 4 contains a list of some of these. Of the reagents listed a number combine with many types of group, but of these some may be used to block a narrower range of groups by careful control of pH and duration of exposure to the blocking agent. Phenyl isocyanate and the diazonium hydroxides are particularly important cases of this type. Certain of the blocking agents, notably Hg, $ClCH_2 \cdot O \cdot CH_3$, $Ph \cdot S \cdot COCl$, $ClCH_2 \cdot O \cdot CH\text{-}CH_2$, $Ph \cdot AsO$, $Ph \cdot HgOH$, can be removed from the groups they protect by very mild treatment with such reagents as dilute acid, dilute alkali, dithioglycerol and heavy metals. Thus a group such as NH_2 may be protected by treatment with $Ph \cdot S \cdot COCl$ against the attack of a reagent used for *e.g.* tyrosine: then when the colour for tyrosine has been developed, treatment with a heavy metal will liberate the NH_2 groups, so that colours for both NH_2 and tyrosine can be developed in the same specimen.

A few specimen procedures will indicate more clearly the way in which these reagents may be used.

Table 3. Aldehyde Reagents under Investigation for the Cytochemistry of -SH and $-NH_2$ Groups

	NH_2	SH
CHO C₆H₄ CHO	+	
CHO C₆H₄ NH_2	+	
CHO C₆H₄ NO_2	+	
CHO C₆H₄ $CO \cdot CH_2Br$		+
CHO C₆H₄ $O \cdot CO \cdot CH_2I$		+
CHO C₆H₄ $NH \cdot CO \cdot CH_2I$		+

VI. Procedures for SH

1. Treat with Hg, PhAsO or PhHgOH to block SH.
2. Treat with phenyl isocyanate to block NH_2 and phenolic OH.
3. Treat with performic acid to destroy S-S.
4. Treat with H_2S or dithioglycerol to release SH.
5. React with dinitrofluorobenzene or phenyl isocyanate, and then proceed as indicated in section III, obtain a colour for SH.

VII. Procedure for SS

1. Treat with phenyl isocyanate to block SH, NH_2 and phenolic OH.
2. Reduce S-S, with HCN for example, to SH.
3. React with dinitrofluorobenzene or phenyl isocyanate, and then proceed as indicated in section III to get a colour at the site S-S.

VIII. Procedure for NH_2

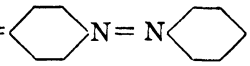
1. Treat with hydrogen peroxide or iodoacetamide to eliminate SH.
2. React with dinitrofluorobenzene to get yellow colour with NH_2, or react with p. nitrophenyl isocyanate and follow the procedure of section III.

IX. Procedure for Tyrosine

1. Treat with ketene, or naphthoquinone, or nitrous acid followed by iodoacetamide, to eliminate SH and NH_2.
2. Treat with dinitrofluorobenzene or p. nitrophenyl isocyanate followed by the procedure of section III to obtain a colour for tyrosine.

X. Nitrogen Mustards

An interesting special case is provided by some of

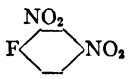
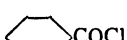
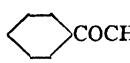
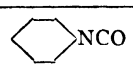
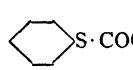
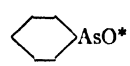
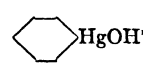
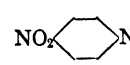
the compounds of the nitrogen mustard series. Those studied have been of the type $R \cdot N(CH_2CH_2Cl)_2$. Where $R = NH_2$  the compound can be localized by diazotisation and coupling to a phenol, as in section III. Where $R =$  it is necessary first to reduce the azo linkage to NH_2 with, for example, stannous chloride, and then proceed as before.

XI. Intensification of Colour

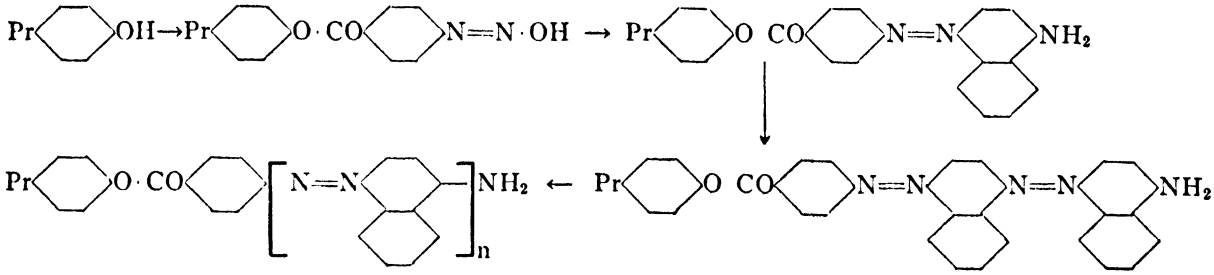
The colour intensity obtained by the methods outlined above is often quite sufficient. But occasionally a more sensitive method is necessary. So far three methods have been studied for obtaining greater sensitivity.

1. In many of the methods given above a diazonium hydroxide has been formed in the tissue section which has been linked to a phenol or aromatic amine to give an azo dye in the section. If the colour intensity so obtained is not sufficiently great, a considerable intensification may be obtained by making the first coupling with an aromatic amine which can itself be diazotised after coupling. Then a second coupling can be carried out with the same, or another, amine. For example,

Table 4. Reagents which May Be Used for Blocking Various Groups in Protein Molecules
Reagents which may readily be removed are marked with an asterisk.

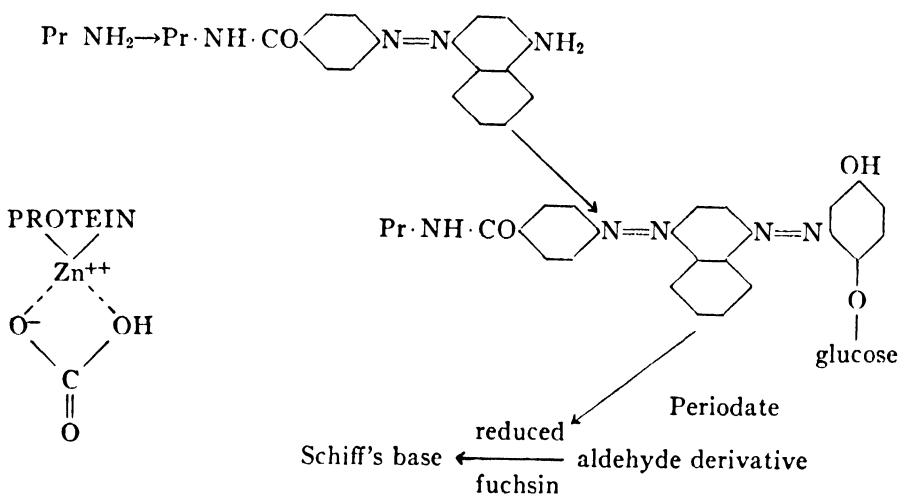
	Tryptophane	Tyrosine	Histidine	SH	S–S	NH_2	CO_2H	CHOH
		+		+		+		
HNO_2						+		
$H\,CO \cdot OOH$	+				+			
	+	+	+	+		+		+
Hg*				+				
$I\,CH_2CONH_2$				+				
				+				
glyoxal						+		
ketene		(+)		+		+		
		+		+		+		
$ClCH_2 \cdot O \cdot CH_3$*							+	
						+		
naphthoquinone				+		+		
methyl naphthoquinone				+				
$ClCH_2 \cdot O \cdot CH{=}CH_2$*						+		
				+				
				+				
	+	+	+					

It is only possible to localise nitrogen mustard molecules which have combined with a tissue component, such as protein or nucleic acid, which can be rendered indiffusible by fixing agents.



As indicated above, this process may be repeated a number of times. The intensity of colour produced is not directly proportional to the number of couplings, and the increment diminishes as n increases. In our experiments we have not found it profitable to couple more than eight times.

2. An alternative procedure is to proceed as in 1 to a given point, and then to couple onto a sugar or related compound which has been linked previously with a phenol in such a way as to have a phenol residue which will react with a diazonium hydroxide. When the coupling is complete, the sugar is oxidised with periodic acid to obtain aldehyde groups where formerly $\alpha\beta$ glycol groups existed. Then the section is treated with reduced fuchsin, which gives a colour with the aldehyde groups. For example,

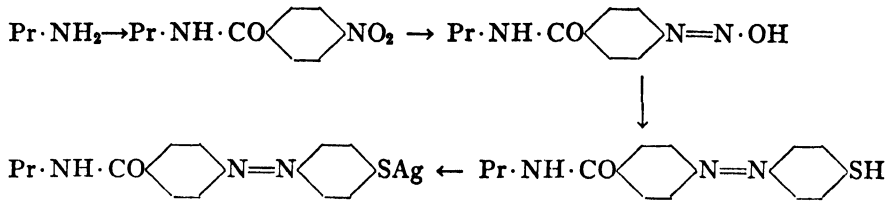


In principle it should be possible to use a polysaccharide in place of a monosaccharide. Then the periodate oxidation would provide a chain of CHO groups capable of reaction with reduced fuchsin. However, this is bound to be complicated by steric factors, and we do not yet know how far this technique can be carried.

3. Another possibility is to link on a phenol or amine which will give a fluorescent compound. Then if the specimen is studied in the fluorescent microscope by photographic means a vast increase in sensitivity is possible in theory. But in practice there are considerable limitations, due on the one hand to imperfections in the optical equipment available, and on the other hand to the fact that biological material is usually self-fluorescent. The use of selective quenching agents may be of value here.

X. Modifications for Electron Microscopy

It is clear that the value of the techniques described above would be greatly extended if they could be modified for use with the electron microscope. This involves the introduction of groups which have a high power for scattering electrons. In some instances a useful increase in electron-absorption may be introduced by the techniques described above, without modification. But the introduction of a heavy metal is likely to prove of far more general applicability. In many of the methods described above this may readily be achieved by substitution of a thiophenol for a phenol as the component to be linked to a diazonium hydroxide. After linking on the thiophenol, exposure to a dilute solution of a heavy metal salt will result in the capture of a metal ion by each thiol group. For example,



If a dithiol is used more secure linking can be obtained with a divalent heavy metal, such as Cu, Pb or Hg. Preliminary experiments have shown that greatly increased intensity of electron absorption may be achieved by methods of this type. At present we are using compounds of the thiophenol type, but it may prove desirable to use phenols with thiol groupings in a side-chain.

Another approach which may be of value is to replace the reduced fuchsin used in aldehyde cytochemistry by a base containing SH groups. It may prove necessary to have the thiol groups in such a compound masked until reaction with the aldehyde group has occurred.

XI. Tests for Artefacts Produced by Diffusion

Freeze-drying is in general the best fixation method for use in cytochemistry, and errors due to diffusion during fixation cannot then be serious. The reactions which have been suggested here all involve the formation of stable covalent bondings, so that no errors can arise from the diffusion of reaction products. Thus the combination of freeze-drying and the methods suggested here eliminates most of the common sources of error in cytochemistry due to diffusion. There remains, however, the possibility that some proteins may remain diffusible despite fixation. There are three methods which can be used to detect artefacts of this origin.

1. This method involves the use of superimposed sections, and is best understood from an example. Suppose that the distribution of NH_2 groups is being studied. One section is placed on a slide, and all the NH_2 groups in it are blocked. The section to be studied is then placed so as to overlap the blocked section, and the cytochemical reaction for NH_2 groups carried with the sections in this position. Then if diffusion of a protein component occurs in such a manner as to produce a displacement of colour due to NH_2, colour will appear in the underlying blocked section.

2. Where protein components can be obtained in a soluble condition, the cytochemical reaction may be carried out with the soluble components, and a tissue section exposed to the coloured soluble components to see if adsorption upon the section will occur.

3. To test whether protein components are diffusing out of a section, the cytochemical reaction may be performed upon say 0.2 gm. of suspended sections, and the fluid phase from each individual step of the reaction tested for protein.

XII. Problems in the Development of Quantitative Methods

Up to now no quantitative studies have been made with the methods described in this paper. There are still a number of difficulties in the interpretation of light absorption by biological materials. For the surmounting of these difficulties much assistance can be derived by the use of the reflecting microscope, and it has seemed wise to postpone study of these problems until reflecting microscopes become available. The main problems at present unsolved, on the theoretical side, are:

1. The magnitude of the scattering correction. Any apparent extinction coefficient needs to be corrected for the loss of light by scattering. This loss is a fraction of the size of the scattering particles, which in systems of the type encountered in cells may range from say 2 mμ to several μ. It is often not easy to decide what the effective particle size is. But the correction becomes increasingly small on moving from the ultraviolet towards the infrared.

2. Commoner (1949; Commoner and Lipkin, 1949) has recently shown that in systems which are optically anisotropic it may be impossible to derive reliable information from extinction coefficients unless the degree of anisotropy can be measured. Optical anisotropy, due for example to orientation of protein molecules, may be either intrinsic in the specimen, or arise as an artefact. Errors of the order of hundreds of percent may arise through neglect of this factor.

3. The dispersion effect. Normally, if one attempts to determine the extinction coefficients of a region of a cell, an area is chosen in which it is supposed that the absorbing material is distributed homogeneously. But the actual distribution of absorbing material in the area studied may profoundly affect the observed extraction coefficient. Consider the following diagram (Fig. 1). The square as a whole is the area whose extinction coefficient is being studied, the hatched areas represent regions rich in absorbing component, and the clear areas regions lacking in absorbing component.

Let us suppose that if a large area had the same composition as the hatched area, practically no light would be transmitted, and if the composition were the same as the clear areas no light would be absorbed. Then the actual absorption of light by an area such as that in Figure 1 depends upon the relationship between the distance d and the wavelength γ at the light used. If $d \gg \gamma$, then the area will transmit 50 percent of the incident light. If

$\gamma \gg d$, then the area will transmit practically no light. These two extremes are well understood. But if d is of the order of $\gamma/4$, it is impossible to predict what proportion of the light will be absorbed. Errors of the order of hundreds of percent in calculated concentration of absorbing component may also arise from this factor.

It seems likely that an experimental study will be the most satisfactory way of investigating these difficulties. And for this study it is desirable to be

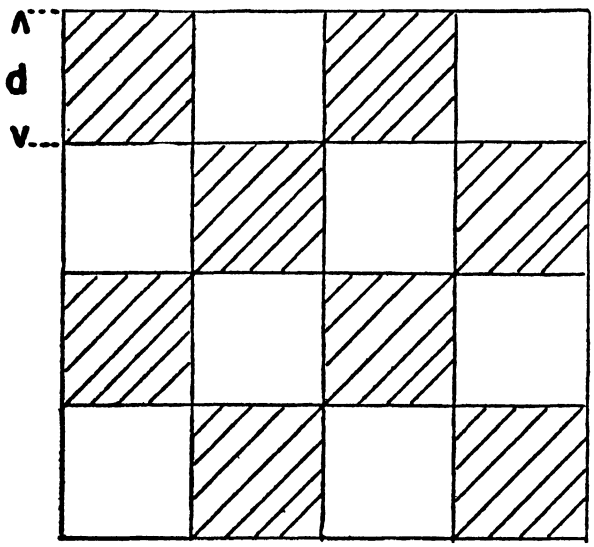


FIG. 1.

able to move freely over the range from the infrared to the ultraviolet.

XIII. Nucleic Acid and Protein Synthesis

In conclusion it seems worth while to consider the possible relationship between nucleic acid and protein synthesis. The views of Caspersson (1947) and of Brachet (1945) on the association of nucleic acid and protein synthesis are well known. It is frequently suggested that the association is due to a template relationship between the two substances. But whilst superficially attractive, such a theory presents great difficulties from a physico-chemical standpoint.

One consequence of adding nucleic acid to a system containing protein is that the protein, by conversion to nucleate, is in effect removed from the system. This suggests that if there is an equilibrium between a protein and its precursors, the presence of nucleic acid in the system would shift the balance in favour of synthesis of protein. We are thus brought to consider schemes such as that illustrated in Figure 2. A gene is the site of protein synthesis and degradation, which is shifted in favour of synthesis by the trapping action of pentose nucleic acid. This equilibrium is itself controlled by the concentration of histone, which will be a better competitor for nucleic acid than will most less basic proteins. Thus increasing pentose nucleic acid in a system should promote protein synthesis, and increasing histone should produce the opposite effect. Thus the action of pentose nucleic acid would be that of a trapping agent, not of a template. Such a scheme leaves ample scope for specificities for nucleic acids. It is also conceivable that a further equilibrium of the same type arises between the proteins and the desoxy sugar nucleic acids.

The work on the synthesis of protein in relation

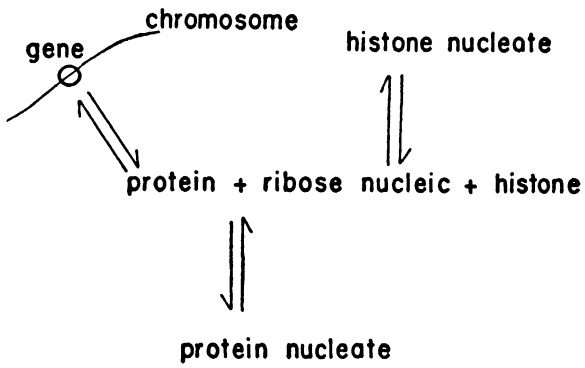


FIG. 2.

to nucleic acid is all, so far as I am aware, compatible with the scheme just discussed. So also is the observation of Stedman and Stedman (1943) that the concentration of histone is lowest in rapidly growing tissues such as tumours and embryos. Correspondingly in red cells, with a high histone content, growth has practically ceased.

Finally, if the scheme suggested were correct, one should find that protamine is a much better inhibitor of protein synthesis than is histone, since protamine as a more basic substance should combine the more vigorously with nucleic acid. Thus in a cell containing much protamine, protein synthesis should be negligible. This is indeed the situation found with spermatozoa. Whether these facts are causally connected however, remains to be discovered.

The work described here has been carried out with the aid and collaboration of a number of colleagues. These include Dr. Sanger and Dr. Ehrensvärd, who have advised me on the use of blocking agents; Mr. Loveless, who has been mainly responsible for methods involving derivatives of iodoacetic acid; and Mr. Bell, who has investigated the nitrogen mustards and techniques for colour intensification. The work on electron microscopy is being carried out with Prof. Randall and Miss Martin. Finally, the author is grateful to Dr. K. B. Warren, from whose editorial vigour this paper has profited.

REFERENCES

BRACHET, J., 1945, Embryologie Chimique. Paris, Masson et Cie.

CASPERSSON, T., 1947, The relation between nucleic acids and protein synthesis. Symposium Soc. Exp. Biol., *1:* 127-151.

COMMONER, B., 1949, On the interpretation of the absorption of ultra-violet light by cellular nucleic acids. Science *110:* 31-40.

COMMONER, B., and D. LIPKIN, 1949, The application of the Beer-Lambert law to optically anisotropic systems. Science *110:* 41-43.

DANIELLI, J. F., 1947, A study of techniques for the cytochemical demonstration of nucleic acids and some components of proteins. Symposium Soc. Exp. Biol. *1:* 101-113.

SANGER, F., 1945, The free amino groups of insulin. Biochem. J. *39:* 507-515.

STEDMAN, E., and STEDMAN, E., 1943, Probable function of histone as a regulator of mitoses. Nature, Lond. *152:* 556-557.

COMPETITIVE REACTIONS AND ANTAGONISMS IN THE BIOSYNTHESIS OF AMINO ACIDS BY NEUROSPORA

STERLING EMERSON
Kerckhoff Laboratories of Biology, California Institute of Technology, Pasadena, California

The interpretation that the only effect of a specific mutation is to alter a single reaction in the synthesis of an essential metabolite has sometimes been understood in too narrow a sense. This idea, usually called the one gene-one enzyme hypothesis, refers to the primary, or direct effect of the gene on its immediate environment. Numerous side effects resulting from the initial impact of the gene are to be expected on this interpretation, and such secondary events may be confused with the primary, gene-controlled reaction itself. Workers in the immediate field have long been aware of the possibility of such confusion, but a more complete statement about events directly and indirectly resulting from the action of genes is desirable. The competitions and antagonisms occurring in the synthesis of amino acids in Neurospora not only offer the opportunity of doing this, but they cannot be readily understood without it.

Many instances have been reported in which the growth of an organism is inhibited by the addition of some one amino acid to the culture medium, and in which the inhibition is counteracted by the further addition of a second amino acid. Such interactions are rare in wild type Neurospora but are not uncommon in mutant strains in which the synthesis of amino acids is blocked. For example, the so-called lysineless mutants—mutants in which the synthesis of lysine is blocked at one or another of several steps—require lysine as a growth factor and are competitively inhibited by arginine (Doermann, 1944). Over a wide range of concentrations, 50 percent inhibition occurs when lysine and arginine are supplied in equimolar proportions. On the other hand, wild type strains are not inhibited by added arginine ten times greater than the concentration of lysine optimal to strains requiring lysine.

A somewhat different example of antagonisms in the synthesis of amino acids was reported before this symposium three years ago by Bonner (1946a, b). In his example, a strain differing from wild type by a single mutant gene requires both isoleucine and valine, and both amino acids must be supplied in a fairly definite ratio to obtain optimal growth. Bonner's interpretation of his results was that a precursor of isoleucine—believed to be the keto acid analog of isoleucine—accumulates because of a genetic block between it and the final product, and that this precursor competes with the corresponding precursor of valine in the final aminating step in the synthesis of that amino acid, thus leading to a double deficiency.

The interactions of methionine and its precursors with threonine and its precursors are among the most complex so far studied in Neurospora, and it is these interrelations that I wish to discuss at this time. In this one system are to be found most of the different types of interactions that are likely to be met. There is competition for a common precursor. There are at least two different competitively antagonistic reactions between the end products of two lines of synthesis, threonine and methionine. There are two instances of competitive antagonism between a catalyst, *p*-aminobenzoic acid, and its analog, sulfanilamide. In one of these sulfanilamide blocks an essential reaction, the methylation of homocysteine to give methionine, whereas in the other it blocks a deleterious reaction, one leading to an apparent threonine deficiency. There are two instances in which several different genes are concerned with what might appear to be a single chemical step—suggesting that some of these genes are involved in the regulation of a reaction different from, but essential to the reaction studied. There is also one example in which the apparent primary effect of a gene is something quite different from the real primary effect of that gene. These last points bring into question the validity of the criteria commonly used in identifying gene controlled reactions—a consideration of some importance to all studies in the field of chemical genetics.

Pathways of Synthesis

In summarizing what is known about the synthesis of threonine and methionine in Neurospora, it soon becomes apparent that a great deal is still to be learned. Not one of the reactions involved is completely understood in the sense of our being able to write the chemical equation for the gene controlled reaction concerned. None-the-less the general plan of these biosyntheses is beginning to clear up. I have attempted to illustrate such a plan in the accompanying diagram (Fig. 1).

Synthesis of methionine from cysteine

A large number of mutant strains of Neurospora which cannot utilize inorganic sulfate, but which will grow on methionine as the sole source of sulfur are being studied by Horowitz (1947a, b, and personal communication) and his associates. Four

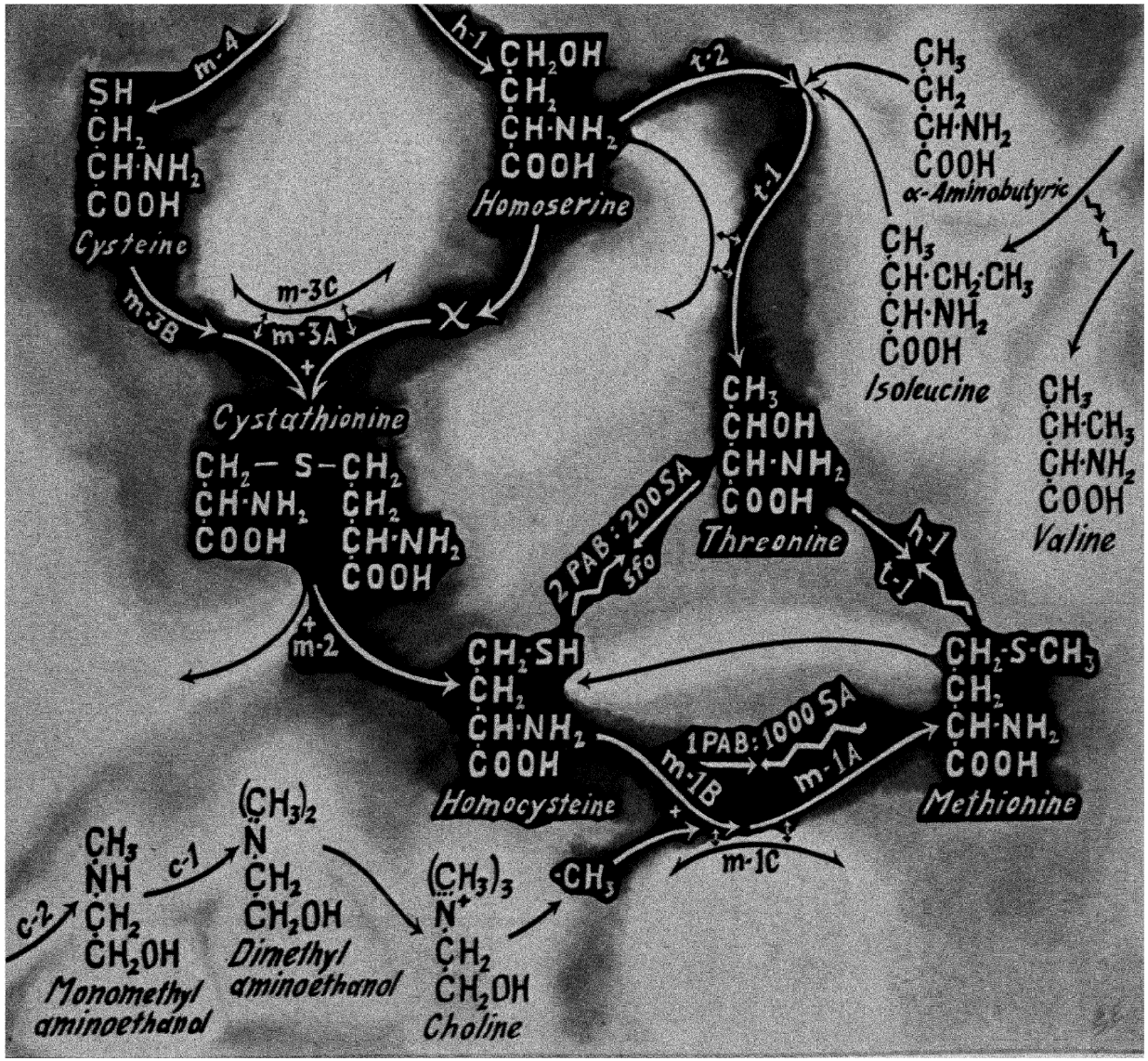


FIG. 1. A diagram intended solely as a visual aid in correlating observations and postulates discussed in the text.

groups of these mutants are important to the present review. One group specifically requires methionine for growth, another can utilize either methionine or homocysteine, the third makes use of both of these and of cystathionine as well, while the fourth can grow on any one of cysteine, cystathionine, homocysteine or methionine. Such evidence suggests that the first three of these substances are intermediates in methionine synthesis in Neurospora, and that they are produced in the order named.

The evidence that cystathionine is actually an intermediate is relatively complete. Horowitz (1947a) found that one mutant strain (H-98), in which methionine synthesis is blocked at step m-2 (Fig. 1), accumulates cystathionine in the mycelium, while mutants (36104, 39103 and 9666) blocked at step m-3, preceding cystathionine formation, can use cystathionine to satisfy the requirement for methionine. Sulfur containing precursors of methionine apparently do not accumulate when methionine synthesis is blocked between homocysteine and methionine (step m-1) or between cysteine and cystathionine (step m-3), or if they do accumulate they are not in a form that can be used by other mutant strains when added in crude extracts to the culture medium.

Strictly speaking, the observation that cystathionine is accumulated by one mutant strain and utilized by another does not in itself prove that it is a primary intermediate in the synthesis of methionine. The possibility still exists that the direct intermediate is a different substance, but one which is readily interconvertible with cystathionine. Only by studies with isolated enzyme systems can this point be definitely resolved. On available evidence, however, it is reasonable to suppose that cystathionine is the intermediate.

Methylation of homocysteine

Horowitz and Fling (personal communication) have found that there are two distinct genetic blocks between homocysteine and methionine (step m-1). We now suspect that there are at least three, and perhaps four different genetically controlled reactions involved in this step. Actually such a situation should not be too surprising since transmethylations in general are complex reactions, involving more-or-less specific methyl donors, and in some cases, energy yielding reactions are required. The black arrow, step m-1c in Figure 1, is intended to suggest such a coupled reaction.

The evidence for three distinct gene controlled reactions is indirect. Challenger and North (1934) had found that selenite ($SeO_3^=$) is reductionally methylated to dimethylselenide in Penicillium. By applying the same method to studies of methionineless mutants of Neurospora, Zalokar (personal communication) was able to distinguish three sorts of mutants each of which blocks the methylation of homocysteine. In his experiments no methylated selenium compounds were recovered—instead the end products seem to be metallic selenium and hydrogen selenide. One mutant strain (36102) with a genetic block between homocysteine and methionine reduces selenite under all conditions, as do mutant strains in which the block to methionine synthesis precedes homocysteine (steps m-2, m-3 and m-4). Another group of mutants (35599, 38706, 44103, 44704 and 68604) blocked between homocysteine and methionine is unable to reduce selenite when the mycelium is starved for methionine, but does bring about the reduction upon the addition of extra methionine. The third group of mutants (29627, 35809 and 48003) blocked at the same step is unable to reduce selenite under any conditions. While there is no evidence that methylation is involved in the reduction of selenite by Neurospora, we can conclude that at least some parts of the mechanism for methylating homocysteine are also involved in the reduction of selenite—otherwise it would be difficult to account for the observed differences in behavior of methionineless mutants. Recent genetic tests, while incomplete, are not in disagreement with the interpretation that three distinct genes are responsible for the differences observed by Zalokar (Fling, personal communication).

Of the mutant strains which cannot methylate homocysteine and which can reduce selenite only in the presence of excess methionine, one (35599) differs from all other methionineless mutants by responding maximally to extremely small amounts of methionine (Horowitz and Fling, personal communication). Most strains require about 0.5 mg methionine per 20 ml culture for maximum growth, whereas this strain gives maximum growth on as little as one microgram. Genetic tests are still not sufficiently extensive to determine possible allelism between this mutant and others belonging to the same group.

We suspect that the methyl group of methionine comes from choline, but this is not certain. Mutant strains blocked at step c-1 (47904) and c-2 (34486) in the synthesis of choline can use methionine as a partial replacement for choline, but do not grow indefinitely on methionine alone (Horowitz and Beadle, 1943; Horowitz, Bonner and Houlahan, 1945; Horowitz, 1946).

The suggestion that there may be a fourth genetic block to the methylation of homocysteine arises from recent studies of a mutant strain (37603) which will grow when either methionine or choline is added as the sole supplement to the medium. One peculiarity of this strain is that it responds to choline and methionine much better at 35° C than at 25°. On solid medium it will eventually make considerable growth on minimal medium, but only after a protracted lag period. When choline is added as the sole supplement, there is also an initial lag phase, but very much shorter than on minimal medium. The addition of methionine, with or with-

out choline, shortens the lag still further, and on some extracts of wild type Neurospora, growth begins without an appreciable lag—suggesting that the principal requirement of this mutant has not yet been discovered. From available data (unpublished results of Zalokar, Fling, Emerson) we are not in a position to say that this mutant is definitely related to the methylation of homocysteine. On the other hand, it can use methionine, but not homocysteine, and like certain other methionineless mutants it is unable to reduce selenite. (Cholineless mutants blocked at steps c-1 and c-2 do reduce selenite.) Genetic tests for allelism with other methionine requiring mutants have not been made.

Para-aminobenzoic acid must also be involved in the methylation of homocysteine since methionine, but not homocysteine, has a sparing action on the *p*-aminobenzoic acid requirement of a mutant strain (1633) which is unable to synthesize this vitamin (Zalokar, unpublished, *cf.* Harris and Kohn, 1941). More will be said about this relationship in connection with antagonistic reactions.

Our present information suggests that the methylation of homocysteine in Neurospora may be somewhat different from that occurring in mammalian liver homogenates (Borsook and Dubnoff, 1947; Dubnoff, 1949). None of the Neurospora mutant strains which require methionine or choline is able to use betaine or dimethylthetin even in a sparing capacity. Homogenates of wild type Neurospora do not methylate homocysteine under conditions similar to those employed with mammalian material (Dubnoff and Horowitz, unpublished). To me this implies that some essential coupled reaction may be limiting in Neurospora homogenates.

Synthesis of methionine from homoserine

The carbon skeleton of methionine is believed to be derived from the four carbon chain of homoserine. There is one mutant strain of Neurospora (51504) which requires homoserine (Teas, Horowitz and Fling, 1948). It will not grow on any other single amino acid, but does grow on mixtures of threonine and methionine, threonine and homocysteine, or threonine and cystathionine. It seems probable that the carbon chains of both threonine and methionine are derived from homoserine, as suggested in Figure 1.

Three distinct genetic blocks (represented by mutant strains 9666, 36104 and 39103, respectively) have been found to be associated with step m-3 in which homoserine and cysteine are coupled to form cystathionine (Buss, 1944). Other than from genetic inference, there is, at present, no evidence that three different biochemical reactions are involved at this step. Horowitz and Fling (personal communication) have found that each of the three mutants blocked at this step accumulates substances which replace both the methionine and threonine requirements of the homoserineless strain (51504). In at least one of these (9666) threonine accumulates, implying that when the pathway to cystathionine is blocked, excess homoserine is shunted into the synthesis of threonine. The same mutant accumulates another substance, represented by the "X" in Figure 1, which satisfies the methionine requirement in the homoserineless mutant. It is presumably an intermediate between homoserine and cystathionine.

Synthesis of threonine from homoserine

Four genetically distinct mutant strains of Neurospora which are unable to synthesize threonine have been studied by Teas (1947, 1948). One of these (35423) has a specific requirement for threonine, and is assumed to be blocked in a terminal step in threonine synthesis—represented as step t-1 in Figure 1. Another (44104) can make use of either threonine or α-aminobutyric acid, and to a lesser extent of isoleucine. The third (51504) is the mutant already discussed in connection with methionine synthesis—it responds either to homoserine alone or to both methionine and threonine. When this mutant is supplied with appropriate amounts of methionine, it resembles a threonineless mutant, having a single requirement under such conditions. The threonine requirement of this mutant can be met by either threonine or homoserine, but not by α-aminobutyric acid or by isoleucine. From the requirements of these three mutants, it is apparent that they can not simply block three successive steps in a chain of reactions leading to threonine. The fourth mutant (46003) in this series is one of that disagreeable kind which makes appreciable growth in the absence of added metabolites. It is enabled to make normal growth upon the addition of any one of threonine, homoserine, α-aminobutyric acid, α-ketobutyric acid, isoleucine, or the keto acid analog of isoleucine. It does not fit into the illustrated scheme (Fig. 1) and will be disregarded, except to remember that it must eventually be accounted for.

Horowitz (personal communication) recently suggested that the requirements of the first three mutant strains could be accounted for by a scheme in which homoserine enters into the synthesis of threonine at two different steps. One such scheme is illustrated in Figure 1. By this scheme homoserine is transformed at step t-2 to some unknown compound which is related in some way to α-aminobutyric acid. This intermediate is then changed to threonine by a reaction, step t-1, which is again dependent upon homoserine. In the diagram, this dependency upon homoserine is represented by the black arrow. The important point in such a scheme is that the postulated intermediates can be transformed into threonine only in the presence of homoserine, which can either be made by the organism or supplied in the medium. There is some additional evidence (Teas, *loc cit*) to support this idea. It was noted earlier that the mutant blocked before homoserine (step h-1)

cannot use either α-aminobutyric acid or isoleucine in place of threonine. However, when homoserine is supplied in suboptimal amounts growth is stimulated by additions of either α-aminobutyric acid or isoleucine.

Utilization of D-*amino acids*

Some additional information about the possible pathways of synthesis of threonine and methionine has come from experiments in which comparisons were made between the unnatural D-isomers and the natural L-isomers of various amino acids as growth promoting substances. Neurospora can use D-methionine just as efficiently as L-methionine, but cannot use D-threonine (Teas, Horowitz and Fling, 1948). Horowitz (1944) found that the D-amino acid oxidase of Neurospora is especially active towards D-methionine, but is unable to use D-threonine as a substrate. Neurospora may thus have a mechanism for changing D-methionine to L-methionine through their α-keto analog, but presumably does not have a similar mechanism for threonine.

Mutants blocked before cystathionine are unable to use any of the unnatural isomers of cystathionine —both asymmetrical carbon atoms must have the L-configuration (Horowitz, 1947a). Similarly the mutant blocked before homoserine (51504) will not grow on D-homoserine alone (Teas, 1947, 1948). On the other hand, Teas obtained one rather curious result with this homoserineless mutant. With increasing concentrations of L-homoserine growth of this strain increases linearly with concentration up to a certain point, beyond which there is little increase in growth. With increasing concentrations of DL-homoserine growth increases linearly over a much larger range of concentrations. While at low concentrations there is only about half as much growth on the racemized mixture as on the pure L-isomer, at high concentrations the DL-mixture supports nearly twice the growth permitted by L-homoserine alone. This result of Teas suggests that although D-homoserine is unable to replace the full homoserine requirement of this strain, it may do so in one of the two pathways, either that leading to threonine or to methionine. Recent tests (Emerson) show that this is true. Methionine and D-homoserine together are inactive, but threonine and D-homoserine together are effective in promoting growth, showing that D-homoserine can substitute for the natural isomer in the synthesis of methionine, but not in that of threonine. This may possibly mean that intermediate "X" (Fig. 1) between homoserine and cystathionine does not have an asymmetrical carbon atom in the alpha position. Furthermore, since pure L-homoserine can satisfy both threonine and methionine requirements, but does not result in as much growth as the racemized mixture, it is possible that the L-isomer is actually inhibitory to some extent. Inhibition by L-homoserine has been observed in threonineless mutants as will be noted later.

Antagonistic Interactions

I have attempted to summarize the amino acid requirements of a number of mutant strains of Neurospora and to point out certain interrelations. I would like now to take up some examples in which amino acids required by one mutant strain suppress growth of another.

Inhibition of threonineless mutants by methionine

Wild type Neurospora is not inhibited by methionine even in relatively high concentrations. In contrast to this, two of three mutant strains which are unable to synthesize threonine are inhibited by relatively low concentrations of methionine. On the scheme illustrated in Figure 1, mutants blocked at step h-1 (51504) and at step t-1 (35423) are competitively inhibited by methionine—growth being restored by increased concentrations of threonine—whereas the mutant blocked at step t-2 (44104) is not inhibited by methionine (Teas, Horowitz and Fling, 1948; Teas, 1948). In the case of the homoserineless mutant (step h-1) it was found that only L-methionine inhibits, the D-isomer having no inhibitory effect.

Competitive inhibitions of this sort are difficult to interpret. We would ordinarily think that an inhibition by methionine which is competitively overcome by additional threonine would indicate that methionine competes in the utilization of threonine. If the inhibition occurred during the synthesis of threonine, then an outside source of threonine should overcome the inhibition in a non-competitive manner. To account for the absence of inhibition in wild type, we might suppose that an exogenous source of threonine is not strictly comparable to the threonine normally synthesized. Then how can we account for the difference in response of mutants blocked at step t-1 and t-2? Neither makes its own threonine, both are supplied with threonine by way of the culture medium, yet one is inhibited by methionine and the other is not. Then again, when the homoserineless mutant (step h-1) is supplied with homoserine it makes its own threonine, yet it is competitively inhibited by methionine under these conditions as well (Teas, 1948). One possible interpretation is that when the synthesis of threonine is blocked at certain steps, some product of the blocked reaction must take part in the inhibitory action of methionine.

Inhibition of threonineless mutants by homoserine

It is even more surprising to find that some threonineless mutants are inhibited by a presumed precursor of threonine. Teas (1948) found that mutants blocked at either step t-1 (35423) or t-2 (44104) were inhibited in a competitive manner by homoserine. The inhibition of the mutant strain involved in step t-1 is especially strong; equimolar proportions of DL-threonine and DL-homoserine result in

complete inhibition. A recent test (Emerson) showed that the inhibition is due to the L-isomer, D-homoserine being ineffective in the range of concentrations tested.

Such strong inhibition of a threonineless mutant by a precursor of threonine suggests to me that this mutant may not be involved in the final step (t-1) of threonine synthesis. It is possible that in this mutant strain, there is a reaction using homoserine say as substrate, which leads to a destruction, or removal of threonine. The deficiency would be overcome by supplying exogenous threonine, but additional homoserine would feed the deleterious reaction, resulting in a loss of the added threonine, and in growth failure. This interpretation was suggested by analogies between the example just described and results previously obtained with the "sulfonamide requiring" mutant strain which will be described later. Some support for the interpretation is to be found in an unpublished experiment of Teas, in which he does not place full confidence, but which he has kindly permitted me to cite. He made the double mutant between the homoserineless strain (51504) and the threonineless mutant under discussion (35423), which should have blocks at both step h-1 and step t-1. Teas found that the presumed double mutant does not grow on a mixture of threonine and methionine, though both parental mutant strains do. On the other hand, it does, unexpectedly, grow on homoserine alone. If small amounts of homoserine are used preferentially for the synthesis of threonine and methionine, and only larger amounts lead to the reaction resulting in deficiency of threonine, then this result is to be expected. More direct tests of the interpretation could not be made due to the loss of the double mutant strain, which has not yet been reconstituted. For the time being, the plausibility of the interpretation must depend on the resemblance between the behavior of this threonineless mutant and that of the sulfonamide-requiring strain.

Relation of p-aminobenzoic acid to methionine-threonine interactions

My laboratory first became interested in interactions involving threonine and methionine from studies involving the mutant strain (E-15172) which was first characterized as having a requirement for sulfanilamide (Emerson, 1947). While making fair growth on minimal medium at 25° C, this strain fails to grow on minimal medium at 35°, but does grow at the higher temperature when sulfanilamide is added as the sole supplement. It was then shown by Zalokar (1948) that this strain is inhibited by the amount of *p*-aminobenzoic acid normally produced by Neurospora—sulfanilamide promoting growth by antagonizing this inhibition. By combining this mutant strain with another (1633) in which the synthesis of *p*-aminobenzoic acid is blocked, it was possible to regulate the *p*-aminobenzoic acid available to the strain by the amount added to the medium. By this means, it was found that the sulfonamide-requiring strain, when unable to synthesize *p*-aminobenzoic acid, can grow in the absence of sulfanilamide if supplied with just the right amount of *p*-aminobenzoic acid in the culture medium. This amount corresponds closely to the minimal requirement of the *p*-aminobenzoicless strain (1633) itself, about 10^{-7} molar. Twice that amount of *p*-aminobenzoic acid results in some inhibition, ten times the amount in complete inhibition, whereas it requires 10,000 times more still to inhibit the *p*-aminobenzoicless strain, or wild type.

It was already known that *p*-aminobenzoic acid is involved in methionine synthesis. Zalokar (unpublished) next looked for a relationship between methionine and the sulfonamide-requiring character. Double mutants were made between the sulfonamide-requiring strain and various methionineless strains. Such double mutants with blocks to methionine synthesis at steps m-2 or m-3 grow in the absence of sulfanilamide when methionine is supplied in relatively low concentrations. At higher concentrations of methionine, growth is inhibited, but can be restored by the addition of sulfanilamide. It appears, therefore, that the sulfonamide-requiring strain is inhibited not only by the amount of *p*-aminobenzoic acid normally produced, but also by the normally occurring quantity of methionine. Inhibition of growth by both *p*-aminobenzoic acid and methionine is overcome by sulfanilamide. However, the antagonism between sulfanilamide and *p*-aminobenzoic acid is competitive, that between sulfanilamide and methionine non-competitive; once the inhibition by methionine is overcome by sulfanilamide, additional methionine has no inhibiting effect. Similar double mutants carrying the sulfonamide-requiring gene and a genetic block to methionine synthesis, but with the block occurring at step m-1, fail to grow in the absence of sulfanilamide no matter what concentration of methionine is added. Since strains blocked at step m-1 can produce homocysteine, and since strains blocked at earlier steps may degrade methionine to homocysteine, it is probable that homocysteine rather than methionine is the inhibitor.

Zalokar (unpublished) then found that the sulfonamide-requiring strain will grow when threonine is added as sole supplement to minimal medium. Threonine antagonizes the inhibition caused by both *p*-aminobenzoic acid and by methionine (or homocysteine). In this case, however, the antagonism between threonine and methionine is competitive (in roughly equimolar proportions), whereas that between threonine and *p*-aminobenzoic acid is non-competitive—once the inhibition has been overcome by added threonine, additional *p*-aminobenzoic acid has no inhibitory effect.

It is now possible to describe the sulfonamide-

requiring strain as one having an aberrant reaction, represented by step sfo in Figure 1, which is catalyzed by *p*-aminobenzoic acid, which requires homocysteine (or methionine) as a substrate, and which results in a threonine requirement. This deleterious reaction is not important until the concentration of *p*-aminobenzoic acid reaches about twice that required for other reactions normally catalyzed by this vitamin. On the other hand, the deleterious reaction is inhibited by sulfanilamide when the molar ratio of sulfanilamide to *p*-aminobenzoic acid is 100:1, whereas inhibition of other reactions (*e.g.* the methylation of homocysteine) occurs only at a ratio of about 1000:1. Similarly, this deleterious reaction is not important until the concentration of methionine is greater than that required by other, essential reactions. Sulfanilamide supports growth of this strain by competing with *p*-aminobenzoic acid, which is an essential catalyst for the deleterious reaction. Once the reaction is blocked by sulfanilamide, additional methionine has no effect, as would be expected if it acts as substrate in a blocked reaction. When the deleterious reaction is counteracted by the addition of threonine, additional *p*-aminobenzoic acid is ineffective, probably because it is produced in such large amounts as not to become limiting; but additional methionine is effective by supplying more material for the reaction.

Zalokar found that the methionine-threonine antagonism occurring in the sulfonamide-requiring strain differs from that occurring in the threonineless strains in a number of respects. In the first place, methionine inhibition in the sulfonamide-requiring strain is especially marked only at high temperatures (35° C) while in the threonineless strains it appears to be independent of temperature. In the second place, the inhibition by methionine in the sulfonamide-requiring strain is catalyzed by *p*-aminobenzoic acid and antagonized by sulfanilamide while in the threonineless strains both *p*-aminobenzoic acid and sulfanilamide are without effect. And finally, inhibition by methionine in the threonineless strains is counteracted by α-aminobutyric acid and isoleucine, but not by valine (Teas, 1948), whereas in the sulfonamide strain it is counteracted to some extent by isoleucine and valine, but not appreciably by α-aminobutyric acid. It follows that inhibition by methionine in the sulfonamide-requiring strain is the result of a different reaction from that responsible for inhibition by methionine in the threonineless strains.

Conclusions

By way of summary, or conclusion, I wish to present a point of view about biosyntheses in general, and about the application of genetic techniques to their elucidation.

In the first place, I believe that the interrelationships just reviewed will not prove to be much more complicated than those to be encountered in the synthesis of almost any other metabolite. As studies of other lines of synthesis progress, we shall probably find very similar interactions there as well.

My principal concern arises from the conclusion that it is going to be much more difficult than we had once supposed to identify the particular biochemical reaction controlled by any one gene. It is certainly possible that a mutant strain may have a specific metabolite requirement due to any one of a number of causes.

While the methylation of homocysteine may be unusually complex, mutant strains which require methionine and cannot use homocysteine may serve as an illustration of this point. If we suppose that a single enzyme is involved in the transmethylation occurring at this step, one possibility is that a particular mutant has lost this enzyme. Another possibility is that the enzyme is still present, but has become modified so that it has lost its affinity for its substrate, homocysteine. For present purposes, let us suppose that the methyl group is transferred from choline. Then the enzyme must also have an affinity for choline, and the loss of this affinity would constitute a third possibility for the failure of the mutant to methylate homocysteine. Further, we know that *p*-aminobenzoic acid has a catalytic effect upon this reaction, possibly by forming part of a co-enzyme. Presumably the co-enzyme must be associated with the enzyme to be effective, and loss of the enzyme's affinity for the co-enzyme would give a fourth possibility. If an enzyme's specificity is determined by a single gene, these alternatives might be realized in different mutant alleles of one gene. Differences have been noted in the response of mutant strains involved in what we suppose is a single step in the methylation of homocysteine—strains belonging to the same sub-group of mutants blocked between homocysteine and methionine. Genetic data are still insufficient to determine the extent of allelism involved.

There are also a number of ways in which the methylation of homocysteine might be blocked without involving the enzyme itself. If the postulated co-enzyme were specific for this reaction, any block in its synthesis would appear as a block in methionine synthesis. Further, if choline becomes dimethylaminoethanol upon giving up its labile methyl to homocysteine, it is possible that it must pick up an electron at the same time. If so, a mutant interrupting this electron transfer would appear as a methionineless mutant.

And still the possibilities are not exhausted. Examples of competitive inhibition have been noted earlier in this review. The idea that an accumulated product resulting from a blocked reaction may inhibit another reaction has previously been suggested by Bonner (1946a, b). It is possible that different mutants would accumulate products which compete with each of the different components of our system: methionine, choline, the co-enzyme, the elec-

tron donor, and so on. Each such mutant would appear as a methionineless strain. It is not even necessary that a reaction be blocked to result in the accumulation of an inhibitor—it might equally well accumulate as a result of a change in the rate of some reaction.

Mutations which do not completely block a reaction, but which change its rate may be fairly common in Neurospora. Many of the so-called temperature mutants (Mitchell and Houlahan, 1946; McElroy and Mitchell, 1946; Horowitz, in press) may be of this sort. Two mutant strains discussed in the present review are probably to be accounted for on the basis of changed reaction rates. In the sulfonamide-requiring strain, the deleterious reaction is much more pronounced at 35° C than at 25°. It was noted that the over-all response of this strain depends not only on this reaction, but also upon the amounts of threonine, methionine and *p*-aminobenzoic acid that are produced. At 25° the sulfonamide-requiring strain grows nearly normally on an unsupplemented medium, which must mean that at this temperature the reactions leading to the production of threonine, methionine and *p*-aminobenzoic acid, and the "deleterious" reaction are so adjusted that a favorable balance between them is attained. The failure of this strain to grow on unsupplemented medium at 35° can be accounted for by the postulate that the reactions concerned have different temperature sensitivities, such that with increasing temperature, they are thrown out of balance—in this instance, presumably due to a rapid increase in the rate of the deleterious reaction. It seems simpler to us to suppose that this so-called deleterious reaction actually is present in wild type, and that the mutant differs from wild type by having an altered temperature sensitivity in this one reaction. The mutant strain (37603) which responds to either methionine or choline, but much better at 35° than at 25° appears to be a similar example, but we know less about it at present.

It should be pointed out that the idea that mutant genes produce their observed effect as a result of changes in the rates of enzyme controlled reactions is not a new one. I believe it was first proposed by Goldschmidt (1916, 1920), and by Wright (1917).

The most strongly felt present need in the Neurospora studies is additional evidence about the enzymes concerned in different synthetic steps. So far only two instances have been reported in which loss of enzyme activity is associated with the genetic loss of a synthetic ability. The first case involves the coupling of β-alanine and pantoyl lactone in the synthesis of pantothenic acid (Wagner and Guirard, 1948). Wagner (1949) has recently reported that the mutant strain which is unable to synthesize pantothenic acid actually still produces the enzyme, but the mutant also accumulates a specific inhibitor which prevents the enzyme from functioning. The second example involves the coupling of indole and serine to form tryptophane (Mitchell and Lein, 1948). The investigation of this system is still in progress.

It now appears that some of us were unduly optimistic in our interpretations of combined genetic and biochemical studies. To me, at least, it seemed that the introduction of a genetic block in a chain of reactions should be the most direct method of identifying a particular reaction. When a reaction is stopped by an outside agent, it is obvious that there may be a number of other reactions affected by that agent—reactions which may have no relation to each other except that they respond to the same outside agent. It is now evident that a genetic block should also be expected to have multiple consequences. On our present information it seems probable that a gene mutation blocks, or otherwise effects, a single step in a chain of biochemical reactions. However, when such a step is blocked by mutation, there should be two primary results: a deficiency of the product normally formed by the blocked reaction, and a piling up of the materials normally taking part in the blocked reaction. Such accumulation of reacting substances can result in quite unpredictable side reactions, thus increasing the number of possible end effects resulting from one mutation.

Once we are aware of the complexities to be expected, however, it seems not improbable that the genetic method will prove more fruitful than ever in arriving at an understanding of the way biochemical reactions take place within the living cell. Complexities in the synthesis of threonine and methionine have been brought to light in part by studying a number of different mutant strains with the same apparent requirement, and in part by making use of double mutants in which there are two genetically blocked reactions. Both lines of investigation hold considerable promise, and I believe that a detailed study of the requirements of different mutant alleles of a single gene will prove profitable.

ACKNOWLEDGMENTS

Professor N. H. Horowitz, Dr. Marguerite Fling, Dr. H. J. Teas and Dr. Marko Zalokar have kindly permitted me to refer to certain unpublished studies. It is a pleasure to acknowledge my debt to them, and to them and other colleagues for many helpful discussions.

REFERENCES

BONNER, DAVID, 1946a, Biochemical mutations in Neurospora. Cold Spr. Harb. Symposia Quant. Biol. *11:* 14-24.

1946b, Further studies of mutant strains of Neurospora requiring isoleucine and valine. J. biol. Chem. *166:* 545-554.

BORSOOK, HENRY, and DUBNOFF, J. W., 1947, Methionine formation by transmethylation in vitro. J. biol. Chem. *169:* 247-258.

BUSS, HELEN R., 1944, The genetics of methionine mutants. Stanford Univ. (Thesis).

CHALLENGER, F. and NORTH, H. E., 1934, The production of organo-metalloid compounds by microorganisms. Part II. Dimethyl selenide. J. Chem. Soc. 1934: 68-71.

DOERMANN, A. H., 1944, A lysineless mutant of Neurospora and its inhibition by arginine. Arch. Biochem. *5:* 373-384.

DUBNOFF, JACOB W., 1949, Non-lability of choline methyl and the role of choline oxidase in transmethylation. Federation Proc. *8:* 195 (Abstract).

GOLDSCHMIDT, RICHARD, 1916, Genetic factors and enzyme action. Science (N.S.) *43:* 98-100.

1920, Die quantitative Grundlage von Vererbung und Artbildung. Vort. EntwMech. Org. *24:* 1-163.

HARRIS, J. S. and KOHN, H. I., 1941, On the mode of action of sulfonamides. II. The specific antagonism between methionine and the sulfonamides in *Escherichia coli.* J. Pharmacol. *73:* 383-400.

EMERSON, STERLING, 1947, Growth responses of a sulfonamide-requiring mutant strain of Neurospora. J. Bact. *54:* 195-207.

HOROWITZ, N. H., 1944, The *d*-amino acid oxidase of Neurospora. J. biol. Chem. *154:* 141-149.

1945, The isolation and identification of a natural precursor of choline. J. biol. Chem. *162:* 413-419.

1947a, Methionine synthesis in Neurospora. The isolation of cystathionine. J. biol. Chem. *171:* 255-264.

1947b, The formation of methionine from cysteine in Neurospora. Federation Proc. *6:* 262-263 (Abstract).

In Press, Biochemical genetics of Neurospora. Advances in Genetics.

HOROWITZ, N. H. and BEADLE, G. W., 1943, A microbiological method for the determination of choline by use of a mutant of Neurospora. J. biol. Chem. *150:* 325-333.

HOROWITZ, N. H., BONNER, D. and HOULAHAN, MARY B., 1945, The utilization of choline analogues by cholineless mutants of Neurospora. J. biol. Chem. *159:* 145-151.

MCELROY, W. D. and MITCHELL, H. K., 1946, Enzyme studies on a temperature sensitive mutant of Neurospora. Federation Proc. *5:* 376-379.

MITCHELL, H. K. and HOULAHAN, MARY B., 1946, Adenine-requiring mutants of *Neurospora crassa.* Federation Proc. *5:* 370-375.

MITCHELL, H. K. and LEIN, J., 1948, A Neurospora mutant deficient in the enzymatic synthesis of tryptophan. J. biol. Chem. *175:* 481-482.

TEAS, H. J., 1947, The biochemistry and genetics of threonine requiring mutants of *Neurospora crassa.* California Institute of Technology. (Thesis)

1948, Threonine requiring mutants of Neurospora. Oak Ridge National Laboratory (Health and Biology). ORNL 164.

TEAS, H. J., HOROWITZ, N. H., and FLING, MARGUERITE, 1948, Homoserine as a precursor of threonine and methionine in Neurospora. J. biol. Chem. *172:* 651-658.

WAGNER, R. P., 1949, In vitro synthesis of pantothenic acid by pantothenicless mutant and wild type Neurospora. Amer. Chem. Soc., Div. Biochem., Abstract no. 24 (San Francisco, March, 1949).

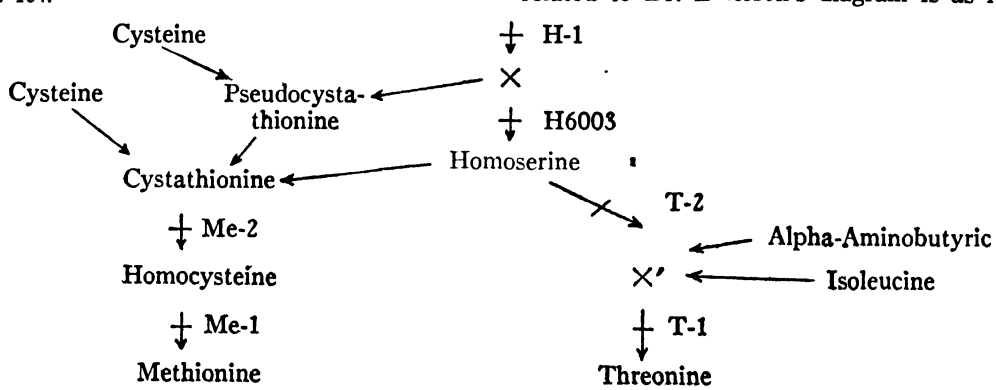
WAGNER, R. P., and GUIRARD, BEVERLY M., 1948, A gene-controlled reaction in Neurospora involving the synthesis of pantothenic acid. Proc. nat. Acad. Sci., Wash. *34:* 389-402.

WRIGHT, SEWALL, 1917, Color inheritance in mammals. J. Hered. *8:* 224-235.

ZALOKAR, MARKO, 1948, The p-aminobenzoic acid requirement of the "sulfonamide-requiring" mutant strain of Neurospora. Proc. nat. Acad. Sci., Wash. *34:* 32-36.

DISCUSSION

TEAS: Dr. Emerson has mentioned another Neurospora mutant which was omitted from his scheme, and I would like to describe this mutant. In experiments with the mutant H-1 Dr. Emerson found that although only the L-isomer of homoserine functions alone, the D-isomer is utilized in the presence of threonine. Since Horowitz reported that only LL-cystathionine functions in Neurospora mutants, there must be an intermediate between homoserine and cystathionine in which the alpha carbon atom lacks asymmetrical groups. This might also be a keto, imino or other group. In the previously omitted mutant, to which Dr. Emerson has referred, number 46003, some growth occurs on minimal medium but optimal growth follows addition of threonine, isoleucine, alpha-aminobutyric acid, or homoserine; however, there is no methionine requirement. The basis for this apparent inconsistency seems to involve alternate pathways for the synthesis of cystathionine. Since mutant 46003 is able to utilize homoserine it must be blocked before homoserine formation. If cystathionine were able to be made directly from homoserine or alternatively from the symmetrical intermediate postulated (that is, through a "pseudocystathionine" with no asymmetric alpha carbon atom on the four carbon portion) the steps involved in these mutants would be understandable. The scheme as it related to Dr. Emerson's diagram is as follows:



In this discussion, Dr. Luck has mentioned the possibility of a CO_2 tension relation in the requirements of Neurospora mutants. The work to which Dr. Luck refers is probably that of Lyman and collaborators who found for certain lactic acid bacteria that high CO_2 tension in the presence of B_6 affected the amino acid requirements. Phenylethylamine and high CO_2 tension eliminated the otherwise absolute requirement for phenylalanine.

THE QUANTITATIVE ANALYSIS OF AMINO ACIDS IN PROTEINS: INSULIN AND LYSOZYME

CLAUDE FROMAGEOT
Laboratoire de Chimie biologique de la Faculté des Sciences, Paris

Recently a method has been described (Fromageot, Jutisz and Lederer, 1948) of fractionating amino acids and eventually peptides arising from protein hydrolysis, into the following groups: basic substances, acidic substances and aromatic neutral substances, separated by adsorption on silica gel, "acidic" alumina and activated charcoal respectively, and non-adsorbed non-aromatic neutral substances. In each case, elution of adsorbed substances permits their quantitative recovery. Such a method constitutes a first series of separations preliminary either to further separations permitting individual isolations, as in the excellent methods of Stein and Moore and of Craig, Gregory and Barry, described in this Symposium, or to specific chemical titrations of each of the amino acids and peptides in each of the groups thus isolated. It seems, indeed, that owing to the great number of amino acids and peptides bound to coexist in one hydrolysate, such a preliminary step is necessary to avoid overlapping in the subsequent operations and to facilitate further quantitative work.

To check the reliability of this method, it was applied (Fromageot, Jutisz and Lederer, 1948) to the fractionation of a total hydrolysate of insulin. The amino acid composition of this protein, is, in fact, quite well known as the result of the extensive work of Brand (Brand, 1946) in the United States, and Chibnall (Chibnall, 1946) and Tristram (Tristram, 1946; Darmon, Sutherland and Tristram, 1948) in England. Considering only each separate groups as a whole, it appeared that the total nitrogen content of each could be accounted for in a satisfactory manner by the data of the British authors.

In the present work, the content of each amino acid in each of the four groups obtained by fractionating a total hydrolysate of insulin has been determined, and a fairly complete analysis of the protein has been obtained; the results show the reliability of the fractionation method and confirm its quantitative character. Furthermore, this technique has been applied to the study of lysozyme, the composition of which was not yet known, with the result that it is possible now to give an almost complete account of its amino acid content.

Insulin

The insulin used is a crystallized sample of bovine insulin with a potency of 27 I.U./mg, which has been kindly given to us by Dr. J. Lens, Organon, Oss, Holland. Its hydrolysis with 5.5N hydrochloric acid, the treatment and the fractionation of the hydrolysate have been already described (Fromageot, Jutisz and Lederer, 1948). The titration of individual amino acids is made in the following ways. *Basic amino acids:* arginine and histidine, by the Sakaguchi and the Pauli reactions respectively, according to the procedure of Macpherson (Macpherson 1946), lysine calculated by difference between totla basic nitrogen and the sum of arginin- and histidine-nitrogen. *Acidic amino acids:* aspartic acid according to Fromageot and Colas (1949), glutamic acid calculated by difference between total acid nitrogen and aspartic acid-nitrogen. *Tyrosine* according to Folin and Marenzi (1929a), *Glycine* according to Alexander, Landwehr and Seligman (1945), *Cystine* according to Folin and Marenzi (1929b), *Threonine* according to Winnick (1942). All the other amino acids by quantitative paper chromatography, after the method given by Fisher, Parsons and Morrison (1948), following the procedure described by Fromageot and Privat de Garilhe (1949). *Proline,* although estimated by the same fundamental technique, is detected by isatin instead of ninhydrin, as proposed by Acher, Fromageot and Jutisz (1949). The details of these determinations are given elsewhere (Acher, Fromageot and Jutisz, 1949).

The results obtained are summarized in Table 1, which illustrates the proportions of the different forms of nitrogen obtained in the present work, compared to those calculated from previous reports.

These results show: first, that there is a satisfactory agreement between the total nitrogen of each of the groups which have been separated by chromatography and the nitrogen of the sum of each of the amino acids contained in each of these groups, this sum being calculated either from the data of previous publications, or from our data; second, that most of the values found for each amino acid agree with previously reported values; the discrepancies are generally too small to affect the number of residues of each amino acid, calculated either from the present data or from the data of Chibnall (1946) and of Tristram (Tristram, 1946; Darmon, Sutherland and Tristram, 1948) for an insulin unit with a molecular weight of 12,000. In the case of glutamic acid, leucine and proline, however, one finds one residue more, one residue more, and one residue less respectively. Considering glumatic acid, the discrepancy between our result (16.0 residues) and Chibnall's result (15.2 residues, rounded to 15) is small, and it is, actually, difficult to decide between

TABLE 1. BALANCE OF THE DIFFERENT FORMS OF NITROGEN IN INSULIN

The data give amino acid nitrogen or group nitrogen in percent of total protein nitrogen. The data of the present work are the average of results obtained with three different hydrolysates.

T = Nitrogen as found experimentally in fractions isolated by chromatography.

C = Nitrogen as calculated from residues of amino acids into the fractions.

Total nitrogen content of the insulin samples: Brand 16.0%, Chibnall and Tristram 15.5%, present work 15.3%.

Form of nitrogen	Brand		Chibnall, Tristram		Present work			Residues
					T	*C*		
Basic fraction	19.1		17.7		18.2	18.2		
Arginine		7.0		6.2			6.0	2
Histidine		9.0		8.5			8.8	4
Lysine		3.1		3.0			3.4	2
Acidic fraction	16.5		15.2		15.8	16.0		
Aspartic acid		4.5		3.8			3.8	5
Glutamic acid		12.0		11.4			12.0	16
Aromatic fraction	10.1		10.4		11.4	10.6		
Phenylalanine		4.2		4.4			4.6	6
Tyrosine		5.9		6.0			6.1	8
Neutral, non-aromatic fraction	40.7		43.9		42.7	44.3		
Glycine		5.4		5.2			5.5	7
Alanine		—		4.4			4.7	6
Serine		4.8		4.5			4.5	6
Cystine		8.0		9.4			9.5	6
Threonine		2.3		1.6			1.8	2
Valine		6.6		5.8			5.8	8
Leucine		8.9		9.3			10.2	13
Isoleucine		1.9		1.7			1.1	2
Proline		2.2		2.0			1.6	2
Amide N	11.0		8.9		9.2	9.2		12
Total nitrogen	97.2		96.1		97.3	98.3		

15 and 16 for the number of residues of glumatic acid in the insulin unit.

As for leucine, our value (13.4 residues, rounded to 13) differs only by 8 percent from the value (12.4, rounded to 12) found by Tristram; it is therefore difficult to decide between these two values also. However, the microbiological titration of leucine made with an artificial mixture of amino acids corresponding to an insulin hydrolysate has given the same result, and we are inclined to think that the value of 13 for the number of residues of leucine is the most probable.

Concerning proline, the discrepancy between our result (2.2 residues, rounded to 2) and Tristram's result (2.7, rounded to 3) is important only by reason of rounding the experimental values to the nearest integer, thus increasing the discrepancy from 20 percent to 50 percent. New investigations, which will be reported elsewhere, seem to show that there are only 2 residues of proline in the elementary unit of insulin.

Apart from these last three amino acids, the results obtained using a method different from that employed by the British authors show the reliability of this fractionation method, and at the same time confirms in quite a satisfactory way the previous results obtained in England for bovine insulin. As for the slight discrepancies which appear between these results and those we have found, on the one hand, and those published by Brand (1946) on the other hand, they are obviously due to differences in the origin of the samples of insulin studied.

LYSOZYME

The method of fractionation in question having been checked, it has been applied to the study of the composition of lysozyme. Lysozyme appears a suitable material for the study of structure and biological activity. It exists in rather large quantities (about 3 g/l.) in egg white, from which it is easily obtained in crystalline state by the method of Alderton and Fevold (1946), and, after 4 to 6 recrystallizations, in what seems to be a chemically pure state. Its enzymatic activity, extensively studied by Meyer and coworkers (Meyer and Hahnel, 1946) is also easy to follow.

Chemically speaking, lysozyme is a protein which differs from insulin in many respects: its trypto-

phane content is very high, it has, on the other hand comparatively little tyrosine and phenylalanine, and, most important of all, it has a high content of basic amino acids, which confers a strong basic character. Two forms of lysozyme have been studied, the carbonate (N = 16.7%) obtained by precipitation of the protein from its acetic acid solution by sodium bicarbonate, and the free isoelectric protein (N = 18.6%) obtained by precipitating it from its acetic acid solution by sodium hydroxide at pH 11.0. Only the data obtained with this last are presented here. I want to express my best thanks to Dr. Fraenkel-Conrat and to Dr. J. C. Lewis whose suggestions concerning these data have been very useful.

After hydrolysis of the protein with hydrochloric acid in sealed tube, and fractional chromatography of the hydrolysate (Fromageot and Privat de Garilhe, 1949), four fractions have been obtained, the nitrogen content of which is given in the first column of Table 3. Without going into technical details which have been presented elsewhere, suffice it to say that most of the amino acids have been determined in each of these fractions. Tryptophane and cystine have been measured separately, the first one either directly in the intact lysozyme molecule, or after alkaline hydrolysis (Graham, Smith, Hier and Klein, 1947), the second after hydrolysis by a 1:1 mixture of 6N hydrochloric acid and pure formic acid as recommended by Miller and du Vigneaud (Miller and du Vigneaud, 1937). Cysteine has been looked for in the non-hydrolyzed molecule by the method of Anson (1941). Suitable corrections have been made to allow for the destruction of some of the amino acids during hydrolysis. Table 2 summarizes the results obtained.

These results suggest the following comments:

The average molecular weight of isoelectric lysozyme, as calculated from the present analytical data, is 14,700 ± 250. Although slightly higher than the

TABLE 2. SULPHUR, PHOSPHORUS AND AMINO ACIDS IN LYSOZYME

The data are the average of results obtained with at least three different hydrolysates.

A = element or amino acid.

N = amino acid nitrogen per 100 lysozyme nitrogen. (For sulphur, weight per 100 lysozyme nitrogen).

T = weight of element or amino acid found per 100 parts of dried and ash-free isoelectric lysozyme.

C = amount of element or amino acid calculated on the assumption that 14,700 g of lysozyme contain the exact number of atoms or residues given under the heading *R*.

$$\text{Dif.} = \frac{(T-C)\cdot 100}{C}$$

R = number of atoms or residues calculated on the basis of the analytical data for 14,700 g lysozyme.

M.W. = molecular weight calculated from experimental data.

A	*N*	*T*	*C*	Dif.	*R*	*M.W.*
Sulphur	13.6	2.53	2.62	− 6	12	15,200
Phosphorus	0.0	0.0	—	—	0	—
Glycine	5.6	5.6	5.61	0	11	14,700
Alanine	5.2	6.1	6.05	0	10	14,600
Serine	5.1	7.2	7.15	0	10	14,600
Cysteine	0.0	0.0	—	—	0	—
Cystine	5.0	8.0	8.15	− 2	5	—
Threonine	3.3	5.3	5.67	− 7	7	—
Methionine	1.2	2.3	2.03	+13	2	—
Valine	4.0	6.2	6.36	− 1	8	14,800
Leucine	4.0	7.0	7.13	− 2	8	14,900
Isoleucine	3.6	6.2	6.25	0	7	14,700
Proline	0.9	1.3	1.56	−20	2	—
Hydroxyproline	0.0	0.0	—	—	0	—
Phenylalanine	1.1	2.3	2.25	—	2	—
Tyrosine	1.6	3.8	3.70	+ 3	3	14,500
Tryptophane	6.2	8.3	8.32	0	6	14,700
Aspartic acid	6.7	11.8	11.8	0	13	14,700
Glutamic acid	1.7	3.3	4.00	−17	4	—
Lysine	6.2	6.0	5.95	0	6	14,600
Histidine	1.53	1.05	1.05	0	1	14,700
Arginine	25.6	14.8	15.4	− 4	13	15,300
Amide *N*	10.0	1.86	—	—	19	—
Total nitrogen	98.5					
Average molecular weight						14,700 ± 250

one obtained from X-ray measurement by Palmer, Ballantyne and Calvin (1948)—13,900 ± 600—it seems actually to agree with it, when margins of error of both methods are taken into account.

Concerning glutamic acid, the number of residues found is 3.3 for a molecular weight of 14,700. This number has been rounded to 4 for the reason that there is a greater chance of getting too low than too high a result, owing to some destruction of glutamic acid in the course of the analysis.

TABLE 3. NITROGEN DISTRIBUTION BETWEEN THE DIFFERENT FRACTIONS AND GROUPS OF AMINO ACIDS

Nitrogen expressed in percent of total nitrogen

Fraction separated by chromatography	Total nitrogen of the fraction I	Nitrogen of amino acids contained in the fraction	
		Found II	Calculated III
Basic	33.3	33.3	34.3
Acidic	15.6	8.4	8.7
Neutral aromatic	5.1	8.9	8.7
Neutral non-aromatic	32.2	37.9	38.5
Amide and miscellaneous nitrogen	10.0	10.0	10.0
Total nitrogen	96.2	98.5	100.2

The nitrogen balance of the different fractions isolated by chromatography, and of the different groups of amino acids, is presented in Table 3. By comparison between column I on one hand and columns II and III on the other hand, it is seen that, on the whole, the agreement observed in the case of insulin is not found here. The agreement between the different data concerning basic nitrogen is obvious, because the lysin-nitrogen has been calculated from these data; but for the data involving aromatic neutral nitrogen, there is a great difference which is due to the destruction of tryptophane. The so-called acidic fraction has decidedly more nitrogen than can be accounted for by the sum: "aspartic acid + glutamic acid." Some acidic nitrogenous product arising from the decomposition of tryptophane, and part of the cystine retained on the alumina could possibly explain this discrepancy. Such a behaviour of the acidic fraction is not generally found in other proteins; it prevents here any calculation of glutamic acid from the total nitrogen of this fraction. As for the fraction corresponding to the neutral non aromatic amino acids, it contains less nitrogen than is calculated from all amino acids belonging to this family. The reason is that tryptophane and cystine have been determined apart, and added afterwards to the others; and we do not know what has been the fate of the two last amino acids during hydrolysis and fractionation, except that they have been more or less destroyed.

If an average molecular weight of 14,700 and a nitrogen content of 18.6 percent be assumed for lysozyme, this protein contains 195 nitrogen atoms. The analysis reported in the present work have determined the nature of the same number, 195, of these. One must not put too much weight on this agreement, which is probably due to some compensation between different errors. But the fact remains that the analysis of lysozyme seems now to be almost complete.

SUMMARY

The fractionation of a total hydrolysate of crystallized bovine insulin by the previously described chromatography method has been achieved, and each of the amino acids in each separated group determined. Using the symbols proposed by Brand, the results may be expressed by the following formula for the unit of insulin:

$Gly_7\ Ala_6\ Ser_6\ [(Cy{\cdot}S_2]_6\ Thr_2\ Val_8\ Leu_{13}\ Ileu_2\ Pro_2\ Phe_6\ Tyr_8\ Asp_5\ Glu_{16}\ Lys_2\ His_4\ Arg_2.$

This agrees with the one found by Chibnall and by Tristram for an insulin from presumably the same species, except for leucine, and proline, for which these authors found 12 and 3 residues respectively. This agreement is considered as establishing the reliability of the chromatographic method used.

The same method has been applied to the study of isoelectric egg white lysozyme, the formula of which appears to be:

$Gly_{11}\ Ala_{10}\ Ser_{10}\ [(Cy{\cdot}S)_2]_5\ Met_2\ Thr_7\ Val_8\ Leu_8\ Ileu_7\ Pro_2\ Phe_2\ Tyr_3\ Tyr_6\ Asp_{13}\ Glu_4\ Lys_6\ His_1\ Arg_{13}\ (NH_2)_{16}.$

with a molecular weight of 14,700.

REFERENCES

ACHER, R., FROMAGEOT, CL., and JUTISZ, M., 1949, Séparations chromatographiques d'acides aminés et de peptides. V. Les acides aminés individuels de l'insuline. Avec une note sur le dosage de la proline. Biochim. Biophys. Acta *3:* in press.

ALDERTON, G., and FEVOLD, H. L., 1946, Direct crystallization of lysozyme from egg white and some crystalline salts of lysozyme. J. biol. Chem. *164:* 1-5.

ALEXANDER, B., LANDWEHR, G., and SELIGMAN, A. M., 1945, A specific micro-method for the colorimetric determination of glycine in blood and urine. J. biol. Chem. *160:* 51-59.

ANSON, M. L., 1941, The sulfhydryl groups of egg albumin. J. gen. Physiol. *24:* 399-421.

BRAND, E., 1946, Amino acid composition of simple proteins. Ann. N.Y. Acad. Sci. *47:* 187-228.

CHIBNALL, A. C., 1946, The contribution of the analytical chemist to the problem of protein structure.—Second Proctor Memorial Lecture. J. int. Soc. Leath. Chem. *30:* 1-20.

DARMON, S. E., SUTHERLAND, G. B. B. M., and TRISTRAM, G. R., 1948, The estimation of the leucine isomers in protein hydrolysates by infra-red analysis. Biochem. J. *42:* 508-516.

FISHER, R. B., PARSONS, D. S., and MORRISON, G. A., 1948, Quantitative paper chromatography. Nature, Lond. *161:* 764.

FOLIN, O., and MARENZI, A. D., 1929 a, Tyrosine and tryptophane determination in one tenth gram of protein. J. biol. Chem. *83:* 89-102.

1929 b, An improved colorimetric method for the determination of cystine in proteins. J. biol. Chem. *83:* 103-108.

FROMAGEOT, CL., and COLAS, R., 1949, Microdosage spécifique de l'acide aspartique. Biochem. Biophys. Acta *3:* 417.

FROMAGEOT, CL., JUTISZ, M., and LEDERER, E., 1948, Séparations chromatographiques d'acides aminés et de peptides. IV. Séparations en quatre groupes. Biochim. Biophys. Acta *2:* 487-498.

FROMAGEOT, CL., and PRIVAT DE GARILHE, M., 1949, La composition du lysozyme en acides aminés. II Acides aminés totaux. Biochim. Biophys. Acta *3:* in press.

GRAHAM, C. E., SMITH, E. P., HIER, S. W., and KLEIN, D., 1947, An improved method for the determination of tryptophane with p-dimethylamino-benzaldehyde. J. biol. Chem. *168:* 711-716.

MACPHERSON, H. T., 1946, Basic amino acid content of proteins. Biochem. J. *40:* 470-481.

MEYER, K., and HAHNEL, E., 1946, The estimation of lysozyme by a viscosimetric method. J. biol. Chem. *163:* 723-732.

MILLER, G. L., and DU VIGNEAUD, V., 1937, The cystine content of insulin. J. biol. Chem. *188:* 101-110.

PALMER, K. J., BALLANTYNE, M., and CALVIN, J. A., 1948, The molecular weight of lysozyme determined by the X-ray diffraction method. J. Amer. Chem. Soc. *70:* 906-908.

TRISTRAM, G. R., 1946, The application of partition chromatography to the determination of the amino acids in proteins. Biochem. J. *40:* 721-733.

WINNICK, T., 1942, Microdiffusion methods based on the bisulfite reaction. III. Determination of threonine by oxidation with periodate. J. biol. Chem. *142:* 461-466.

DISCUSSION

CASPARI: In connection with Dr. Brand's remarks on strain differences in amino acid composition in dogs I should like to mention that we have had similar experiences with the mealmoth, Ephestia.

We compared the tryptophane content of a normal wild type strain with an eye color mutant strain which had previously been outcrossed to the wild type strain for 8-9 generations. The strains had therefore relatively few genetic differences besides the eye color gene. It was found that the mutant strain contained consistently more tryptophane than the wild type strain, and that the larger part of this excess tryptophane was found in the proteins. The proteins were obtained both by digestion with pepsin and trypsin, and by precipitation with trichloroacetic acid. With both procedures a larger amount of tryptophane was found in the proteins of the mutant strain than in those of the wild type strain. This difference has been found both in adult moths and in old larvae. By fractionated alcohol precipitation, seven fractions of larval proteins were obtained which differed clearly among each other in their solubilities and in their tryptophane content. Differences between the two strains were found in three out of these seven fractions. They consisted in every case in an increased amount of tryptophane in the mutant strain.

These results indicate either that qualitatively different proteins are formed in the mutant strain, or that certain proteins rich in tryptophane are increased at the expense of some proteins poor in tryptophane. In either case, the protein makeup of the organism must be different in these two genetically similar strains. This is borne out by the observation that the rate of autolysis is different in larval proteins derived from the two strains, the mutant proteins releasing both nitrogen and tryptophane at a considerably slower rate than the wild type proteins.

It appears likely that the difference in tryptophane content between the two strains is due to the eye color gene which differentiates the two strains. The results agree with the studies quoted by Dr. Brand in suggesting that the amino-acid composition of proteins even in the same species may vary under the influence of different genes.

BENESCH: In view of Dr. Fromageot's experiences on the increased reactivity of aromatic amino acids at very alkaline pH, it may be pertinent to recall the recent papers of du Vigneaud *et al.*, which showed that when proteins are treated with mustard gas at pH 5 and then brought to pH 8 and reacted with Folin's phenol reagent the hydroxyl groups seemed to have disappeared. However, when the experiment was repeated using radioactive mustard gas, and the mustard gas treated protein was treated with duponol P. C. at pH 5 and subsequently adjusted to pH8, reaction of the hydroxyl groups with the phenol reagent was restored in spite of the fact that no mustard gas had been removed from the protein, as judged by radioactivity measurements. It might therefore be possible to achieve the same result in Dr. Fromageot's experiments with a detergent as has been observed by means of adjustment to the region of pH 12.

NEURATH: It may be well to recognize that the differences in the analytical values of the amino acid composition of insulin, so clearly emphasized by Dr. Fromageot, need not be related to experimental errors but rather to the differences in the composition of the preparations of this protein, used by the various investigators. Crystalline Zn-insulin often contains an impurity which has been shown by Dr. Sutherland of the Washington University, St. Louis, to be a protein of *hyper*glycemic activity. While the exact amount of this impurity is unknown, estimates range from 1 or 2 percent all the way to 10 percent of the total protein. It is claimed that some of the insulin preparations from European manufacturers contain little if any *hyper*glycemic activity but the specific *hypo*glycemic activity (in

units per mg. of protein) is no reliable measure of this factor. The present argument of the purity of the protein is, of course, pertinent to every protein analysis and could be clarified in part if the analysts could only agree to use identical, standardized preparations of a protein.

THE METABOLISM OF PEPTIDES

JOSEPH S. FRUTON AND SOFIA SIMMONDS
Yale University, New Haven, Connecticut

As structural intermediates between the amino acids and the proteins, peptides have long been assumed also to be metabolic intermediates in the biological interconversions between these two groups of substances. Thus, in the catabolic degradation of proteins in the gastrointestinal tract of higher animals, the action of proteinases such as pepsin, trypsin, and chymotrypsin clearly involves the formation of peptides which may be absorbed as such, or may be broken down further to amino acids by the pancreatic and intestinal peptidases. The experimental evidence for the formation of peptides on partial enzymatic hydrolysis of proteins is largely based on analytical data in which the number of peptide bonds broken by the action of a proteinase is compared with the number of such bonds cleaved by acid hydrolysis. Although such results can give little information concerning the particular peptide linkages split by a given proteinase, they may be used to calculate the average chain length of the fragments formed by enzymatic action. Data of this kind have been presented in the literature many times, and an example is offered by the recent publication of Beloff and Anfinsen (1948) in which it was shown that the action of crystalline pepsin on egg albumin leads to the formation of fragments that contain, on the average, seven amino acid residues. In view of the complexity of the substrate employed, and the complex nature of the enzymatic action, the limits between which the size of these fragments may vary is difficult to ascertain. As a possible lower limit, it is necessary to consider the free amino acids, since it is known that crystalline pepsin will hydrolyze certain peptide bonds at the ends of peptide chains (Fruton and Bergmann, 1939) and, in special cases, will even hydrolyze dipeptides (Harington and Pitt Rivers, 1944; Dekker *et al.*, 1949). The upper limit is even more uncertain; the data of Tiselius and Eriksson-Quensel (1939) indicate, however, that in the digestion of crystalline egg albumin by crystalline pepsin, no appreciable amount of split products of molecular weight greater than about 1100 and smaller than that of the intact protein, are formed. These authors have suggested, therefore, that the hydrolysis of the first sensitive linkages leads to extremely rapid enzymatic action at the other points in the protein molecule that are subject to enzymatic attack.

Of decisive importance in the study of the enzymatic degradation of proteins is the determination of the sites of enzymatic action and the chemical structure of the split products that are formed. Some information as to the possible loci of proteinase action has been provided by the finding of simple synthetic substrates for the protein-splitting enzymes crystallized by Northrop and his associates (*cf.* Northrop *et al.*, 1948). Since the results of these specificity studies have been summarized at earlier symposia in this series (Fruton, 1938, 1941), and elsewhere (Bergmann and Fruton, 1941), it may be sufficient to recall that each of the crystalline proteinases exhibits a striking selectivity of action with regard to the nature of the amino acid residues joined by a sensitive peptide bond. Thus, trypsin was found to be specifically adapted to the hydrolysis of peptide bonds involving the carbonyl group of an arginine or lysine residue (Bergmann *et al.*, 1939; Hofmann and Bergmann, 1939), while chymotrypsin specifically hydrolyzed amide bonds involving the carbonyl group of a tyrosine, phenylalanine, or tryptophan residue (Bergmann and Fruton, 1937; Kaufman and Neurath, 1949). Several of the synthetic substrates hydrolyzed by well-defined proteinases are also split by crude enzyme preparations from animal tissues, such as spleen (Fruton *et al.*, 1941) or kidney (Gutmann and Fruton, 1948), and it has been concluded, therefore, that such tissues contain protein-splitting enzymes with specificities similar to those of trypsin, chymotrypsin, and pepsin. It may be hoped that the use of simple substrates will facilitate the difficult task of purifying these tissue proteinases, or as they are frequently termed, the cathepsins. Similar considerations apply to the enzymatically heterogeneous preparations of the plant proteinases such as papain and ficin (*cf.* Dekker *et al.*, 1949b).

While the studies of the specificity of proteinases by means of synthetic peptides and peptide derivatives have provided valuable clues as to the points of enzymatic attack on proteins, it is clear that, in themselves, such studies give only limited information as to the nature of the peptides that are formed. Such knowledge must come primarily from the isolation and characterization of peptides following the action of a single pure proteinase on a single pure protein. Assuming that adequate criteria of purity can be met for both enzyme and substrate, there still remains the formidable task of separating the resulting complex mixture of peptides. The recent developments in chromatography, and in the use of partition methods, encourage the hope that a solution of this problem may be forthcoming. At present, however, the application of these beautiful techniques to the separation of peptide mixtures of unknown composition appears to require more information as to the behavior of synthetic peptides of known structure in such fractionation procedures.

It may also be suggested that, pending the further development of new methods for the separation of peptides, the study of the products formed on enzymatic degradation of proteins would profit greatly from the extension of our knowledge of the action of proteinases on synthetic substrates. Although the methods of peptide synthesis are sufficiently broad in scope to permit the synthesis of nearly any desired peptide or peptide derivative (Fruton, 1949), there still remain many new peptides to be made and to be tested as possible substrates for proteinases. For example, the recent synthesis of an extensive series of peptides of methionine has brought to light the fact that, of the available synthetic substrates of ficin, those containing an L-methionine residue were most sensitive to the action of this enzyme (Dekker *et al.*, 1949b). Similarly, studies currently in progress on the synthesis of new peptides of L-serine may bring forth new synthetic substrates for known proteinases, and thus indicate new peptide linkages as possible points of enzymatic attack in proteins. It will be especially important to examine the behavior of serine derivatives in which the β-hydroxyl group of this amino acid is linked to the carboxyl group of another amino acid. Many years ago, Emil Fischer (1906) suggested the possibility that such ester linkages might be involved in the union of amino acids in proteins (*cf.* also Chibnall, 1942), and the important discovery recently made by Neurath and his associates that trypsin and chymotrypsin can act at ester linkages (Schwert *et al.*, 1948; Kaufman *et al.*, 1948) emphasizes the need for the renewed study of Fischer's suggestion by means of suitable synthetic models. An initial effort in this direction has recently been made by Kaufman and Neurath (1949) who have reported that the mixture of the diastereoisomeric O-(benzoyl-DL-phenylalanyl)-N-benzoyl-DL-serine ethyl esters is resistant to the action of chymotrypsin.

In searching for the possible points of cleavage of a protein by proteinases, it is well to bear in mind the suggestion that, in addition to the cleavage of peptide bonds, there may occur transamidation reactions which lead to peptides having an amino acid sequence different from that originally present in the intact protein molecule. The occurrence of such reactions has been demonstrated by means of synthetic models in the case of papain (Bergmann and Fraenkel-Conrat, 1938), but the generality of this rearrangement under the influence of enzymes still remains to be investigated. The question may justifiably also be raised as to the possible occurrence of such transpeptidation reactions in the partial hydrolysis of proteins by means of acids (*cf.* Synge, 1943).

It is clear, therefore, that despite the extensive work already done on the conversion of proteins to peptides by proteinases, the most important questions as to the nature of the split products still remain to be answered, and it may be hoped that the further development of techniques for the separation of peptides and of studies with synthetic substrates will contribute to this end.

It is generally agreed that a significant portion of the proteins broken down in the gastrointestinal tract of higher animals is absorbed in the form of peptides (London and Kotchneva, 1935), and that these peptides enter the nitrogen metabolism. Little information is available, however, as to the fate of these peptides, and the dearth of knowledge concerning the metabolism of peptides in higher animals applies to other biological systems as well. Of the several possibilities which may be envisaged, four seem especially worthy of consideration: 1) The peptides may be rapidly hydrolyzed by the peptidases of the blood and tissues to form the component amino acids which mix with the amino acids already present in the organism; 2) The peptides may in part be incorporated into tissue proteins, without prior enzymatic hydrolysis; 3) The energy of the peptide bonds of certain peptides may be used for the synthesis of new peptide bonds by transpeptidation reactions; 4) The peptides may undergo transformation, without prior scission of peptide bonds, either at α-amino or α-carboxyl groups, or at reactive groups in the side chains. It is fair to say that experimental evidence is available only for the first of these possibilities, since the ubiquitous presence of peptidases of varied specificity provides the enzymatic apparatus for the hydrolysis of peptides to amino acids.

In the face of the fragmentary knowledge concerning the metabolism of peptides in any biological system, it has seemed desirable to embark on a systematic investigation of this question using microorganisms, since these permit the rapid accumulation of data on a large number of peptides. The initial objective of these studies was to examine the manner in which microorganisms which exhibit characteristic growth requirements for particular amino acids would respond to a series of peptides containing such amino acids. The number and variety of organisms that could be used for such experiments are very great; for our purpose, however, the artificially produced amino acid requiring mutants of *Escherichia coli* appeared to offer unusually desirable features in the simplicity of the medium required for growth and also in the fact that the mutation provided a valuable marker for the stability of the bacterial strain. It was most fortunate for these studies that they began with the active collaboration of Dr. E. L. Tatum, who first developed the mutants of *E. coli* by irradiation with X-rays or by treatment with N-methyl, bis (β-chloroethyl)amine (Tatum, 1945). One of the mutants obtained from the wild-type strain K-12 exhibits a characteristic nutritional requirement for biotin; further treatment of this strain (No. 58) yields a series of mutants which also require particular amino acids for growth. One of these double mu-

tants is a *phenylalanineless* strain (No. 58-278); another is a *tyrosineless* strain (No. 58-5030). Similarly, double mutants showing a requirement for proline and threonine (No. 679-183) and for leucine and threonine (No. 679-680) were developed. In what follows, the last two mutants will be referred to as the *prolineless* and *leucineless* strains respectively, and the minimal media used for studies with these two mutants were routinely supplemented with sufficient threonine to support growth. The amino acid requirement of each of these four strains is completely specific; thus, the *phenylalanineless* strain grows in the presence of L-phenylalanine, but when phenylalanine is absent, it does not grow in the presence of L-tyrosine.

In testing the growth-promoting activity of peptides and other derivatives of an essential amino acid, the appropriate organism was inoculated into 10 cc. of minimal medium containing graded amounts of the test compound. The extent of bacterial growth in each tube after a given time (24-48 hours) was determined by turbidimetric measurement, and the data so obtained were used to draw a curve which relates the extent of growth (expressed as a density reading) to the amount of test substance added to the minimal medium. From the curve an estimate was made of the amount of substance required for the attainment of half-maximal growth, and the relative growth-promoting activity of the various test compounds is expressed in terms of micromoles, per 10 cc., required to produce half-maximal growth in a given time.

One may consider first some of the results of experiments in which the growth of the *phenylalanineless* mutant was followed in the presence of peptides of L-phenylalanine, instead of the parent amino acid (Simmonds *et al.*, 1947a). It will be seen from Table 1 that, on a molar basis, the peptides glycyl-L-phenylalanine and L-phenylalanylglycine are approximately as effective growth promoters as is L-phenylalanine. L-glutamyl-L-phenylalanine is less effective, and substitution of the amino group with a carbobenzoxy residue or the conversion of the carboxyl group to an amide further reduces the growth-promoting activity. The data in Table 1 also indicate that DL-phenylalanine is exactly one-half as active as the L-form, and it would appear, therefore, that at these concentrations the D-form is not utilized and does not inhibit bacterial growth. This conclusion as to optical specificity has been found to apply to all the mutants studied thus far, with the exception of the *methionineless* mutant (Simmonds, 1948).

In the right-hand half of Table 1 are given an analogous set of data for the utilization of tyrosine derivatives by the *tyrosineless* mutant of *Escherichia coli*. It will be seen that the comparable substitution of the two aromatic amino acids results in a very similar effect on the growth-promoting activity toward the appropriate mutant.

These data permit the formulation of a simple working hypothesis to explain the differences in the relative effectiveness of various peptides or peptide derivatives containing phenylalanine and tyrosine residues. It is important to note that none of the compounds tested is more effective than the parent amino acid, and this suggests that these derivatives must be split by the bacterial enzymes to yield the free amino acid, which then serves as the growth factor. If this is so, then the variation in the growth-promoting activity of the individual phenylalanine or tyrosine derivatives may be interpreted as a reflection of the rate at which they are subjected to enzymatic hydrolysis. The fact that the relative efficacy with which a given derivative replaces its parent amino acid as a growth factor is rather similar for the phenylalanine and tyrosine series, is consonant with our knowledge of the specificity of proteolytic enzymes. In all cases studied thus far where the specificity of an enzyme is directed to the hydrolysis of a peptide bond involving one of the two aromatic amino acids, it also includes an action on peptide bonds involving the other amino acid (Bergmann and Fruton, 1941).

Support for the view that the relative efficacy of the various phenylalanine derivatives is a reflection

TABLE 1. EFFECT OF PHENYLALANINE AND TYROSINE DERIVATIVES ON GROWTH OF MUTANT STRAINS OF *Escherichia coli*

Phenylalanineless strain		*Tyrosineless* strain	
Compound	Quantity (μM) per 10 ml. for half maximal growth in 24 hours	Compound	Quantity (μM) per 10 ml. for half maximal growth in 24 hours
L-Phenylalanine	0.1	L-Tyrosine	0.09
DL-Phenylalanine	0.2		
Glycyl-L-phenylalanine	0.1	Glycyl-L-tyrosine	0.13
L-Glutamyl-L-phenylalanine	0.3	L-Glutamyl-L-tyrosine	0.19
L-Phenylalanylglycine	0.15	L-Tyrosylglycine	0.13
L-Phenylalaninamide	3.3	L-Tyrosinamide	3.1
Glycyl-L-phenylalaninamide	95	Glycyl-L-tyrosin amide	64
Carbobenzoxyglycyl-L-phenylalanine	401	Carbobenzoxyglycyl-L-tyrosine	235

of the relative rates of their hydrolysis by the bacterial enzymes has been provided by unpublished work of our colleague, Dr. J. E. Ziegler, who has found that crude saline extracts of the *phenylalanineless* mutant hydrolyzed glycyl-L-phenylalanine and L-phenylalanylglycine most rapidly, while L-phenylalaninamide was split rather slowly. Of the phenylalanine derivatives that were found to be split by the bacterial extracts, glycyl-L-phenylalaninamide and carbobenzoxyglycyl-L-phenylalanine were hydrolyzed at the slowest rate. Since, as is known from the work of Berger *et al.* (1938) and of Maschmann (1939), several of the bacterial peptidases depend for their maximal activity on the presence of specific metal ions, this qualitative correlation between the growth-promoting activity of the phenylalanine derivatives and their hydrolysis by bacterial extracts can be offered only as a first approximation.

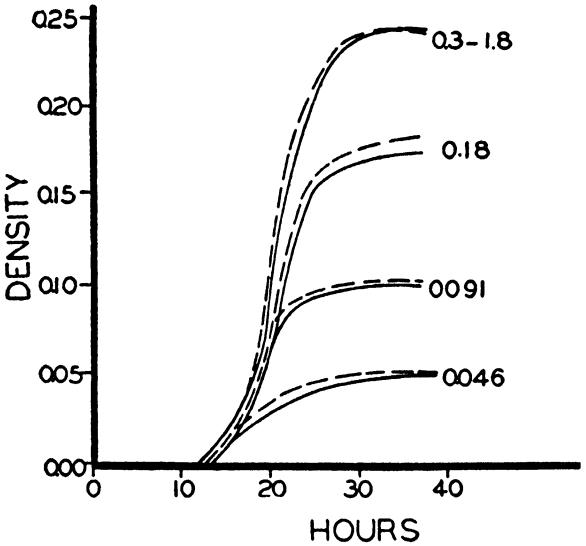


FIG. 1. Representative growth curves for the *phenylalanineless* mutant in the presence of L-phenylalanine (solid lines) and of glycyl-L-phenylalanine (dash lines). The figures to the right of the curves denote the concentrations of the test compounds in μM per 10 ml.

It was shown in Table 1 that glycyl-L-phenylalanine is as effective in promoting the half-maximal growth of the *phenylalanineless* mutant as is L-phenylalanine itself. Since these data were obtained by measurement of the density of the bacterial culture after incubation for 24 hours, they gave no information as to any possible differences in the shape of the growth curve when the peptide was used in place of the amino acid. It seemed of importance to determine whether the duration of the lag phase and the slope of the logarithmic phase of the growth curve was unchanged. That this was indeed the case is shown in Figure 1, and it will be seen that for each concentration of amino acid and peptide, the two growth curves are identical in all respects (Simmonds and Fruton, 1949a). It follows, therefore, that if the utilization of the dipeptide for growth depends upon its prior enzymatic hydrolysis, this splitting must proceed at an extremely rapid rate and cannot be the limiting factor in the duration of the lag phase or on the rate of logarithmic growth of the mutant.

It would appear, therefore, that the results with the *phenylalanineless* and *tyrosineless* mutants support the first of the several possibilities raised earlier with regard to the conceivable metabolic pathways of peptides, namely, that they are first hydrolyzed to the component amino acids. In extending our studies to other mutants of *Escherichia coli*, however, it was found that the *prolineless* strain gave more growth in the presence of proline peptides than in the presence of the free amino acid (Simmonds and Fruton, 1948). For example, the peptide L-prolyl-L-glutamic acid was about three times as effective as proline in promoting half-maximal growth of the mutant in 48 hours, while an equimolar mixture of proline and glutamic acid produced a growth-promoting effect similar to that of proline alone. In fact, all peptides of L-proline that were tested with the *prolineless* mutant exhibited this enhanced activity. One of these peptides, glycyl-L-proline, which was about twice as effective as proline, was used for a more detailed study of the bacterial response. As will be seen from Figure 2, at low concentrations of the test substances, the duration of the lag phase and the slope of the logarithmic portion of the growth curves are the same with the peptide as with the amino acid; the only difference which can be noted is that the maximal density of the culture obtained with any given concentration of glycyl-L-proline is about twice that found with an equimolar concentration of L-proline (Simmonds and Fruton, 1949a).

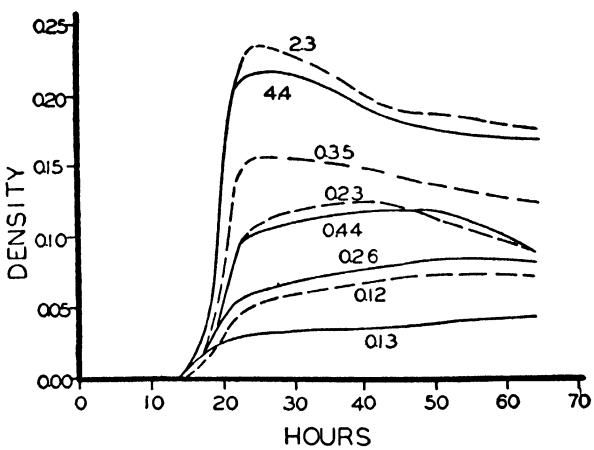


FIG. 2. Representative growth curves for the *prolineless* mutant in the presence of L-proline (solid lines) and of glycyl-L-proline (dash lines). The concentration of test compound, in μM per 10 ml., is indicated for each curve.

In the case of the prolineless mutant, therefore, it is necessary to assume a more complex metabolic pathway for the utilization of peptides than prior enzymatic hydrolysis. One possibility which must be considered is that a part of the proline supplied to the organism as the free amino acid is converted, in the course of bacterial growth, to products which are not utilized for growth, while the proline supplied in the form of peptides is less readily converted to such products. This explanation finds an analogy in the report of Gale (1945) who showed that arginine peptides formed by partial acid hydrolysis of salmine exhibit a greater growth-promoting activity for Group D streptococci than does arginine itself, presumably because of the extensive enzymatic degradation of the free amino acid by "arginine dihydrolase."

The question must be left open, however, as to whether any of the other possibilities as to the metabolic fate of peptides applies in the case of the *prolineless* mutant. The view that the proline peptides may be incorporated into the bacterial proteins as such is made less attractive by the fact that all the proline peptides tested were more effective than was proline itself. *A priori,* it would appear unlikely that all combinations involving proline would have an arrangement of amino acids that might be needed for the synthesis of bacterial proteins. For this reason, one must also consider the possibility that the presence of proline in a peptide linkage may make the amino acid more readily available for protein synthesis, perhaps by means of as yet unknown enzymatic systems which cause transpeptidation reactions. It is to be hoped that further work with the *prolineless* mutant, using isotopic peptides, will furnish more decisive evidence concerning the enhanced growth-promoting activity of the proline peptides.

Pending the elucidation of the behavior of the *prolineless* mutant, it will be important to seek out other biological systems in which peptides appear to be used more effectively for protein synthesis than are the component amino acids. For this reason, we have welcomed the recent chance discovery of a bacterial strain, tentatively classified as a member of the genus Alcaligenes, which grows well in a solution of L-leucylglycine in 0.9 percent NaCl, but does not grow in the presence of a mixture of L-leucine and glycine, and grows only poorly in the presence of L-leucine (Simmonds and Fruton, 1949b). Since this organism grows in a medium in which the peptide provides the sole source of carbon and nitrogen for the synthesis of the bacterial protein, it appears to be suitable for studies, involving isotopic techniques, of the possible metabolic pathways of peptides. It does not seem unlikely that a systematic search for peptide-requiring microorganisms will yield a rich harvest, for there are many indications that numerous bacterial strains grow better in partial hydrolysates of proteins than in the amino acid mixtures formed by complete hydrolysis with acid. In the case of *Lactobacillus casei,* the enhanced growth-promoting activity of such partial hydrolysates has been attributed to a peptide-like growth factor, termed "strepogenin," and it has been found that L-seryl-glycyl-L-glutamic acid exerts an effect similar to that of the growth factor (Woolley, 1946; Krehl and Fruton, 1948).

It is appropriate at this point to recall that recent research in the field of antibacterial agents has shown that peptides may, in some cases, be powerful inhibitors of bacterial growth. Chemical studies of these peptide antibiotics, such as gramicidin and tyrocidine (Hotchkiss, 1944) and the polymixins (Jones, 1948), have shown them to yield, on hydrolysis, non-protein amino acids such as ornithine or α, γ-diaminobutyric acid, as well as the D-isomers of some protein amino acids. It is attractive to consider the possibility that these peptide antibiotics may exert their inhibitory action on bacterial growth by interference with the enzymatic mechanisms involved in the normal protein metabolism of sensitive organisms (Work, 1948). It would appear that further light on this question could be gleaned from studies in which synthetic peptides of known structure were tested as possible inhibitors of bacterial growth. To our knowledge efforts to synthesize simple peptides or peptide derivatives related to the natural antibiotics, and which have the bacteriostatic or bactericidal properties of the more complex molecules, have been unsuccessful (*cf.* Work, 1948; Fruton, 1948). It is hoped, however, that these will be continued, since, if successful, they may provide extremely valuable reagents for the study of protein metabolism.

In connection with this question of the inhibition of bacterial growth by peptides, it may be of interest to consider some results obtained with the *leucineless* mutant of *Escherichia coli.* All the L-leucine peptides which were tested served as growth factors for the mutant in place of L-leucine, and, on a molar basis, glycyl-L-leucine, L-leucylglycine, and L-leucylglycylglycine were approximately as effective in promoting half-maximal growth as was the parent amino acid (Simmonds *et al.,* 1947b). Other peptides and peptide derivatives, if active, were in all cases less effective than L-leucine. The results with the *leucineless* mutant thus resemble those obtained with the *phenylalanineless* and *tyrosineless* strains and suggest that here also there is involved enzymatic hydrolysis of the peptide prior to its utilization for bacterial growth.

An important difference in the case of the *leucineless* mutant, however, emerged when the concentration of the dipeptides glycyl-L-leucine and L-leucylglycine were raised to values much greater than those required for maximal growth. As will be seen from Figure 3, at these high concentrations of dipeptides, there was a marked inhibition in the extent of growth observed in 24 hours. Control

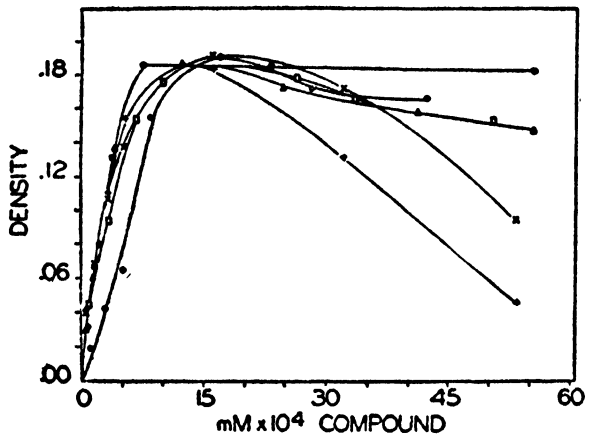


FIG. 3. The effect of leucine peptides on the growth of the *leucineless* mutant. ● L-leucine; ,+ glycyl-L-leucine; × L-leucylglycine; △ L-leucylglycylglycine; □ diglycyl-L-leucylglycine; ○ triglycyl-L-leucylglycine.

experiments showed that this inhibitory effect was a property of the dipeptides *per se,* since acid hydrolysis abolished it completely, and the growth attained was that to be expected on the basis of the L-leucine present in the hydrolysate.

To study this inhibition more closely, the relative growth-promoting action of L-leucine and of glycyl-L-leucine was studied as a function of time. It will be seen from Figure 4 that, in contrast to the results with the *phenylalanineless* and *prolineless* mutants, the duration of the lag phase of the *leucineless* strain was greater in the presence of the peptide than with the amino acid, and the initiation of rapid growth in the presence of increasing concentrations of the peptide required increasingly longer

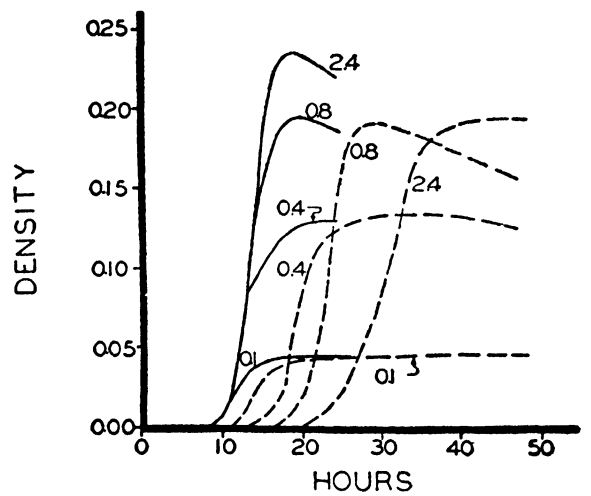


FIG. 4. Representative growth curves for the *leucineless* mutant in the presence of L-leucine (solid lines) and of glycyl-L-leucine (dash lines). The concentration of test compound, in μM per 10 ml., is indicated for each curve.

time intervals (Simmonds and Fruton, 1949a). Thus, if the activities of L-leucine and of glycyl-L-leucine were compared only after 24 hours, the higher concentration of the dipeptide would appear to exert an inhibitory effect on growth. However, the extent of bacterial growth ultimately attained in media containing higher concentrations of glycyl-L-leucine was essentially the same as that observed in the presence of corresponding amounts of L-leucine. In these experiments, therefore, glycyl-L-leucine exerted a significant inhibitory effect only upon the initiation of rapid growth.

The possibility had to be considered that this prolongation of the lag phase may have been due to the necessity for the development of adaptive enzymes essential for the conversion of the dipeptide to a compound more readily utilized for growth. This was rendered unlikely by the fact that no change in the relative growth responses to glycyl-L-leucine and L-leucine was observed when inocula were taken from cultures previously grown in the presence of the dipeptide. As was to be expected, the duration of the lag phase depended on the age of the culture used for inoculation; actively multiplying cultures exhibited a shorter lag period than did "older" cultures, and this effect was especially striking with respect to glycyl-L-leucine. At high concentrations of the test substances, however, the difference in the duration of the lag period was still marked.

A clue to the mechanism of the inhibitory action of the dipeptide was provided by experiments in which the *leucineless* mutant was allowed to grow in media containing a constant amount of L-leucine and varying amounts of glycyl-L-leucine. As will be seen from Figure 5, for each dipeptide concentration, the duration of the lag phase was the same as that previously found in the absence of L-leucine. The maximum of each curve, however, depended upon the total concentration of L-leucine, both as such and in the form of the dipeptide. In another set of experiments in which media containing a constant amount of dipeptide (0.4 micromoles) and increasing amounts of L-leucine were used, the duration of the lag phase was the same as that noted in the presence of 0.4 micromoles of glycyl-L-leucine without added amino acid. Here again the position of the maxima was determined by the total concentration of L-leucine, as amino acid and peptide.

These results suggest that glycyl-L-leucine inhibits one or more of the enzymatic systems involved in the utilization of leucine for growth and that rapid growth cannot ensue until the concentration of the dipeptide has been reduced, possibly by enzymatic hydrolysis, to a level that is no longer inhibitory. If this explanation is correct, the above results offer an example of a simple peptide which interferes with the protein metabolism of a biological system and whose inhibitory action can be overcome by enzymatic cleavage. It is tempting to

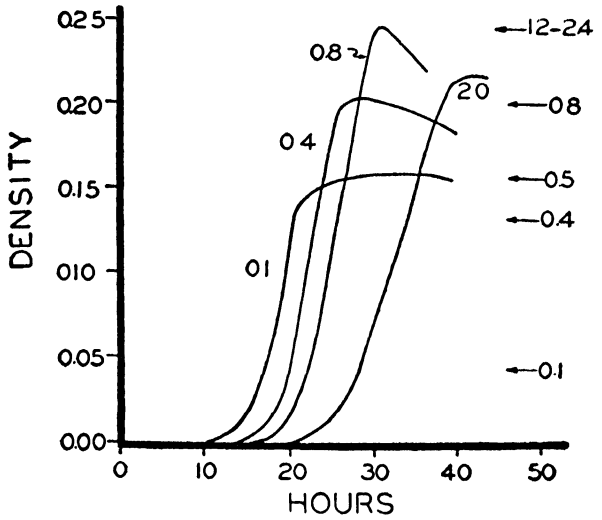


FIG. 5. Representative growth curves for the *leucineless* mutant in media containing 0.4 μM of L-leucine and varying amounts of glycyl-L-leucine. The concentration of the dipeptide, in μM per 10 ml., is indicated for each curve. The arrows at the right of the figure denote the maximum of growth curves obtained with media containing the indicated concentrations of L-leucine in the absence of the dipeptide.

speculate that the bacteriostatic action of peptides such as gramicidin and of glycyl-L-leucine may have something in common, but that the differences in their effects lie in the fact that the former is resistant to hydrolysis by the enzymes of sensitive organisms, while the latter is readily hydrolyzed to the component amino acids. At the present time, this idea can be offered only as a working hypothesis, whose correctness will be tested by further studies.

Thus far, consideration has been given to some aspects of the problem of the formation of peptides from proteins, and the possible metabolic pathways which peptides might follow when they enter a biological system. There remains still another important question for discussion, namely, the possible mechanisms for the enzymatic synthesis of peptides *in vivo*. It appears very likely that, in the case of naturally occurring peptides such as glutathione or carnosine, they are built up from smaller units rather than formed by the degradation of proteins. This would seem to be the situation also in the biosynthesis of the tri-L-glutamine apparently present in marine algae such as *Pelvetia fastigiata* (Dekker *et al.*, 1949a). There have been many discussions as to the possible biochemical reactions involved in peptide bond synthesis (*cf.* Bergmann and Fruton, 1944). One view is that the biological formation of peptide bonds involves the synthetic capacity of the proteolytic enzymes, which are still the only catalysts known to have the requisite specificity for this purpose. The experimental evidence in favor of this suggestion has been presented at an earlier symposium in this series (Fruton, 1941) and is based largely on work with model reactions in which peptide synthesis catalyzed by proteolytic enzymes was effected by operating in heterogeneous systems. Although there has been much discussion of this idea during recent years, there has been no further experimental evidence either for or against it. It may be pointed out, however, that the view that the proteolytic enzymes are involved in biological peptide bond synthesis does not necessarily require the assumption that they act alone in such reactions without a source of energy to make an endergonic reaction proceed to an appreciable extent (*cf.* Greenberg *et al.*, 1948). While it is reasonable to assume that such energy is transferred from the oxidation of metabolites by means of specific "energy carriers," our ignorance is still quite complete as to the possible chemical nature of such carriers. The recent work of Lipmann (1945) and of Cohen and McGilvery (1947) has suggested that prior phosphorylation may be involved in peptide synthesis (*cf.* also Chantrenne, 1948; Winnick and Scott, 1947; Sciarini and Fruton, 1949).

In addition to the general suggestions that biological peptide synthesis may involve the synthetic action of the proteolytic enzymes, or the intervention of phosphorylated intermediates, a number of hypotheses have been offered based largely on reactions discovered in the organic chemical laboratory. For example, it was shown by Bergmann and Grafe (1930) that acid amides can react with keto acids to form substituted dehydroamino acids. If an amino acid amide would react in this manner with a keto acid, there would be formed a peptide containing an α, β-unsaturated amino acid residue. Such substances are termed dehydropeptides, and they can readily be converted to saturated peptides by catalytic hydrogenation. These facts would remain solely a matter of interest to the organic chemist were it not for the fact that many cells contain enzymes specifically adapted to the hydrolysis of dehydropeptides and hence named dehydropeptidases (Greenstein, 1948). It has seemed of interest, therefore, to test the possibility that dehydropeptides might be intermediates in the biosynthesis of peptides, and for this purpose the mutant strains of *Escherichia coli* appeared especially suitable. It was noted above that in the case of the *phenylalanineless* mutant, glycyl-L-phenylalanine was utilized for growth as effectively as L-phenylalanine. If the enzymatic capacities of this organism permitted the reduction of the double bond of the dehydropeptide glycyldehydrophenylalanine, this substance should also have served as a growth factor in the absence of L-phenylalanine. It was found, however, that the dehydropeptide did not support the growth of the mutant, nor did it spare the requirement for phenylalanine (Simmonds *et al.*, 1947a). Similar negative results were obtained with derivatives of dehydrotyrosine and of dehydroleucine, when these were

tested with the *tyrosineless* and *leucineless* mutants respectively.

It would appear, therefore, that the dehydropeptides do not serve as intermediates in the biosynthesis of peptides by *Escherichia coli*, but this by no means excludes the possibility that derivatives of dehydroamino acids may be intermediates in other metabolic reactions involved in the biological transformation of peptides. Indeed, recent experiments with *E. coli* support the conclusion that the growing organism rapidly converts acetyldehydrotyrosine to a metabolic product which may be a derivative of 6-hydroxyindole (Fruton *et al.*, 1947). Under the same experimental conditions acetyl-DL-tyrosine was not converted to this product by *E. coli*, thus excluding the direct hydrogenation of the α, β-double bond in the metabolic conversion of acetyldehydrotyrosine.

The fact that glycyldehydrophenylalanine did not promote the growth of the *phenylalanineless* mutant also indicates that, if dehydropeptidase is present in the organism, it did not hydrolyze the peptide sufficiently rapidly to form phenylpyruvic acid, which can replace L-phenylalanine in the medium of this organism (Simmonds *et al.*, 1947a).

Additional evidence against the assumption that derivatives of dehydroamino acids are to be considered as intermediates in peptide synthesis *in vivo* is offered by experiments with rats which received injections of acetyldehydrotyrosine labeled with N^{15}. It was found, in collaboration with Dr. Henry D. Hoberman, that negligibly small amounts of N^{15} appeared in the tyrosine isolated from the liver proteins of the animal, and analysis of the urinary urea and ammonia showed the presence of only small amounts of the isotope. In view of the demonstration (Bloch and Rittenberg, 1947) that acetylamino acids are readily metabolized by the rat, this result argues against the enzymatic reduction of derivatives of dehydroamino acids in this animal. This conclusion is in agreement with the recent data of Sprinson and Rittenberg (1948) who found that the metabolism of leucine labeled with deuterium in the α- and β-carbon atoms, and with N^{15} in the amino group, leads to a more extensive exchange of the hydrogen at the α-carbon than at the β-carbon.

The mutants of *Escherichia coli* were used to test still another possible mechanism for the biosynthesis of peptides. It has been suggested that ketoacylamino acids (*e.g.*, phenylpyruvylglycine) might be converted to a dipeptide by transamination with glutamic acid (Linderstrøm-Lang, 1939). Since such postulated intermediates were available (Fruton and Bergmann, 1946), it was possible to test for their possible conversion to peptides which had been found to promote the growth of one of the *Escherichia coli* mutants. However, when phenylpyruvylglycine was offered to the *phenylalanineless* mutant in place of L-phenylalanine, it could not be utilized for growth (Simmonds *et al.*, 1947a). It would appear, therefore, that the amination of the ketoacylamino acid to yield phenylalanylglycine, which is an active growth factor, does not occur at a rate sufficient to promote bacterial growth.

The principal purpose of this review of the problem of peptide metabolism has been to emphasize the fragmentary nature of the available knowledge in this field. Although it is clear that the decisive discoveries still lie ahead, it may be hoped that continued efforts, by means of enzymatic, isotopic, and microbiological techniques, along the lines suggested above, will provide useful information, not only about the metabolism of peptides, but concerning some of the larger questions of protein metabolism as well.

References

Beloff, A., and Anfinsen, C. B., 1948, The products of proteolysis of some purified proteins. J. biol. Chem. *176:* 863-872.

Berger, J., Johnson, M. J., and Peterson, W. H., 1938, The proteolytic enzymes of bacteria. I. The peptidases of *Leuconostoc mesenteroides*. J. biol. Chem. *124:* 395-408.

Bergmann, M., and Fraenkel-Conrat, H., 1938, The enzymatic synthesis of peptide bonds. J. biol. Chem. *124:* 1-6.

Bergmann, M., and Fruton, J. S., 1937, On proteolytic enzymes XIII. Synthetic subtrates for chymotrypsin. J. biol. Chem. *118:* 405-415.

1941, The specificity of proteinases. Advances in Enzymology *1:* 63-98.

1944, The significance of coupled reactions for the enzymatic hydrolysis and synthesis of proteins. Ann. N.Y. Acad. Sc. *45:* 409-423.

Bergmann, M., Fruton, J. S., and Pollok, H., 1939, The specificity of trypsin. J. biol. Chem. *127:* 643-648

Bergmann, M., and Grafe, K., 1930, Synthese eines Peptids und anderer Abkömmlinge der α-Aminoacrylsäure (Dehydroalanin) aus Brenztraubensäure. Z. physiol. Chem. *187:* 187-195.

Bloch, K., and Rittenberg, D., 1947, The metabolism of acetylamino acids. J. biol. Chem. *169:* 467-476.

Chantrenne, H., 1948, Mixed anhydrides of benzoic and phosphoric acids. Compt. rend. trav. lab. Carlsberg *26:* 297-314.

Chibnall, A. C., 1942, Amino acid analysis and the structure of proteins. Proc. Roy. Soc. London *131B:* 136-160.

Cohen, P. P., and McGilvery, R. V., 1947, Peptide bond synthesis III. On the mechanism of p-aminohippuric acid synthesis. J. biol. Chem. *171:* 121-133.

Dekker, C. A., Stone, D., and Fruton, J. S., 1949a, A peptide from a marine alga. J. biol. Chem. *181:* 719-731.

Dekker, C. A., Taylor, S. P., Jr., and Fruton, J. S., 1949b, Synthesis of peptides of methionine and their cleavage by proteolytic enzymes. J. biol. Chem. in press.

Fischer, E., 1906, Untersuchungen über Aminosäuren, Polypeptide und Proteine. Berlin: Julius Springer.

Fruton, J. S., 1938, Protein structure and proteolytic enzymes. Cold Spr. Harb. Symposium Quant. Biol. *6:* 50-57.

1941, Proteolytic enzymes as specific agents in the formation and breakdown of proteins. Cold Spr. Harb. Symposium Quant. Biol. *9:* 211-217.

1948, A cyclic peptide derivative related to gramicidin. J. Amer. Chem. Soc. *70:* 1280.

1949, The synthesis of peptides. Advances in Protein Chemistry *5:* in press.

FRUTON, J. S., and BERGMANN, M., 1939, The specificity of pepsin. J. biol. Chem. *127:* 627-641.

1946, Phenylpyruvyl derivatives of amino acids. J. biol. Chem. *166:* 449-453.

FRUTON, J. S., IRVING, G. W., JR., and BERGMANN, M., 1941, On the proteolytic enzymes of animal tissues. III. J. biol. Chem. *141:* 763-774.

FRUTON, J. S., SIMMONDS, S., and SMITH, V. A., 1947, The action of *Escherichia* coli on acetyldehydroamino acids. J. biol. Chem. *169:* 267-276.

GALE, E. F., 1945, The arginine, ornithine and carbon dioxide requirements of streptococci (Lancefield Group D) and their relation to arginine dihydrolase activity. Brit. J. Exp. Path. *26:* 225-233.

GREENBERG, D. M., FRIEDBERG, F., SCHULMAN, M. P., and WINNICK, T., 1948, Studies on the mechanism of protein synthesis with radioactive carbon-labeled compounds. Cold Spr. Harb. Symposium Quant. Biol. *13:* 113-117.

GREENSTEIN, J. P., 1948, Dehydropeptidases. Advances in Enzymology *8:* 117-169.

GUTMANN, H. R., and FRUTON, J. S., 1948, On the proteolytic enzymes on animal tissues. VIII. An intracellular enzyme related to chymotrypsin. J. biol. Chem. *174:* 851-858.

HARINGTON, C. R., and PITTRIVERS, R. V., 1944, The synthesis of cysteine-(cystine)-tyrosine peptides and the action thereon of crystalline pepsin. Biochem. J. *38:* 417-423.

HOFMANN, K., and BERGMANN, M., 1939, The specificity of trypsin. II. J. biol. Chem. *130:* 81-86.

HOTCHKISS, R. D., 1944, Gramicidin, tyrocidin, and tyrothricin. Advances in Enzymology *4:* 153-199.

JONES, T. S. G., 1948, The chemical basis for the classification of the polymixins. Biochem. J. *43:* xxvi-xxvii.

KAUFMAN, S., and NEURATH, H., 1949, Structural requirements of specific substrates for chymotrypsin II. Arch. Biochem. *21:* 437-453.

KAUFMAN, S., SCHWERT, G. W., and NEURATH, H., 1948, The specific peptidase and esterase activities of chymotrypsin. Arch. Biochem. *17:* 203-205.

KREHL, W. A., and FRUTON, J. S., 1948, The utilization of peptides by lactic acid bacteria. J. biol. Chem. *173:* 479-485.

LINDERSTRØM-LANG, K., 1939, Proteolytic enzymes. Ann Rev. Biochem. *8:* 37-58.

LIPMANN, F., 1945, Acetylation of sulfanilamide by liver homogenates and extracts. J. biol. Chem. *160:* 173-190.

LONDON, E. S., and KOTCHNEVA, N. P., 1935, The form of absorption of protein from the alimentary tract and the further fate of the products of absorption. Arch. Sci. biol. (U.S.S.R.) *37:* 3-25. (*cf.* also Chem. Abstr., 1936, *30:* 7183).

MASCHMANN, E., 1939, Uber Bakterienproteasen. XII. Zur Kenntnis der Dipeptidasen anaerober Bakterien. Biochem. Z. *302:* 332-368.

NORTHROP, J. H., KUNITZ, M., and HERRIOTT, R. M., 1948, Crystalline enzymes. 2nd ed. New York: Columbia University Press.

SCHWERT, G. W., NEURATH, H., KAUFMAN, S., and SNOKE, J. E., 1948, The specific esterase activity of trypsin. J. biol. Chem. *172:* 221-239.

SCIARINI, L. J., and FRUTON, J. S., 1949, Derivatives of N-phosphorylated amino acids. J. Amer. Chem. Soc. *71:* 2940.

SIMMONDS, S., 1948, Utilization of sulfur-containing amino acids by mutant strains of *Escherichia coli.* J. biol. Chem. *174:* 717-722.

SIMMONDS, S., and FRUTON, J. S., 1948, The utilization of proline derivatives by mutant strains of *Escherichia coli.* J. biol. Chem. *174:* 705-715.

1949a, The utilization of amino acids and peptides by mutant strains of *Escherichia coli.* J. biol. Chem. *180:* 635-646.

1949b, A microorganism exhibiting a growth requirement for peptides. Science *109:* 561-562.

SIMMONDS, S., TATUM, E. L., and FRUTON, J. S., 1947a, The utilization of phenylalanine and tyrosine derivatives by mutant strains of *Escherichia coli.* J. biol. Chem. *169:* 91-101.

1947b, The utilization of leucine derivatives by a mutant strain of *Escherichia coli.* J. biol. Chem. *170:* 483-489.

SPRINSON, D., and RITTENBERG, D., 1948 The metabolism of the α-, γ-, and β-hydrogen atoms of L-leucine. Federation Proc. *7:* 191.

SYNGE, R. L. M., 1943, Partial hydrolysis products derived from proteins and their significance for protein structure. Chem. Rev. *32:* 135-172.

TATUM, E. L., 1946, Induced biochemical mutations in bacteria. Cold Spr. Harb. Symposium Quant. Biol. *11:* 278-284.

TISELIUS, A., and ERIKSSON-QUENSEL, I., 1939, Some observations on peptic digestion of egg albumin. Biochem. J. *33:* 1752-1756.

WINNICK, T., and SCOTT, E. M., 1947, Phosphorylated amino acids. Arch. Biochem. *12:* 201-208.

WOOLLEY, D. W., 1946, Strepogenin activity of seryl glycyl glutamic acid. J. biol. Chem. *166:* 783-784.

WORK, T. S., 1948, D- and L-amino acids in antibiotics. Biochemical Society Symposia *1:* 61-77.

DISCUSSION

HOTCHKISS: It seems to me unlikely that the energy content of the peptide bond will provide a quantitatively significant boost in the metabolism of an amino acid. In experiments with Staphylococcus, which does not synthesize most amino acids, we find that when amino nitrogen is coupled into peptide form by washed bacteria, the amount of phosphorus converted into esterified form is correspondingly diminished by one mole of phosphate for every two moles of nitrogen. If now we may assume from the experiments of Ochoa with mammalian tissues, that oxidation may result in esterification of three phosphorus atoms per atom of oxygen, it can be supposed that in bacteria also, the maximal oxidation of one molecule of glucose would account for 12 oxygen atoms used, or 36 phosphorus atoms esterified, or 72 amino groups converted into peptide linkage. Accordingly, there is some reason for thinking that the energy content of preformed peptide bonds would be only a very small part of the available energy during most growth conditions. This conclusion depends upon the correlation of data from rather diverse experimental conditions, but it should be noted that no estimations of free energy levels or changes in level are involved.

On the other hand, work with the Staphylococcus suggests that a kind of "propinquity-factor" operates when several different amino acids are provided. The magnitude of amino acid assimilation is greatly increased when a greater variety of amino acids is present at the same time. Similar findings are recently reported for mammalian nutrition. It may be that peptides provide this advantage of propinquity for bacteria, and even better than simple mixtures do, for those organisms which can synthesize and degrade amino acids.

FRUTON: While it may be true that the "energy content of preformed peptide bonds" may be a "small part of the available energy during most growth conditions," there is no evidence available at present to indicate that it may not be important in the metabolic transformations of peptide bonds. That oxidation of metabolites, and the concomitant formation of phosphorylated intermediates, may be involved in peptide bond synthesis appears plausible by analogy with the results of studies of intermediary carbohydrate metabolism. It seems important to reiterate, however, that the data on this question thus far have been suggestive rather than conclusive. Until the various postulated phosphorylated intermediates in biological peptide formation are synthesized and tested in biological systems, no clear-cut decision as to their role appears possible.

Dr. Hotchkiss' suggestion that the enhanced growth-promoting activity of some peptides may be due to the cooperative action of the several amino acids of which they are composed may perhaps apply to organisms having complex nutritive requirements such as Staphylococcus. This explanation cannot be invoked, however, to explain our studies with the *prolineless* mutant of *Escherichia coli*. In every case where the growth-promoting activity of a proline peptide was compared with that of the mixture of the component amino acids, the same relationship was found as that noted between the peptide and proline alone. Furthermore, proline peptides such as L-prolyl-L-tyrosine and L-prolyl-L-leucine were utilized by the *tyrosineless* mutant and the *leucineless* mutant respectively, but to a lesser extent than the equimolar quantities of L-tyrosine or L-leucine. It must be concluded, therefore, that in the case of some organisms, peptides may be metabolized by pathways which do not involve prior hydrolysis to the component amino acids.

X-RAY ANALYSIS AND PROTEIN STRUCTURE

DOROTHY CROWFOOT HODGKIN
Laboratory of Chemical Crystallography, University Museum, Oxford, England

This appears to be a critical moment at which to try to write an account of the evidence from X-ray analysis on protein structure. Since the contribution that Fankuchen wrote for the Cold Spring Harbor Symposium in 1941, X-ray measurements have been extended to cover, not only more crystalline proteins of molecular weights varying from 10,000,000 to 12,000, but also certain groups of crystalline peptides and one or two more amino acids. We can now give a list of compounds examined by X-ray methods, increasing in molecular weight by multiples of about 3 or 4, from amino acids such as glycine and alanine to the bushy stunt virus. At every molecular weight step it would be useful to observe an added variety of crystal structures, and survey work on the different possible types is still of the greatest importance. Nevertheless, one can see that the concentration of effort of all the main groups of X-ray workers in the field has now shifted very markedly, away from further surveying, towards the detailed interpretation of the X-ray data already available. There are at least three series of calculations now in progress in different laboratories which might, in the very near future, provide definite evidence on the configuration of peptide chains in proteins. One is tempted to wait for the outcome of these calculations before writing more. But in case the solution of our problem is not so near as it sometimes seems to be, it may prove to be of great value to try now to look at the present position of protein X-ray analysis as a whole.

I have chosen to restrict my present account almost entirely to X-ray evidence on crystals of the different types of compound mentioned. But it must be realised that not all the crystals examined are equally good, considered simply as crystals. A crystalline material is essentially one that consists of identical groups of atoms repeating regularly, according to the symmetry operations present, at definite intervals in three dimensions. Both the size of the repeating unit, its symmetry and the regularity of the structure can be recognised by the observation of diffraction effects when X-rays are passed through the crystals. But in many cases to be discussed here, only very limited X-ray diffraction effects can be detected, and the crystals concerned are clearly very disordered structures. In such cases it ceases to be possible either to measure the repeating unit exactly, or to be certain that this has identical composition throughout the crystal. Very roughly, the observed spacing limit of the crystal planes that contribute to the X-ray reflections does provide an indication of the degree of regularity of the structure, though other factors also affect the spacing limit—size of molecule is one. For crystals of low molecular weight compounds, such as amino acids, the spacing limit may be 0.77 A or less, *i.e.*, smaller than interatomic distances; for protein crystals it is always larger than this.

In Table 1 I have listed certain X-ray data on crystals of a number of proteins, including the virus nucleo-proteins, of peptides and of amino acids. The compounds here recorded are a selection only of those which have been examined; they have been chosen, where possible, from those on which new measurements have been made recently. They illustrate a variety of points about protein crystal structure and behaviour on which I should like now to comment in more detail. For purposes of this discussion, I have divided them into four main groups between which the molecular weight changes by powers of 10. These include molecules of weight (1) above 100,000, (2) 10,000-100,000, (3) 1,000-10,000, (4) 0-1,000.

Proteins of molecular weight more than 100,000

The first three proteins in Table 1 all have molecular weights, variously estimated, of 1,000,000 or more. They also all have been observed in the electron microscope as well as by X-ray diffraction, and X-ray diffraction data exist both for the wet and dry crystals. We start therefore with the great advantage that we can see from the electron microscope pictures that the crystals are built up of roughly spherical particles; and the particles have dimensions which agree well with inter-molecular distances calculated from X-ray data (Table 2). But as we make the comparison, a number of complexities appear, of kinds which will recur throughout our survey.

First, the bushy stunt virus, the largest of all, crystallises in body centred cubic packing with two "molecules" of the order of magnitude shown in the electron microscope in the crystal unit cell. It ought not to be difficult to find the actual molecular weight of the particles present, but the data are, in fact, conflicting. A direct estimate, based on the lattice constant of the wet crystal and measured loss of water on drying, gave a value of 10,800,000, in good agreement with the value, 10,600,000, deduced from ultracentrifuge measurements by Neurath and Cooper (1940). But this mass and the observed particle diameter of 276 A would require a particle density of 1.7, altogether too high if the particles were actually spherical. It seems most likely that the particle mass is lower, of the order of 8,600,000 ($\pm$500,000) which corresponds to a particle density of about 1.38 derived from Pirie's measurement of

TABLE 1

Substance	State	a	b	c	β	ρ	Space group	Mol. wt. asymmetric unit-solvent	Mol. wt. in soln.
I. Bushy Stunt virus	wet	386				1.286	I	10,800,000	10,600,000
	dry	314				1.35		?12	
Turnip yellow mosaic virus	wet	706				1.33	F	3,500,000	3,500,000
	dry	528						c. ?12	
Turnip yellow mosaic virus	wet	725							
	dry	550							
Tobacco necrosis protein	wet	179*	176*	179*	α 83° β109° γ119°		PI		
	dry	157	154	147	α100° β110° γ120°	1.29		1,800,000	1,800,000
Ferritin	wet	131.5	131.5	186		1.45	$P2_12_12_1$		
	dry	109	109	154					460,000
Excelsin	dry	86		208.2		1.285	R3	305,800	294,000
								3	
II. Horse methaemoglobin	wet	109	63.2	54.4	111°	1.160	C2		66,700
	dry	102	51.4	47	130°	1.270		33,400	
Adult sheep methaemoglobin	wet	164	70	66	94.5°	1.225	C2	c. 68,000	c. 68,000
Foetal sheep methaemoglobin	wet	112	108	56			$B22_12$	c. 34,000	68,000
	dry	78	99	54					
pepsin	wet	67.9		292			$C6_12$		c. 36,000
	dry	60.5		268					
β lactoglobulin I	wet	69.29	70.42	156.47		1.144	$P2_12_12_1$	35,400×2	34,000
	dry	60.7	61.0	112.4		1.259		35,600×2	40,000
Horse metmyoglobin	wet	57.3	30.8	57.0	112°		$P2_1$	c. 17,000	17,000
	dry	51.5	28.0	37.0	98°				
Lysozyme I	dry	71.1		31.3		1.305	$P4_12_1$	13,900	14–17,000
Ribonuclease I	wet	30.85	38.72	53.83	106°		$P2_1$	c. 13,700	13–15,000
	dry	31.93	30.08	51.03	114°				
Insulin	wet	83.0		34.0		1.28	R3		
	dry	74.8		30.9		1.31		11,900	12–48,000
III. Gramicidin A	dry	26.1	31.8	26.1		1.15–1.18	$P2_12_12_1$	3,800	2,800–5,000
Gramicidin B	dry	25.7	13.4	62.8	101°	1.19	$P2_1$	7,500	
Gramicidin S hydrochloride	wet	26.15	35.30	37.90			C 222_1	c. 2×1,400	
	dry	24.10	34.65	18.95		1.217	C 222	2×1,374	1,248
Gramicidin S hydriodide	wet	28	41	39			I 222	c. 2×1,400	
	dry	23.2	41.2	38.8			$P2_12_12_1$	c. 4×1,400	1,396
N-acetylgramicidin S	wet	26.75		57.8			$C6_12$		
	dry	26.20		57.8		1.165	$C6_12$	3×665	1,236

TABLE 1.—*Continued*

Substance	State	*a*	*b*	*c*	*β*	*ρ*	Space group	Mol. wt. asymmetric unit-solvent	Mol. wt. in soln.
IV. glutathione	dry	8.74	5.57	28.4		1.46†	$P2_12_12_1$	305	305
β-glycylglycine	dry	17.89	4.62	17.06	125°10′	1.52†	A 2/a	132	132
diketopiperazine	dry	5.19	11.50	3.96	83°		P 2_1/a	57	114
DL-alanine	dry	12.04	6.04	5.81		1.40	P na	89	89
glycine	dry	5.10	11.96	5.45	111°38′	1.607	P 2_1/n	75	75

* Packing unit, not unit cell.
† Calculated density.

the partial specific volume. This fits well with data on the air dried crystal if all the space not occupied by spherical particles is calculated as filled with with water of crystallisation. But alternatively, the particle might be heavier if it had a shape more closely approximating to the space filling cubo-octahedron, which would appear roughly spherical in the electron microscope.

The early X-ray data obtained by Bernal and Fankuchen (1941) consisted of powder lines only, and very few of these, and so could give no evidence on the crystal symmetry. Lately, remarkable photographs of stationary wet single crystals have been obtained by Carlisle and Dornberger (1948) which give evidence of literally millions of X-ray reflections extending to spacings of 7.5 A; these still give no evidence of crystal symmetry. If the symmetry is genuinely cubic, two alternative conclusions are possible; the molecules may be oriented statistically in all directions to simulate spherical symmetry or they may consist of 12 n identical submolecules, each of weight, say, 750,000/n. The simplest kind of structure that could be devised to fit these facts would be one of 12 spheres of radius 45 A, arranged as in cubic close packing, but this structure would be too open to account for the particle density. It is more likely therefore that the submolecules are not spheres, and that there are more of them; in particular the X-ray data do not exclude there being one submolecule at the center of each particle, additional to the number required by symmetry.

When the crystals are exposed to the air, they shrink, and the intermolecule distance changes by 62 A. At the same time they become much disordered and show very few X-ray reflections. Two explanations are possible: either the molecules remain identical throughout, in contact in the dry crystal and separated by 60 A of liquid in the wet, or the submolecules may themselves separate somewhat in the wet crystal structure so that the layers of water between them are less deep. So far the interpretation of the X-ray data has not been carried far enough to distinguish between these hypotheses.

The second system, that provided by the turnip yellow virus, is even more remarkable. Both X-ray and electron microscope data show that the crystals, which are isotropic octahedra, have the diamond structure (Bernal and Carlisle, 1948; Cosslett and Markham, 1948). The co-ordination here is very open, each molecule being surrounded by only four others, and this in itself is perhaps not so notable. An early speculation of Bernal's proposed that protein molecules in crystals might be expected to show just this open type of co-ordination, since the forces between the molecules, and between the molecules and the salt solution enclosing them, would be of the same order of magnitude. But in this speculation, it was hardly envisaged that there might be very deep layers of water separating the tetrahedrally disposed molecules. The actual change of interparticle distance in passing from wet to dry turnip yellow crystals is nearly 80 A. Again, somewhat similar conditions exist here to those found for the bushy stunt virus. Symmetry conditions (which here receive support from "still" X-ray photographs, at least as far as the presence of 4-fold axes is concerned) suggest the subdivision of the particle of mass 3,500,000 into 12 n submolecules of mass 300,000/n. These may conceivably float apart in the wet crystal and reduce the actual depth of the individual water layers. Even so, in the air dried crystal, there would still be volumes

TABLE 2

Substance	Particle diameter in electron microscope	Interparticle distance in dry crystal	Interparticle distance in wet crystal	Change in interparticle distance wet to dry
Bushy stunt virus	255–270	272	334	62
Turnip yellow virus	193–220	228	306	78
Tobacco necrosis protein (hexagonal layer)	130–166	157	179	22

FIG. 1. Tobacco necrosis protein. Patterson-Fourier projection along the *c* axis. The vector distribution shown here is of the simplest form: the peaks correspond only to vectors between the large molecules and provide no evidence of internal structure.

of liquid or space of the order of 200 A across, between the virus molecules. And, in addition, it is very curious that this virus protein can lose nucleic acid amounting to 28 percent of its weight without alteration of crystal structure, beyond possibly a slight *expansion* of the lattice.

With the third crystal structure in this group, that of the tobacco necrosis protein (Rothamsted strain), first isolated by Bawden and Pirie, we reach a system much more like those found in the next group of protein crystals. The crystals are triclinic and very weakly birefringent and the particle arrangement can be thought of as a rather distorted version of cubic close packing. In one face the molecules do lie in an array which is exactly close packed within the limits of experimental error, but succeeding layers are slid a little from the regular positions. This close packed face is the one observed by Wyckoff (1948) in the electron microscope; it is remarkable how closely the layer structure observed by him agrees with the external form of the crystal (Plate 1). The X-ray data on the air dried crystal are here single crystal data, but much limited by the disorder in the crystals—disorder which is obvious to the eye on the electron microscope pictures. The shortest spacing observed is 58 A while that from the wet crystals is right down to 2.8 A. Although, unfortunately, the data obtained some time ago on the wet crystals (Crowfoot and Schmidt, 1945) are not complete, they can be fitted quite well with that on the air dried crystals. The simplest interpretation is that the molecules move apart to positions on a slightly expanded lattice with layers of water 20 A across between them. But in this lattice, the X-ray reflections show that the unit cell corresponding to the air dried unit cell is a pseudo cell only; *i.e.* that there is some crystallographic difference between molecules in two succeeding lines in the structure. The reflections also indicate detailed structure within the protein molecules.

There are far too many atoms in molecules of weight 1,000,000 or more to think of finding by calculation their actual positions in space. But we can, at this stage, summarize the information we can get from the X-ray data on these crystals by the use of a particular mathematical treatment first described by A. L. Patterson (1935). The relative intensities of the X-ray reflections indicate the existence of atoms at definite distances apart in space in particular directions in the crystal. By the use of Patterson's series we can map the distribution of these vector distances in one, two or three dimensions in space according to the number of X-ray reflections we employ. Figure 1 shows a map obtained for the projection on a particular crystal plane of the vector distribution in a dry crystal of the tobacco necrosis protein. This looks rather like the electron microscope pictures, and shows that all the X-rays tell us is that there are very roughly spherical masses closely packed in the crystal of the tobacco necrosis protein. The peaks correspond to vectors just between one mass and the next. The corresponding pattern from the wet crystal would be more complex, judging by the many X-ray reflections observed. It would show that there is structure within the molecule and that this is different in different directions. The molecule here appears therefore, by its very lack of symmetry, to have a specific individual character which determines its arrangement in the crystal.

It is rather interesting that the same can be said of ferritin. This protein crystallises in most beautiful deep red octahedra containing a high proportion of iron. The arrangement of units within the crystals is close to face centered cubic—the original powder lines observed by Fankuchen (1943) could be indexed on a cubic lattice. But large single crystals given by Dr. Michaelis proved to have quite marked birefringence and the actual crystal symmetry is orthorhombic, corresponding to the existence of molecules without crystallographic symmetry. The octahedra shrink regularly on drying (the change in lattice constant is 32 Å) to form a very disordered air dried crystal. But data on this are not yet sufficient to provide even a molecular weight measurement to compare with the ultracentrifuge value of 460,000. Nor have crystals of apoferritin, the iron free protein, large enough for single crystal work yet been obtained; the comparison of these with ferritin ought to be most valuable.

I have inserted excelsin next in my list of proteins with some hesitation since the measurements originally obtained by Astbury, Dickinson and Bailey (1935) were concerned only with air dried crystals and were very incomplete. But they bear on my theme in two directions. First we have, once again, the phenomenon that the asymmetric unit in the crystal is a submultiple, here one third, of the molecular weight found in the ultracentrifuge, and

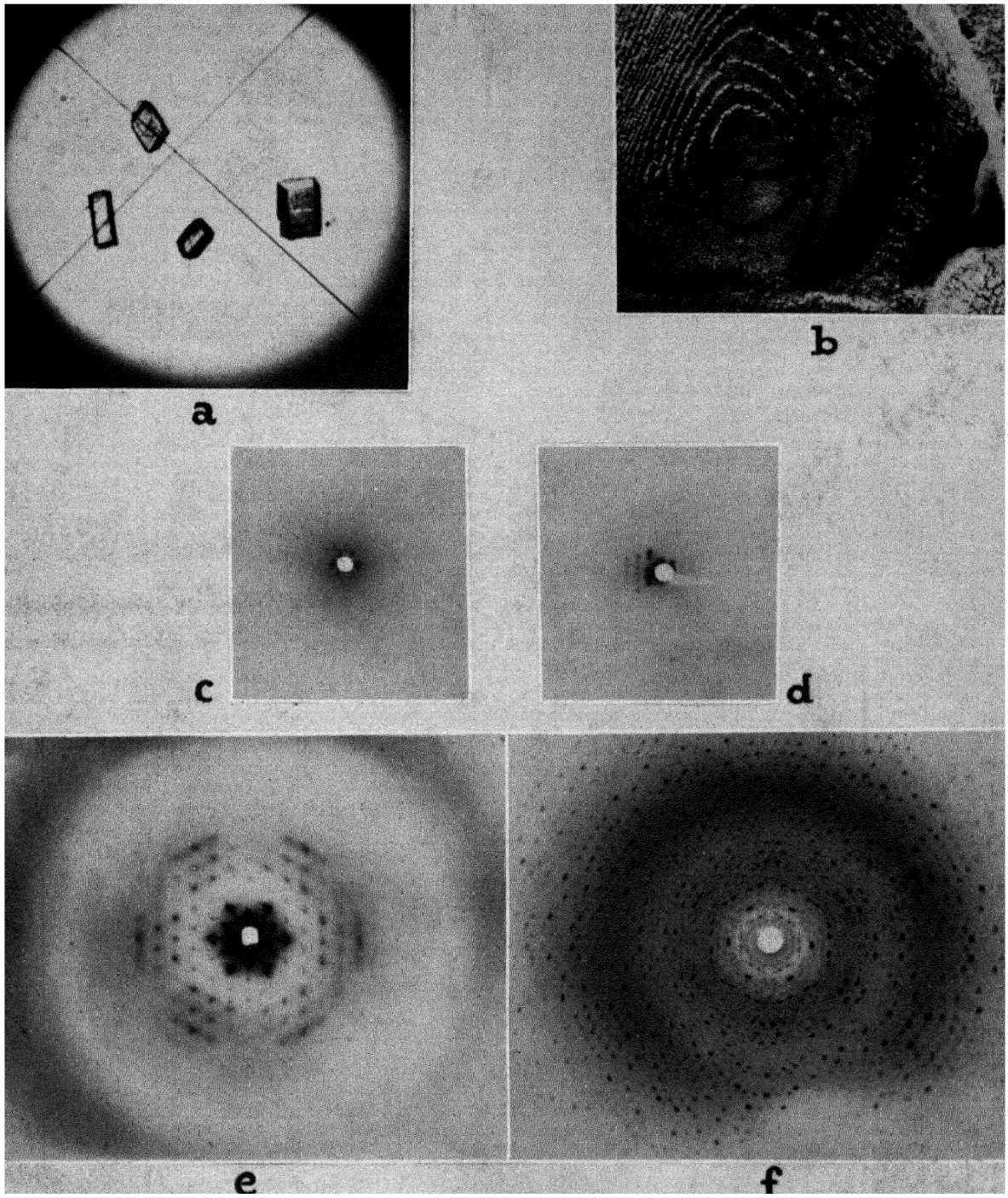


PLATE I

a. Dry crystals of tobacco necrosis protein.

b. Electron microphotograph of crystal of tobacco necrosis protein (from Wyckoff).

c. X-ray diffraction photograph of single dry crystal of tobacco necrosis protein. Cr radiation, oscillation range 4°.

d. X-ray diffraction photograph of single wet crystal of tobacco necrosis protein. Cu radiation, oscillation range 2°.

e. X-ray photograph of a single dry crystal of pepsin. The crystal was oscillating ± 2½° from a position with the beam parallel to the hexagonal axis. Film distance, 10 cm. Cu radiation (from Perutz, 1949).

f. X-ray photograph of a single wet crystal of pepsin. The crystal was stationary, with the beam nearly parallel to the hexagonal axis. Film distance, 6 cm. Cu radiation (from Perutz, 1949).

this is now correlated with physical chemical evidence—that excelsin under certain conditions, particularly in urea solution, splits into 3n subunits. This gives us more confidence in the argument from crystal symmetry. And secondly we have, for the first time, some evidence of the existence and arrangement of peptide chains within the protein crystal. Superimposed on the X-ray reflections characteristic of the excelsin single crystal was an oriented fibre pattern showing spacings not identical with, but rather similar to, those found in β keratin. Such a fibre pattern suggests that the molecules in the protein crystals are built up of a parallel arrangement of peptide chains.

Molecular Weight Group 10,000-100,000

Ferritin and excelsin provide something of a bridge in my list between the very large virus molecules and the much more intensively studied range of small protein molecules, from haemoglobin down. Here we are moving out of the range to which electron microscope observation has reached. Instead we have a much greater wealth of already collected X-ray data (*cf.* Plate I d) which permits us to deal more precisely with many of the problems raised earlier, particularly that of the rigidity or otherwise of the units in the crystal, the nature of the water layers between molecules in wet crystals and the units of which the protein molecules are built.

That, in this size range, the protein molecules are rigid entities, which move apart as the crystals swell and together as they shrink, was first established by the comparison of vector maps obtained from X-ray data on wet and dry insulin (Crowfoot, D., and Riley, D. P., 1939), lactoglobulin (Riley, D. P., 1942) and horse methaemoglobin (Boyes-Watson, Davidson and Perutz, 1947). Two sets of these vector patterns for haemoglobin and for insulin, are shown in Figure 2. These are all of considerable complexity. They show peaks at distances corresponding in magnitude with those which must exist between masses of atoms in the protein molecules. The numbers of atoms involved are still too great to permit the easy deduction of the actual atomic arrangement from the derived vector patterns. Nevertheless, the fact that the calculated vector pattern, in the neighbourhood of the crystallographic origin, is sensibly constant as we pass from a wet to a dry crystal does establish that there is the same arrangement of groups within some unit in the two crystal structures, of a size which depends on the area of constancy. This unit must be identified with a rigid molecule.

In horse methaemoglobin, this effect has been traced by Perutz (1946) through a very remarkable series of crystallographic transformations. The crystals are monoclinic and on slow drying they shrink step wise, predominantly in one direction, normal to the crystallographic *c* plane. The intermediate stages differ from each other by intervals which correspond closely with the removal of one or two layers of water molecules at a time. From the normal wet cell to the air dried state, the intervals are 4.6 A, 3.8 A and 6.3 A. And in ammonium sulphate solution of concentration more than 2.8 M or in acid at pH 5.4 the crystal expands, again suddenly, by a step of 4.2-3.7 A. The actual thickness of a puckered hexagonal layer of water molecules in ice is 3.7 A and the total change in *c* spacing in haemoglobin from acid expanded to air dried crystals, 18.5 A, corresponds to a thickness of five such layers.

The evidence is therefore that the water molecules have a quite definite, though probably not quite regular, distribution, in haemoglobin crystals. One may imagine that in the neighbourhood of the surface of the protein molecule there will be a certain amount of fitting of water molecules into crannies between projecting side chains and rounding off corners before the layer structure can develop. Water molecules in this situation may constitute the "bound" water, impermeable to ions, of the presence of which certain measurements of Perutz give evidence. In addition, in air dried crystals, there is also probably still some water packing between the molecules, which like the "bound" water can be removed on more drastic drying for analytical purposes. In different crystals rather different arrangements of the water molecules may obtain. In some of the cubic crystals, for example, there may be polyhedral spaces filled with water as in the phosphotungstic acids. In these, a group of 29 H_2O exists in the crystal enclosed by anions and occupying a space 11 A across. 17 of the water molecules form a compact group; the remainder are arranged in six-sided puckered rings and fitted into spaces between the main water molecule system and the anions (Bradley and Illingworth, 1936). But in all the protein crystals so far studied the water is essentially still liquid; it is penetrated with rapidity by ions and also by quite large dye molecules.

It is largely the existence of the water layers in haemoglobin crystals and the possibility of varying their thickness and composition that has made it possible to draw some tentative conclusions from the mass of X-ray evidence available about the shape and structure of the molecule. First, the crystal symmetry again requires a unit smaller than the molecule found in the ultracentrifuge. The crystallographic unit, of weight 33,350, is half that of 66,700 shown by ultracentrifuge measurements to be present in solution, and there is evidence here also that splitting of the 66,700 unit into two occurs in solution. In another haemoglobin, sheep foetal haemoglobin (Kendrew and Perutz, 1948) the crystallographic evidence suggests that the molecule is built up of four identical or near identical parts, each of weight 17,000, and molecular weight measurements in solution confirm that splitting is indeed into units of

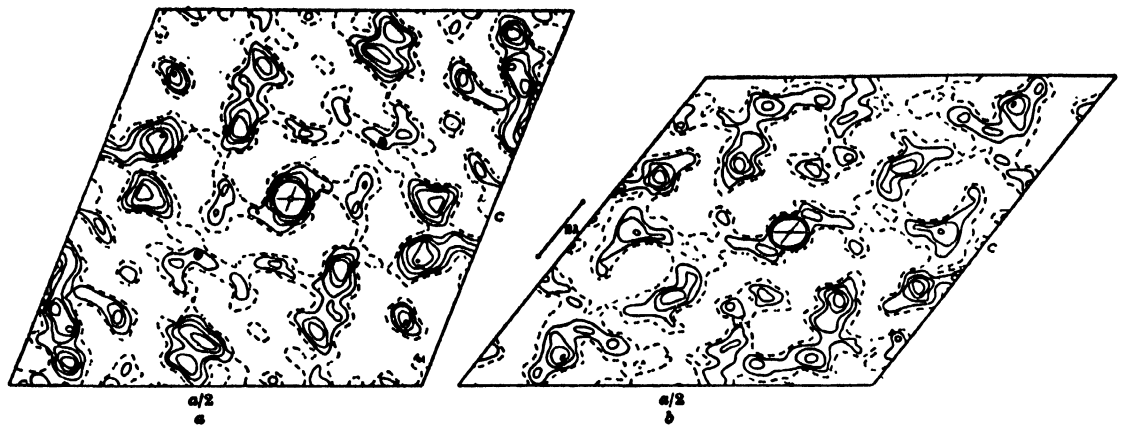


FIG. 2

i) Horse methaemoglobin. Patterson-Fourier projection along the *b* axis for a) the normal wet crystal, and b) a partly shrunk crystal.

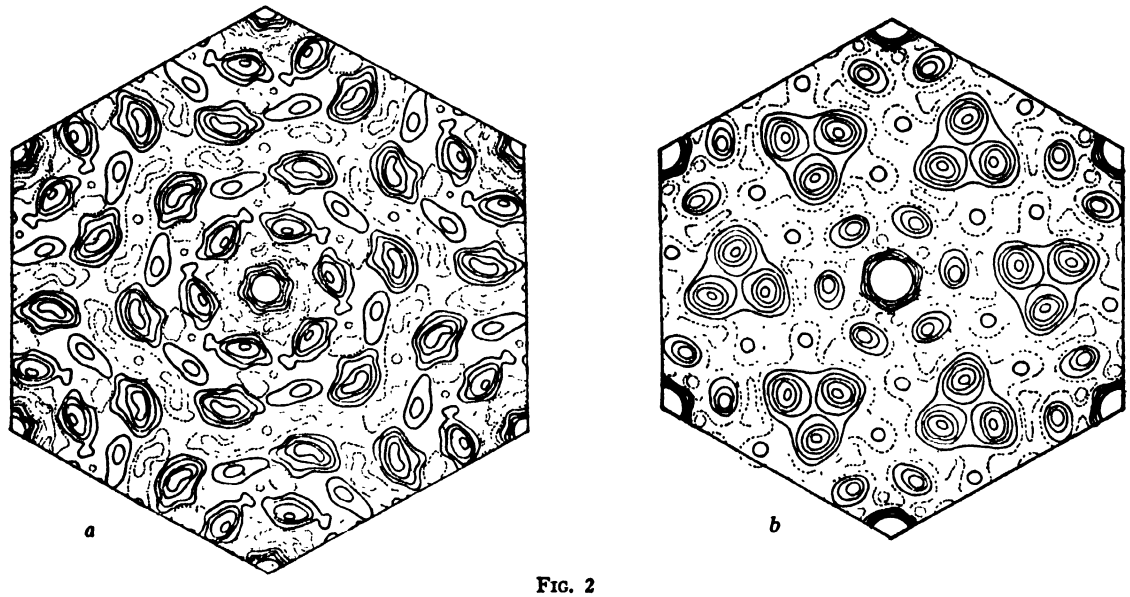


FIG. 2

ii) Insulin. Patterson-Fourier projection along the *c* axis for a) the normal wet crystal, and b) the air dried crystal. The vector distribution here shows considerable detail: this detail remains sensibly constant around the origin in the different shrinkage stages. Fig. 2i is taken from Perutz, 1942, Nature, Lond. *149:* 491.

this size, of the same weight as the molecule of myoglobin. Perutz's view of the structure of haemoglobin provides an easy interpretation of these facts. The molecule is considered to be cylindrical in shape and built of four layers, two of which are identical, or very nearly identical, with the other two (Figs. 3 and 4). Within each layer there is evidence of a chain arrangement which is shown in Figure 5. This evidence is derived from a three dimensional vector distribution calculated from some 7,000 X-ray reflections (Perutz, 1948). It shows long ridges parallel to the long axis of the crystal (Fig. 6) and distributed roughly in hexagonal packing in a plane at right angles to this axis (Fig. 7). It is very significant that there is, in the vector distribution around the origin, a series of peaks at 5 A distance, and along the ridges, peaks at roughly 5 A intervals. These suggest a correlation between the chain configuration here and the α keratin structure found by Astbury. The interval between the chains of about 11 A is closely similar to the second keratin spacing of about 10 A.

The crystal structure of myoglobin, investigated by Kendrew (1948), provides good confirmation of some of Perutz's conclusions. The system is much more simple, the molecule consists of only one unit of 17,000 weight and packing considerations suggest that this is a flat disk which lies parallel to one crystal plane, the *b* plane. A vector projection on this plane is astonishingly similar to that on the

corresponding plane of haemoglobin and the peptide chains must accordingly have a very similar arrangement within one protein layer in the two crystals.

There are certain obvious difficulties about Perutz's model for haemoglobin which it may be worth stating straight away. First, the arrangement of chains in a single circular layer, as drawn out in Figure 5 is not very convincing. If there are chains with a definite orientation, as the evidence indicates, one would expect the individual layers not to be circular in shape. I think this is most likely the case. The cylindrical shape for the haemoglobin molecule was favoured by a number of considerations of which one of the most important was the close packed nature of the structure. A rather more rectangular form would meet this need equally well. Secondly, the packing of the chains together at right angles to the layer plane in Figure 5 is not very convincing and does not fit altogether well with the vector map shown in Figure 7. For example, this shows no peak corresponding to the vector A′-A of Figure 5. This difficulty could in part be surmounted if the chain cross section were, in fact, very far from spherical.

The correct interpretation of the haemoglobin vector distribution is all the more important because there are definite resemblances between the patterns here and those found for some members of the next group of small proteins of molecular weight 12,000-13,000, which includes lysozyme, ribonuclease and insulin. The analysis of ribonuclease, started by Fankuchen and continued by Carlisle (1941) is particularly promising. In the modification now being studied, the crystals shrink comparatively little

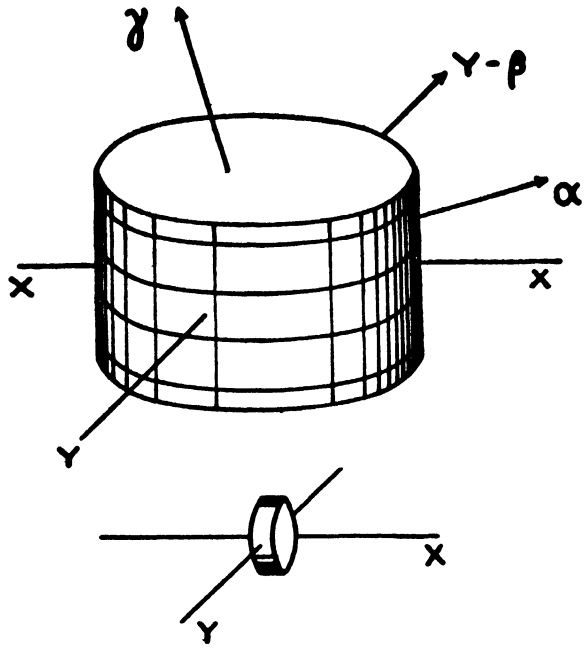


FIG. 3. Diagrammatic model of haemoglobin molecule, showing its orientation with respect to the crystal axes. *Y* is the diad axis. The small disk underneath represents a haem group drawn on the same scale and in its correct orientation with respect to the crystal axes. The four lines on the cylinder surface indicate the positions of the concentrations of scattering matter deduced from the Fourier projections. The directions of the principal refractive indices are indicated by arrows.

Taken from Boyes-Watson, Davidson and Perutz, 1947, Proc. Roy. Soc. *191:* 123.

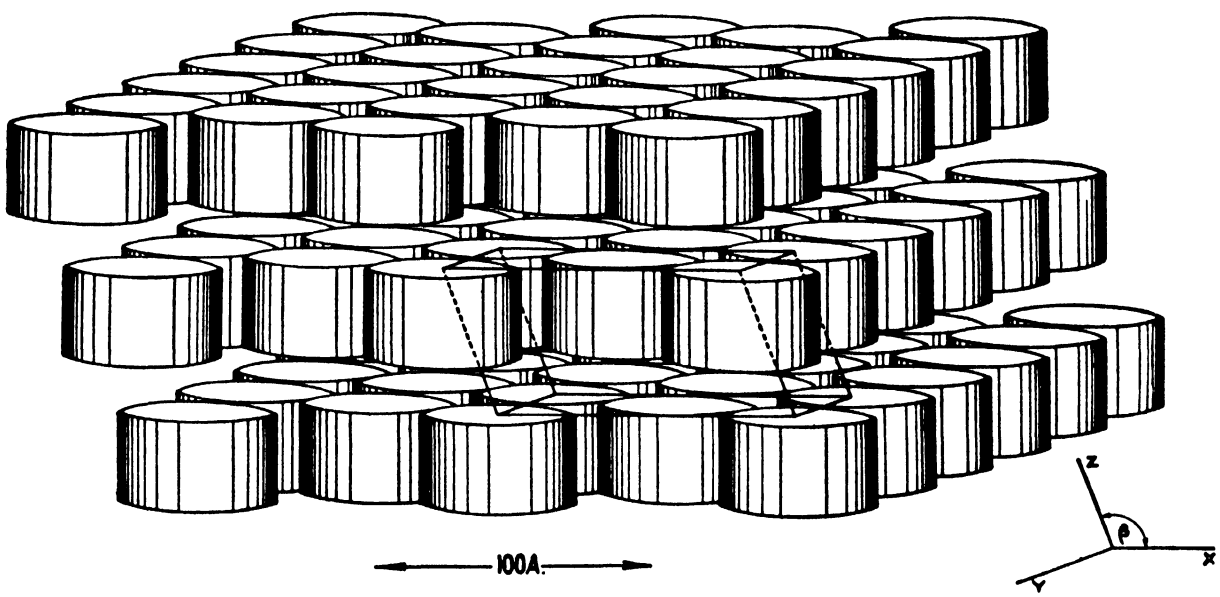


FIG. 4. Packing of haemoglobin molecules in the crystal structure, showing layers of close-packed molecules separated by liquid. One unit cell is shown in the foreground on the right.

Taken from Boyes-Watson, Davidson and Pertuz, 1947, Proc. Roy. Soc. *191:*123.

on drying (the maximum change in cell dimension is 8 A) and there is little in the way of guide to sort out the shape of the molecules in the crystal. But in one projection, there is a remarkably similar distribution of peaks to that observed on the *a* plane in haemoglobin, looking down the proposed chains. And on one of the planes at right angles to this, it is also possible to trace (though not very obviously) a system of chains with peaks at intervals of 5 A. This suggests that the molecule is composed of possibly 5 or 6 chains arranged in roughly hexagonal packing, as in haemoglobin.

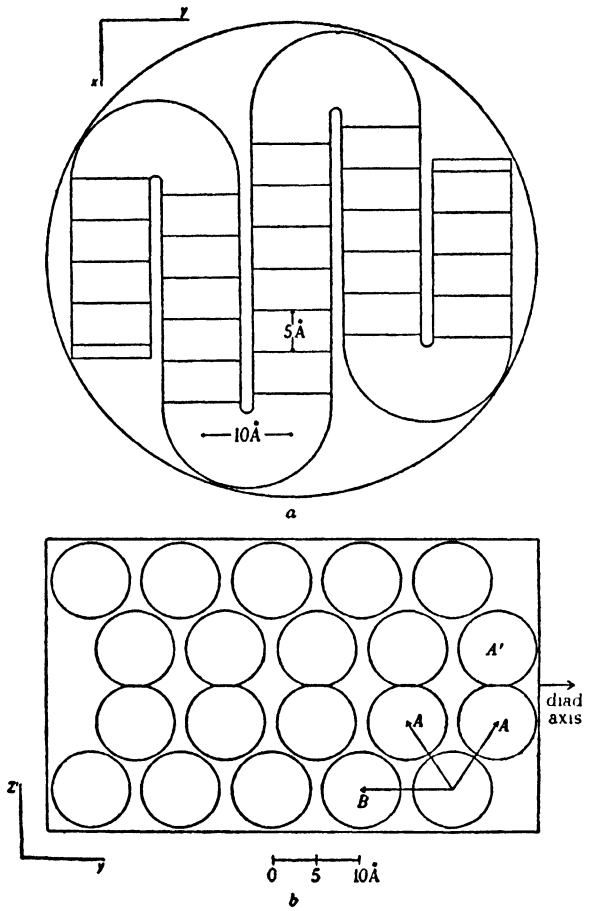


FIG. 5. Suggested structure of haemoglobin. a) Basal section showing folded polypeptide chain. b) Vertical section showing packing of chain. Taken from Perutz, 1949, Proc. Roy. Soc. *195:* 474 (Fig. 23).

It is somewhat illogical of me to put insulin in this particular group on grounds of molecular weight. The crystal unit cell contains 36,000 molecular weight units of protein and the trigonal symmetry requires this to be subdivided into three identical units of weight 12,000 (Crowfoot, 1938). In solution the molecular weight has been shown to vary from 48,000 to 12,000 under different experimental conditions (Gutfreund, 1948). So far the situation is very like that of haemoglobin, which was classed above according to its maximum solution molecular weight. Such difference as exists here is chiefly due to the weight of the chemical analytical evidence which is much more complete for insulin than for any other protein. This provides strong confirmation for the view that the 12,000 unit is the chemical unit and that we can properly treat this as a separate individual to assist in interpreting the X-ray evidence.

There is, in insulin, a strong resemblance between the distribution of vectors in the *c* plane projection and those normal to the proposed chain directions in haemoglobin and ribonuclease. If we correlate this with the chemical evidence that there are four chains in each 12,000 unit in insulin, we are led to propose a projected chain distribution of the form shown in Figure 8. This has many advantages, but also presents some difficulties. The number of amino acid residues in each of the four chains must be of the order of 20-30. In the α keratin configuration three amino acids occupy 5.1 A and the length occupied by 25 residues in this arrangement should therefore be 42 A—much longer than the *c* dimension of 30 A. From the appearance of the vector pattern it is very likely that a different chain configuration is involved here from that found in α keratin. But other explanations are also possible. Two of the chains may be longer than the other two, for example, and be folded back—the whole pattern having a different projection, corresponding closely with the vector map itself. Or the chains may take an altogether different direction in the structure.

From this group of protein crystals, taken together, one gets the impression that the characteristic feature of each protein is the particular mode of folding and packing the polypeptide chains which is adopted in each case. There are general similarities but also marked differences between the crystal structures of the different proteins. Crystallographically, these must be so correlated with large scale differences in chain arrangement, rather than with small differences in amino acid composition.

Molecular Weight Group 1,000-10,000

Since we have now specifically raised the problem of the arrangement of peptide chains in protein molecules it is natural to try to examine next the arrangement of peptide chains in peptide crystals. Unfortunately only a very limited number of compounds have yet been examined in this group and those which have, gramicidins A, B and S, are probably not typical of peptides as they occur in proteins. All contain one or more amino acids of unnatural configuration, probably all are cyclic, and all crystallise from organic solvents. Though the crystals do contain solvent of crystallisation when first they form, which they lose on exposure to the air, they do not shrink very much on drying. The

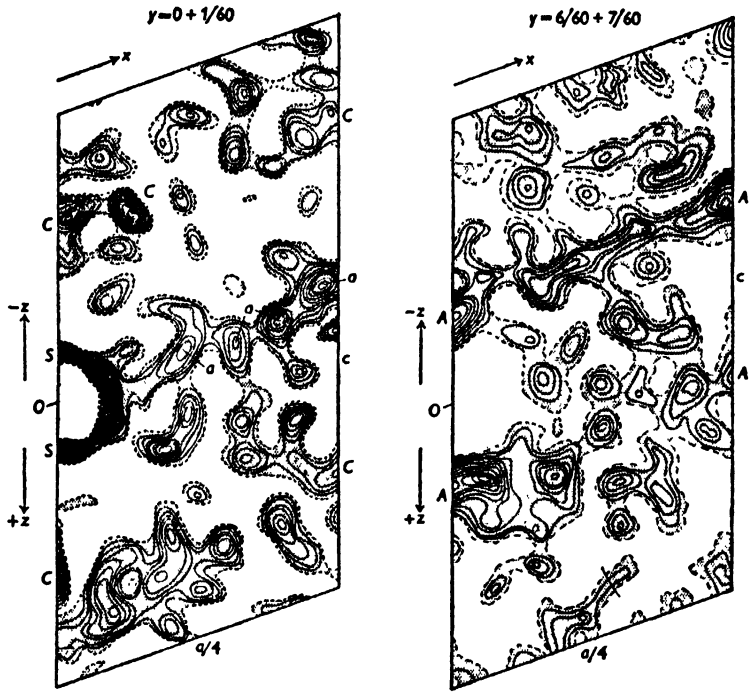


FIG. 6. Sections through vector structure for horse methaemoglobin. The shaded and the blank areas indicate regions of high and low vector density respectively. Taken from Perutz, 1949, Proc. Roy. Soc. *195:* 474 (Fig. 17, 18).

solvent—alcohol or acetone—appears to pass out, probably leaving holes in the structure of a somewhat disordered crystalline phase. Sometimes this phase has a different crystal symmetry from the original wet crystal, or one unit cell dimension halved.

Specimens of gramicidin A and B were given by Dr. Craig to Dr. Synge, who kindly passed them on to us for examination (Cowan and Crowfoot, unpublished). Gramicidin A crystallised in beautiful small orthorhombic prisms which gave quite good X-ray photographs. The calculated molecular weight from the X-ray data is 3,800; it is possible that the molecule has a weight twice this only if it has a two-fold axis of symmetry. Gramicidin B (Gregory and Craig, 1948) crystallised in much smaller, less well formed, monoclinic plates. The crystal symmetry here formally requires only two "molecules" of weight about 7,500 in the cell but there is very strong pseudo symmetry which suggests a packing unit of half this size (*cf.* the situation in say, lactoglobulin, where the packing unit corresponds in weight to the molecule found in solution, Senti and Warner, 1948). For both compounds, therefore, the X-ray data favour a molecular weight of about 3,800 and this appears to be compatible with the chemical analyses carried out by Moore and Stein. But it should be realised that the X-ray data here are rather limited in character. The crystals are a good deal disordered and the X-ray reflections do not go out beyond a spacing of

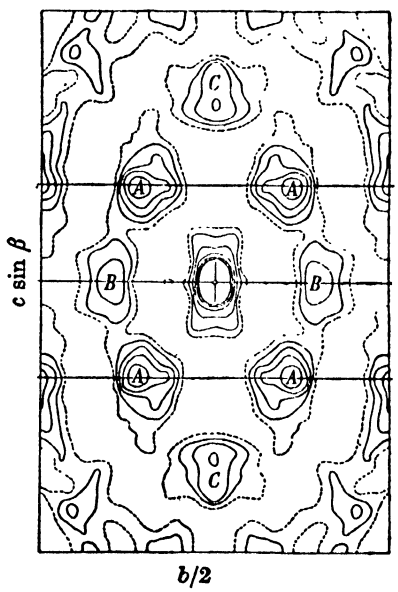


FIG. 7. Horse methaemoglobin. Patterson-Fourier projection along the *a* axis. The peak distribution corresponds to the projection of the chain systems of Fig. 6 looked at in a direction roughly down their length. Taken from Perutz, 1949, Proc. Roy. Soc. *195:* 474 (Fig. 21).

1.8 A. It is quite possible that the units present of weight 3,800 are not all absolutely identical; there might, for example, be two such units differing in

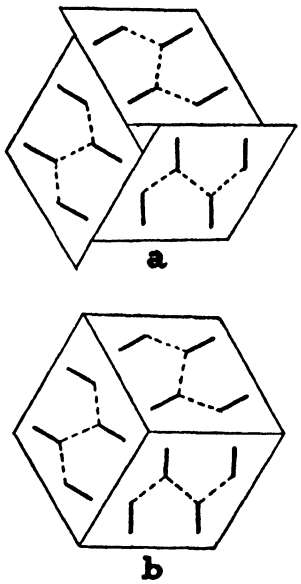


FIG. 8. Possible mode of packing of chains in the crystal structure of insulin. Four "chains" are here seen end on, packed to form a molecule of weight 12,000, the molecules being then packed together in the unit cell of the rhombohedral crystal. They can be treated as independent units and in projection are probably better packed as in a) rather than b).

one or two amino acids. Statistically distributed, these would give the observed symmetry. The two units might even be combined together to form a single molecule with pseudo twofold symmetry. It is fairly clear that in crystals of this kind the X-ray reflections are sensitive only to the main outlines of the molecule present, and are not noticeably affected by impurities or variations if these are small in comparison with the molecular size.

The crystal structure of gramicidin B is interesting in another way. One cell dimension, *b*, is the shortest we have yet encountered among compounds of this kind—13.4 A. A vector projection on the *b* plane (Fig. 9) shows a marked chainlike structure, again with peaks along it at somewhere near 5 A intervals. The situation is, however, complicated and it seems unlikely that we have a normal α keratin chain here. The interval per amino acid residue is still approximately three residues to 5.1 A but there is the problem of forming a cyclic structure. From the projection it is likely that this, if present, consists of two closely packed parallel chains.

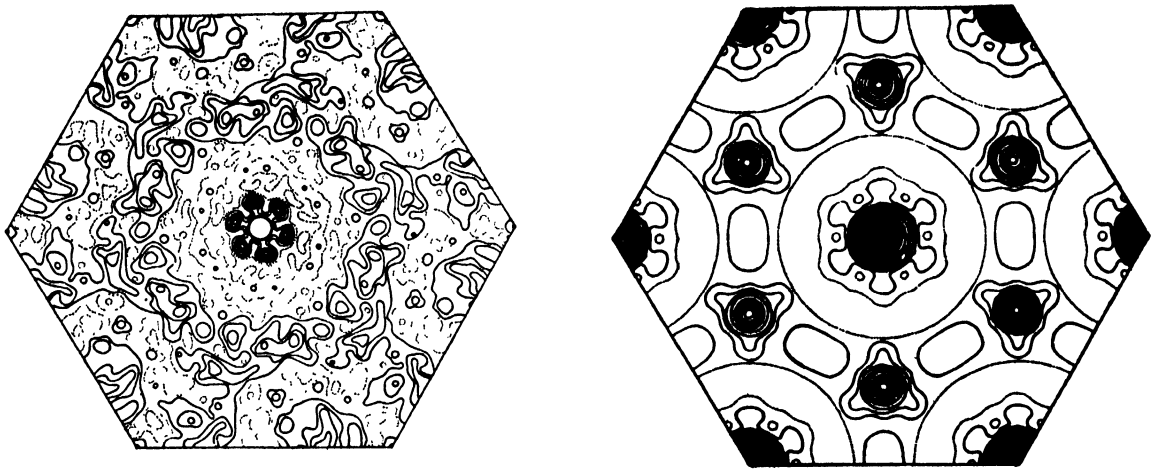
With gramicidin S there is a curious reversal of the situation that we have found in almost all other classes of compound examined. Chemically the peptide contains only five amino acids, L-valine, L-ornithine, L-leucine, D-phenylalanine and L-proline, probably in this order. But although Dr. Synge has prepared nine different derivatives for us, which we have examined in some twenty different crystal structures (including wet and dry forms) the crystallographic unit is never as small as the pentapeptide and often much larger (Crowfoot and Schmidt, unpublished). It seems quite likely that the smallest observed unit, the decapeptide, is the molecule, two pentapeptide units being joined in a ring. There is some evidence too, derived from comparing different crystal structures, that the decapeptide unit has twofold symmetry. But the impression that the crystal structures as a group leaves with one is that there is a number of non crystallographic ways of putting these peptide units to fit into the repeating crystal pattern. One cannot help feeling that these ought to provide some kind of model for packing peptide units into proteins—and if so, there is quite a variety of models that can be followed even in this one group of crystals.

As might be expected, there are certain definite similarities between vector patterns calculated for gramicidin S derivatives and protein crystals. For example, one section of the vector distribution for *N*-acetylgramicidine S shows a marked ring of peaks at 5 A distance from the origin, very like the rings found in insulin and haemoglobin (Fig. 10). But these probably represent little more than the general chemical similarity of the two classes of compound.

Molecular Weight Group 0-1000

From the point of view of X-ray analysis, the situation changes radically between this group and the preceding groups. It is reasonably certain that the structure of any sufficiently large crystal of an amino acid or peptide in this molecular weight group could be solved in detail, *i.e.* the actual positions of the atoms in space could be found with a high degree of certainty. The accuracy might vary from say $\pm.02$ A at the low molecular weight end to, perhaps, $\pm.5$ A at the high molecular weight end. But we are definitely out of the region where we can so far calculate little more than the F^2 series to show us vector distribution, into a region where we can find, by well-established methods, the actual electron density distribution in the crystals. And we can expect, from all crystals in this group (and possibly from some in the group above) enough X-ray data to show us separate resolved atoms in our calculated electron density distribution. But so far this

FIG. 9. Patterson-Fourier projection along the *b* axis for gramicidin B, showing vector distribution of a chain-like character.



a b

FIG. 10. Sections through the vector structure parallel to the *c* plane for a) insulin, b) n-acetylgramicidin S. Both of these show a marked ring of peaks around the origin at 5 A away. (The scale of *a* is about half that of *b*).

statement has only been proved for a very limited group of compounds.

Glutathione, I have put into the table as the highest molecular weight compound of this group although the X-ray analysis is only as yet in the early stages. Enough has been done already at Birkbeck College, London, to show that the crystal structure is soluble. But so far exact analyses have been published only for the simpler amino acids and peptides, glycine (Albrecht and Corey, 1939), D-L alanine (Levi and Corey, 1941), and β glycylglycine (Hughes and Moore, 1942) to which we might add diketopiperazine (Corey, 1938). These have been lately reviewed by Dr. Corey (1948) and I will not here make more than one or two general comments.

This group of crystal structures provides us, in the first place, with a knowledge of the normal bond distances and bond angles we are likely to find in the peptide chains of proteins. These are summarised briefly in Figure 11. Small variations may occur, but they are not likely to have any great importance in determining the general lay-out of proteins. Although complete data are not available for β-glycylglycine, it is clear that this molecule has the extended configuration expected for a β-keratin type of peptide chain. All the atoms in the molecule are coplanar except the terminal nitrogen atom which lies about 0.7 A from the plane of the rest of the atoms.

In the second place, these structures show us that, wherever possible, amino or imino groups and carbonyl or carboxyl groups are so arranged that they can interact with one another in the crystal, either through normal ionisation leading to zwitter ion structures or through the formation of hydrogen bonds or both. This leads to shortened interatomic distances and, in these structures, where there are few projecting R groups, to relatively high crystal densities. Also, since the molecules are of simple shapes, it is no longer necessary to pack intervening spaces with solvent molecules.

These points may be illustrated by reference to the crystal structure of D-L alanine. The molecules are fitted intricately together so that both internal and external binding relations are satisfied. But in the result a simple plan appears. The molecules follow one another in a chain in the crystal structure and the chains are very roughly close packed. If we look at the alanine molecule as a unit, then the centres of gravity of these units are closely 5 A apart in space and there are maxima in the F^2 series, as might be expected corresponding to overlapping vectors of approximately this length. The prevalence of 5 A spacings in protein and peptide systems appears as a necessary consequence of the fact that the structures are largely determined by the packing of nonbonded groups, the side chains, at mean distances which must be the same as those found when packing separate amino acid molecules in crystals.

CONCLUSION

We have now, in theory at least, unrolled protein molecules to the end and finished by breaking them into pieces. In the process some conclusions have emerged about their organisation which may help us to proceed further in building the pieces together again. In each of the four main groups of compounds we have examined, we can see that the molecules present can in many cases be recognized as consisting of subunits. These subunits are of the same order of magnitude as the molecules of the next lower molecular weight group. So that in rebuilding again, we start naturally with the prob-

lem of the organisation of amino acid residues into peptide chains and then of the arrangement of the peptide chains into protein molecules.

There are two general methods of procedure from this point on, both of which may usefully continue side by side. In the first place, by making use of the existing evidence on simple amino acid structures, a number of models of peptide chains which are stereochemically plausible have been constructed. These can be directly compared with the vector patterns already obtained for peptides and proteins. The one which is being most intensively examined at the moment is the threefold spiral model described

synthesis of peptide chains proceeds. We need to examine crystallographically, in all the laboratories now engaged in the X-ray analysis of proteins, numbers of peptides, preferably built up of say between three and six residues to start with, and containing a variety of side chains. In this way we might hope to find whether there are, in fact, only a few commonly adopted methods of chain arrangement or whether there is a large number. We should also get to know how the different modes of grouping would appear if we blurred the electron density patterns as they must appear blurred when derived from protein crystals.

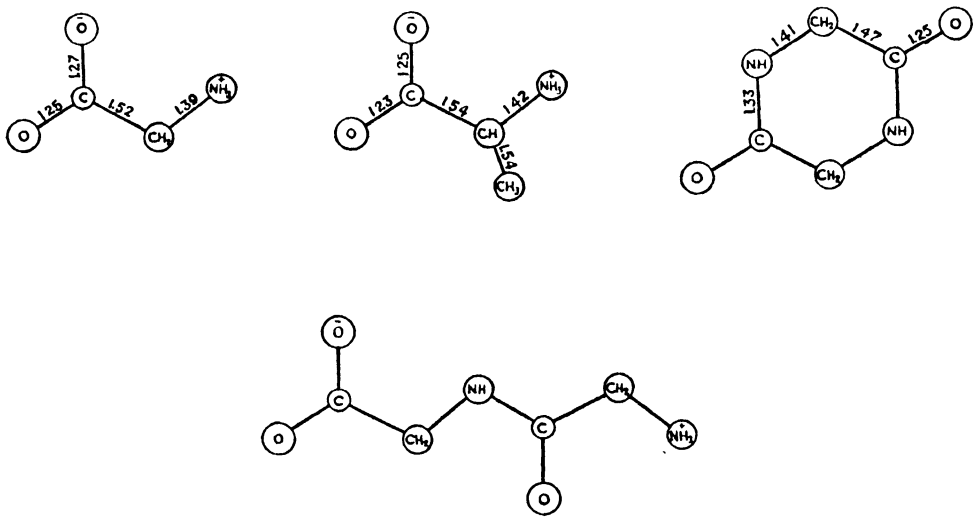


FIG. 11. Interatomic distances and atomic arrangement found in glycine, alanine, diketopiperazine and β-glycylglycine.

a long time ago, first by H. S. Taylor (1941) and then by Huggins (1943). It has many advantages, including internal hydrogen bonds and packing of the side chain residues at 5.1 A intervals. Vector patterns derived from it do show marked similarities with those observed in haemoglobin and ribonuclease. But at present it is not possible to exclude the likelihood that other models would give equally good agreement, and we are faced with a problem about the nature of our evidence and of proof. If we choose a high molecular weight protein, such as haemoglobin, we cannot observe enough X-ray reflections to show us atoms within the structure and so provide us with our usual criterion of a correct crystal analysis. We ought to consider very carefully what new criteria may be available. (This may not be so important as it seems. There is something about a correct solution to a crystal structure problem to anyone who has gone through the process of finding it, which is often convincing long before the details have been properly worked out).

In the meantime, there is the second method of approach which is becoming far more practically possible as chemical work on the breakdown and

So far I have treated the problem of protein structure as one depending mainly on the organisation of peptide chains. There is, from the X-ray data, some evidence in favour of the existence of peptide chains in proteins, but, I must admit that I might not have accepted this so easily without much stronger evidence, provided by work in other fields which are covered by this Symposium. At the same time, one must recognise that the problem of the organisation of many protein molecules is more complex than the organisation of peptide chains alone and that other chemical molecules, carbohydrates and nucleic acids, for example, may be intimately involved. Detailed X-ray analyses of these molecules and particularly of these molecules combined with peptide fragments are therefore also likely to be most important in providing us with a knowledge of the structure of the protein molecule as a whole.

REFERENCES

ALBRECHT, G., and COREY, R. B., 1939, The crystal structure of glycine. J. Amer. Chem. Soc. *61:* 1087-1103.

ASTBURY, W. T., DICKINSON, S., and BAILEY, K., 1935, The

X-ray interpretation of denaturation and the structure of the seed globulins. Biochem. J. *29:* 2351-2360.

BAWDEN, F. C., and PIRIE, N. W., 1942, The preliminary description of preparations of some of the viruses causing tobacco necrosis. Brit. J. Exp. Path. *23:* 314-328.

BERNAL, J. D., and CARLISLE, C. H., 1948. Unit cell measurements of wet and dry crystalline turnip yellow mosaic virus. Nature, Lond. *162:* 139-140.

BERNAL, J. D., and FANKUCHEN, I., 1941, X-ray and crystallographic studies of plant virus preparations. J. Gen. Physiol. *25:* 111-165.

BRADLEY, A. J., and ILLINGWORTH, J. W., 1936. The crystal structure of $H_3PW_{12}O_{40}\cdot 29H_2O$ Proc. Roy. Soc. *157:* 113-131.

BOYES-WATSON, J., DAVIDSON, E., and PERUTZ, M. F., 1947, An X-ray study of horse met-haemoglobin. Proc. Roy. Soc., A *191:* 83-132.

CARLISLE, C. H., and DORNBERGER, K., 1948, Some X-ray measurements on single crystals of Tomato Bushy-Stunt Virus, Acta Cryst. *1:* 194-196.

COREY, R. B., 1938, The crystal structure of diketopiperazine. J. Amer. Chem. Soc. *60:* 1598-1604.

1948, X-ray studies of amino acids and peptides. Advances in Protein Chemistry *4:* 385-406.

COSSLETT, V. E., and MARKHAM, R., 1948, Structure of turnip yellow mosaic virus in the electron microscope. Nature, Lond. *161:* 250-252.

CROWFOOT, D., 1938, The crystal structure of insulin, I. Proc. Roy. Soc., A *164:* 580-602.

CROWFOOT, D., and RILEY, D. P., 1939, X-ray measurements on wet insulin crystals. Nature, Lond. *144:* 1011.

CROWFOOT, D., and SCHMIDT, G. M. J., 1945, X-ray photographs of a single crystal of tobacco necrosis virus derivative. Nature, Lond. *155:* 504.

FANKUCHEN, I., 1941, X-ray and crystallographic study of ribonuclease. J. Gen. Physiol. *24:* 315-316.

GREGORY, J. D., and CRAIG, L. C., 1948, Counter current distribution of gramicidin. J. Biol. Chem. *172:* 839-840.

GUTFREUND, H., 1948, The molecular weight of insulin and its dependence upon pH, concentration and temperature. Biochem. J. *42:* 544-548.

HUGGINS, M. L., 1943, The structure of fibrous proteins. Chem. Rev. *32:* 195-218.

HUGHES, E. W., and MOORE, W. J., 1942, The crystal of β-glycylglycine. J. Amer. Chem. Soc. *64:* 2236-2237.

KENDREW, J. C., and PERUTZ, M. F., 1948. A comparative X-ray study of foetal and adult sheep haemoglobins. Proc. Roy. Soc., A *194:* 375-398.

KENDREW, J. C., 1948, Preliminary X-ray data for horse and whale myoglobins. Acta Cryst., *1,* 336.

LEVI, H., and COREY, R. B., 1941, The crystal structure of DL alanine. J. Amer. Chem. Soc. *63:* 2095-2108.

NEURATH, H., and COOPER, G. P., 1940, The diffusion constant of tomato bushy stunt virus. J. Biol. Chem. *135:* 455-462.

PALMER, K. J., BALLANTYNE, M., and GALVIN, J. H., 1948, The molecular weight of lysozyme determined by the X-ray diffraction method. J. Amer. Chem. Soc. *70:* 906-908.

PATTERSON, A. L., 1935, A direct method for the determination of the components of interatomic distances in crystals. Z. Krist. *A, 90:* 517-554.

PERUTZ, M. F., 1946, The composition and swelling properties of haemoglobin crystals. Trans. Far. Soc. *42 b:* 187-195.

1948, An X-ray study of horse methaemoglobin II. Proc. Roy. Soc. *195:* 474-499.

1949, X-ray studies of crystalline proteins. Research *2:* 52-61.

PRICE, W. C., WILLIAMS, R. C., and WYCKOFF, R. W. G., 1946, Electron micrographs of crystalline plant viruses. Arch. Biochem. *9:* 175-185.

RILEY, D. P., 1942, The crystal structure of some proteins. Ph.D. Thesis. Oxford.

SENTI, F. R., and WARNER, R. C., 1948, X-ray molecular weight of β-lactoglobulin. J. Amer. Chem. Soc. *70:* 3319-3320.

TAYLOR, H. S., 1941, Large molecules through atomic spectacles. Proc. Amer. Phil. Soc. *85:* 1-12.

WYCKOFF, R. W. G., 1948, The electron microscopy of macromolecular crystals. Acta Cryst. *1:* 292-295.

DISCUSSION

HAUROWITZ: Should it be possible to say something about the amount and the kind of bound water from a comparison of the X-ray diagram of hemoglobin and myoglobin? The difference between the dimensions of one hemoglobin layer and 4 myoglobin layers could possibly give some information on the mode of linkage of the water.

HODGKIN: In the crystal structure of myoglobin there is very little shrinkage normal to the layers when the wet crystal passes into the air dried state. On the other hand, the distance between the protein layers is 14 A in the air dry crystal, 5 A more than the distance between the layers within a haemoglobin molecule. It seems likely that, in the myoglobin crystal, the layers are held apart by haem groups projecting at their ends. Water layers, probably two, may then be trapped between the protein layers by the haem groups, and so be unable to pass out when the wet crystal dries.

LOW: During the course of their recent ultracentrifuge and diffusion studies on insulin in acid solution Oncley and Ellenbogen have crystallized insulin in the absence of zinc, and in the presence of sulphuric acid at pH values between 2 and 3.5. I have begun an X-ray study of this material, which is proceeding.

The crystals are well formed orthorhombic dipyramids and on exposure to air rapidly lose water and become opaque. Photographs of the dry crystal give for the three orthogonal axes $a = 44$ A, $b = 51.4$ A, $c = 30.4$ A. The space group $P2_12_12$. $n = 4$ has been tentatively assigned. The axes have been named so that the *c* axis is approximately equal to the *c* axis of the hexagonal dry 36.000 unit cell (Crowfoot, 1938). The volume of the unit cell which has four asymmetric units is therefore about 69×10^3 A^3. Compare 50×10^3 A^3 dry insulin with one 36,000 or three 12,000 units.) If we assume a density of approximately 1.3 this gives a unit molecular weight of approximately 13,500. According to Ellenbogen (*ibid*) each molecule of insulin in the crystal is associated with twelve sulphate groups, and the air dried cell contains approximately 5 percent water by weight.

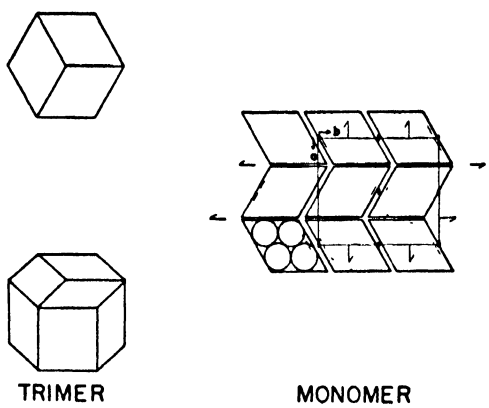
Further X-ray measurements on the wet cell, and

accurate density determination, together with accurate estimation of the water content of both wet and dry crystals, should give a more accurate value for the unit molecular weight. At this stage the preliminary study definitely establishes the presence of a molecule of molecular weight approximately 12,000 or some sub-multiple of this.

One of the alternative forms of the 36,000 unit suggested by Crowfoot (1941) is represented by a prism of hexagonal base 43 A across and height 30 A. The unit of this "trimer" is a right prism based on a 60° rhombus. Four such units appear to give a suggestive fit in the orthorhombic dry cell of the insulin monomer crystals. The height of the prism corresponds to the axial length *c*, as shown in the accompanying drawing which gives the "trimer" model chosen for this comparison and a possible packing of the insulin sulphate monomer.

The work of Sanger (*q.v.*) demonstrates the presence of four polypeptide chains in the 12,000 unit linked together by disulphide bonds. This suggests a model for the monomer unit in which four polypeptide chains approximately 10 A in diameter are close packed in the *c* plane as illustrated in the lower left hand corner of the drawing.

CRYSTALLINE INSULIN



TRIMER MONOMER

This model requires that the individual chains should be more puckered or folded along *c* than an α-keratin chain, in order to accommodate the required number of amino acid residues.

As Dr. Crowfoot pointed out this model appears, in general, to be equally applicable to the molecules in the rhombohedral 36,000 unit.

(References. Crowfoot, D. M., 1938, Proc. Roy. Soc. *164A:* 580; 1941, Chem. Rev. *28:* 215; Ellenbogen, E., 1949, Ph.D. Thesis, Harvard University.)

RANDALL: I wish to amplify Dr. Hodgkin's remarks concerning the orientation of the haem group in haemoglobin. The early work was carried out by Perutz (Nature, 1939, *143:*731). More recently, in collaboration with Barer and Jope, Perutz has used a reflecting microscope designed by Burch in the University of Bristol. This microscope is built from mirrors at least one of which is aspherized; the instrument is achromatic and can thus be used over a wide spectral range without change of focus. The instrument is specially suited to the study of small crystals and may be expected to be of value in the determination of orientation of particular molecular groups.

The absorption spectrum of the crystal is determined by projecting the enlarged image produced by the microscope onto a spectrograph slit. The maximum at 400 mμ is an absorption band of haem, which substance absorbs light when a component of the electric vector is parallel to the plane of the ring system. The marked dichroic effect shows that the haem is oriented with its plane normal to the *a*-axis. In similar manner the above authors have also shown that the plane of the indol group of tryptophane is also normal to the *a*-axis. The X-ray data on haemoglobin show the polypeptide chains to be parallel to the *a*-axis. It follows therefore that the planes of both the haem and the indol rings are normal to the polypeptide chain direction.

As a more general comment, the reflecting microscope is likely to be of great use in the study of biologically interesting crystals. Wilkins and Seeds, in my laboratory (Nature, 1949, *164:*228), have developed a simple instrument, using spherical mirrors, with a numerical aperture of 0.5; this instrument has a working distance of 9 mm. It is thus possible to introduce, for example, a cooling device on the stage, which enables crystals to be studied at low temperatures. It may be expected that valuable results will accrue from studies in the ultraviolet and from further developments in the use of infrared; work in the latter region has already been briefly reported by Thompson and Barer (Nature, 1948).

HODGKIN: I entirely agree with Professor Randall's view that these recent developments in the examination of crystal optics are of the greatest assistance in the study of structure. For example, the observation he mentions, that the plane of the indole groups of tryptophane is at right angles to the chain length in haemoglobin, limits considerably the kind of folded configuration the chains may have. And other limitations are imposed by the observations on the infrared dichroism of certain peptide systems, made by Ambrose (1949, Nature, Lond. *163:*859-862).

PROTEIN MERCAPTIDES[1,2]

WALTER L. HUGHES, JR.
Department of Physical Chemistry, Harvard Medical School, Boston, Massachusetts

The strikingly insoluble precipitates which form upon the addition of heavy metal salts to protein solutions have led to extensive investigation of these complexes during the last century (Galeotti, 1903). The lack of any demonstrable stoichiometry, or even constant composition of these complexes, no doubt markedly strengthened the concept of proteins as colloidal aggregates not markedly different from hydrated ferric oxide or other inorganic colloids. As knowledge of the complexity and lability of protein molecules increased, these early experiments appeared valueless, since the proteins used were not sufficiently pure and the conditions of interaction caused grave doubts that the original structure had been preserved. Nevertheless, some of these reagents, such as lead acetate, have occasionally proven useful in preparative protein chemistry.

At the present time the availability of purer and better characterized proteins and the greater certainty of the structure of metal complexes has permitted a reinvestigation of several of these interactions with very promising results. It has been found that by the choice of a suitable reagent, such as $(PO_3^-)_x$ (Perlmann, 1941), or HgI_3^- (Lewin, unpub.; see also Baudouin, Lewin and Hillion, 1944), a close stoichiometric relationship can be demonstrated between the ion and the total number of basic groups in the protein. Furthermore, while such saturated complexes are extremely insoluble, with smaller amounts of these reagents soluble complexes are formed which may be investigated by the methods already described by Klotz (1949). Many of these soluble complexes are stable enough to be very useful for preparative purposes. Thus, Astrup has used anionic reagents such as thiosalicylic acid in the step-wise precipitation and fractionation of proteins (Astrup and Birch-Anderson, 1947), and Cohn and collaborators (1950) have used Zn^{++} in ethanol-water systems. Many complexes may be obtained in crystalline form (Lewin, in preparation; Perlmann, 1938), making possible additional conditions for further fractionation and purification of the proteins. Protein-metal complexes should prove most valuable for structural studies of proteins in the solid state by X-ray diffraction and other means.

It would appear likely that for most of these reagents there exist many binding sites per molecule of approximately equal affinity, so that if the protein molecule is not saturated with the reagent (and this is frequently the case), its distribution over the protein molecules may follow a statistical pattern. Nevertheless, it should be possible to find reagents with markedly different affinities for different classes of binding sites so that it would be possible to saturate one type of site selectively. Protein mercaptides are an example of such specific reaction, as the mercury-sulfur bond is much stronger than the bond between mercury and any other grouping commonly found in proteins.

In view of this fact, the literature on protein mercaptides is surprisingly limited. This probably follows from the small numbers of thiol groups in proteins and the avid affinity of mercury for other groups once the thiol groups are saturated. Consequently, the early workers probably missed the first subtle changes coincident with mercaptide formation. The pharmacological role of mercaptide formation in mercurial antisepsis had been the subject of interesting speculation (Smith *et al.*, 1936; Miller and Rose, 1939; Fildes, 1940). The first stoichiometric studies of the biochemical reaction of mercurials appear to have been those of Hellerman (1937, 1943) with urease using *p*-carboxylphenylmercury chloride. Singer and Barron extended the use of this reagent to other enzymes (1945), and Anson (1941) also used it for estimating thiol groups in denatured egg albumin. Recently, Benesch has used silver mercaptide formation in analysis by amperometric titration with $Ag(NH_3)_2^+$ (Benesch and Benesch, 1948).

Hellerman's reagent with but one reactive mercury bond should be decidedly more specific than divalent mercuric salts. Nevertheless, Warburg has observed a simple stoichiometry in the isolation of enolase as a crystalline mercuric salt (Warburg and Christian, 1942). These crystals have been shown to contain one atom of mercury per mole of protein (M = 66,000) (Bucher, 1947). While no evidence has been presented, their mercaptide nature must be suspected. Kubowitz and Ott (1943) have similarly crystallized another enzyme as a mercuric salt.

[1] This paper is Number 80 in the series "Studies on Plasma Proteins" from the Harvard Medical School, Boston, Massachusetts, on products developed by the Department of Physical Chemistry from blood collected by the American Red Cross.

[2] This work was originally supported by grants from the Rockefeller Foundation and from funds of Harvard University. It was aided early in 1941 by grants from the Committee on Medicine of the National Research Council which included a grant from the American College of Physicians. From August, 1941, to July, 1946, it was carried out under contract, recommended by the Committee on Medical Research, between the Office of Scientific Research and Development and Harvard University. Since then it has been aided by a grant recommended by the Panel on Hematology of the National Institutes of Health.

For greatest specificity, it would seem desirable to have the mercurial as mono-functional as possible, and with this in mind the carboxyl group in Hellerman's reagent would seem definitely contra-indicated. Recently Bennett, following such reasoning, has developed mono-functional colored mercurials for the histological identification of sulfhydryl groups (Bennett, 1948; Bennett and Yphantis, 1948). Where colored derivatives are not needed, it would seem desirable to choose as small a molecule as possible, in view of the many reports in the literature of the difficult accessibility of some protein sulfhydryl groups. (For a recent review of this subject and its relation to protein denaturation, see Anson, 1945.) The above considerations suggest the use of methylmercury salts.

Mercaptalbumin

The development of methylmercury salts as thiol reagents has followed the isolation and characterization of mercaptalbumin, a suitable protein test reagent (Hughes, 1947 and in preparation). This fraction of serum albumin, containing one sulfhydryl group per molecule (the remaining serum albumins appearing to have no sulfhydryl groups), can be separated from the total serum albumins by crystallization as its mercuric salt. The maximal yield is obtained when an amount of mercuric chloride just equivalent to the total sulfhydryls (approximately ⅓ mole per mole protein) is added, and the protein mercaptide after repeated recrystallizations is found to contain precisely ½ atom of mercury per albumin molecule, suggesting that the unit is an albumin dimer. Ultra-centrifugal analysis of a fresh solution of these crystals confirms this hypothesis (Fig. 1), the normal albumin component ($S_{20} = 4.6$) —curve C—having been replaced by a new component ($S_{20} = 6.5$)—curve A. On standing in dilute solution, this partially dissociates to the original monomer—curve B.

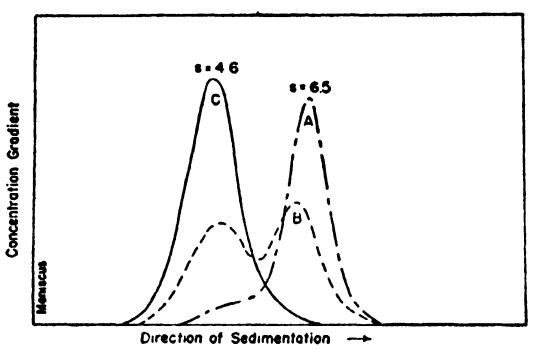


FIG. 1. Sedimentation diagrams of one per cent solutions of mercuric serum albumin at pH 7.2 in 0.05Γ/2 sodium phosphate and 0.1Γ/2 sodium chloride buffer, in the analytical untracentrifuge, 100 minutes after reaching full speed (54,000 RPM). Distance from meniscus to axis of rotation = 5.7 cm.
A = Analysis immediately following solution of the Crystals of 5 x recrystallized mercuric serum albumin.
B = Analysis of the same solution after seven days at 4°C.
C = Solution prepared similarly to (A) with the addition of one mole of mercuric chloride per mole of albumin. Curves identical with (C) have been obtained following the addition of a like amount of BAL to mercuric albumin and are characteristic of normal human serum albumin before crystallization with mercuric chloride.
(These analyses were performed by C. Gordon and P. Baker under the direction of J. L. Oncley.)

The formation of this derivative may therefore be expressed as follows:

$$\begin{array}{l} \text{AlbSH} + \text{HgCl}_2 \rightleftharpoons \text{AlbSHgCl} \\ \qquad\qquad\qquad\quad + \\ \qquad\qquad\qquad\quad \text{AlbSH} \\ \qquad\qquad\qquad\quad \uparrow\downarrow \\ \qquad\qquad\qquad\quad (\text{AlbS})_2\text{Hg} \end{array}$$

The dimerization reaction, involving the diffusion of two large molecules, is the slow reaction and can be readily followed by suitable physico-chemical methods such as nephelometry (Edelhoch and Edsall, in preparation). Both reactions are reversible so that both the kinetics and the equilibria may be followed. However, while the second reaction may be readily reversed by various reagents which form undissociated mercury complexes (*e.g.*, the halides), the first reaction has only been reversed by reagents such as cysteine, which form equally stable mercury derivatives. Thus, mercaptalbumin can be regenerated from its mercuric salt by dialysis, or by ethanol precipitation, in the presence of cysteine.

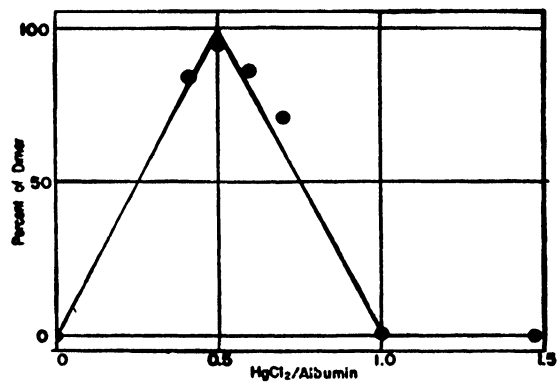


FIG. 2. Interaction between serum mercaptalbumin and $HgCl_2$ as determined by light scattering measurements. The points indicate the amount of dimer found in ten percent aqueous solutions of serum mercaptalbumin following equilibration with the indicated molar ratios of $HgCl_2$. (These measurements were carried out by H. Edelhoch.)

Mercuric albumin dimer is rapidly dissociated by the stoichiometric amount of $HgCl_2$ (½ mole/mole alb) as would be predicted from the above equation, offering presumptive evidence that mercaptalbumin contains only one available thiol group. This is beautifully illustrated by light-scattering studies on aliquots of mercaptalbumin to which have been added increasing amounts of mercuric chloride (Fig. 2). The maximal increase in turbidity occurs when

exactly 0.5 mole $HgCl_2$ has been added per mole of protein, after which the turbidity decreases, returning to its initial value when 1.0 mole of $HgCl_2$ has been added.

When dissolved at pH 10 in 5M guanidinium bromide, mercuric albumin gives no nitroprusside color, indicating that no further thiol groups are liberated on denaturation. Thus, with but a single SH group, and this readily available, situated on the surface of the molecule, mercaptalbumin would appear as an ideal model for protein mercaptide studies.

Methylmercury Iodide Reaction

Methylmercury salts react with mercaptans according to the following scheme:

$$CH_3HgX + PrSH \rightleftharpoons CH_3HgSPr + H^+ + X^-$$

The specificity can therefore be controlled by variation of the concentrations of the reactants. Thus the hydrogen ion concentration may be varied within the stability limits of the protein; the X^- ion concentration may be increased to any desired extent above that produced by the reaction alone; and the protein or the methylmercury salt concentration may be varied from near saturation to the lower limits imposed by analysis.

Furthermore, by taking advantage of the variation in degree of dissociation of several methylmercury salts, the equilibrium constant itself can be varied. For this purpose the methylmercury halides appear particularly suitable.

Thus a proper choice of the possible permutations of these variables will permit the investigation of mercaptide formation for any protein under a variety of conditions. The specificity of the reaction can be determined by studying the breadth of variation in concentrations of the reactants over which the same stoichiometry exists, and evidence for the mercaptide nature of the linkage may be obtained from determination of the equilibrium constant, in comparison with the values obtained for simpler mercaptides of known structure.

The metal-catalyzed oxidation of sulfhydryl groups (Michaelis, 1929), which, while very slow for proteins, is more troublesome in comparative studies on model substances such as cysteine, has been readily prevented by adding small amounts of ethylene diamine tetra-acetic acid as a chelating agent (Schwarzenbach, 1946). Effective concentrations (approximately 10^{-4}M) give evidence of only slight interaction with the mercurial, for which allowance can be made by appropriate control experiments. In general, estimations of protein sulfhydryl with and without this reagent have given identical results.

The determination of CH_3HgX concentration at equilibrium has been accomplished by taking advantage of the favorable distribution coefficient of the halides between water and immissible organic solvents. Toluene has proven particularly useful because of its low volatility, low solubility in water, and the relatively high solubilities of CH_3HgX in it.

Equilibration between the organic phase and the aqueous phase has been accomplished with a minimum of emulsification by slow, horizontal rotation of a stoppered bottle about its long axis, permitting the two phases to be in contact over a large area. If the inner surface of the bottle is entirely wetted with the protein phase before the organic phase is added, the latter will roll smoothly over the former without breaking the glass-water interphase.

Under favorable conditions equilibrium is established within thirty minutes. The attainment of equilibrium can be verified by approach from opposite sides. An experiment in which methylmercury halide is added in the organic phase is compared with one in which a highly dissociated methylmercury salt, such as CH_3HgNO_3, and the corresponding alkali halide, are added to the aqueous phase. The same result is attained by either procedure.

Dithizone Titration

Methylmercury iodide, at a concentration of approximately 10^{-4}M, may be conveniently estimated in organic solvents by titration with dithizone ($\O NH \cdot NH \cdot CS \cdot N{:}N\O$) provided a proton acceptor of the proper affinity is present to remove the hydrogen ion formed in the reaction. For this purpose, amyl amine in the presence of an excess of acetone has proved satisfactory. At the end point, the color changes quite sharply from orange-yellow to green when an excess of free dithizone is present. The most sensitive end point (the first appearance of a greenish hue) may be reproduced with an accuracy of 1 percent by comparison with a standard. Since the end point is rather unstable, it has proved best to titrate in the cold.

Equilibrium Constant

Mercaptalbumin, known to contain one sulfhydryl group by independent methods, as already described, has been investigated by the above technique and the equilibrium constant determined (Table 1). The measurements were made at room temperature for convenience in pH measurements. Known amounts of the reactants were added:—mercaptalbumin, NaI, H_2SO_4 or NaOH to give the desired pH, and water to a constant volume of 2 cc., followed by 2 cc. of CH_3HgI in toluene. After equilibration, the concentration of CH_3HgI in the toluene phase was determined. From these data, the concentrations of all the other reactants were calculated (except the hydrogen ion which was measured after equilibration with a glass electrode). The non-specific binding of CH_3HgI was determined under similar conditions using mercaptalbumin which had been treated with 3 equivalents of iodine to oxidize the sulfhydryl group (Hughes and Straessle, 1950). (After this treatment, the nitroprusside reaction was negative.) The non-specific binding was small and proportional to the CH_3HgI concentration and so was treated as an altered distribution coefficient.

TABLE 1. THE EQUILIBRIUM CONSTANT FOR THE REACTION: $CH_3HgI + AlbSH \rightleftharpoons CH_3HgSAlb. + H^+ + I^-$

Concentrations at equilibrium ($M \times 10^4$)					Equilibrium constant ($-\log_{10}$)	
Toluene phase CH_3HgI	Aqueous phase				pK	pK corr.*
	$CH_3HgSAlb.$	AlbSH	I^-	pH		
2.8	2.0	2.5	12	4.71	4.18	4.53
3.8	1.1	3.5	51	4.80	4.18	4.41
4.1	0.8	3.8	250	5.44	4.33	4.45
3.1	1.8	2.8	52	5.30	4.27	4.44
2.5	2.4	2.2	53	5.73	4.38	4.50
1.8	3.1	1.4	54	6.15	4.34	4.43

* Corrected for iodide bound to protein as explained in text.

From these data the log of the equilibrium constant has been calculated. In this calculation, the activities of the protein and its mercaptide were considered as equal to their molar concentrations; the hydrogen ion activity was taken as the pH determined by the glass electrode; the methylmercury iodide activity was taken as equivalent to its concentration in the toluene phase. (When reliable values for the distribution coefficient of CH_3HgI between water and toluene exist, it would seem desirable to replace this last by its concentration in the aqueous phase, so that the equilibrium constant would give a truer picture of the actual molecular distribution.)

The iodide ion presented a special problem. In the first calculation of the equilibrium constant, its activity was considered equal to its concentration as calculated from the amount of NaI added plus the amount of I^- formed by the reaction. In the presence of albumin the activity of I^- is reduced by binding to the protein, in addition to the effect of interionic attraction. The concentration of free unbound iodide was kindly calculated by Scheinberg from studies on ion binding (Scatchard and Black, 1949; Scatchard, Scheinberg and Armstrong, 1950a and b) and this concentration was multiplied by the appropriate activity coefficient. The last column of Table 1 gives the pK recalculated with these corrections. It will be seen that less variation in the equilibrium constant now occurs. A similar determination of the equilibrium constant for cysteine-mercaptide formation gave a constant differing from the above by only 0.1 in logs. This offers a strong confirmation for the mercaptide nature of the reaction.

STOICHIOMETRY

From the equilibrium constant, the conditions necessary for quantitative reaction can readily be calculated. With concentrations of reagents similar to those used in the above study (10^{-3} to 10^{-4}M), it appears that, for stoichiometric reaction, measurements must be made above pH 7. Table 2, giving comparative studies at pH 7.3 and 9.7, indicates that this condition is sufficient.

TABLE 2. STOICHIOMETRY

Amount of CH_3HgI bound by a constant amount of mercaptalbumin under varying conditions

Excess CH_3HgI at end point	pH	CH_3HgI bound
6%	9.7	(1.00)
20	9.7	1.00
7	7.3	.98
16	7.3	.97
23	7.3	.97

It should be possible to force the reaction to completion at a lower pH by the use of higher activities of CH_3HgI. This can be done without decrease in precision of measurement by using a smaller volume of organic phase or a poorer organic solvent such as octane. However, when this was attempted, a nonspecific reaction apparently involving the adsorption of CH_3HgI by the protein occurred. An alternative method of studying sulfhydryl groups at lower pH would involve shifting the equilibrium constant by the use of the more dissociated CH_3HgBr.

OTHER PROTEINS

The rapid dimerization of serum albumin with mercuric chloride indicates that its sulfhydryl grouping must be on the protein surface and readily available. However, other workers, as reviewed by Anson (1945), have shown that in some proteins the thiol groups can be detected only with difficulty. Preliminary studies on other proteins with methylmercury iodide have shown a similar effect. Thus, crystallized human oxyhemoglobin has been found to react readily at pH 7.5 with 2 moles of methylmercury iodide per mole of hemoglobin. However, with this protein, equilibration takes appreciably longer than with mercaptalbumin. This is not the case when equilibrium is approached from the opposite side (*i.e.*, by the addition of aqueous CH_3HgNO_3 and NaI to the protein phase followed by equilibration with pure toluene). With crystallized ovalbumin, the effect is even more marked: while CH_3HgNO_3 and NaI still react rapidly, the attainment of equilibrium, starting with CH_3HgI in toluene requires a day or longer.

Therefore, it would appear that the thiol groups in native ovalbumin, previously shown to be reactive to iodine (Anson, 1945) and to chloropicrin (Fredericq and Desreux, 1947), are also reactive to methylmercury salts. A further investigation of this mild, specific reaction with this and other proteins should yield further information as to the nature and "availability" of protein sulfhydryl groups. Its

reversibility not only permits more definite characterization of the thiol groups, but equally important, it permits the restoration of the original protein so that the absence of denaturative changes may be proven.

It is hoped that the study of protein mercaptides, in addition to contributing to the knowledge of protein structure, may serve as a model for the study of other protein groups, where the greater numbers involved and the less specific reactions available make both the experimental investigations and their interpretation more difficult.

References

Anson, M. L., 1940, The reactions of iodine and iodoacetamide with native egg albumin. J. Gen. Physiol. *23:* 321-331.

1941, The sulfhydryl groups of egg albumin. J. Gen. Physiol. *24:* 399-421.

1945, Protein denaturation and the properties of protein groups. Advances in Protein Chemistry *2:* 361-386. New York: Academic Press.

Astrup, T., and Birch-Anderson, A., 1947, Fractional precipitation of serum proteins. Nature, Lond. *160:* 637.

Benesch, R., and Benesch, R. E., 1948, Amperometric titration of sulfhydryl groups in amino acids and proteins. Arch. Bioch. *19:* 35-45.

Bennett, H. S., 1948, Abstract. Anat. Rec. *100:* 640 (Suppl.).

Bennett, H. S., and Yphantis, D. A., 1948, 1-(4-chloromercuriphenylazo)-naphthol-2. J. Amer. chem. Soc. *70:* 3522.

Bucher, T., 1947, Über das Molekulargewicht der Enolase. Biochimica Biophysica Acta *1:* 467-476.

Cohn, E. J., Gurd, F. R. N., Gillespie, J. M., Mittelman, D., Derouaux, G., Mouton, R. F., Liu, C. H., Uroma, E., Lever, W. F., Kahnt, F. W., Barnes, B. A., Brown, R. K., Schmid, K., and Surgenor, D. M., Quantitative Procedures for the Separation of Protein Components. I. A System for the Separation of the Plasma Protein Components of Human Blood. J. Amer. Chem. Soc. In press.

Edelhoch, H. and Edsall, J. T., in preparation.

Fildes, Paul, 1940, Mechanism of the antibacterial action of mercury. Brit. J. exp. Pathol. *21:* 67.

Fredericq, E., and Desreux, V., 1947, Recherches sur les fonctions sulfurées des protéines II. Bull. Soc. Chim. Biol. *29:* 100-105.

Galeotti, G., 1903, Über die sogenannten Metallverbindungen der Eiweisskörper nach der Theorie der Chemischen Gleichgewichte. Physiol. Chem. *40:* 492-549.

Hellerman, Leslie, 1937, Reversible inactivations of certain hydrolytic enzymes. Physiol. Rev. *17:* 454-484.

Hellerman, Leslie, Chinard, F. P., and Deitz, V. R., 1943, Protein sulfhydryl groups and the reversible inactivation of the enzyme urease: The reducing groups of egg albumin and of urease. J. biol. Chem. *147:* 443-462.

Hughes, W. L., Jr., 1947, An albumin fraction isolated from human plasma as a crystalline mercuric salt. J. Amer. chem. Soc. *69:* 1836.

Methylmercury iodide as a reagent for thiols, in preparation.

Hughes, W. L., Jr., and Straessle, R., Preparation and Properties of Serum and Plasma Proteins. XXIV. Iodination of Human Serum Albumin. J. Amer. Chem. Soc., in press.

Klotz, I., 1949, The nature of some ion-protein complexes. Cold Spr. Harb. Symposium Quant. Biol. *14:* 97-112.

Kubowitz, F., and Ott, P., 1943, Isolierung und Kristallization eines Gärungsfermentes aus Tumoren. Biochem. Zt. *314:* 94-117.

Lewin, J., unpublished (see also Baudouin, A., Lewin, J., and Hillion P., 1944, Bull. Soc. Chim. Biol. *26:* 123. Preparation and properties of serum and plasma proteins XXVIII. Crystalline derivatives of serum albumin and of certain other proteins. In preparation.

Michaelis, L., 1929, The mechanism of the catalytic effect of iron on the oxidation of cysteine. J. biol. Chem. *84:* 777-787.

Miller, R. E., and Rose, S. B., 1939, The influence of agar on mercurial antiseptics. J. Bact. *38:* 539.

Perlmann, G. E., 1938, On the preparation of crystallized egg albumin metaphosphate. Biochem. J. *32:* 931-932.

1941, Combination of proteins and metphosphoric acid. J. biol. Chem. *137:* 707-711.

Scatchard, G., and Black, E. S., 1949, The effect of salts on the isoionic and isoelectric points of proteins. J. phys. coll. Chem. *53:* 88-89.

Scatchard, G., Scheinberg, I. H., and Armstrong, S. H., Jr., 1950a, Physical chemistry of protein solutions, IV. The combination of human serum albumin with the chloride ion. J. Amer. chem. Soc. *72:* 535.

1950b, Physical chemistry of protein solutions, V. The combination of human serum albumin with the thiocyanate ion. J. Amer. chem. Soc. *72:* 540.

Schwarzenbach, G., 1946, New methods for determining certain cations. Helv. Chim. Acta *29:* 1338.

Singer, T. P., and Barron, E. S. G., 1945, Sulfhydryl enzymes in fat and protein metabolism. J. biol. Chem. *157:* 241-253.

Smith, Dorothea E., Czarnetzky, E. J., and Mudd, Stuart, 1936, The mechanism of inactivation of mercurial antiseptics by serum, and its implications regarding the possibility of intravenous antisepsis. Amer. J. med. Sci. *192:* 790.

Warburg, O., and Christian, W., 1942, Isolierung und Kristallization des Gärungsferments Enolase. Biochem. Zt. *310:* 384-421.

Discussion

Benesch: I would like to mention several points which have arisen from the work of Mrs. Benesch and myself in our study of protein sulfhydryl groups by the amperometric silver method. Like Dr. Hughes, we have been impressed by the high degree of stoichiometry in protein mercaptides. It was found, for example, that treatment of solutions of serum albumin with standard *p*-chloromercuribenzoate always diminished the amount of silver required for complete reaction by the calculated amounts. It is also of interest that we obtain the same result on bovine serum albumin by the silver method as Dr. Hughes with methyl mercuric iodide, that is, one sulfhydryl group per mole; we also agree as to the easy accessability of the —SH groups of serum alubumin which is, for example, illustrated by the fact that we observe no difference in the —SH groups titratable in water and in various concentrations of ethanol. This is in contrast

to results obtained with egg albumin where a linear increase of titratable —SH groups was observed with increasing ethanol concentrations. It was possible, however, to discern a distinct, if slow, release of —SH groups by the silver method even in water. In view of Dr. Hughes' finding that human oxyhemoglobin contains two —SH groups per mole, it may be mentioned that we found recently that dog oxyhemoglobin contains four such groups per molecule. In conclusion, I would like to draw attention to a recent publication by Schoberl, who describes a new method for the introduction of sulfhydryl groups into proteins using polythioglycolides. This procedure should be of great value since it might permit the construction of protein sulfhydryl models of more or less known structure *in situ*.

LOCALIZATION OF CELLULAR PROTEINS BY ENZYMATIC HYDROLYSIS[1]

BERWIND P. KAUFMANN, HELEN GAY AND MARGARET R. McDONALD
Department of Genetics, Carnegie Institution of Washington, Cold Spring Harbor, N.Y.

Identification of types of cellular proteins and determination of their patterns of distribution and association with nucleic acids may be regarded as essential to an understanding of the processes involved in duplication of the living cell. A more specific evaluation of these processes with respect to replication of the gene complex requires precise information concerning the organization of the chromosome and the arrangement of its constituent materials during the various phases of mitosis. Inferences concerning the chemical composition of chromosomes derived from the earlier observations of descriptive cytologists have been supplemented in recent years by information provided by a variety of experimental procedures. Descriptions of patterns of linear organization in terms of achromatic and chromatic materials have accordingly been replaced by descriptions in terms of proteins and nucleic acids. The total information available, however, is still too fragmentary to permit precise definition of all types of constituent materials and their patterns of association. In an attempt to supplement the essential basic information, studies were initiated about three years ago using purified enzymes in combination with various staining procedures to determine the distribution and interrelations of proteins and nucleic acids in plant and animal cells (Kaufmann *et al.*, Carnegie Institution of Washington Year Books Nos. 45-47). This relatively simple experimental procedure is based on the premise that precise information concerning fundamental patterns of cellular organizaion can be obtained if the enzymes used are free of contaminants possessing in themselves hydrolytic activity, and if an adequate series of controls is maintained to measure the influence of all variables capable of modifying the cytochemical reactions. In essence, this approach is an extension of methods of experimental cytology used during the later years of the past century and the early years of the present one, with substitution of purified enzymes for the crude extracts then available, and maintenance of more rigid control of experimental variables. This article presents experimental evidence that these purely qualitative procedures serve as valid guides to the location of proteins and nucleic acids as they exist independently and in combination in the cell.

In this connection no effort should be spared to emphasize the fact that the chemical conditions existing in living cells can only be inferred from studies on fixed preparations. Nevertheless, cytologists have accumulated in the course of years a wealth of information indicating that certain fixatives preserve cellular components without gross distortion of the spatial relations that exist in the living state (for example, see the dark-field studies of Strangeways and Canti, 1927, of the reaction of living cells to various fixing agents). Observations of sections or smears of preserved cells have provided the basis of factual material on which present concepts of cellular organization and chromosome behavior are largely formulated. It now remains to be determined to what extent these fixed preparations may be utilized in analysis directed toward an understanding of submicroscopic patterns of organization. One method of analysis at this level is the enzymatic approach that we have adopted. Studies using this method have been carried out on smears and sections of plant and animal tissue (root tips and anthers of Allium, Lilium, Tradescantia, Trillium, and Vicia; salivary-gland chromosomes of Drosophila and Chironomus; spermatogenous cells of Chortophaga and other Orthoptera) preserved in various cytological preservatives (fluids of Carnoy, Flemming, Navashin, Bouin, Zenker, and Helly). The enzymes used were trypsin (McDonald and Kunitz, 1946), chymotrypsin (Kunitz and Northrop, 1935), pepsin (Northrop, 1946), ribonuclease (McDonald, 1948), and desoxyribonuclease (Kunitz, 1948).

The validity of this approach lies in the specificity of action of the enzyme preparations utilized. Since the proteases and nucleases used in early experiments were crude extracts, it was not possible to formulate unequivocal interpretations concerning the nature of the substrate materials attacked. With the crystallization of pepsin by Northrop in 1930, and the crystallization in succeeding years of other proteases, and also of nucleases, by Kunitz and his associates, agents having more specific properties were made available for cytochemical studies. These enzymes have been used extensively in a number of studies during the past ten years.

In evaluating these results it appears that too often crystallinity *per se* has been accepted as a valid criterion of purity of an enzyme, although proteins are noted for the ease with which they form solid solutions or are adsorbed on each other. With few exceptions, the enzymes were used cytochemically without the preliminary chemical assay which would determine whether impurities capable of modifying the reactions were present. Without such assay it is apparent that no unequivocal interpretation could be formulated concerning the nature of

[1] Progress of these studies was facilitated by a grant from the National Institutes of Health, U.S. Public Health Service.

the substrate the enzyme was attacking. Crystalline ribonuclease (Kunitz, 1940), for example, is one of the enzymes that has been used most extensively in recent years in cytochemical studies of nucleoproteins. Kunitz, in his original publication, noted the possible presence of small amounts of impurities in the sample of ribonuclease used in his solubility tests. Despite this warning, the enzyme was used extensively in cytochemistry without determination of the nature of these impurities and their possible influence in modifying the results obtained. When it was demonstrated by Cohen (1945) and Schneider (1946) that some samples of crystalline ribonuclease carry measurable traces of proteolytic activity, it was evident that precise information concerning the cytochemical action of this enzyme could be obtained only when these contaminants were removed. In the studies reported here, the various enzymes used have been carefully assayed for all possible interfering impurities. Only when such assays have been negative have the enzymes been used cytochemically.

The results obtained using these enzymes serve to refute the contention that their applicability in cytochemistry is promising in theory but of little practical importance. Granting the objection that complete purity of an enzyme cannot be demonstrated (Danielli, 1946, 1947), it is nevertheless possible to purify an enzyme with respect to its specificity of action on one of two alternative classes of substances; thus, nucleases can be freed of measurable traces of proteases—as determined by chemical criteria involving careful assay—so that they do not degrade proteins, and proteases can be freed of nuclease activity so that they do not degrade nucleic acids. Another conceivable difficulty in the cytochemical application of enzyme preparations—that their access to underlying strata might be blocked by superficial layers of resistant materials—has proved to a large extent illusory, as can be demonstrated by a comparison of smears and sections of the same type of cell.

Fulfillment of the second requirement of this experimental approach, namely, the control of all variables capable of modifying the cytochemical reactions, is more difficult of attainment. In such experiments two sets of conditions must be carefully controlled: those capable of influencing the process of enzymatic hydrolysis, and those modifying the staining reactions. These two procedures depend in turn on a variety of factors, some of which are indicated in Table 1.

Replication of results depends to a large measure on maintenance of rigid experimental conditions and reduction of variability in histological procedures. Time schedules must be faithfully maintained in preparing material for hydrolysis and during the postdigestion staining processes. Stainability of enzyme-treated tissues will depend on the action of the enzyme in "removing" stainable substances, or in altering their capacity for combination with the dyes. However, other conditions that are independent of the action of the enzyme may influence colorability. Electrolytes can either intensify or reduce staining reactions markedly (Stowell and Zorzoli, 1947; Kaufmann *et al.*, 1947). Hence, it is important that effects such as these, likely to mask the true enzymatic action, be eliminated if possible. Our experiments (with the exception of those utilizing pepsin) have accordingly been carried out to a large extent by using salt-free enzymes in aqueous solutions at pH 6. Even though water disturbs substrate materials less than many other solvents, it reduces basophilia when used at temperatures approaching or exceeding 60° C (see Brachet, 1940, and Panijel, 1947). Consequently, all enzymatic hydrolyses were performed at 37° C. For those

TABLE 1. CONDITIONS INFLUENCING ENZYMATIC HYDROLYSIS AND STAINING REACTIONS

1. Temperature	6. Condition of Substrate
2. Time	a. Fixation
3. Concentration	b. Affixative
4. Stability	c. Physiological state
5. Activators	d. Competitors
7. pH	
8. Solvent	

enzymes whose optimum temperature or pH are not met by these conditions, an increase in concentration or prolongation of the digestion period is necessary. However, if an enzyme is unstable under these conditions, it is useless to lengthen the time of hydrolysis unless a check is made chemically to determine the rate of loss of activity. Performance of such tests as routine procedures during the course of enzymatic hydrolysis has enabled us to renew the solutions whenever 25 percent of the original activity had been lost. For some enzymes, such as trypsin this occurs rapidly, in less than one hour, whereas in others, such as ribonuclease, activity is maintained for twelve hours or longer.

It has long been known that the nature of the chemicals used in fixing tissues affects profoundly the course and extent of enzymatic hydrolysis (Stowell and Zorzoli, 1947; Tulasne and Vendrely, 1947). Mixtures of acetic acid and alcohol (Carnoy's fluid) preserve nucleic acids and proteins in such condition that they are readily degraded by proteases and nucleases. All of our experiments have involved as routine procedure digestion of materials fixed in Carnoy's fluid. This fixative is inadequate, however, for preserving details of chromosome structure especially when tissues are embedded subsequently in paraffin. For such studies a mixture of chromic, osmic, and acetic acids (Flemming's fluid) has well-established superiority, but it does modify the substrate so that hydrolysis is only effected by prolonged digestion with high concentra-

tions of the enzyme. Materials preserved in Flemming's fluid—and in other standard fixatives such as those of Navashin, Bouin, Zenker, and Helly—have been used to supplement and augment the observations obtainable from the Carnoy-preserved tissue.

After the period of enzymatic hydrolysis, experimental and control preparations in any single experiment were stained simultaneously in dishes designed to carry a series of slides or cover slips. Subsequently the stained sections were rinsed thoroughly over long periods of time (12 to 24 hours) to insure removal of all dye that was not firmly combined in the tissues (following the recommendation of Michaelis, 1947). Comparisons of color values to determine the effect of enzymatic hydrolysis were usually made between adjoining sections of the same cell, or, if that was not possible, between adjacent sections of the same piece of tissue.

An adequate appraisal of these many variables indicates that observable differences in stainability between enzyme-treated and control cells are due to the specific action of the enzyme. Experimental application of the method of enzymatic hydrolysis not only has confirmed results obtained using other cytochemical and chemical procedures, but also has furnished additional information, especially with respect to the *in situ* localization and patterns of association of cellular materials during the various phases of mitosis (Kaufmann, McDonald, and Gay, 1948; Kaufmann *et al.*, 1948). This statement is not meant to imply that the method is self-sufficient. Each of the cytochemical and chemical procedures commonly employed has its limitations and advantages with respect to identification and localization of cellular components (Danielli, 1946; see also the discussions in Volume 12 of these Symposia, 1947, and the article by Danielli in this volume). Results obtained by applying the method of enzymatic hydrolysis to analysis of chromosome structure are indicated in succeeding paragraphs.

Proteins of the Chromosome

In an approach to an understanding of the nature and specificity of action of genes in higher organisms, attention has been directed to the organization of the chromosome. It has often been suggested that the fundamental structural component of the chromosome is protein. This concept stems in part from earlier observations of descriptive cytologists that the basophilic material diminishes, and often seems to disappear, at certain stages in the cycle of mitosis, whereas the oxyphilic component persists to maintain structural continuity. Use of a crude nuclease preparation obtained from spleen (Mazia and Jaeger, 1939) reduced Feulgen-stainability of salivary-gland chromosomes of Drosophila but did not impair their coloration with the ninhydrin reagent. Conversely, treatment with solutions of trypsin (Caspersson, 1936; Mazia, 1941) led to disruption of salivary-gland chromosomes. Such results suggest that the chromosome possesses a continuous protein framework, the integrity of which does not depend upon the presence of nucleic acid (Mazia and Jaeger, 1939; Mirsky, 1943), or that "the reticulum consists of a protein framework in which varying quantities of nucleic acids are embedded" (Frey-Wyssling, 1948). Darlington (1947) notes that "during the cycle of mitosis, desoxyribose nucleic acid is attached at specific loci to the polypeptide chain which represents the permanent chromosome fibre." Experiments using purified enzymes should provide further evidence concerning the nature of these proteins, and the dependence of structural continuity on their presence.

Digestion by trypsin: That treatment with purified trypsin can lead to cellular dissolution is shown in Figure 1. Two adjoining sections of a root tip of onion fixed in acetic alcohol are represented. The section to the left, which served as the control, remained in 0.05 M phosphate buffer (postassium hydrogen phosphate-potassium dihydrogen phosphate) pH 6, at 37° C for a period of 15 minutes, during which the section shown on the right was digested at the same temperature with trypsin at pH 6 (0.1 mg/ml in 0.05 M phosphate buffer). Both sections were then washed thoroughly in water, stained simultaneously by the Feulgen reaction, and counterstained in fast green. In the enzyme-digested section there is evidence of nuclear dissolution, as indicated by the dispersal of the Feulgen-positive desoxyribonucleic acid. Prolonged action of the enzyme under these conditions leads to more complete breakdown of structure, so that cell walls alone are discernible. Mazia, Hayashi, and Yudowitch (1947) attribute such disintegration to the action of trypsin on peptide linkages involving basic amino acids, which would be abundant in histones and present in more acid proteins. They also report that tests on both soluble and fibrous protein substrates confirm the expectation that trypsin will digest both histone and nonhistone proteins.

Comparable action of chymotrypsin on cells of an onion root tip fixed in acetic alcohol is shown at the right in Figure 2. The section represented in this photograph was digested in chymotrypsin (0.1 mg/ml in 0.05 M phosphate buffer at pH 6 for 15 minutes at 37° C, whereas the control shown at the left remained during this period in the buffer at the same temperature. Differences with respect to potency and rate of action were detected between solutions of trypsin and chymotrypsin of the same concentration, but the ultimate effect of digestion in each case was complete dissolution of cellular contents.

In all these cases the enzymes were used in the presence of electrolytes. Because of the modifying action of buffers on cytochemical reactions—as indicated on page 86—it seemed essential to determine whether electrolytes play a major role in the process

of disintegration effected by trypsin or chymotrypsin. The results obtained with trypsin serve to illustrate the findings.

When a section of an onion root tip fixed in acetic alcohol was digested in an aqueous solution of trypsin (0.1 mg/ml in water at pH 6) for one hour at 37° C, there was no perceptible disruption of nuclear organization, or loss of Feulgen-stainable material, as compared with an adjoining section treated with water at pH 6 for the same length of time at the same temperature (Fig. 3). It is thus apparent that in the absence of electrolytes (except for the traces necessary to adjust the pH) trypsin does not lead to cellular disintegration. This finding immediately raises the question whether trypsin is active under such conditions. That it is, is suggested by the observation that cells treated with aqueous solutions of trypsin show increased affinity in chromosomes, nucleoli, and cytoplasm for the basic dye, pyronin. One possible explanation of this result (as will be apparent subsequently) is that trypsin effects a degradation of ribonucleoprotein to provide an increase in ribonucleic acid, which is stainable with pyronin. Aqueous solutions of trypsin have also been found to degrade protein released by the action of ribonuclease (see page 89). Another line of evidence that aqueous solutions of trypsin are enzymatically active is provided by experiments in which sections were placed in veronal or phosphate buffer, or sodium chloride solutions, after digestion in aqueous trypsin (0.1 mg/ml for 1 hour). Under these conditions cellular dissolution occurs, as is shown in Figure 4. The rate of the reaction is dependent on the concentration of salt, proceeding progressively slower in 0.033 M, 0.011 M, and 0.0037 M, than in 0.1 M sodium chloride. If, however, the use of either buffer or salt solution precedes digestion in aqueous trypsin, there is no gross disturbance of cellular morphology, and the cells have an increased affinity for basic dyes (Fig. 4).

Since in all these experiments the sections were washed thoroughly in distilled water before staining, it remained to be determined to what extent water was facilitating the dissolution process. Some light was cast on this question by observing salivary-gland chromosomes of Chironomus with the phase-contrast microscope. Glands fixed in 45 percent acetic acid were rinsed in several changes of distilled water, flattened beneath a cover, and then digested at 37° C in a 0.1 percent aqueous solution of trypsin at pH 6, which was continuously renewed over a period of two hours. Such treatment did not lead to any perceptible distortion of the pattern of banding of the chromosomes. When phosphate buffer replaced the trypsin solution, there was an immediate swelling of the chromosomes and lateral separation of the bands into component chromomeres. If the buffer was then replaced by aceto-orcein (a solution of orcein in 45 percent acetic acid) the chromosomes contracted to reveal again the precise pattern of banding that characterizes the individual chromosomes. If, on the other hand, thorough rinsing in water followed the phosphate buffer, there was marked swelling and interchromosomal adhesion. Subsequent staining with acetic orcein revealed only shrunken, distorted residues of disintegrating structures. It thus appears that the dissolution commonly associated with the action of trypsin is attributable in large measure to the removal by water of the degradation products of digestion after their dissociation by salts. An additional demonstration of this effect is given in Figure 5. The two sections of a root tip of onion fixed in acetic alcohol shown in this figure were digested simultaneously in aqueous trypsin, then immersed in phosphate buffer. The section on the left was then transferred to 0.1N hydrochloric acid before washing in water, whereas the section shown on the right was washed in water before transferring to hydrochloric acid. The latter sequence leads to loss of nuclear contents, whereas no such loss occurs when acid precedes washing in water. The significance of these results with respect to the action of trypsin on histone and nonhistone proteins remains to be determined, but from these preliminary experiments it appears that the action of trypsin in itself does not lead to disintegration of the chromosome.

Digestion by pepsin: Pepsin is another proteolytic enzyme that has been used extensively in chemical and cytochemical studies. Miescher (1871) used pepsin-hydrochloric acid to free nuclei of cytoplasmic material. This was a crude extract, as were also those used by Kossel and Matthews (1898) in their chemical studies, and by Zacharias (1909), Jörgenssen (1913) and others in the earlier cytochemical applications. No conclusions about the specificity of action of this enzyme could be derived from these studies.

Solutions of purified pepsin in hydrochloric acid effect considerable shrinkage of salivary-gland chromosomes without, however, impairing their structural continuity (Mazia and Jaeger, 1939; Mazia, 1941; Mazia *et al.* 1947; Frolova, 1944). Figure 6 shows the progress of digestion of salivary-gland chromosomes of Chironomus in a solution of crystalline pepsin (3 mg/ml in 0.02N HCl at pH 1.6) as detected with the phase-contrast microscope. Digestion by pepsin reduced these chromosomes in two hours at 37° C to about one-tenth of their original volume. No corresponding shrinkage was seen in the boiled-pepsin controls, or in other controls placed in either 0.02N hydrochloric acid or 0.3 percent egg albumin dissolved in 0.02N hydrochloric acid. Mazia (1941) attributed such shrinkage to removal of a "matrix protein" containing many acidic groups. Mirsky (1943) suggested that shrinkage could also be explained by the breakdown of histone to histopeptone, as reported by Kossel (1906). The fundamental question to be resolved in interpreting these digestion experiments with pepsin

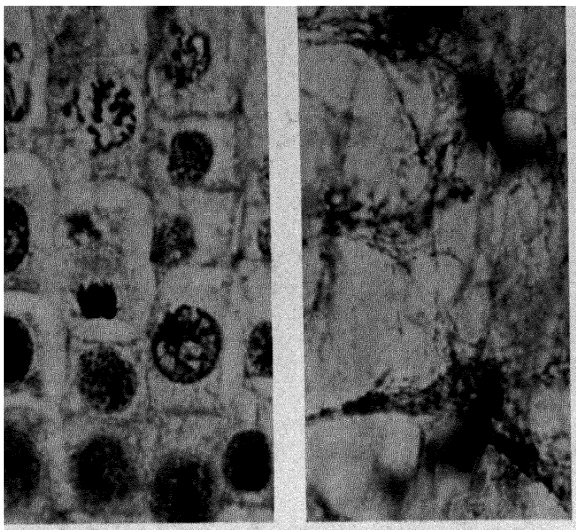


FIG. 1. Nuclear disintegration following use of crystalline rypsin in buffer. Stain: Feulgen and fast green. Left: phosphate buffer control. Right: crystalline trypsin dissolved n buffer.

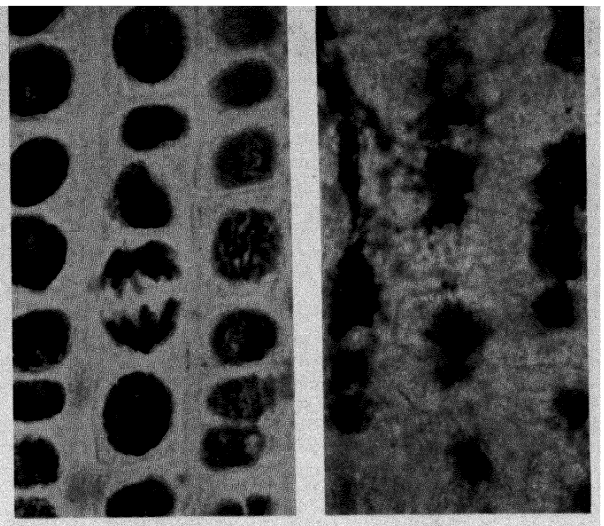


FIG. 2. Nuclear disintegration following use of crystalline chymotrypsin in buffer. Stain: Feulgen and fast green. Left: phosphate buffer control. Right: crystalline chymotrypsin in buffer.

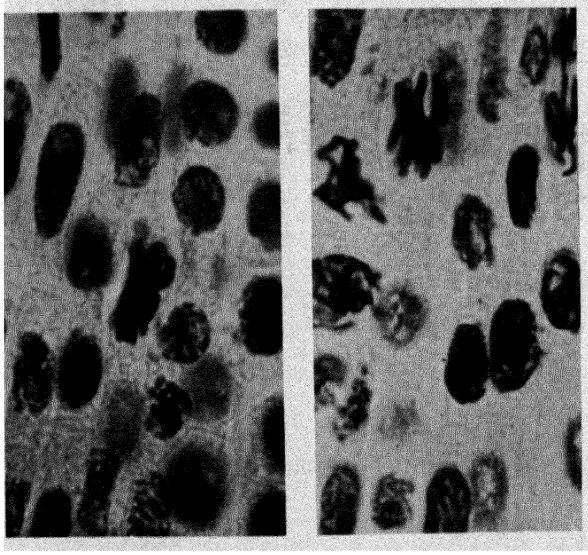


FIG. 3. Absence of nuclear distintegration following use of aqueous solution of trypsin. Stain: Feulgen. Left: water control. Right: trypsin in water.

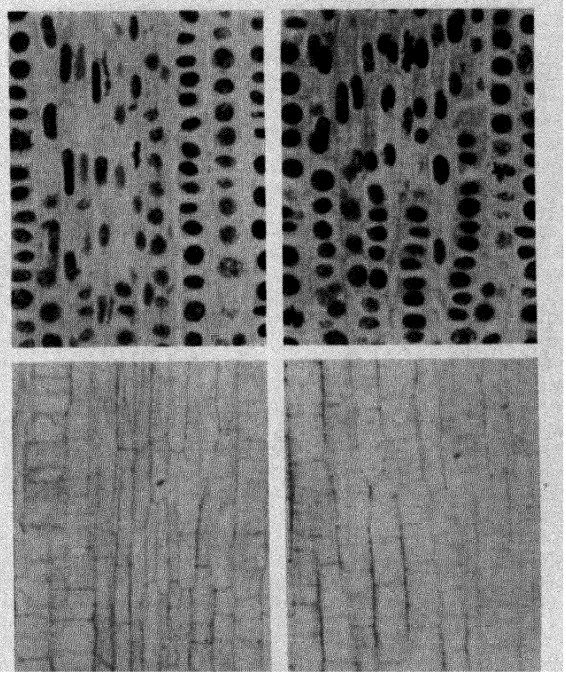


FIG. 4. Cellular dissolution produced by digestion in aqueous solution of trypsin followed by treatment in 0.05 M phosphate buffer. Upper left: buffer control. Lower left: crystalline trypsin in buffer. Upper right: buffer followed by trypsin in water. Lower right: trypsin in water followed by buffer.

FIGS. 1-4. *Allium cepa*, l.s. root tip meristem, 7 μ sections. Acetic-alcohol fixation. Prints made from Kodachrome transparencies.

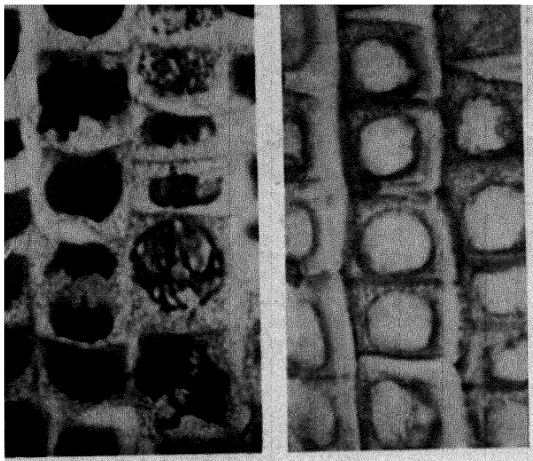


FIG. 5. Effect of water after treatment with aqueous solution of trypsin followed by buffer. Stain: methyl green-pyronin. Left: trypsin in water, followed by phosphate buffer, followed by 0.1N HCl, followed by water. Right: trypsin in water, followed by buffer, followed by water, followed by acid.

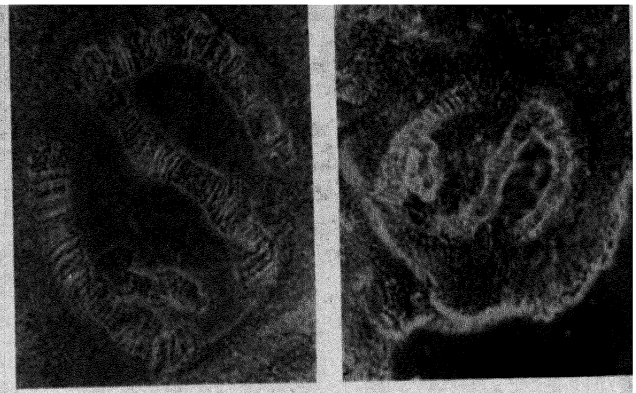


FIG. 6. Chromosomal shrinkage caused by crystalline pepsin i 0.02N HCl. Left: before treatment. Right: after 2 hours in pepsi solution.

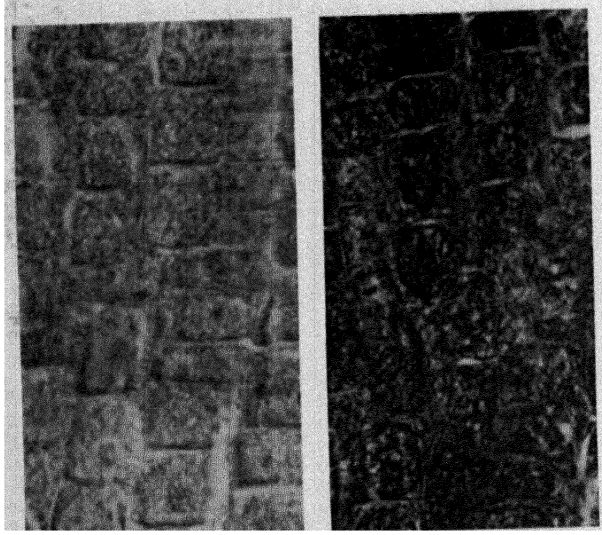


FIG. 7. Intensified stainability with acidic dyes after digestion with ribonuclease. Stain: acid fuchsin. Left: water control. Right: ribonuclease treated.

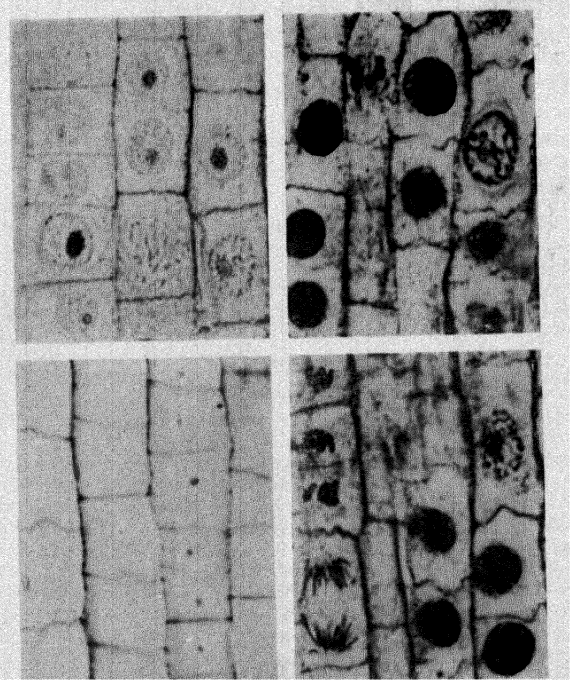


FIG. 8. Intensification of stainability with acidic dyes after extraction of nucleic acids with hot trichloracetic acid; and subsequent nuclear dissolution effected by pepsin. Stain: fast green. Upper left: TCA, 0° C, 15 mins. Upper right: TCA, 90° C, 15 mins. Lower left: TCA, 90° C, 15 mins. followed by pepsin, 2 hrs. Lower right: TCA, 90° C, 15 mins. followed by boiled pepsin, 2 hrs.

FIGS. 5, 7, 8. *Allium cepa*, l.s. root tip meristem, 7 μ sections. Acetic-alcohol fixation. Prints from Kodachrome transparencies. FIG. 6. Chironomus salivary-gland chromosomes, squash preparation in 45% acetic acid, unstained, photographed using phase-contrast microscope.

is whether under the conditions employed histones as well as nonhistone proteins are being removed from the cell. It was suggested by Mazia and Jaeger (1939) that histopeptones might be expected to combine with nucleic acids and to remain insoluble. More recently Mazia *et al.* (1947) have reexamined this problem and report that pepsin fails to digest fibers of histone or nucleohistone, but digests fibers of more acidic proteins. Emphasis is placed by these authors on the observation that enzymes may fail to digest fibers composed of the classes of proteins believed to be present in chromosomes, whereas these enzymes degrade similar proteins in solution. In this connection the possibility must be evaluated that digestibility of chromosomes may be determined by their physical organization as well as by their chemical structure. Inferences concerning cytochemical reactions, derived from test-tube experiments, must also take into account those changes effected in cellular materials in preparing the tissue for study. Studies carried on in this and other laboratories have amply demonstrated that the ability of an enzyme to degrade substrate materials is influenced to a large extent by fixation and subsequent treatment.

For these various reasons it appears that the action of pepsin that is recognizable cytochemically in the fixed salivary-gland nuclei illustrated in Figure 6 concerns primarily the nonhistone proteins. Accepting this assumption, it follows that shrinkage by digestion in pepsin is due in large measure to removal of this type of protein, which would therefore constitute a considerable share of the chromosomal protein. This interpretation has recently been emphasized by Pollister and Leuchtenberger (1949), who point out that the amount of this type of protein is much larger than would be expected from analysis of chemically isolated nuclei or chromosomes.

Use of nucleases and hot trichloracetic acid in the localization of chromosomal proteins: Ribonuclease and desoxyribonuclease have been used to determine patterns of association of proteins and nucleic acids in the chromosome. Stainability of chromosomes with basic dyes, such as pyronin, toluidin blue, and safranin is markedly decreased after their treatment with ribonuclease (Kaufmann, McDonald, and Gay, 1948; Kaufmann *et al.*, 1948). Analysis of the results indicates that under the experimental conditions employed these dyes serve as a guide to the location of ribonucleic acid, as has been suggested by Brachet (1940) and others. The rate of the reaction and the intensity of coloration are dependent to some extent on the purity of the enzyme preparation, but the decrease of basophilia by ribonuclease is in general not masked by the presence of small amounts of proteolytic contaminants (Kaufmann, Gay, and McDonald, in press). On the other hand, the use of "protease-free" ribonuclease has permitted detection of increased stainability of chromosomes with acidic dyes, which is not readily discernible when using protease-contaminated samples (Carnegie Institution of Washington Year Book No. 47). This type of response is shown in Figure 7, which presents photographs of adjoining sections of a root tip of onion stained in acid fuchsin. The section on the left, which served as a control, remained in water at pH 6 for two hours at 37° C, whereas the other section was digested during this period in an aqueous solution of ribonuclease (1 mg/ml) at pH 6.

Questions then arise concerning the nature of the material that shows greater affinity for acidic dyes after treatment of cells with ribonuclease. It will be seen from Figure 7 that increased coloration occurs in both cytoplasm and nucleus. This suggests that the stainable material is not primarily histone. Furthermore, its distribution within the cell parallels that of the blue color developed by histochemical application of Bates' modification of the May-Rose test for tryptophane (Kaufmann *et al.*, 1946, 1947). Finally, stainability with acidic dyes can be reduced by using pepsin, trypsin, or chymotrypsin after ribonuclease. These factors indicate that ribonuclease degrades a complex ribonucleoprotein to release a tryptophane-containing protein, which stains with acidic dyes.

Another line of evidence leading to the same conclusion involves treatment of cells successively with desoxyribonuclease and ribonuclease. Treatment with desoxyribonuclease does not provide any discernible increase in stainability of cytoplasm and nucleoli with acidic dyes, although the chromosomes are more intensely colored than in the controls. Digestion with ribonuclease after digestion with desoxyribonuclease provides further increase in coloration of chromosomes, and a marked increase in color of cytoplasm and nucleoli. Structural details are clearly defined in these digested and stained cells, indicating that the continuity of the chromosome is not impaired by enzymatic degradation of both types of nucleic acid.

A similar result is obtained using the method of Schneider (1945) for removal of both types of nucleic acid with hot trichloracetic acid. Extraction of a section of an alcohol-acetic-fixed root tip with 5 percent trichloracetic acid at 90° C for 15 minutes leads to increased stainability with acidic dyes, as is shown by a comparison of the two upper photographs of Figure 8. On the basis of the experience gained by use of the two nucleases in sequence, this result may be attributed to the degradation by trichloracetic acid of a ribonucleoprotein. (The increase in stainability with acidic dyes effected by hot trichloracetic acid also serves to remove any suspicion that the similar response obtained following the use of ribonuclease is attributable to adsorption of the enzyme by substrate materials.) When pepsin is used after hot trichloracetic acid, dissolution of nuclear material ensues. This is shown in the photo-

graph at the lower left in Figure 8, in which the only detectable nuclear materials remaining are the more refractory nucleoli. Assuming that the action of pepsin is primarily on the tryptophane-containing protein, it follows that any histone remaining after degradation of nucleohistone by hot trichloracetic acid is not sufficient to maintain the chromosome as a structural entity.

This observation might in turn seem to emphasize the importance of the tryptophane-containing protein as the major structural component. On the other hand, the action of pepsin over periods of time sufficiently long to degrade this protein does not lead to dissolution of the chromosome. Neither histone nor tryptophane-containing protein alone seems to be the only material assential to maintenance of structural continuity. One further observation is relevant in this connection: when digestion with pepsin follows digestion with desoxyribonuclease, there is loss of nuclear materials similar to that obtained when pepsin is used after hot trichloracetic acid. This result in itself might seem to point to the structural independence of two separable complexes, one desoxyribonucleohistone, and the other ribonucleotryptophane-protein. Degradation of these complexes does not destroy the chromosome, however, as was demonstrated in those experiments in which ribonuclease and desoxyribonuclease were used in combination, and in those employing hot trichloracetic acid.

Evaluation of the results obtained in this series of experiments therefore leads to the conclusion that the chromosome represents an integrated fabric, in which no single protein or nucleic acid may be regarded as the primary structural component.

Summary

Enzymatic hydrolysis was used in combination with various staining procedures to determine the types of chromosomal proteins and their patterns of association with nucleic acids. The validity of this procedure rests upon the use of highly purified and carefully assayed enzymes, and the meticulous control of all variables capable of modifying the cytochemical reactions.

Smears and sections of fixed plant and animal cells were digested with trypsin, chymotrypsin, pepsin, ribonuclease, and desoxyribonuclease, used independently and in succession. Hot trichloracetic acid was also employed for extraction of nucleic acids.

None of the purified enzymes, when used alone, effects complete degradation of the chromosome. The dissolution usually attributed to the specific action of trypsin has been shown to be due to the removal by water of degradation products produced by the action of trypsin in combination with electrolytes.

Successive treatments with proteases and nucleases, or with proteases and hot trichloracetic acid, lead to dissolution of chromosomes. The results obtained indicate that the chromosome represents an integrated fabric, in which no single protein or nucleic acid may be regarded as the primary structural component.

References

Brachet, J., 1940, La détection histochimique des acides pentosenucléiques. C. R. Soc. Biol. Paris *133:* 88-90.

Caspersson, T., 1936, Über den chemischen Aufbau der Strukturen des Zellkernes. Scand. Arch. Physiol. *73,* Suppl. 8.

Cohen, Seymour, 1945, The enzymatic degradation of thymus nucleohistone. J. biol. Chem. *158:* 255-264.

Danielli, J. F., 1946, Establishment of cytochemical techniques. Nature, Lond. *157:* 755-757.

1947, A study of techniques for the cytochemical demonstration of nucleic acids and some components of proteins. Symposium Soc. Exp. Biol. (Cambridge) *1:* 101-113.

1949, Studies on the cytochemistry of proteins. Cold Spr. Harb. Symposium Quant. Biol. *14:* 32-39.

Darlington, C. D., 1947, Nucleic acid and the chromosomes. Symp. Soc. Exp. Biol. (Cambridge) *1:* 252-269.

Frey-Wyssling, A., 1948, Submicroscopic morphology of protoplasm and its derivatives. Elsevier Publishing Co., Inc. New York.

Frolova, S. L., 1944, Study of fine chromosome structure under enzyme treatment. J. Hered. *35:* 235-246.

Jörgenssen, M., 1913, Zellenstudien I. Morphologische Beiträge zum Problem des Eiwachstums. Arch. Zellforsch. *10:* 1-126.

Kaufmann, B. P., Gay, H., and McDonald, M. R., Enzymatic degradation of ribonucleoprotein. In press.

Kaufmann, B. P., McDonald, M. R., and Gay, H., 1946, Organization of the chromosome. Carnegie Institution of Washington Year Book *45:* 158-162.

1948, Enzymatic degradation of ribonucleoproteins of chromosomes, nucleoli, and cytoplasm. Nature, Lond. *162:* 814-815.

Kaufmann, B. P., McDonald, M. R., Gay, H. Wilson, K., and Wyman, R., 1947, Organization of the chromosome. Carnegie Institution of Washington Year Book *46:* 136-146.

Kaufmann, B. P., McDonald, M. R., Gay, H., Wilson, K., Wyman R., and Okuda, N., 1948, Organization of the chromosome. Carnegie Institution of Washington Year Book *47:* 144-155.

Kossel, A., 1906, Über das Histone. Hoppe-Seyl. Z. *49:* 314-321.

Kossel, A., and Matthews, A., 1898, Zur Kenntnis der Trypsinwirkung. Hoppe-Seyl. Z. *25:* 190-194.

Kunitz, M., 1940, Crystalline ribonuclease. J. gen. Physiol. *24:* 15-32.

1948, Isolation of crystalline desoxyribonuclease from beef pancreas. Science *108:* 19-20.

Kunitz, M., and Northrop, J. H., 1935, Crystalline chymotrypsin and chymo-trypsinogen. I. Isolation, crystallization, and general properties of a new proteolytic enzyme and its precursor. J. gen. Physiol. *18:* 433-458.

McDonald, M. R., 1948, A method for the preparation of "protease-free" crystalline ribonuclease. J. gen. Physiol. *32:* 39-42.

McDonald, M. R., and Kunitz, M., 1946, An improved

method for crystallization of trypsin. J. gen. Physiol. *29:* 155-156.

MAZIA, D., 1941, Enzyme studies on chromosomes. Cold Spr. Harb. Symposium. Quant. Biol. *9:* 40-46.

MAZIA, D., and JAEGER, L. 1939, Nuclease action, protease action and histochemical tests on salivary chromosomes of Drosophila. Proc. Nat. Acad. Sci. Wash. *25:* 456-461.

MAZIA, D., HAYASHI, T., and YUDOWITCH, K., 1947, Fiber structure in chromosomes. Cold Spr. Harb. Symposium. Quant. Biol. *12:* 122-130.

MICHAELIS, L., 1947, The nature of the interaction of nucleic acids and nuclei with basic dyestuffs. Cold Spr. Harb. Symposium Quant. Biol. *12:* 131-141.

MIESCHER, F., 1871, Chemische Zusammensetzung der Eiterzelle. Hoppe-Seyler's Med. chem. Untersuch. 441-460.

MIRSKY, A. E., 1943, Chromosomes and nucleoproteins. Advances in Enzymology *3:* 1-34.

NORTHROP, J. H., 1930, Crystalline pepsin I. Isolation and tests of purity. J. gen. Physiol. *13:* 739-766.

1946, Crystallization of pepsin from alcohol. J. gen. Physiol. *30:* 177-184.

PANIJEL, J., 1947, Contribution a l'étude biochimique de la fécondation chez *Ascaris megalocephalus.* Bull. Soc. Chim. biol., Paris *29:* 1098-1106.

POLLISTER, A. W., and LEUCHTENBERGER, C., 1949, The nucleoprotein content of whole nuclei. Proc. Nat. Acad. Sci. Wash. *35:* 66-71.

SCHNEIDER, W. C., 1945, Phosphorus compounds in animal tissues. I. Extraction and estimation of desoxypentose nucleic acid and of pentose nucleic acid. J. biol. Chem. *161:* 293-303.

1946, The action of ribonuclease on the succinoxidase system. J. biol. Chem. *164:* 241-247.

STOWELL, R. E., and ZORZOLI, A., 1947, The action of ribonuclease on fixed tissues. Stain Tech. *22:* 51-61.

STRANGEWAYS, T. S. P., and CANTI, R. G., 1927, The living cell in vitro as shown by darkground illumination and the changes induced in such cells by fixing reagents. Quart. J. micr. Sci. *17:* 1-14.

TULASNE, R., and VENDRELY, R., 1947, Demonstration of bacterial nuclei with ribonuclease. Nature, Lond. *160:* 225-226.

ZACHARIAS, E., 1909, Die chemische Berchaffenheit von Protoplasma und Zellkern. Progr. rei Botan. *3:* 67-258.

THE APPLICATION OF THE ISOTOPIC DERIVATIVE METHOD TO THE ANALYSIS OF PROTEINS

ALBERT S. KESTON and SIDNEY UDENFRIEND
New York University College of Medicine, New York, and Washington University, St. Louis, Missouri

The isotopic derivative method (Keston, Udenfriend and Cannan, 1946, 1949; Keston, Udenfriend and Levy, 1947 and in press) is a general method for the estimation of organic compounds in the form of isotopically labeled derivatives. A mixture is treated with a reagent containing a stable or radioactive isotope under such conditions that the components which are to be estimated are quantitatively converted into isotopic derivatives of the reagent.

The components of the mixture can then be determined by the incomparably sensitive and specific techniques of radiochemistry or other isotopic methods. If a mole of the isotopic derivative of the compound of interest prepared with the isotopic reagent has "y" units of isotope, and if "b" units are found in the unknown mixture in the derivative of interest, then the moles of substance of interest originally present is b/y. The problem of analysis reduces itself to the determination of the number of units "b" of isotope in the isotopic derivative of the compound of analytical interest.

Where the derivatives of interest can be quantitatively separated from the other isotopic compounds present, as for example by chromatography (Keston, Udenfriend and Cannan, 1949), or counter-current distributions between solvents, a direct determination of "b" can be made.

In most mixtures of biological origin, quantitative separations may be inconvenient or extremely difficult procedures. It may thus become desirable to use techniques such as carrier and indicator techniques which do not require the quantitative isolation of pure derivatives. The use of these techniques in the analyses of the isotopic derivatives will be discussed in this article.

There is in theory practically no lower limit to the sensitivity of the isotopic derivative method, and it has been operated far below the microgram level.

The isotopic reagent selected was p-iodophenyl sulfonyl chloride (pipsyl chloride) labeled with I^{131} or S^{35}. This reagent will react with organic compounds containing amino or hydroxyl groups. The pipsyl derivatives of the amino acids are very stable under the conditions of acid hydrolysis of proteins. The reagent, therefore, has been utilized in analysis of end groups of proteins (S. Udenfriend, S. F. Velick, private communication). It has been demonstrated to react quantitatively with amino acids and peptides (Levy and Slobodiansky, 1949), and analyses of these types of compounds can readily be performed with the aid of this reagent.

The validity and precision of the method which has been outlined depend primarily on, (a) the completeness of the reaction between the isotopic reagent and the compound which is to be estimated, (b) the rigorous purification from radioactive contaminants of the derivative of interest, and (c) the precision of the measurements of the radioactivity of the derivatives and the standards.

TABLE 1. RECOVERY OF AMINO ACIDS AS THE MONOPIPSYL AND DIPIPSYL DERIVATIVES*

Amino acid	Amount added, mg.	% recovered as monopipsyl	% recovered as dipipsyl	Total, %
Glycine	1.000	100.2	—†	100.2
	1.000	99.8	—†	99.8
D,L-Alanine	1.000	98.7	1.3	100.0
	1.000	99.1	1.2	100.3
D,L-Isoleucine	2.000	98.0	1.5	99.5
L-Proline	3.232	98.6	—‡	98.6

* These recoveries are from analyses of the individual pure amino acids, using the standard procedure.

† No dipipsyl derivative of glycine could be isolated for use as carrier.

‡ Proline is a secondary amine and can form only the monosubstituted derivative.

COMPLETENESS OF REACTION BETWEEN AMINO ACIDS AND PIPSYL CHLORIDE

The reaction of amino acids with pipsyl chloride occurs in two stages. In the first stage, a monopipsyl derivative is formed and this reacts to a much lesser extent to form dipipsyl derivatives. Under conditions where a series of successive partial reactions with intermittent removal of the monopipsyl derivative were carried out it has been consistently found possible to recover 98 to 100 percent of the amino acid in the form of the monopipsyl derivative (Keston, Udenfriend and Cannan, 1949). (See Table 1.)

In most of our analyses of amino acid mixtures and protein hydrolysates only the monopipsyl derivatives are determined, and when necessary small corrections are made for the dipipsyl derivative formation.

CARRIER TECHNIQUE

In the carrier method, an overwhelming excess (W moles) of the unlabeled derivative of the desired constituent is added to the mixture of isotopic derivatives, as prepared by quantitative reaction of the amino acid mixture with excess of isotopic reagent.

The derivative added (the carrier) is then separated and purified to constant molal isotope concentration, C_c. If C_r be the molal isotope concentration of pure isotopic derivative which has been prepared with the same sample of the reagent, then the amount of the isotopic derivative (w) which was present in the mixture is

$$w = C_c (w + W)/C_r$$

When relatively large amounts of the carrier are added, this equation reduces to

$$w = WC_c/C_r$$

The carrier principle gives the method wide scope and flexibility. It makes it possible to carry out estimations of very minute amounts of the compound of interest and yet to operate with milligram quantities of material. It is, moreover, unnecessary to seek a quantitative recovery of the compound in pure form. What is essential is the rigorous purification of the carrier from significant amount of isotopic impurities. To accomplish this, large losses of the product may be accepted. Finally, in the analysis of optically active compounds, the use of a large excess of a racemic carrier ensures the estimation of the total of both of the active forms. This is a significant advantage in the analysis of proteins in which indeterminate degrees of racemization may have occurred during hydrolysis. If, on the other hand, one wishes to estimate only one of the isomers, the corresponding carrier may be used.

Purification of Carriers

The purification of the carriers from those other radioactive pipsyl derivatives which might be present turned out to be a difficult problem. Inasmuch as the quantity of carrier is huge compared with the radioactive contaminants, it was thought that if one exceeded the solubility of the carrier, the radioactive contaminants whose solubilities were not approached would stay in solution. In all the cases tried, however, it required a very large number of crystallizations to purify the carrier; and in some cases the rate of loss of impurity was too slow to depend upon for purification. It is common in isotopic tracer work to add suspected impurities in non isotopic form to dilute the impurity, and to conduct the isolation of the desired substance in a milieu of diluted impurities. This technique proved entirely unsuccessful in the purification of the carriers. The fraction of the impurity precipitated, and hence the number of counts due to impurity, was independent, within very wide limits, of the amount of non isotopic impurity added. The impurity distributed itself between the carrier in the solid phase and the carrier in solution as though it were being distributed between two liquid phases. The relation

$$\frac{\left\{\frac{\text{Amount of impurity}}{\text{Amount of carrier}}\right\}\text{Supernatant}}{\left\{\frac{\text{Amount of carrier}}{\text{Amount of impurity}}\right\}\text{Crystals}} = K_D$$

describes the phenomenon. It is obvious that the purification achieved by a crystallization depends on the value of constant K, as well as on the amount sacrificed. Even where K is large, if too large a yield is taken, the degree of purification is small. Where K is small, crystallization alone is a tedious and almost hopeless method of achieving purity. As an example, when K is 1.5, and where 5 percent of the material is lost during a single crystallization, 188 recrystallizations are required to reduce the ratio of impurity to carrier by a factor of 100. When small amounts of material are sacrificed, the rate of change of isotope concentration may, in some cases, be so slow that one may be misled into believing, at a point when significant amounts of impurities are still present, that constant specific activity has been attained.

Table 2. Coprecipitation of Glycine with Alanine*

Sample	Alanine	Glycine	Impurity	K_D
	Micrograms	Micrograms	Percent	
1	7,350	2.4	0.03	1.9
2	7,350	4	0.05	1.8
3	7,350	24	0.3	1.8
4	8,350	75	0.9	1.6
5	8,350	100	1.2	1.8
6	8,350	125	1.5	1.6
7	8,350	190	2.3	1.7
8	8,350	305	3.7	1.6
9	8,350	600	7.2	1.6
10	8,350	1,075	12.9	1.8

* Keston, Udenfriend, Howell and Levy, in press.

Distribution coprecipitation has been known in the field of inorganic chemistry for a long time. Its application to organic chemistry and its implications in the problem of purification of substances encountered in biological isotopic tracer work has not been clearly recognized. It seems to be a general phenomenon. We have observed it in our laboratory with the pipsyl derivatives of the amino acids, peptides, and pyrimidines, and, to a similar degree, with the amino acids themselves. Table 2 illustrates the phenomenon for the glycine-alanine system. In biological tracer work, as well as analytical work (isotope dilution method and isotopic derivative method), it is desirable to test the ability to purify a compound by attempting its purification from known mixtures of the isotopic substances expected as impurities.

In many cases, it was possible to purify carriers sufficiently rigorously by repeated recrystallizations, treatments with activated charcoal, and extractions with solvents. When impurities not readily removed by these means were encountered, recourse was had to counter-current distributions between solvents (Craig, 1944). Table 3 shows the distribution coefficients of some pipsyl derivatives. After preliminary separation of the impurities by this means, purifica-

TABLE 3. DISTRIBUTION COEFFICIENTS OF SEVERAL PIPSYL AMINO ACIDS BETWEEN $CHCl_3$ AND 0.2 M HCl*

Amino acid	C_{chlor}/C_{H2O}	
Aspartic acid	0.04	Group I
Glutamic acid	0.04	
Serine	0.04	
Hydroxyproline	0.11	
Threonine	0.16	
Glycine	0.67	
Alanine	3.25	
Amino isobutyric	7.0	
Valine	20.0	
Proline	28.0	
Leucine	43	
Isoleucine	50	

* Keston, Udenfriend and Cannan, 1949.

tion was continued by means of recrystallizations and treatments with activated charcoal. Incidentally, several carriers, easily separable by countercurrent distribution, may be added to the same aliquot, thus permitting the analysis of several amino acids on the same aliquot. We accepted as pure a carrier whose isotope concentration remained unchanged after repeated recrystallizations and treatments with activated charcoal, attended by large losses in material. In many cases, the homogeneity of the ostensibly purified carrier was investigated by the use of countercurrent distribution techniques. Table 4 shows the recoveries of amino acids from mixtures.

Amino acid determinations in several proteins have also been carried out (Table 5), employing carrier technique (see Keston, Udenfriend and Cannan, 1949).

CHROMATOGRAPHIC TECHNIQUES

If any one of the derivatives can be quantitatively separated from the other derivatives present, it can be estimated without the addition of a carrier. The isotopic derivative method is, in such cases, completely independent of isotope dilution techniques. It has been found possible to separate some of these pipsyl derivatives by means of paper chromatography (Keston, Udenfriend and Levy, 1947). By placing ammonium salts of the pipsyl derivatives as a thin line on a long strip of filter paper and developing the chromatogram with solvents such as n-butanol, radioactive bands corresponding to each derivative are found on the paper. The bands are easily located by radioautographic techniques, or by passing the paper under a lead slit beneath a Geiger counter. By cutting successive transverse strips and measuring each with a Geiger counter, the total count in a resolved band T_c may be determined:

$$T_c/C_r = \text{moles of substance present.}$$

Figure I shows counts of successive 5 mm strips of a chromatogram, prepared from a mixture of proline, serine, and threonine. The darkened areas represent a radioautograph of the same chromatogram.

Any band may be identified, and its degree of purity established, by adding appropriate noniso-

TABLE 4. RECOVERY OF ADDED AMOUNTS OF GLYCINE, ALANINE AND PROLINE FROM A MIXTURE* OF AMINO ACIDS††

Added γ	Glycine carrier remaining at various stages of purification %	Amino acid found at corresponding stages γ	Added γ	Alanine† carrier remaining at various stages of purification %	Amino acid found at corresponding stages γ	Added γ	Proline carrier remaining at various stages of purification %	Amino acid found at corresponding stages γ
0.00	40.5	24.5	0.00	15.6	18.5	0.00	27.5	4.60
	30.0	12.0		12.2	5.0		20.6	0.48
	17.6	4.2		10.0	3.0		16.5	0.00
	11.1	1.2		8.4	1.55			
	9.1	0.11		7.5	0.67			
				6.2	0.08			
50.0	17.9	50.5	99.5	24.1	97.5	80.8	42.0	81.1
	14.7	50.1		18.4	96.4		30.5	82.9
	11.3	50.6		14.7	99.3		19.5	80.6
Average		50.4			97.7			81.5
50.0	18.5	51.0	99.5	23.9	102.1	80.8	42.0	81.1
	13.8	50.7		19.8	102.1		33.1	81.7
	9.9	50.4		16.3	99.6		23.1	80.9
Average		50.7			101.3			81.2

* The amino acid mixture contained 150 μg of leucine and of lysine, 110 μg of glutamic acid, 100 μg of isoleucine and of phenylalanine, 99.5 μg of alanine, 80.8 μg of proline, 50 μg of threonine, of glycine and of valine and 25 μg each of methionine and aspartic acid. In all cases (except glycine) the synthetic racemic forms were used. In the control analyses the glycine, alanine and proline were omitted from the mixture.

† The value for alanine is corrected for 1.3% dipipsylalanine.

†† Keston, Udenfriend and Cannan, 1949.

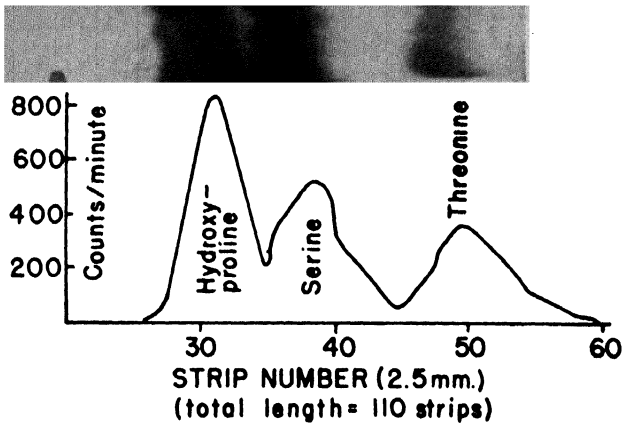


FIG. 1. Radioautograph of a chromatogram prepared from a mixture of proline, serine, and threonine, and a plot of the counts of successive 5 mm strips of the chromatogram.

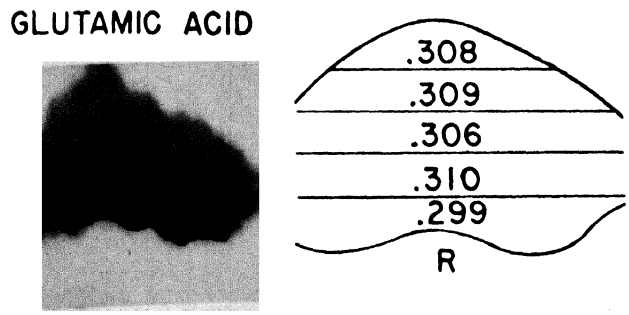


FIG. 2. Radioautograph of part of a chromatogram in the region of the glutamic acid band, and the values of R for the strips indicated.

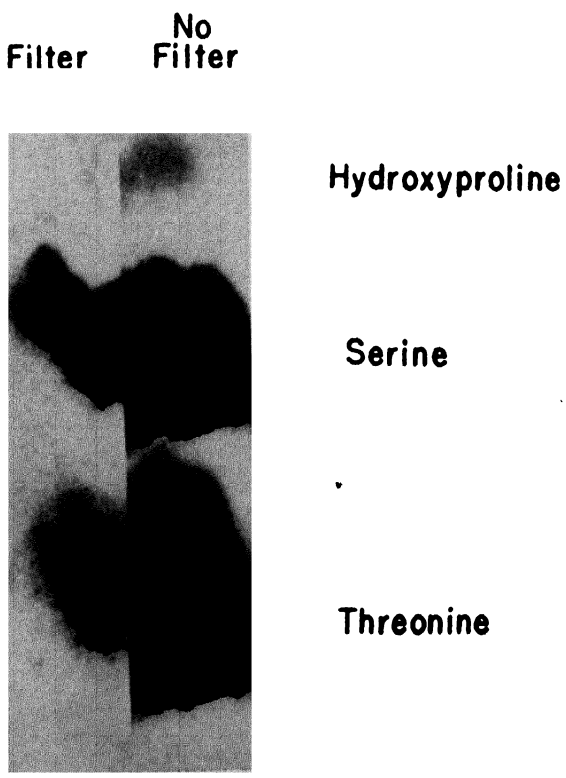


FIG. 3. Radioautograph of a chromatogram of an amino acid mixture not containing hydroxyproline.

topic carrier to the paper strips and eluting. If the band is homogeneous, the isotopic composition of the carrier remains unchanged after rigorous purification.

Estimations may be made which are independent of complete resolution of bands by adding indicators (either unlabeled derivative or derivative labeled with a second isotope) before chromatography. Estimation then depends on isotope ratios in any pure sample. For example, in the determination of glutamic acid in a protein hydrolysate reacted with I^{131} pipsyl chloride, S^{35} pipsyl glutamic acid (B counts) are added, and the chromatogram developed. Eluates of selected successive strips in the glutamic acid band (identified by S^{35} counts) showed constant ratios, $R = I^{131}/S^{35}$. The S^{35} counts are 99.7 percent removed by a 0.003″ Al filter which passes 56 percent of I^{131} counts. The moles of glutamic acid originally present equals RB/C_r. Figure 2 is a radioautograph of part of a chromatogram in the region of the glutamic acid band, together with the values of R for the strips indicated.

In the event that any portion of the band were found to contain only S^{35} and no I^{131}, it would be possible to state with certainty that glutamic acid was absent from the original mixture. If only I^{131} is found in a region of the chromatogram, and S^{35} indicators for all the known amino acids have been added, the possible occurrence of an unknown component is indicated.

Figure 3 illustrates the manner in which indicators may be used for a qualitative test for a substance. The mixture contained no oxyproline but did contain serine and threonine. When treatment with I^{131} pipsyl chloride had been followed by addition of the S^{35} indicators for all three amino acids, radioautographs were made, with and without an intervening Al filter, which would remove all the S^{35} radiations but leave most of the more penetrating I^{131} radiations. The failure to darken the plate at the oxyproline band when a filter was used indicated the absence of I^{131} pipsyl oxyproline, and thus the absence of oxyproline, while the other amino acids were found to be present by this technique.

TABLE 5. COMPOSITION OF PROTEINS*

Protein	Grams of N per 100 grams of protein N				
	Glycine	Alanine	Proline	Isoleucine	Oxyproline
β-lactoglobulin	1.86 (1.81) (1.7M)	7.11 (6.7)P (6.3M)	3.78 (3.2M) (4.2P)	—	—
Human hemoglobin	4.94	9.25 (9.2C)	3.57	<0.2 (0.5B) (0M) (0.2M)	—
Bovine serum albumin	—	—	3.86 (3.9M) (4.2M)	—	0.00
Aldolase-rabbit skeletal muscle	6.23 (6.80M)	8.04 (6.77M)	4.19 (4.85M)	—	—
Phosphoglyceraldehyde dehydrogenase rabbit skeletal muscle	6.88 (7.00M)	6.46 (5.71M)	2.74 (3.08M)	—	—

I—isotope dilution. M—microbiological assay. P—partition chromatography. C—chemical assay.

* Keston, Udenfriend, and Cannan, (1949).

When too many substances are present on the same chromatogram, the bands tend to overlap. In some cases, dependence on retardation factors do not guarantee that mistakes will not be made in the identification of bands. Even with the powerful means available for identification of isotopic bands, it is desirable to limit the number of possible substances on any particular chromatogram. We are developing a method for the systematic analysis of protein hydrolysates which depends on the sharp separation of the amino acid derivatives into several groups (Keston, Udenfriend and Levy, in press). Countercurrent distribution between solvents has been used to achieve this. (See Table 3.) If S^{35} labeled pipsyl derivatives are added as indicators, quantitative isolations are unnecessary and extraction techniques which are especially simple are required to effect suitable separations. The first group investigated consists of the most hydrophilic of the pipsyl derivatives—those of oxyproline, serine, threonine, aspartic acid, and glutamic acid. Very satisfactory recoveries have been obtained in known mixtures of amino acids, and several proteins have

been analyzed for Group I, using a combination of indicators, group separations, and paper chromatography.

It is germane to point out that there is no necessity of obtaining quantitative yield of the isotopic derivative during the reaction with the isotopic reagent, as well as in subsequent isolations, if a known amount of the compound to be analyzed for, labeled with a second isotope, is introduced into the mixture as an indicator. The recovery of the second isotope present in any pure isolated sample of the isotopic derivative may be used to correct for both the lack of quantitative yield and recovery. If "f" be the fraction recovered of the second isotope and "R" be the number of moles of the substance of interest found in the sample isolated, the number of moles present, B, is equal to R/f.

Table 6. Recovery of Group I Amino Acids from a Mixture*

Amino acid†	Added micrograms	Found micrograms
Glutamic acid	0	0.01
	99.3	98
Aspartic acid	0	0.07
	50.5	51.8
Hydroxyproline	0	0
	37.7	38.1
Serine	0	0.03
	49.6	48.3
Threonine	0	0
	33.6	33.8

† For other constituents of the mixture see text. In the zero added experiments each of the other listed amino acids was present.

* Keston, Udenfriend and Levy, in press.

Indeed when indicators are used in this fashion, the method may be applied to substances which do not react with the isotopic reagent. For example, if one wishes to determine acetic acid, a trace of C^{14} labeled acetic acid may be added as an indicator, the mixture then chlorinated, the product treated with ammonia to form glycine, and the glycine then treated with pipsyl chloride labeled with I^{131}. None of the reactions need be quantitative; the yield of C^{14} in any isolated sample of pipsyl glycine serving to correct for lack of quantitative yields in reactions or in recoveries.

Indicator techniques may also be used as a powerful means for the identification of substances formed in biological tracer studies. To establish whether radioactive serine, for example, might have been formed from a radioactive precursor in a biological experiment, the mixture of amino acids and other substances is reacted with nonisotopic pipsyl chloride, so that any C^{14} serine is converted to C^{14} pipsyl serine. I^{131} pipsyl serine is added in small amount as an indicator and the mixture of pipsyl derivative subjected to paper chromatography after suitable chemical separations. If the pipsyl serine band, identified by the presence of I^{131}, contains C^{14} and the C^{14}/I^{131} ratio is constant in different parts of the chromatogram, the presence of C^{14} serine is established with a high degree of certainty.

Chromatography of isotopic derivatives leads to especially reliable and interpretable results when combined with group separations and indicator techniques. The group separations limit the number of possible substances present on a chromatogram. Lack of constancy of isotope ratios along a bond shows up impurities when they are present. The presence of the indicator prevents errors in the identification of the bands. Table 6 shows analytical values obtained for recoveries of amino acids from mixtures and Table 7 shows values for the Group I amino acids in various proteins. These results were obtained by employing a combination of the isotopic derivative method, group separations, and indicator technique.

Table 7. Group I Amino Acids in Several Proteins††

Amino Acid	β lacto-globulin	Bovine Serum Albumin	Human Serum Albumin†
	grams of Amino acid per 100 grams protein		
Glutamic acid	18.2	16.3	17.0
Aspartic acid	10.5	9.6	8.95
Hydroxyproline	0.05	0.0006	0.0007
Serine*	3.85	4.03	2.97
Threonine*	4.95	5.49	4.35

* These values are not corrected for loss in hydrolysis.

† For literature values *c.f.* Brand, E., 1946, Ann. N. Y. Acad. Sci. *47*:187.

†† Keston, Udenfriend and Levy, in press.

References

Craig, L. C., 1944, Identification of small amounts of organic compounds by distribution studies. II. Separation by counter-current distribution. J. biol. Chem. *155:* 519-534.

Keston, A. S., Udenfriend, S., and Cannan, R. K., 1946, Microanalysis of mixtures (amino acids) in the form of isotopic derivatives. J. Amer. Chem. Soc. *68:* 1390.

1949, A method for the determination of organic compounds in the form of isotopic derivatives. I. Estimation of amino acids by the carrier technique. J. Amer. Chem. Soc. *71:* 249-257.

Keston, A. S., Udenfriend, S., and Levy, M., 1947, Paper chromatography applied to the isotopic derivative method of analysis. J. Amer. Chem. Soc. *69:* 3151-3152.

1949, J. Amer. Chem. Soc., in press.

Levy, M., and Slobodiansky, E., 1949, The application of the isotopic derivative technic to the study of protein structure. Cold Spr. Harb. Symposium Quant. Biol. *14:* 113-116.

THE NATURE OF SOME ION-PROTEIN COMPLEXES

IRVING M. KLOTZ
Department of Chemistry, Northwestern University, Evanston, Illinois

Introduction

In many of its aspects, the problem of ion-protein interactions is one to which biologists have devoted their attention for almost a century. Physiologists, in particular, have long been interested in the excitatory or inhibitory effects of specific anions and cations on nerve activity, or on other biological responses. Similarly, in enzymatic reactions, the biochemist has found the concept of enzyme-substrate combinations a very fruitful model for the interpretation of metabolic reactions. *A priori,* there is no reason to expect, of course, that all of these interactions embody a common mechanism or similar types of forces. Nevertheless, many of these phenomena can be described in terms of a picture involving actual binding of the small ion by the macromolecular protein. There has been a great deal of interest, therefore, particularly in recent years, in investigations which might give some insight into the nature of these ion-protein complexes.

In approaching this problem, it is advantageous at the outset to set up a very schematic diagram, such as that in Figure 1, to represent an ion-protein

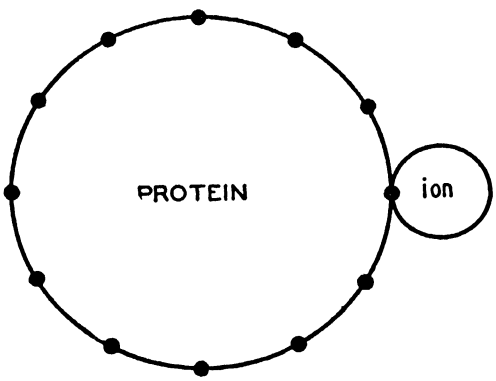


Fig. 1. Schematic diagram of ion-protein complex.

complex. In this diagram, each constituent of the complex is indicated simply by a sphere, and possible sites of combination on the protein are represented by dots. With this model, several fundamental questions can be formulated in a precise manner. In this paper we shall consider the following six:

(1) How many ions are held by a given protein molecule, under specified environmental conditions?
(2) What is the maximum number of sites on a protein molecule available to a given ion?
(3) What is the strength of the bond between the protein and any given type of ion, that is, what is the energy of combination?
(4) How does the environment affect the extent or the energy of combination?
(5) What structural features in an ion favor combination with the protein?
(6) What is the molecular and configurational nature of the site on the protein at which a given ion is bound?

Complete answers to all of these questions cannot be given as yet. Nevertheless, a great deal of information has been accumulated in the past few years, and much of it can be used to convert the bare schematic representation of Figure 1 into a much more detailed picture of ion-protein complexes.

I. Stoichiometric Relations

A. Determination of Number of Bound Ions

The first two questions which have been posed may be considered together, since they are very

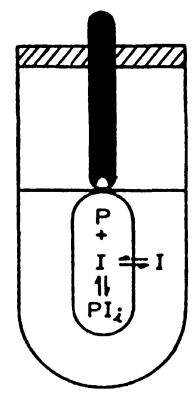


Fig. 2. The dialysis-equilibrium method for determining the extent of binding of an ion by a protein.

closely related. The very earliest experiments in this connection showed that more than one ion may be bound by a single protein molecule. It is convenient, therefore, to consider the maximum number of available sites, as the limiting value of the number of bound ions, as the free ion concentration is made larger and larger.

Several methods are available for the determination of the number of ions bound by a protein molecule under specified conditions. Perhaps the simplest of these is the dialysis-equilibrium method (Fig. 2) in which a bag with a known quantity of protein is immersed in a solution of the ion, the sys-

tem permitted to reach equilibrium, and the concentration of free ion determined analytically. With an appropriate control to establish the quantity of free ion which is present in the absence of protein, it is a simple matter to calculate the quantity of ion bound by the protein. If the protein is one of known molecular weight, the results may be expressed in moles of bound ion and moles of total protein. The ratio of the former to the latter gives the *average*

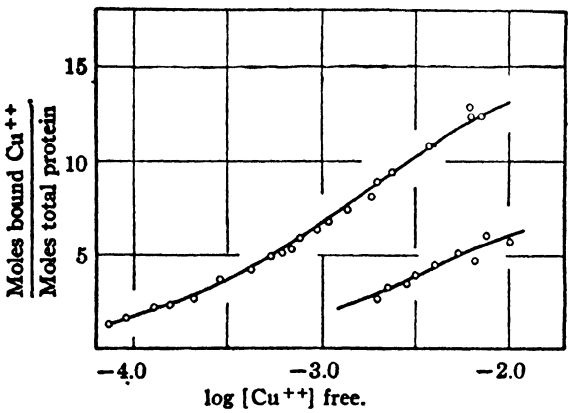


FIG. 3. Binding of cupric ions by bovine serum albumin at 25°C.: upper curve at pH 4.83; lower curve at pH 4.00.

number of ions bound by each protein molecule.

The results of a typical experiment of this type, carried out for a wide series of free ion concentrations, are illustrated in Figure 3. In this example, complexes of cupric ion with crystallized bovine serum albumin have been studied (Klotz and Curme, 1948). The graph is a concise summary of the average number of copper ions per molecule of albumin, as a function of (the logarithm of) the concentration of free ion. Thus at a copper concentration of 10^{-3} molar (log = −3), the number of bound ions per protein averages about 7. If the free copper is 10^{-2} molar (log = −2), the number of bound ions rises to about 13 per protein molecule.

B. Distribution of Ions Among Several Complexes

It is necessary to recognize, however, that in the solution of 10^{-2} molar, for example, complexes exist which have fewer than 13 cupric ions on each protein molecule, and that there are some also present that have more than 13. No method has been devised yet for measuring by a direct experiment the concentration of each type of complex in a given solution. Nevertheless, these concentrations can be calculated from the fundamental equations of the law of mass action, if adequate data are available on the variation of the average number of bound ions with changes in concentration of free ion.

An example of the results of such a calculation, for a simpler case than copper-albumin complexes, is illustrated in Figure 4. In this hypothetical situation, it has been assumed that ten sites are available on the protein molecule, P. The fraction of each specific ion-protein complex, PI_1, PI_2 . . . can be calculated, by methods to be described shortly, as a function of the concentration of free ion, I. In Figure 4, the fraction of protein in the form of any specific complex is given by the ratio of the vertical distance between two adjacent curves, to the height of the entire ordinate. Thus at a concentration of 10^{-5} molar free ion, approximately 0.9 of the protein is non-complexed and 0.1 is in the form PI_1. As the free ion concentration is increased, higher complexes are formed. At a concentration 10^{-4} molar free I, the fraction of free P has dropped to 0.38, whereas that of PI_1 has risen to 0.39 and approximately 0.17 of PI_2 has been formed. In this solution, the average number of bound ions, calculated by methods to be described later, is 0.9 moles per mole of total protein. It is thus apparent that even though the average number of bound ions is less than 1 per protein molecule, a substantial quantity (about 17 percent) of the protein is in the form of PI_2. The results cited in this simple hypothetical example, are typical of all actual ion-protein complexes.

C. Available Sites on the Protein Molecule

Proceeding to the problem of determining the maximum number of ions which can be held by a single protein molecule, we find that the answer can be obtained occasionally by direct experiment, but more often, it requires some extrapolation proce-

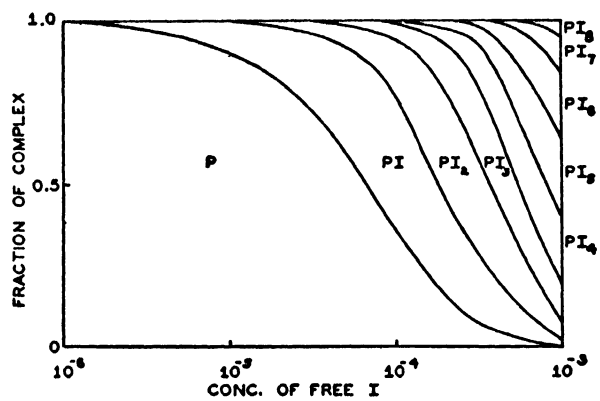


FIG. 4. Different types of complex formed in a hypothetical case in which ten sites are available on a protein and the intrinsic binding constant is 10^5.

dure. In principle, we would expect that the binding curve, as shown in Figure 3, should become horizontal as the concentration of free ion is increased. In practice, however, such a plateau is rarely reached at concentrations of free ion which are experimentally feasible. One of these relatively rare cases is exemplified in lysozyme complexes with the dye anion Orange II, illustrated in Figure 5, where it is apparent that binding essentially stops after 11

ions are taken up by each protein molecule. In most cases, however, the strength of binding is too weak to enable one to find the plateau directly. Thus even for copper-albumin complexes, a relatively favorable case, it is difficult to tell from Figure 3 just what the maximum number of bound copper ions can be. It is necessary, therefore, to resort to some extrapolation procedure.

Two different methods of extrapolation have been proposed (Klotz, Walker and Pivan, 1946; Scatchard, 1949). Both, nevertheless, are based on the same fundamental equations derived from the law of mass action, to describe multiple equilibria of the type under consideration here.

1. Law of mass action for multiple equilibria. When one is dealing with a molecule, such as a protein, P, which is capable of combining with more than one ion, I, the successive equilibria which are obtained may be represented by a set of equations of the form:

$$\begin{aligned} P + I &= PI \\ PI + I &= PI_2 \\ PI_2 + I &= PI_3 \\ &\cdots\cdots\cdots \\ PI_{i-1} + I &= PI_i \end{aligned} \qquad (1)$$

For each one of these equilibria, one can then formulate an equilibrium constant. For the set of equations in (1), the constants are defined in the following way:

$$\begin{aligned} \frac{(PI)}{(P)(I)} &= k_1 \\ \frac{(PI_2)}{(PI)(I)} &= k_2 \\ \frac{(PI_3)}{(PI_2)(I)} &= k_3 \\ &\cdots\cdots \\ \frac{(PI_i)}{(PI_{i-1})(I)} &= k_i \end{aligned} \qquad (2)$$

Mathematical analysis indicates (Klotz, 1946a), that the average number of bound ions per mole of total protein, r, can then be expressed as a function of these equilibrium constants and of the concentration, (I), of free ion by the equation

$$r = \frac{k_1(I) + 2k_1k_2(I)^2 + 3k_1k_2k_3(I)^3 + \cdots + i(k_1 \cdots k_i)(I)^i + \cdots}{1 + k_1(I) + k_1k_2(I)^2 + k_1k_2k_3(I)^3 + \cdots + (k_1 \cdots k_i)(I)^i + \cdots} \qquad (3)$$

2. Equations for extrapolation to determine available sites. This general equation can be simplified greatly if a bound ion exerts no electrostatic influence on the succeeding bindings, and if each ion is attached to the same type of group on the protein molecule. In such a situation, the successive equilibrium constants of equation (2) are no longer independent and may be replaced by a single intrinsic constant, k, and certain statistical factors.

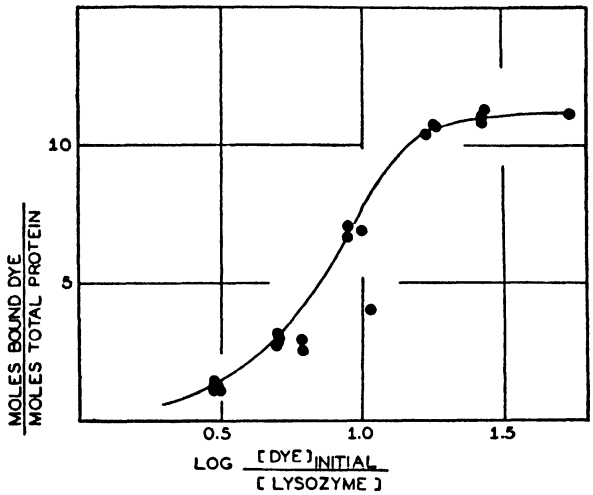


FIG. 5. Binding of Orange II anions by lysozyme, in a phosphate buffer at pH 7.6.

Equation (3) may then be reduced to the expression

$$\frac{1}{r} = \frac{1}{kn}\frac{1}{(I)} + \frac{1}{n} \qquad (4)$$

where n is the maximum number of ions of the specified type. In a graph of $1/r$ versus $1/(I)$, the intercept on the ordinate is $1/n$; its reciprocal gives n. A typical example of this type of graph and of the required extrapolation is illustrated in Figure 6, prepared from experimental data on calcium-casein complexes (Chanutin, Ludewig and Masket, 1942). The intercept on the ordinate, slightly over 0.06, corresponds to an n value of 16.

By considerations similar to those just described it is also possible to derive an equation (Scatchard, 1949) fundamentally equivalent to (4) but of the form

$$\frac{r}{(I)} = kn - kr \qquad (5)$$

With this equation, the appropriate graph for the determination of n is one with $r/(I)$ as ordinate and r as abscissa. The intercept on the abscissa (at which point $r/(I)$ is zero) gives n. A typical example of this type of graph and of the required extrapolation is illustrated in Figure 7, prepared from

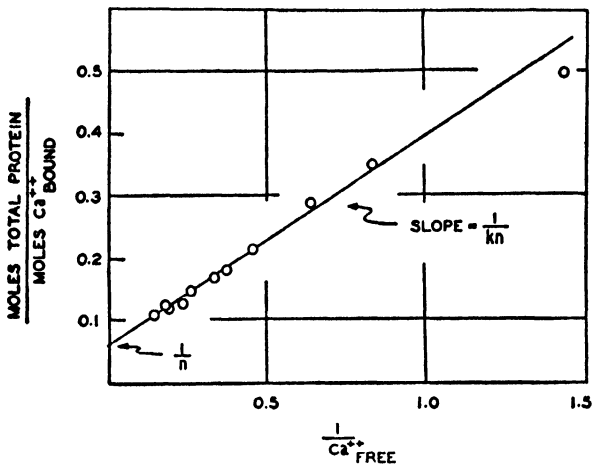


FIG. 6. One type of extrapolation to determine the maximum number of calcium ions which may be bound by casein.

the same experimental data as were used for Figure 6.

Equation (5) has an advantage over (4) in that the length of the extrapolation, (Fig. 7), is emphasized and one is not likely to be deceived about

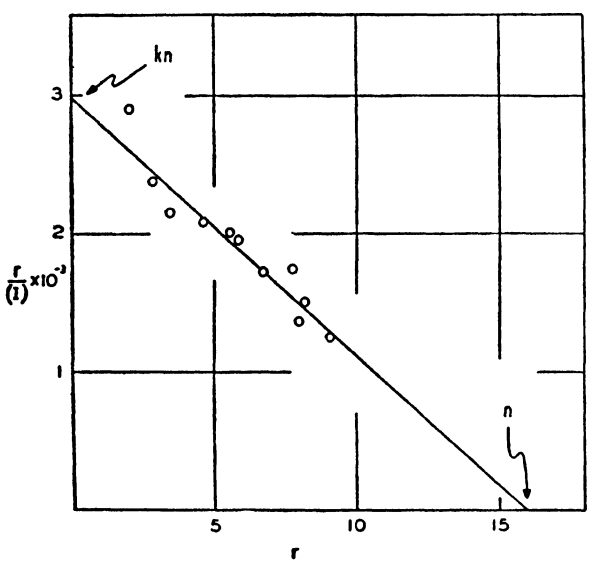


FIG. 7. Another type of extrapolation to determine the maximum number of calcium ions which may be bound by casein.

the relative reliability of the value found for n. The graph, Figure 6, corresponding to equation (4), seems to have less uncertainty in the intercept from which n is calculated, but since the numerical value of this intercept is generally near zero, a very small error in the extrapolation to $1/n$ may be reflected into a large uncertainty in n.

3. Number of sites on albumin available to different ions. For one protein, serum albumin, the maximum number of sites available on a single molecule has been reported for several different anions and cations. Some typical values have been assembled in Table 1.

Even among the anions, no common value of n is apparent. Such a dependence of n on structure of anion may reflect differences in configuration of the amino acids around the sites at which binding occurs. On the other hand, a great deal of doubt has often been cast (Scatchard, 1949; Karush and Sonenberg, 1949) on the reliability of n values ob-

TABLE 1. MAXIMUM NUMBER OF SITES AVAILABLE ON SERUM ALBUMIN*

Ion	n	Reference
Chloride	10	(1)
Dodecyl sulfate	14	(2)
Phenyl butyrate	24	(3)
o-Nitrophenolate	6	(3)
m-Nitrophenolate	24	(3)
p-Nitrophenolate	25	(3)
Methyl orange	22	(4)
Cupric	16	(5)

(1) Scatchard, Scheinberg and Armstrong, 1949; (2) Karush and Sonenberg, 1949; (3) Teresi and Luck, 1948; (4) Klotz, Walker and Pivan, 1946; (5) Klotz and Curme, 1948.

* No distinction is made in this table between albumins of different origin, that is, whether human, bovine or equine.

tained by methods of extrapolation. In no case has an unequivocal plateau been observed in a graph of r versus log (I) for a soluble albumin complex. Measured values of r as high as 40 have been found in this laboratory for albumin complexes with the dye anion, Orange II. Similarly, Scatchard and Black (1949) have reported measured values well over 40 for the number of thiocyanate ions bound by a single albumin molecule. It may be possible that n changes with the quantity of bound ion, due to changes induced in the protein. In any event, it is apparent that the question of the number of sites available for binding in soluble complexes is still an open one.

II. BINDING ENERGY

A. Fundamental Significance

If we are dealing with a series of equilibria such as those represented by equation (1), it is convenient to define binding energies in terms of the changes in standard free energy, ΔF°, for each reaction. Thus, if R represents the gas law constant and T the absolute temperature, ΔF° is dependent on the equilibrium constant, k, for a reaction, according to the equation

$$\Delta F^\circ = -RT \ln k \qquad (6)$$

Consequently, for a series of complexes PI, PI_2, PI_3 . . . , PI_i, we may define corresponding binding ener-

gies, $\Delta F_1°$, $\Delta F_2°$, $\Delta F_3°$ . . . , $\Delta F_i°$, by the equations paralleling (1) and (2):

$$\begin{aligned} \Delta F_1° &= -RT \ln k_1 \\ \Delta F_2° &= -RT \ln k_2 \\ \Delta F_3° &= -RT \ln k_3 \\ &\cdots\cdots\cdots\cdots \\ \Delta F_i° &= -RT \ln k_i \end{aligned} \qquad (7)$$

Focusing attention on the first equilibrium, leading to the formation of PI, we can see from equation (2) that the stronger the bond between P and I, the greater will be the concentration (PI) [and the smaller will be free (P) and free (I)] and hence the larger will be k_1. In equation (7), in turn, the larger k_1, the larger does $\Delta F_1°$ become in magnitude, though always in the negative direction. Thus the stronger the bond, the more negative is the binding energy, if it is defined in terms of $\Delta F°$.

The actual evaluation of $\Delta F°$'s from experimental data involves relatively extensive computations (Klotz, Walker and Pivan, 1946; Scatchard, 1949; Karush and Sonenberg, 1949) which it would not be profitable to discuss here. A typical series of results, nevertheless, is illustrated in Table 2 for copper complexes with serum albumin, at each of two temperatures.

TABLE 2. THERMODYNAMICS OF BINDING OF Cu^{++} BY BOVINE SERUM ALBUMIN
pH 4.83

i	$\Delta F°$ at 0° cal./mole	$\Delta F°$ at 25° cal./mole	$\Delta S°$ cal./mole/deg.	$\Delta H°$ cal./mole
1	−5,179	−5,908	29.2	2,780
2	−4,708	−5,399	27.6	2,840
3	−4,391	−5,058	26.7	2,900
4	−4,135	−4,784	26.0	2,960
5	−3,910	−4,544	25.4	3,020
6	−3,704	−4,325	24.8	3,080
7	−3,509	−4,117	24.3	3,140
8	−3,319	−3,915	23.8	3,200
9	−3,131	−3,715	23.4	3,260
10	−2,941	−3,513	22.9	3,320
11	−2,746	−3,305	22.4	3,380
12	−2,540	−3,086	21.8	3,440
13	−2,315	−2,846	21.2	3,500
14	−2,059	−2,572	20.5	3,560
15	−1,742	−2,231	19.6	3,620
16	−1,271	−1,722	18.0	3,680

B. Variation in Binding Energy with Number of Bound Ions

It is of interest to note in Table 2 that $\Delta F°$ becomes smaller in absolute magnitude as more and more Cu^{++} ions are bound by the protein. In other words the strength of binding decreases as successive ions are added to the protein. This trend in $\Delta F°$ is characteristic of all ion-protein complexes. It is of interest, therefore, to examine the explanation at the molecular level for this trend.

For this purpose, it is helpful to represent the energy of binding in terms of a well, whose depth is proportional to the absolute magnitude of $\Delta F°$ and whose downward direction corresponds to the negative sign of the observed binding energy (Figure 8). Energetically speaking, the complex, *e.g.* PCu, is at

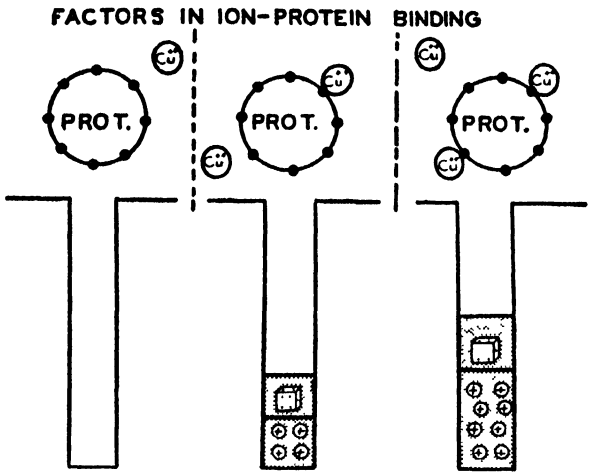


FIG. 8. Schematic representation of binding energies in connection with their dependence on number of bound ions.

the bottom of the well, the dissociated components, in this example P and Cu, at the ground level. Thus it takes energy to raise the complex from the bottom of the well to the surface, that is, to dissociate a Cu^{++} ion from the complex. If the strength of binding of a second Cu^{++}, to give PCu_2, is less than that of the first, then the depth of the well representing the second binding energy (Figure 8), should not be as great. The question which we wish to raise is, what factors tend to make the second well less deep than the first, in other words, what factors fill in the well.

Careful analysis indicates that the well is filled in, in most cases, by two effects, the first of which may be called a statistical one, the second, electrostatic.

The statistical effect may be appreciated by comparing the situation facing a Cu^{++} ion as it approaches P (Figure 8), with that when it approaches PCu. In the second case, the number of sites open to the Cu^{++} is only 7, whereas 8 were available[1] to the first Cu^{++} approaching P. Thus the chances of a successful collision in the second case are only 7/8 that in the first, and hence the probability of forming the second complex, PCu_2, is less than that of forming PCu. Another statistical effect working against PCu_2, as compared to PCu, arises even after the complexes have been formed. With PCu_2,

[1] The number 8 was chosen merely for diagrammatic simplicity. Actually 16 sites on a serum albumin molecule seem to be available to Cu^{++}.

there are two positions at which a single Cu^{++} may be lost, since the first Cu^{++} as well as the second may come off. Therefore, the chance of PCu_2 dissociating a single Cu^{++} is twice as much as that of PCu losing a single Cu^{++}. Again, then, we have a probability factor which tends to reduce the concentration of (PCu_2). Referring to equation (2), we can see that k_2 will be reduced if (PI_2) is reduced. Consequently, ΔF_2° in equation (7) will not be as negative a number as it would have been if the probability factor were not operative. In terms of our schematic diagram (Figure 8), the second well will be filled in partially by the probability factor.

TABLE 3. BINDING ENERGIES FOR COMPLEXES WITH SERUM ALBUMIN

Ion	ΔF_1°	Reference
Chloride	− 3,700	(1)
p-Nitrophenolate	− 4,865	(2)
Methyl orange	− 6,050	(3)
Dodecyl sulfate	−11,000	(4)

(1) Scatchard, Scheinberg and Armstrong, 1949; (2) Teresi and Luck, 1948; (3) Klotz and Urquhart, 1949; (4) Karush and Sonenberg, 1949.

In addition to the statistical effect, there is another factor which tends to reduce the relative stability of the second complex as compared to the first. Since ions are being bound, we are dealing with charged units, and, furthermore, species of like charge. It is to be expected, then, that the first bound copper ion would exert an electrostatic repulsion toward an approaching second ion. Again, the tendency of the second Cu^{++} to be bound by PCu would be reduced, and hence the strength of the bond decreased. Diagrammatically speaking, the well would be filled in to some extent by the electrostatic repulsion.

These descriptive considerations can be expressed also in quantitative terms and explicit mathematical relations derived for the change in ΔF° with increase in number of bound ions (Klotz, Walker and Pivan, 1946; Klotz and Curme, 1948; Scatchard, 1949). Except in a few cases (Karush and Sonenberg, 1949), electrostatic and statistical factors can adequately account for the observed binding energies in ion-protein complexes. In the exceptions, the difficulties may lie in variation in n with amount of bound ion, or perhaps in the heterogeneous nature of the sites at which the ions are bound.

C. *Binding Energies for Some Typical Anions*

Binding energies may be compared also for different types of ion, as well as between different numbers of the same ion on a given protein. Some typical values of ΔF_1°, the binding energy for the first complex, PI, have been assembled in Table 3 for serum albumin complexes. Strictly speaking, compounds should be compared with respect to their intrinsic binding constants, k, rather than k_1, but since the question of variable n's for different ions is still open, it seems preferable to use k_1 which can be evaluated more precisely without a knowledge of n (Klotz and Urquhart, 1949). Some relations between structure and binding energy will be considered in one of the following sections.

III. EFFECT OF ENVIRONMENT

Investigations of the effects of changes in the medium have contributed greatly to the interpretation of the binding process at the molecular level. Among the variables which have been studied, pH changes produce the most pronounced effects. However, changes in pH are usually associated with modifications in the nature of the buffer. Since these two effects are so closely inter-linked they will be considered together.

A. *pH and Nature of Buffer*

Disregarding temporarily the different types of buffer used, we can examine the overall effect of pH in the results for methyl orange-albumin complexes summarized by Figure 9. It is apparent that only minor differences are encountered in the region below pH 9. In more basic solutions, however, there is a distinct drop in extent of complex formation,

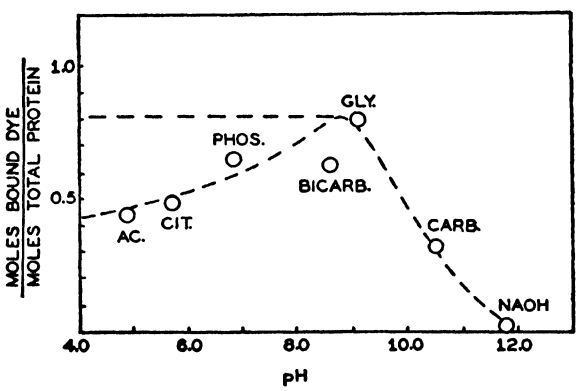


FIG. 9. Effect of pH on binding of methyl orange by bovine albumin. The concentration of free anion in each case is 1.00×10^{-5} molar.

and complete disappearance of binding ability occurs above about pH 12.

It is apparent that some fundamental change must occur in the protein molecule in this pH region, since the anion does not undergo any modification. The change, however, is not a denaturation process, for the phenomenon is a reversible one, this is, if the protein at pH 12 is returned to pH 7, it behaves in a manner identical with that of a sample maintained throughout at pH 7. It seems, therefore, that some group on the protein which undergoes modification

between pH 9-12 is intimately involved in the binding process. This behavior, combined with other evidence to be cited later, points strongly toward lysine and perhaps arginine residues as the focal points of attachment for anions.

An evaluation of specific buffer effects on the binding process can be made most reliably by working at a single pH. Examples of such effects are then obtained readily. In Figure 10, acetate and phthalate buffers are compared in their effects on the degree of binding of methyl orange by serum albumin at pH 5.0. It is quite obvious that phthalate interferes with binding to a greater extent than acetate. Many other buffers have been examined also (Klotz and Urquhart, 1949a) and their relative influences are compared in Figure 11. At pH's above the isoelectric point (*ca.* 5) of serum albumin, phosphate, bicarbonate and glycinate buffers cause least interference with binding of anions by this protein.

The mechanism of specific buffer effects seems almost certainly due to competition between the buffer anion and the primary anion for loci on the protein. Thus the severity of the buffer effect depends on the structure of the anion in the same way as structure determines extent of binding by the protein. The buffer effect is also a direct function of the buffer anion concentration, and the results can be correlated quantitatively by equations derived from the law of mass action for competitive binding (Klotz, Triwush and Walker, 1948). Similarly the effects of these buffers on the electrophoretic behavior (Alberty, 1949; Longsworth and Jacobsen, 1949) and on the isoionic point (Scatchard and Black, 1949) of albumin are in the direction to be expected if the anion is bound to the protein.

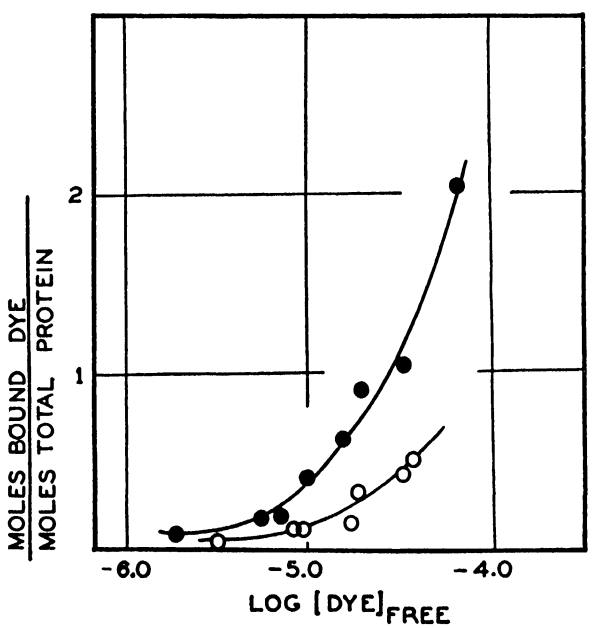


FIG. 10. Comparison of effect of acetate buffer, ●, with that of phthalate, ○, on the extent of binding of methyl orange anions by bovine serum albumin at pH 5.0.

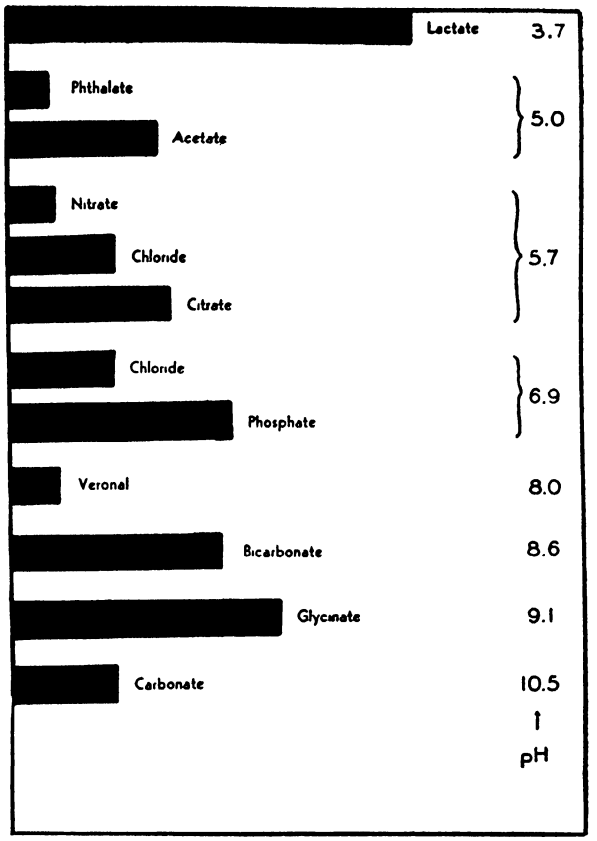


FIG. 11. Effects of various buffers on binding of methyl orange by serum albumin. The longer the bar, the more methyl orange is bound, that is, the smaller the interference by the buffer anion.

Anions other than those found in common buffers may also compete with the binding of the reference ion by the protein. Thus even chloride, and especially nitrate ions, can interfere very markedly with the binding of methyl orange by albumin (Klotz and Urquhart, 1949a). For chloride ions Scatchard, Scheinberg and Armstrong (1949) have demonstrated directly that binding occurs with serum albumin.

B. Effect of Ion Concentration

The dependence of degree of binding on the concentration of free ion present has been discussed adequately in preceding sections. It is merely desirable to mention it again at this point, to emphasize that the ion concentration is one of the factors characterizing the medium.

C. Effect of Protein Concentration

In comparison with the effect of changing the ion concentration, the consequences of increasing the

protein concentration are only second order. Thus a five-fold (0.2-1.0 percent) increase in albumin concentration produces only a very small change in the extent of binding of methyl orange (Figure 12).

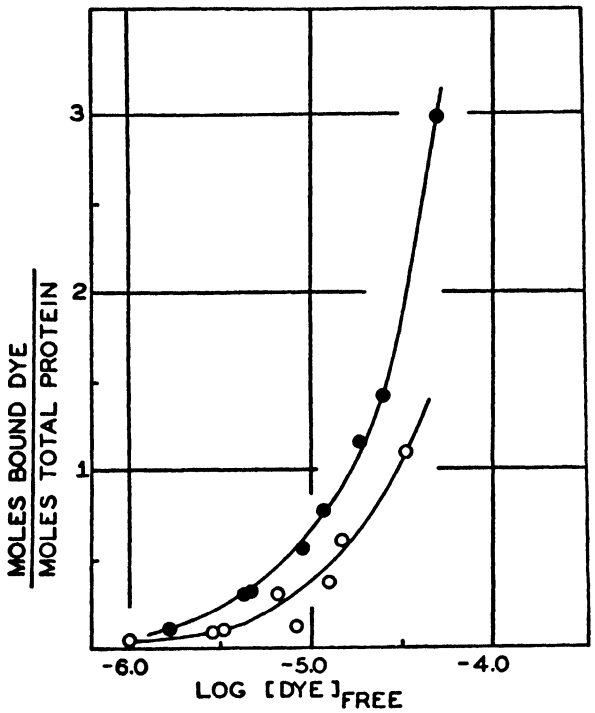


FIG. 12. Dependence of extent of binding of methyl orange on the concentration of serum albumin: ●, 0.2 percent; ○, 1 percent.

With the alkyl sulfates, no detectable change in extent of binding has been observed even with a ten-fold (0.05-0.5 percent) change in albumin concentration (Karush and Sonenberg, 1949). If the protein concentration is not too high, we would expect the extent of binding to be independent of protein concentration, for equation (3) for r does not have (P) in it at any point. When the protein concentration becomes high, however, it is not unlikely that the binding constants, k_i, are affected by changes in properties of the medium, and hence, the protein concentration begins to manifest itself, although only in an indirect manner.

D. Effect of Temperature

Many investigations have demonstrated that temperature changes do not affect greatly the extent of binding of ions by serum albumin. In his exploratory work on sulfonamide-plasma protein complexes, Davis (1943) observed no marked change in degree of binding with rise in temperature. Similarly, Putnam and Neurath (1945) in their electrophoretic studies of complexes between albumin and sodium dodecyl sulfate obtained no indication of a temperature dependence. Boyer, Ballou and Luck (1947) also have reported a zero temperature coefficient in the binding of caprylate by serum albumin.

The quantitative results of more recent investigations (Klotz and Curme, 1948; Klotz and Urquhart, 1949b; Karush and Sonenberg, 1949; Scatchard, Scheinberg and Armstrong, 1949) substantiate the general observations of the earlier qualitative work, although in several cases, the temperature coefficient, though small, is measurable (Figure 13).

For any equilibrium reaction which is not sensitive to temperature, it is a thermodynamic consequence that the heat of the reaction (or enthalpy change ΔH) must be small. It also follows from the thermodynamic equation

$$\Delta F = \Delta H - T\Delta S \qquad (8)$$

that if ΔH is small, the magnitude of ΔF at any fixed temeperature, T, is determined primarily by the value of ΔS, the entropy change in the reaction. The favorable free energy of binding which has been observed (Table 3) for many ion-albumin complexes thus seems to be a result of a favorable entropy change during binding, rather than any favorable

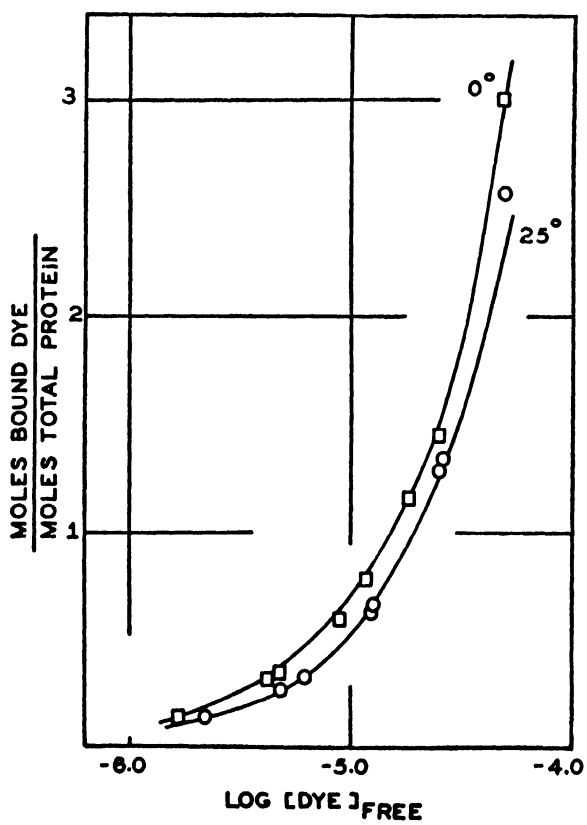


FIG. 13. Effect of temperature on binding of methyl orange by bovine serum albumin at pH 6.8.

heat effect. The significance of these characteristics in molecular terms will be considered in subsequent discussions.

IV. Molecular Nature of the Binding Process

So far little effort has been made to use the data presented to elaborate on the binding process from the molecular viewpoint. Since there are wide variations in binding ability among various proteins, as well as between different anions, we have in these investigations a very sensitive probe for the general problem of molecular structure in proteins, as well as in their complexes.

From the molecular viewpoint, anions and cations show pronounced differences in behavior in the formation of protein complexes. Since the greater amount of information has been obtained on anion-protein complexes, the present discussion will be restricted to these combinations.

A. Evidence for Participation of Positively-Charged Residues

Although the suggestion had been made earlier, the first indication that anions may be attracted largely by positive loci on the protein came from comparisons of the degree of binding of pairs of

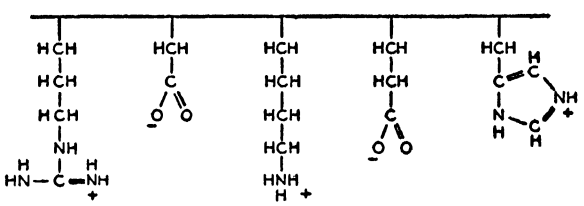
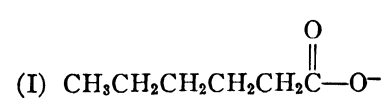
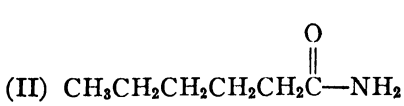


Fig. 14. Schematic diagram of charged side-chains on a protein molecule.

compounds of practically identical structure, but with one carrying a negative charge. Thus it was observed (Boyer, Lum, Ballou, Luck and Rice, 1946) that caproate ion (I)



(I) $CH_3CH_2CH_2CH_2CH_2C(=O)-O^-$



(II) $CH_3CH_2CH_2CH_2CH_2C(=O)-NH_2$

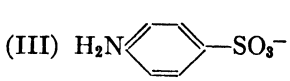
is bound much more strongly by albumin than the analogous caproamide (II). Similarly it was found (Klotz, 1946b) that sulfanilate ion (III) forms far stronger complexes with serum albumin than does sulfanilamide (IV). In each case the negative charge seems to supply the missing binding energy.

At first glance it is perhaps surprising that the presence of a negative charge should assist in the formation of an ion-protein complex, since the proteins investigated themselves carry a net negative charge. It is necessary to keep in mind, however, that a net negative charge on a protein molecule merely implies an excess of negative over positive residues. At pH's of interest, in the vicinity of 7, the protein has a very substantial number of positively-charged residues (Figure 14).

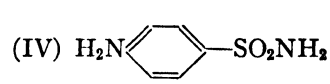
Several lines of evidence seem to substantiate the viewpoint that cationic centers on the protein are intimately involved in the binding process. These centers presumably consist of the basic amino acid residues, histidine, lysine and arginine (Figure 16). It would be expected, then, that any factor which removes the positive charge from these loci would reduce the extent of binding of anions markedly.

One of the simplest methods for removing the positive charge from the basic amino acid residues is to increase the pH. As has been pointed out already (Figure 9) binding ability does indeed disappear above pH 12. This loss of binding ability is not due to denaturation of the protein, for a sample of albumin returned from pH 12 to 7 behaves just like a control maintained throughout at pH 7. Evidently then, binding affinity does decrease under conditions in which the positively-charged loci become electrically neutralized.

Further evidence that cationic loci on the protein are essential for binding has been obtained by chemical removal of some of the charged centers. Within recent years, many investigators have developed methods of treating proteins under very mild conditions with reagents which are quite specific in the groups with which they will react, and which apparently will not injure the structural framework of the macromolecule. These methods have been reviewed by Olcott and Fraenkel-Conrat (1947) and by Herriott (1947). One of the reac-

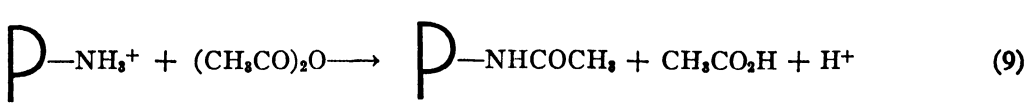


(III) $H_2N-C_6H_4-SO_3^-$



(IV) $H_2N-C_6H_4-SO_2NH_2$

tions used, treatment with acetic anhydride, allows one to block a lysine residue on the protein as indicated by the following equation:



$$P\text{—}NH_3^+ + (CH_3CO)_2O \longrightarrow P\text{—}NHCOCH_3 + CH_3CO_2H + H^+ \qquad (9)$$

One would expect then that acetylated albumin should show a decreased binding ability, compared to the parent, unacetylated protein, if lysine residues are normally involved in the binding. The data in Figure 15 corroborate this interpretation, for the sample of acetyl-albumin (with approximately 90 percent of its amine groups acetylated) shows much less binding ability than the parent material.

A similar experiment has been carried out in which the *absence* of differences in behavior fits in with the postulated picture. It is also possible to

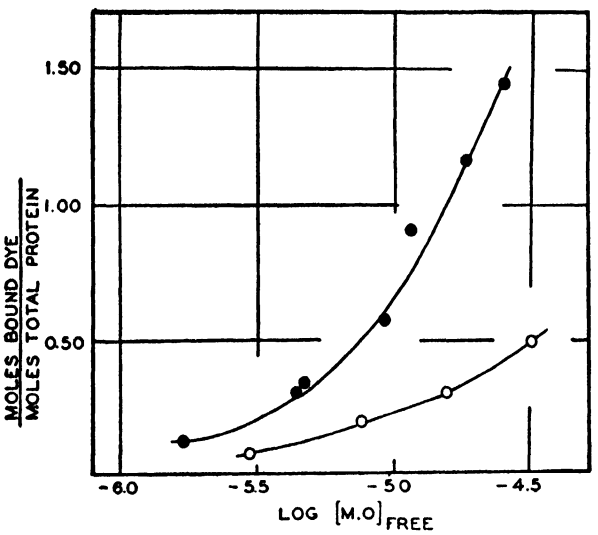


FIG. 15. Comparison of the binding of methyl orange by bovine albumin, ●, and by acetyl-amino bovine albumin, ○.

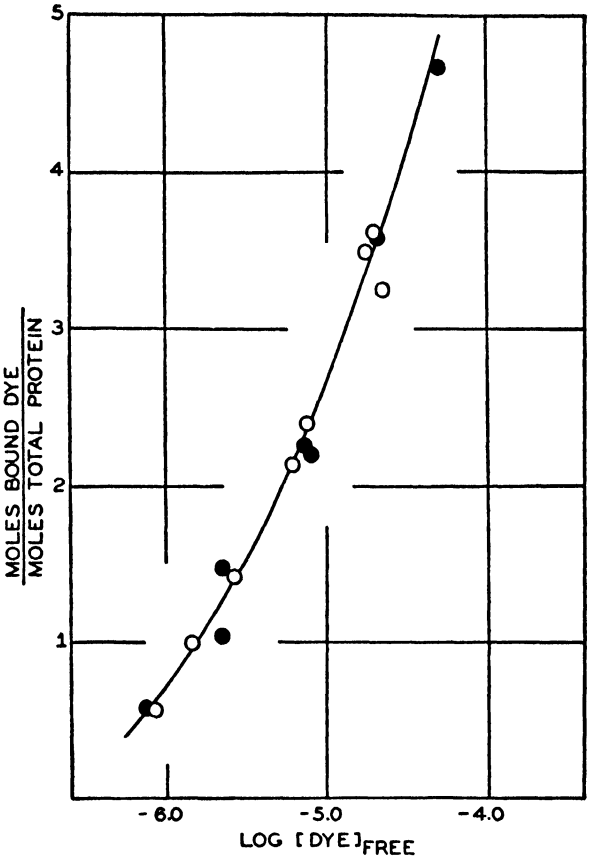


FIG. 16. Comparison of the extent of binding of methyl orange by human serum albumin, ●, and by guanidinated human serum albumin, ○.

change the ammonium groups of the lysine residues to guanidinium groups, by the following reaction:

$$P\text{—}NH_3^+ + CH_3O\text{—}C(=NH)\text{—}NH_2 \longrightarrow P\text{—}NH\text{—}C(=NH_2^+)\text{—}NH_2 + CH_3OH \qquad (10)$$

In this case, a positive locus is retained though it is changed in character. One might expect, therefore, that binding ability should be retained by the guanidinated albumin. As Figure 16 demonstrates, the behavior of the guanidinated protein is not detectably different from that of the parent material.

For the reasons cited, therefore, it seems reasonable to assume that cationic loci on the protein are basically involved in the binding of anions. Confirmatory evidence is also available from the small ΔH's of binding mentioned earlier, for heats of reaction near zero are characteristic of combinations between oppositely-charged species, as will be shown later.

It is possible of course that other residues also take part. In particular, much of the observed behavior is quite consistent with some binding by tyrosine residues. Several investigators (Davis, 1946; Davis and Dubos, 1947) have also suggested that lipophilic residues, such as leucine, may contribute an important portion of the binding energy. Strong evidence for the participation of these other residues has yet to be provided, however.

Since the importance of positively-charged residues seems to be well established, there are two questions in connection with the molecular picture which have yet to be considered. First, it is necessary to provide an explanation for the variation in binding affinity of certain anions with change in structure. In addition there remains the problem of specificity; all proteins have positively-charged residues within them, but only a few, among which serum albumin is the outstanding example, show general anion-binding properties. Both of these

questions are interdependent, and a completely satisfactory answer has not yet been developed. Some clear features seem to be emerging, nevertheless, from recent quantitative investigations.

B. Dependence of Binding Affinity on Structure of Anion

It has been known from several lines of investigation (Steinhardt, 1941; Boyer, Ballou and Luck, 1947) that increasing the molecular weight of an anion in a series of related substances increases the degree of binding by a protein. In the series of aliphatic carboxylic acids, for example, binding affinity increases with chain length. The most obvious explanation for this trend is in terms of increased Van der Waals forces which accompany increased molecular size. Recent quantitative investigations, however, indicate that Van der Waals forces do not play a major role in increasing binding affinity.

It has been pointed out earlier that the temperature dependence of binding is quite small, and that, as a thermodynamic consequence, the heat of the reaction must also be small. For several ions, the heat of reaction has actually been calculated, and following that the entropy change in the reaction determined. Some representative results are shown in Table 4. Two features in this table are highly significant. First, ΔS for each reaction is a positive quantity. Secondly, increased binding affinity, for example in going from chloride to dodecyl sulfate, is accompanied by small changes in ΔH, but by large changes in ΔS (and hence in $-T\Delta S$).

The positive values of ΔS are in themselves surprising, for the reactions as written in equation (1) are association reactions for which one would expect negative entropy changes. If one keeps in mind the charged nature of the anion, I^-, however, a reasonable explanation can be formulated. Though the anion may be represented merely by the symbol I^-, it is recognized that this ion has several polarized water molecules "frozen" to it in the aqueous solution. Similarly, the protein molecule is highly hydrated, particularly perhaps around the charged loci of cationic nitrogen atoms which seem to be so directly involved in the binding process. The formation of a bond between these two oppositely charged species would release some of the "frozen" water molecules. Thus, at a molecular level, there would be an increase in the number of molecular species upon formation of the anion-protein complex, rather than a decrease as indicated by equation (1), and hence one would reasonably expect an increase in entropy for the process.

It is of interest in connection with this interpretation that Lundgren (1945) has observed an increase in the volume of the solution when an alkylsulfonate anion is bound by egg albumin.

TABLE 4. THERMODYNAMICS OF BINDING* OF SOME ANIONS BY SERUM ALBUMIN

Ion	t	$\Delta F°$ cal./mole	$\Delta H°$ cal./mole	$\Delta S°$ cal./mole/deg.	Reference
Chloride	25°	−2,220	400	8.7	(1)
Octyl sulfate	25°	−5,010	0	16.7	(2)
Decyl sulfate	25°	−6,030	−2,000	13.3	(2)
Dodecyl sulfate	25°	−7,220	0	24.0	(2)
Methyl orange	25°	−6,410	−2,100	14.5	(3)
Azosulfathiazole	25°	−7,150	−2,000	17.1	(3)

(1) Scatchard, Scheinberg and Armstrong, 1949; (2) Karush and Sonenberg, 1949; (3) Klotz and Urquhart, 1949.
* The quantities listed refer to the intrinsic constant k for the four ions and to k_1 for the remaining two ions.

This view of the mechanism of the binding process makes it analogous to the ionization process of organic carboxylic acids. Written as an association reaction, the ionic changes in these acids may be represented by the equation

$$H^+ + A^- = HA \qquad (11)$$

For the aliphatic organic acids the entropy change accompanying the reaction shown is around 20 cal./mole/deg. and the enthalpy change of the order of several hundred cal./mole (Harned and Owen, 1943). Here too then, there is an increase in entropy in a reaction which superficially involves a reduction in the number of molecular species. When one recognizes, however, that the disappearance of the free ions is accompanied by the release of fixed water molecules, the positive entropy for reaction (11) becomes reasonable.

On the basis of this water-release picture, the increase in affinity of albumin for larger anions, with its associated increase in entropy change, is to be attributed to greater release of water molecules, greater dehydration, therefore, of the complex containing the large ions. The situation is thus analogous to the ionization of acids just described. In going from formate to acetate ion, for example, there is an increase in affinity for the hydrogen ion and a corresponding increase in the absolute magnitude of $\Delta F°$. The change in $\Delta F°$, however, is due almost entirely to the accompanying change in $\Delta S°$. Again, the difference between two carboxylate ions may be attributed to the lesser degree of dehydration of the neutral molecule with the single carbon atom.

Even where quantitative data of the type presented in Table 4 are lacking, it is still known that the temperature effect on binding ability by albumin is small. It is evident, therefore, that these ion-protein complexes all follow a general pattern in that their heats of reaction are small. Hence in all these cases ΔF is determined by the magnitude of ΔS (see equation 8). It becomes apparent, therefore, that in comparisons of changes in binding energy with variations in structure of the anion, greater attention should be paid to the ability of the ion to release water molecules from the ion-protein complex, rather than to increased Van der Waals interactions.

C. Dependence of Binding Affinity on the Type of Protein

A survey of the results obtained in binding studies with a variety of native proteins (Klotz and Urquhart, 1949c) indicates that proteins may be arranged roughly in a vertical array with respect to binding ability toward a specified anion. When this anion is one which is not bound too strongly, as for example methyl orange, the order of the proteins can be discerned quite clearly (Table 5). Serum albumin stands far above any of the others. β-Lactoglobulin also binds a substantial quantity of methyl orange. Although pure hemocyanin has not been examined, studies with the plasma of *Limulus polyphemus* indicate that this protein is also capable of forming complexes with methyl orange at pH's near the isoelectric point. Egg albumin and casein, in turn, show no affinity for methyl orange at pH's basic to their isoelectric points, but on the acid side, they do give evidence of binding the anion. Apparently the additional attraction of the net positive charge on the protein supplies the missing energy required for a stable complex. Finally there is a host of proteins, typified by serum γ-globulin which give no evidence at all of affinity toward methyl orange.

TABLE 5. ORDER OF BINDING AFFINITY OF PROTEINS

Serum albumin
β-Lactoglobulin
Hemocyanin
Egg albumin, casein
Insulin, ribonuclease, pepsin, trypsin, γ-globulin

On the other hand, with an anion which is bound very strongly, such as the detergent dodecyl sulfate, even proteins low in the scale of binding affinities will form complexes. Thus electrophoretic investigations (Lundgren, 1945) and ultracentrifugal studies (Miller and Andersson, 1942) have revealed complexes of dodecyl sulfate with egg albumin and with insulin.

All of these proteins, with strong or with negligible affinity for the methyl orange anion, have appreciable quantities of the basic amino acids, histidine, lysine and arginine, and hence of cationic loci. The question naturally arises, then, why should there be such marked differences in affinity toward anions?

Since the presence of cationic groups in a protein is not a sufficient condition for binding properties, it has been suggested frequently that one or more other residues, particularly leucine (Davis, 1946), must be in proper juxtaposition to the positively-charged nitrogen to supply the additional specific attraction to insure binding. Although this explanation cannot be ruled out entirely, it seems unlikely on several counts. Since anions of all sizes and shapes, organic and inorganic form associations with serum albumin, it does not seem likely that a particular secondary residue need be at a specific distance from the cationic locus. Furthermore, the replacement of the ε-ammonium group of lysine by the more extended guanidinium group does not produce a decrease in binding affinity of albumin. Finally, as has been pointed out previously, the more favorable free energy of binding of larger ions is not due to the ΔH term but to the ΔS term, that is, Van der Waals type interactions implied by the suggestion of secondary residues, do not seem to play an important role.

Since the presence of a specific secondary amino acid to reinforce the attraction by the cationic group does not seem to be a rigid requirement for binding ability by a protein, the alternative possibility may

TABLE 6. AMINO ACID COMPOSITION AND BINDING ABILITY OF SOME PROTEINS

Protein	$\Sigma(-COO^-)$	$\Sigma(-OH)$	$\Sigma(=NH^+)$	$\frac{\Sigma(=NH^+)}{\lvert\Sigma(OH)-\Sigma(COO^-)\rvert}$
Bovine albumin	133	128	145	29
β-Lactoglobulin	141	118	105	4.6
Egg albumin	98	134	90	2.5
Casein	99	146	100	2.1
Insulin	63	150	72	0.83
Ribonuclease	48	233	128	0.69
Bovine γ-globulin	67	216	99	0.66
Pepsin	109	248	18	0.13

be explored, that the *absence* of a particular pattern might be conducive toward the attraction of anions. In particular, the possibility suggests itself that functional groups within the protein molecule might act as competitive agents for the cationic groups, just as buffer ions have been observed to compete with methyl orange anions (Klotz and Urquhart, 1949a).

To examine this proposal, available amino acid analyses on the proteins of present interest have been converted to a common unit weight (10^5 grams). The total number of cationic nitrogen residues, Σ ($\equiv NH^+$), and of free carboxyl groups, (*i.e.* corrected for amide nitrogens), $\Sigma(-COO^-)$, could then be calculated on a comparable basis for each protein, and the sums obtained (Klotz and Urquhart, 1949c) have been placed in Table 6 .It is apparent from the data, however, that competition between internal carboxylate groups and external anions for cationic centers cannot account for the differences between the proteins, for γ-globulin, or ribonuclease, has a much smaller ratio of $-COO^-$ to $\equiv NH^+$ groups than does the strongly binding serum albumin.

If the sums of the hydroxyamino acids, $\Sigma(-OH)$, are also tabulated for the proteins under consideration, a very interesting trend becomes apparent (Table 6). On the whole, the proteins in the upper part of the list, that is, those with more pronounced anion affinities, have a smaller content of hydroxyamino acids than do the proteins of the bottom half which exhibit little anion affinity.

The rough correlation between content of hydroxyamino acids and binding ability is an experimental fact. Its significance, and interpretation on a molecular basis, however, may be a subject of much discussion. The following molecular picture offers one interpretation which is in accord with the observed behavior and is consistent with present concepts of molecular interactions.

All of the three groups under consideration, $-COO^-$, $-OH$, and $\equiv NH^+$, can form hydrogen bridges, such as are indicated in Figure 17. Not all of these are of equal strength, however. Bond energies for the different types of bond have been estimated from information given by Pauling (1945) and by Davies (1946), combined with a correction for the effect of charges (Pressman, Grossberg, Pence and Pauling, 1946). The results are listed in Figure 17.

It is apparent that the hydroxyamino acids should form hydrogen bonds with carboxyl groups preferentially to ones with the cationic nitrogen. The net result of this interaction would be to decrease the number of $-OH$ and of $-COO^-$ groups which can combine with the $\equiv NH^+$ loci. Consequently, if a protein has a relatively small number of $-OH$ residues, accompanied by a nearly equal number of $-COO^-$ groups, there would be a very great chance that the cationic $\equiv NH^+$ residues would be free and available for attraction of anions. On the other hand, if the protein has a large excess of $-OH$ over $-COO^-$ groups, many of the hydroxyl groups would be available for coupling with the cationic nitrogens even after all the carboxyl groups had been saturated. In such a situation, binding of outside anions by the $\equiv NH^+$ would be greatly hindered.

A molecular interpretation of comparative binding abilities of proteins thus seems possible in terms of internal interactions between polar groups. This qualitative picture can be converted in fact into a quantitative one. In the hydroxyl-carboxyl interaction, it is evident that if the number of either one of these functional groups exceeds that of the other, the excess number of residues

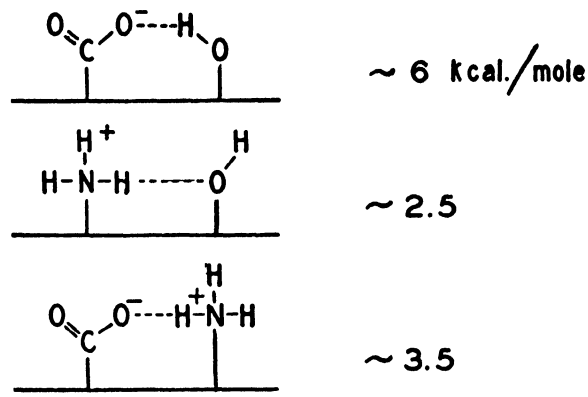


FIG. 17. Some hydrogen bonds between polar residues in a protein molecule, and their dissociation energies.

$|\Sigma(-OH)-\Sigma(-COO)|$, would be available to combine with the cationic nitrogens, and thereby to decrease the ability of the proteins to bind anions. Thus a measure of binding ability might be the ratio $\Sigma(\equiv NH^+)/|\Sigma(-OH)-\Sigma(-COO^-)|$, for the more cationic nitrogens the greater the possibility of binding, whereas the greater the difference between total hydroxyl the total carboxyl groups (regardless of the sign of this difference), the greater the interference with binding.

The correlation between this type of ratio and binding affinity is very satisfactory, as can be observed from the last column of Table 6. Thus serum albumin stands out among the proteins. β-Lactoglobulin, in turn, is clearly superior to its successors in the table, though far below serum albumin. Egg albumin and casein fall next into line. With lower values of the ratio $\Sigma(\equiv NH^+)/|\Sigma(-OH)-\Sigma(-COO^-)|$, these two proteins evidently have such weak binding properties that an observable effect cannot be obtained with methyl orange in solutions basic to the isoelectric point. In solutions acid to the isoelectric point, however, opposite charges on protein and anion tend to increase the concentration of anion at the surface of the protein, and hence, supply the additional free energy needed to produce an ob-

servable degree of binding. With a protein such as γ-globulin, finally, the residue ratio, and hence, the binding affinity, is so low that even the additional contribution of electrostatic attraction at pH's acid to the isoelectric point is inadequate to produce binding.

In view of the excellent agreement between binding ability and the value of the ratio $\Sigma(\equiv NH^+)/|\Sigma(-OH) - \Sigma(-COO^-)|$, it seems worthwhile to use this ratio as a "binding index" from which to predict whether a given protein will show a measurable affinity for anions. The index, at present, is purely a relative one. Nevertheless, it offers a criterion on the basis of which one can place a new protein in its proper position among those in the list in Table 6. With an ion which is bound only weakly, such as methyl orange, only proteins with a high binding index will exhibit observable effects. On the other hand, with an anion, such as dodecyl sulfate, which is bound much more strongly, even proteins with low values of the binding index may form complexes. Dodecyl sulfate is such a strong competing agent, that it is capable of breaking up the internal interactions between cationic nitrogens and the other functional groups on the protein. Thus with anions which can be bound very strongly by albumin, even proteins low in the scale in Table 6 will form stable complexes.

Conclusions

From this survey, it is perhaps apparent that a great deal of progress has been made in filling details into the bare schematic representation of ion-protein complexes shown in Figure 1. The existence of multiple equilibria is well established. Methods of evaluating binding energies have been worked out, and the effects of several important environmental factors have been determined. Although still somewhat in the speculative stage, a molecular picture of the nature of the binding sites is definitely emerging, and is forming a very suggestive model for the design of further experiments.

It is, of course, a long way from the study of ion-protein complexes in relatively simple, artificial systems to the prediction of the effects of ions under physiological conditions. Nevertheless, the elucidation of nature of such interactions under artificial conditions cannot help but serve as a guide to the behavior of these molecules in living systems.

References

Alberty, R. A., 1949, A study of the variation of the average isoelectric points of several plasma proteins with ionic strength. J. phys. colloid Chem. *53:* 114-126.

Boyer, P. D., Lum, F. G., Ballou, G. A., Luck, J. M., and Rice, R. G., 1946, The combination of fatty acids and related compounds with serum albumin. I. Stabilization against heat denaturation. J. biol. Chem. *162:* 181-198.

Boyer, P. D., Ballou, G. A., and Luck, J. M., 1947, The combination of fatty acids and related compounds with serum albumin. III. The nature and extent of the combination. J. biol. Chem. *167:* 407-424.

Chanutin, A., Ludewig, S., and Masket, A. V., 1942, Studies on the calcium-protein relationship with the aid of the ultracentrifuge. I. Observations on calcium caseinate solutions. J. biol. Chem. *143:* 737-751.

Davies, M., 1946, Physical aspects of the hydrogen bond. Rep. Progr. Chem. *43:* 5-30.

Davis, B. D., 1943, The binding of sulfonamide drugs by plasma proteins. A factor in determining the distribution of drugs in the body. J. clin. Invest. *22:* 753-762.

1946, Physiological significance of the binding of molecules by plasma proteins. Amer. Scientist *34:* 611-618.

Davis, B. D., and Dubos, R. J., 1947, The binding of fatty acids by serum albumin, a protective growth factor in bacteriological media. J. exp. Med. *86:* 215-228.

Harned, H. S., and Owen, B. B., 1943, The physical chemistry of electrolytic solutions. New York: Reinhold Publishing Corp.: 514.

Herriott, R. M., 1947, Reactions of native proteins with chemical reagents. Advances in Protein Chem. *3:* 169-225.

Karush, F., and Sonenberg, M., 1949, Interaction of homologous alkyl sulfates with bovine serum albumin. J. Amer. chem. Soc. *71:* 1369-1376.

Klotz, I. M., 1946a, The application of the law of mass action to binding by proteins. Interactions with calcium. Arch. Biochem. *9:* 109-117.

1946b, Spectrophotometric investigations of the interactions of proteins with organic ions. J. Amer. chem. Soc. *68:* 2299-2304.

Klotz, I. M., and Curme, H. G., 1948, The thermodynamics of metallo-protein combinations. Copper with bovine serum albumin. J. Amer. chem. Soc. *70:* 939-943.

Klotz, I. M., Triwush, H., and Walker, F. M., 1948, The binding of organic ions by proteins. Competition phenomena and denaturation effects. J. Amer. chem. Soc. *70:* 2935-2941.

Klotz, I. M. and Urquhart, J. M., 1949a, The binding of organic ions by proteins. Buffer effects. J. phys. colloid Chem. *53:* 100-114.

1949b, The binding of organic ions by proteins. Effect of temperature. J. Amer. chem. Soc. *71:* 847-851.

1949c, The binding of organic ions by proteins. Comparison of native and of modified proteins. J. Amer. chem. Soc. *71:* 1597-1603.

Klotz, I. M., Walker, F. M., and Pivan, R. B., 1946, The binding of organic ions by proteins. J. Amer. chem. Soc. *68:* 1486-1490.

Longsworth, L. G., and Jacobsen, C. F., 1949, An electrophoretic study of the binding of salt ions by β-lactoglobulin and bovine serum albumin. J. phys. colloid Chem. *53:* 126-135.

Lundgren, H. P., 1945, Synthetic protein fibers from protein-detergent complexes. Textile Research J. *15:* 335-353.

Miller, G. L., and Andersson, K. J. I., 1942, Ultracentrifuge and diffusion studies on native and reduced insulin in duponol solution. J. biol. Chem. *144:* 475-486.

Olcott, H. S., and Fraenkel-Conrat, H., 1947, Specific group reagents for proteins. Chem. Rev. *41:* 151-197.

Pauling, L., 1945, The nature of the chemical bond. Ithaca: Cornell University Press: 333-334.

Pressman, D., Grossberg, A. L., Pence, L. H., and Pauling, L., 1946, The reactions of antiserum homologous to the p-azophenyl-trimethyl-ammonium group. J. Amer. chem. Soc. *68:* 250-255.

PUTNAM, F. W., and NEURATH, H., 1945, Interactions between proteins and synthetic detergents. II. Electrophoretic analysis of serum albumin-sodium dodecyl sulfate mixtures. J. biol. Chem. *159:* 195-209.

SCATCHARD, G., 1949, The attractions of proteins for small molecules and ions. Ann. N. Y. Acad. Sci. in press.

SCATCHARD, G., and BLACK, E. S., 1949, The effect of salts on the isoionic and isoelectric points of proteins. J. phys. colloid Chem. *53:* 88-99.

SCATCHARD, G., SCHEINBERG, I. H., and ARMSTRONG, S. H., JR., 1949, Physical chemistry of protein solutions. IV. The combination of human serum albumin with the chloride ion. J. Amer. chem. Soc. *71:* in press.

STEINHARDT, J., 1941, Participation of anions in the combination of proteins with acids. Annals N.Y. Acad. Sci. *41:* 287-320.

TERESI, J. D., and LUCK, J. M., 1948, The combination of organic anions with serum albumin. VI. Quantitative studies by equilibrium dialysis. J. biol. Chem. *174:* 653-661.

DISCUSSION

FREDERICQ: We are studying in Dr. Neurath's laboratory some interactions between insulin and thiocyanate ions. We have been able to show that insulin can bind a considerable amount of thiocyanate. Moreover, the state of aggregation of insulin is greatly influenced by low concentrations of thiocyanate, the effect of this ion being to promote association of the insulin subunits. Solubility and electrophoretic mobility are likewise strongly affected. In this connection, I would like to ask Dr. Klotz if he has any evidence of similar cases of relations between the binding of small ions by proteins and important effects on their physical properties and also if the fact that insulin binds large amounts of thiocyanate can fit into his classification of proteins with respect to their binding ability?

KLOTZ: There are several other examples of changes in physical properties of a protein which has bound one or more small ions. To mention one which we have investigated very recently, I might refer to the changes in absorption spectra and in extent of ionization of iodinated serum albumin when it forms complexes with dodecyl sulfate ions. Many laboratories have also shown changes in electrophoretic properties. Putnam and Neurath, for example, have demonstrated marked modifications in electrophoretic properties of serum albumin in the presence of dodecyl sulfate. Similarly, Velick earlier this year published some data showing a shift of several pH units in the isoelectric point of aldolase when it binds phosphate ions.

Insofar as the binding of thiocyanate ions by insulin is concerned, the results mentioned by Dr. Fredericq seem quite reasonable from our viewpoint. Thiocyanate ion is bound fairly strongly, although not as well as methyl orange, by serum albumin. Scatchard and Black have published data which show roughly 10 thiocyanate ions per albumin molecule at a concentration of 10^{-2} molar thiocyanate, and over 20 bound ions in 10^{-1} molar solutions. Qualitatively speaking one would expect that an ion bound so strongly by serum albumin should also be bound to some extent by insulin, despite the smaller binding index of the latter protein. It may be pertinent to point out also, that binding would be increased in solutions acid to the isoelectric point, where I believe Dr. Fredericq has carried out his studies, because of known electrostatic influences.

HAUROWITZ: We determined the ratio cationic detergent/hemoglobin at different pH and found that more than 200 detergent molecules were bound per protein molecule. This is more than the whole number of acidic groups. Apparently micelles of the detergent were bound. Similar observations have been published by Neurath and others, who investigated anionic detergents. The detergent used in our experiments was dodecyltoluyl-dimethylammonium salt.

KLOTZ: It is quite possible that micelles, as well as single ions, are bound by proteins when the ions are long-chain detergent molecules. It is necessary to keep in mind though that investigations which indicate binding of micelles have been carried out very close to, or above, the critical concentration for micelle formation of the detergent. With ions which do not form micelles readily, a stoichiometric relationship with the number of cationic, or anionic groups, on the protein has been observed in many cases. Dr. Perlmann, for example, has shown that the binding of metaphosphate by different proteins is stoichiometric with the number of basic amino acids. Similar observations with certain dye ions have been reported by Fraenkel-Conrat. In investigations with lysozyme, we have found that with sufficiently small anions, the number of bound ions does not exceed the number of cationic groups on the protein, but with large anions, which may form micelles readily, the number bound may be greater than the number of positively charged sites.

LUCK: I would like to congratulate Dr. Klotz on this very stimulating discussion of various aspects of the problem of protein-ion interactions. With respect to the theory of anion binding which he presented in conclusion, two comments would appear to be in order.

a) Application of this theory to human serum albumin would yield a binding coefficient of only 4 (calculated from data of Brand) instead of 29 for bovine serum albumin. This would suggest a very low binding capacity, whereas, in our experience at Stanford, the binding of anions by human serum albumin is fully equal to that by bovine serum albumin.

b) Does the theory also not require a juxtaposition of hydroxy amino acids and dicarboxylic amino acids on a scale that could hardly be expected to be realized? It would seem, according to the theory, that all of the OH groups or of the COO^- groups have to be blanked out by interformation of hydrogen bonds which could only be realized if a large

scale contiguity of such groups existed. Our own theory, which is less amenable to experimental testing than that of Dr. Klotz, suffers from a similar defect in that again a close juxtaposition of COO^- groups with certain other groups is required. However, we would ask only for an adequacy of nonpolar side chains and since the number of valyl, leucyl, isoleucyl and phenylalanyl residues in these serum albumins is about twice that of the OH groups we think the chances of establishing the desired juxtaposition with nonpolar side chains is greater than with OH groups.

KLOTZ: In comparing the binding abilities of two proteins with very high binding indices, as calculated by our methods, it is necessary to proceed with caution. The very high binding index for bovine albumin arises from the fact that it involves, in part, the difference between two large numbers, the content of free carboxyl and of hydroxyl groups, respectively. A very small error in the analytical values for the corresponding amino acids would produce a marked change in the binding index. New data on bovine albumin just published by Stein and Moore show substantial differences in contents of these critical amino acids from those given by Brand. From Stein and Moore's data, we find the binding index for bovine albumin to be 7.5 instead of the 29 calculated from Brand's results. A value of 7.5 for bovine albumin does not compare too unfavorably with 4.6 for human albumin, especially since newer analyses may also modify the latter slightly. It seems most interesting to us, at present, that the binding index predicts high affinities for both bovine and human albumins, as compared to the other proteins examined, and that these two albumins are indeed the proteins with the greatest ability to bind anions.

With regard to Dr. Luck's second comment, molecular models which we have made indicate that an unusual juxtaposition of dicarboxylic and hydroxy amino acids is not required to produce the hydrogen bonding we have postulated. If we use the interatomic distances given in Corey's recent summary of X-ray work on amino acids and peptides, and assume an extended chain model for the protein, we find that a lysine side-chain (as an example) has sufficient free rotation that it could extend over to hydroxyl side-chains in the same backbone chain even if the hydroxyl groups were two amino-acid residues removed in either direction along the backbone. The lysine group could also couple with one of several side-chains on either of the two neighboring backbone chains in the same plane as its backbone chain. There remains in addition the possibility of interaction with hydroxyl side-chains from backbone chains above and below that of the lysine. Thus the lysine residue has a wide sphere of possible interaction. With all of these possibilities for encountering a hydroxyl side-chain, it seems likely that the postulated hydrogen bridges could occur without any special juxtaposition of the polar amino acids.

TAYLOR: The ratio suggested by Dr. Klotz is 3.7 for the enzyme aldolase and 2.6 for the enzyme d-glyceraldehyde-phosphate dehydrogenase from rabbit muscle. Dr. Velick has shown that aldolase binds phosphate quite strongly. I believe that it binds methyl orange much less strongly than serum albumin does.

KLOTZ: Dr. Taylor's comments are of particular interest to us since the results mentioned are in agreement with predictions made from the binding indices of the proteins described.

THE APPLICATION OF THE ISOTOPIC DERIVATIVE TECHNIC TO THE STUDY OF PROTEIN STRUCTURE

MILTON LEVY AND EVELYN SLOBODIANSKY
New York University College of Medicine, New York

A conservative opinion of the structure of proteins is that they are essentially polypeptides formed by amide condensations between alpha amino acids. The Fischer-Hofmeister peptide hypothesis has the support of most chemists concerned with proteins. The great number of possible isomers stemming from the existence of more than 20 different amino acids was recognized early. We shall adopt the view that the order of the amino acids in the polypeptide chains is of primary importance, both for the physical nature of the individual proteins and for their biological properties. At the conclusion of this paper, some other structural principles will be discussed, but the data and experiments which are reported in the body of the paper are concerned with the organic chemistry of silk fibroin from *Bombyx mori*. This was chosen for study because of the presence in it of large quantities of two amino acids, glycine and alanine, and because of the previous studies of its partial hydrolysis products by Fischer and Bergell (1903), Fischer and Abderhalden (1906a, b, 1907), Abderhalden (1909a and b, 1910, 1922, 1923a and b), Abderhalden and Suwa (1910) and Stein, Moore and Bergmann (1944) which tend to show that alanylglycine and glycylalanine are present in large amounts in properly prepared hydrolysates.

The extent and nature of the problem which the polypeptide hypothesis poses for the organic chemist finds its closest analogy in written language. Using about 30 symbols it is possible to express millions of ideas and relationships, and to produce in words, sentences, paragraphs and books a great many functional entities (functioning to convey or conceal ideas) with a wide variety of nuances. It is even possible for the human mind to grasp this almost infinite variety of functional arrangements and to derive from them not only their functional significance but also to reduce them to rules of grammar, of spelling and of good usage. One may inquire whether it would be possible for one who knew nothing of the meaning of this symbolism to decipher it and to discover its working principles. Perhaps the first step such a novice would take would be to count the relative occurrences of the various symbols (having first recognized their differences). This would correspond to the analysis of proteins for their amino acids. The similarity of the analysis of language to the analysis of a protein is, if we use a modification of the symbolism advocated by Dr. Brand, rather striking. Thus a certain sentence has the composition A_{10} B_2 C_{10} D_5 E_{13} F_5 G_2 H_{10} I_{12} L_4 M_4 N_9 O_{12} P_2 Q_2 R_{10} S_{10} T_{21} U_7 Y_6 sp_{28} co_1 pe_1. There are 186 residues in this composition, so that it would correspond to a protein with a molecular weight of about 17,000; there are 23 kinds of residues just as there are about this many amino acids. The analysis is perfect, since it was done by counting. The sentence is probably unique in composition compared to other sentences. On the other hand, sentences of quite other composition would fulfill almost the same function (*i.e.*, convey nearly the same meaning) and nearly the same composition could very well convey a different meaning if the arrangement of the symbols were different. All of the above statements can be made because we know the functions of the symbols of written language.

However, if we do not know the functions of the residues nor their relationship to the overall properties of the proteins we can at least attempt to find their order in the chain and to recognize in the chain those combinations of residues which often occur together. This represents about the position at present attained concerning the arrangement of amino acids in proteins. We know the amount of each residue in some cases and can hope to determine this in each case whenever we wish, using the elegant methods which have been described at this symposium. In some instances we have recognized combinations of two or more residues, and have even penetrated from the ends of the chain for a few residues. The magnitude of the task before us is perhaps evident when we realize that the 186 residues in the sentence analyzed above is less than the number of residues in an ordinary protein. A paragraph is more nearly representative of the complexity of a protein. One can hardly doubt that several hundred paragraphs (a book) represents the complexity of a cell and certainly quite a large library is no more complex than some higher organisms.

Attempts to solve the problem of arrangements of amino acids in proteins were reviewed by Synge (1943). This excellent review expresses the intent of studies of partial hydrolytic products and reviews the data which had been accumulated up to its date.

When the isotopic derivative method showed its potentiality and accuracy in the first studies made by Keston, Udenfriend and Cannan (1946, 1949), it occurred to us that the method could be used for the accurate estimation of the amounts of the peptides known to be present in partial hydrolysates of silk. We therefore prepared the carriers for the *para* iodobenzenesulfonyl (pipsyl) derivatives of glycylglycine, alanylglycine and glycylalanine. We analyzed known mixtures of the peptides with each

TABLE 1. COMPOSITION OF SILK FIBROIN

	Percent of N	Percent of peptide bonds	Ratios	Percent N per residue
Glycine*	42.5	43.3	81	0.534
Alanine*	28.2	28.7	54	0.531
Serine†	9.4	9.6	18	0.533
Aspartic acid†	1.54	1.57	3	0.523
Glutamic acid†	1.07	1.09	2	0.504
Hydroxyproline†	<.05	—	—	—
All others (by difference)		15.94	30	0.531
Total residues (minimum)			188	0.532

* Analysis by carrier technic (Keston, Udenfriend and Cannan, 1949).

† Analysis by indicator technic (Keston, Udenfriend and Levy, 1948).

other and with alanine and glycine for each of the five components by the isotope derivative technic as described by Keston, Udenfriend and Cannan (1949). The recoveries were satisfactory and "zero" experiments, in which the substance determined was omitted, gave satisfactorily low values (less than 0.5 percent of the total N present).

A sample of silk fibroin was prepared (degummed by autoclaving). It contains 19.5 percent of nitrogen on an ash and moisture free basis, and the ash was less than 0.05 percent. The composition as determined on a complete hydrolysate is shown in Table 1. The analysis is given in terms of the percentage of the total nitrogen in each amino acid. Assuming that the alpha amino nitrogen of the hydrolysate is 98 percent of the total nitrogen (Stein, Moore and Bergmann, 1944) the fraction of the total peptide bonds derived from each of the amino acids listed has been calculated and entered in the third column of Table 1. The fourth column of figures is calculated on the assumption that the polypeptide conforms to the laws of ordinary organic chemistry, that is, that residues are present in whole number ratios to one another. The extent of agreement for the groups estimated is shown in the fifth column, which gives the nitrogen per residue. It may be pointed out that, while all of the ratios are expressible as 2^n3^m (except "all others"), it is not possible to combine the glycine and alanine simultaneously into a periodic arrangement. The figures therefore do not support the hypothesis of Bergmann and Niemann (1936).

Five grams of the same silk were dissolved in 10 ml. of 12N HCl and held at 39° in a constant temperature bath. At 16 hours, 24 hours and 48 hours, samples were removed, diluted with water and neutralized with NaOH. A drop of 1 percent mercuric chloride solution was added and the mixture diluted to 25 ml. Nitrogen was estimated in each sample as well as the amounts of the three peptides mentioned and of alanine and glycine with the results shown in Table 2. The significant points of this table are the gradual increases of the amounts of glycine, alanine, alanylglycine and glycylalanine and the small amounts of glycylglycine found. The amount of alanylglycine found is much greater than the amount of glycylalanine. In the 48 hour hydrolysate practically all of the alanine is accounted for as free alanine or in the two peptides of alanine estimated. Almost half of the alanine in this solution is in the single dipeptide alanylglycine.

The simplest and perhaps most naïve assumption we can make about the arrangement of the amino acids in silk is that the arrangement is completely random. Within the limits of the composition, peptide bonds are as likely to be formed between any pair of amino acids as between any other. On this basis we may calculate the maximum possible yields of each dipeptide from the composition. This is simply the product of the fraction of each component of the dipeptide with the fraction of the other. Thus, for glycylglycine we would predict $0.425 \times 0.425/0.98 = 0.182$ (the 0.98 is introduced to correct for the non-peptide nitrogen). For alanylglycine and for glycylalanine the fraction 0.122 is predicted on this basis. Since more than twice the calculated amount of alanylglycine is found, in spite of the obvious hydrolysis of some of the dipeptide, it is clear that the arrangement cannot be random. Furthermore it seems rather unlikely that much glycylglycine was ever present since it is hydrolysed by acids only about 60 percent more rapidly than is alanylglycine. The near absence of glycylglycine from the hydrolysates is good evidence that glycine does not form an extensive polyglycine chain in silk.

Nothing is known of the kinetic statistical factors, ordinarily expressed as rate constants for the incorporation of amino acids in a protein chain. It is therefore not possible to introduce and use the interesting suggestion first made to us by our colleague Dr. A. S. Keston that, after allowing for these rates, the actual incidence of a specific amino acid at a particular point in a particular protein molecule is a matter of chance. Thus if the relative

TABLE 2. ANALYSIS OF PARTIAL HYDROLYSATES OF SILK FIBROIN

One gram of silk and two ml of 12N HCl incubated at 39°; entries are percent of total nitrogen in each compound

	16 hrs.	24 hrs.	48 hrs.
Glycine	4.4	7.1	12.9
Alanine	3.1	6.3	10.5
Alanyl glycine	16.9	23.3	27.0
Glycylalanine	5.4	9.0	8.3
Glycylglycine	0.1	lost	1.8
	% of amino acid recovered		
Glycine	37	55	77
Alanine	50	79	98.5

rates of incorporation of alanine and glycine after the arrangement —XGXGAG— are 100 to 1, then the chances are 100 to 1 that the sequence —XGXGAGA— will be formed rather than —XGXGAGG—. Both arrangements might be equally useful and essentially subserve the same function in silk. Thus, two peptide chain molecules might not have the same arrangement of amino acids. Deductions such as that made above from the yields of dipeptides do not remove this kind of randomness as a possible principle in determining the order of amino acids in proteins.

The composition of another set of silk hydrolysates, this time with one gram of silk to 5 ml of 12N HCl, is shown in Table 3. In this case again the yield of alanylglycine is greater than that which random arrangement would allow. The accumulation of dipeptides in this hydrolysis is not as marked as in the first case but this is due to the slower rate of production of the dipeptides and not to a greater rate of liberation of the amino acids. In fact, if one examine the rate of liberation of free glycine and alanine in the two instances, they are found to follow the kinetics of a first order reaction with the constant independent of the ratio of silk to hydrochloric acid but different for the two amino acids. The first order constants for the different times of hydrolysis are given in Table 4 (logs to the base 10 are used and the time is in hours).

At first it seemed puzzling that so complex a set of reactions as those that lead to the liberation of an amino acid by hydrolysis should follow such a simple law. However the characteristics of a series of consecutive reactions are determined largely by the characteristics of the slowest step in the series. From the fact of accumulation of large amounts of dipeptides, the hydrolysis of the latter must be the slowest step in the hydrolysis of the protein, and therefore the rates of Table 4 are determined by the rates of hydrolysis of the glycine and alanine peptides respectively. The fact that the glycine rate is less than the alanine rate indicates that a glycyl peptide of greater stability than glycylalanine or alanylglycine is present in the hydrolysates. The

TABLE 3. ANALYSIS OF PARTIAL HYDROLYSATES OF SILK FIBROIN

One gram silk and five ml of 12N HCl incubated at 39°; entries are percent of total nitrogen in each compound

	8 hrs.	24 hrs.	48 hrs.	72 hrs.
Glycine	2.5	7.8	13.0	17.4
Alanine	2.1	6.1	9.6	13.7
Alanylglycine	10.8	11.9	17.4	13.7
Glycylalanine	3.1	6.0	5.0	1.9
Glycylglycine	2.4	0.5	0.3	1.6
	% of amino acid accounted for			
Glycine	27.7	37.6	57.5	63.6
Alanine	32.4	52.6	73.7	77.5

TABLE 4. RATE OF LIBERATION OF GLYCINE AND ALANINE FROM SILK FIBROIN

Entries are the values of $(1/t) \log_{10} A/A-x$ for t in hours, A the total amino acid in the fibroin and x the amount free at time t

One gram silk to two ml 12N HCl

t (hours)	16	24	48
Glycine	0.0030	0.0033	0.0033
Alanine	0.0031	0.0046	0.0042

One gram silk to five ml 12N HCl

t (hours)	8	24	48	72
Glycine	0.0032	0.0036	0.0032	0.0032
Alanine	0.0042	0.0044	0.0040	0.0040

relationship of the charges on the groups adjacent to a peptide bond to its resistance to hydrolysis is discussed by Synge (1943). The situation with respect to the liberation of alanine and glycine suggests that all of these two amino acids liberated in the hydrolysis passed through dipeptides.

The finding that about twice as much alanylglycine as glycylalanine is present in the hydrolysates suggests to us that most of the alanine in the polypeptide is in the arrangement —AGAG—. By hydrolysis of this combination, two moles of alanylglycine can be formed but only one mole of glycylalanine. This tetrapeptide group must be adjacent on each side to an amino acid other than alanine or glycine. The analysis indicates that there are two alanyl residues per three glycyl residues and that these five residues are present with two others which are neither glycine nor alanine. The following arrangement accounts for these facts and might be called a minimum repeating peptide arrangement.

—XAGAGXG—

From the analysis again, one third of the Xs are serines and one can venture to write the peptide as:

—XAGAGXGSAGAGXGSAGAGXGXAGAGXGS—

Of the remaining X's, one of each 12 is aspartic acid and one of each 18 is glutamic acid.

What speculation can we base upon these findings with respect to the structure of silk? The small repeating unit of a tetrapeptide (—AGAG—) in silk suggests that silk and probably other proteins are built from peptides, rather than by adding amino acids to the ends of preëxisting long chains.

A problem just as important as the order of the amino acids in the peptide chain but not so clearly understood involves the cross linkages between or within polypeptide chains, whether these are ordinary covalent bonds such as the disulfides or the less fixed hydrogen bonds. One is tempted to call the latter "slippery" bonds, both because they seem

somewhat elusive to the investigator of proteins and because they are kinetically active and the pairs of donor and acceptors seem to be able to slip into new partnerships with considerable ease (as in denaturation). It may be noted that hydrogen bonds are as strong as or stronger than peptide bonds. The free energies of formation of hydrogen bonds vary from −1000 to −10,000 calories per mole (Pauling, 1939) whereas peptide bonds have free energies of formation which are positive and between 1000 and 4000 calories per bond. Even if we note that a potent competitor in hydrogen bond formation (water) is present, it seems likely that some of the hydrogen bonds would be stronger than some of the peptide bonds. Let us assume that the hydrogen bonds are formed first between polypeptide chains smaller than the proteins and at the site of protein formation. The resulting lattice of hydrogen bonded polypeptides might be considered a crystallite. This crystallite which might have the essential configuration of the native protein is then stabilized by the formation of peptide bonds between the ends of the polypeptide chains. The kinetic-stability of the peptide bond adds to the strength of the hydrogen bond to give a structure which is sufficiently metastable to survive the viscissitudes of the life of a protein molecule. The relationship of this proposed mode of formation to the concept of a globular protein as a monomolecular crystal (Wu, 1931) is obvious.

Perhaps through finding what the pieces are we will be able to discover the scheme of arrangement of amino acids in proteins and discover whether or not the above hypothesis is reasonable; but at any rate it is necessary to know the order of the amino acids in the peptide chains to interpret the properties of proteins.

Acknowledgments

One of us (E. Slobodiansky) held a fellowship from the Stanley Tausend Foundation during the course of the work described. The work was further supported by grants made to R. K. Cannan and to M. Levy by the American Cancer Society, on recommendation of the Committee on Growth of the National Research Council. The radioactive iodine and sulfur used was supplied by the Oak Ridge National Laboratories on allocation by the Atomic Energy Commission.

References

Abderhalden, E., 1909a, Weiterer Beitrag zur Kenntnis der bei der partiellen Hydrolyse von Proteinen auftretenden Spaltprodukte. Hoppe-Seyl. Z. *62:* 315-321.

1909b, Hoppe-Seyl. Z. *63:* 401-404.

1910, Hoppe-Seyl. Z. *65:* 417-419.

1922, Beitrag zur Kenntnis der Zusammensetzung des Seidenfibroins und seiner Struktur. Hoppe-Seyl. Z. *120:* 207-213.

1923a, Ueber die Struktur der Proteine. Hoppe-Seyl. Z. *128:* 119-128.

1923b, Weitere Studien uber den stufenweisen Abbau von Eiweissstoffen. Hoppe-Seyl. Z. *131:* 284-295.

Abderhalden, E., and Suwa, A., 1910, Weiterer Beitrag zur Kenntnis der bei der partiellen Hydrolyse von Proteinen auftretenden Spaltprodukte. Hoppe-Seyl. Z. *66:* 13-18.

Bergmann, M., and Niemann, C., 1936, On blood fibrin. A contribution to the problem of protein structure. J. biol. Chem. *115:*77-85.

Fischer, E., and Abderhalden, E., 1906a, Bildung eines Dipeptids bei der Hydrolyse des Seidenfibroins. Ber. chem. es. *39:* 752-760.

1906b, Bildung von Dipeptiden bei der Hydrolyse der Proteine. Ber. chem. Ges. *39:* 2315-2320.

1907, Bildung von Polypeptiden bei der Hydrolyse der Proteine. Ber. chem. Ges. *40:* 3544-3562.

Fischer, E., and Bergell, P., 1903, Ueber die Derivate einiger Dipeptide und ihr Verhalten gegen Pankreasfermente. Ber. chem. Ges. *36:* 2592-2608.

Keston, A. S., Udenfriend, S., and Cannan, R. K., 1946, Microanalysis of mixtures (amino acids) in the form of isotopic derivatives. J. Amer. chem. Soc. *68:* 1390.

1949, a method for the determination of organic compounds in the form of isotopic derivatives. I. Estimation of amino acids by the carrier technique. J. Amer. chem. Soc. *71:* 249-257.

Keston, A. S., Udenfriend, S., and Levy, M., 1947, Paper chromatography applied to the isotopic derivative method of analysis. J. Amer. chem. Soc. *69:* 3151-3152.

Pauling, L., 1939, The nature of the chemical bond and the structure of molecules and crystals. Ithaca, Cornell University Press.

Stein, W. H., Moore, S., and Bergmann, M., 1944, Aromatic sulfonic acids as reagents for peptides. Partial hydrolysis of silk fibroin. J. biol. Chem. *154:* 191-201.

Synge, R. L. M., 1943, Partial hydrolysis products derived from proteins and their significance for protein structure. Chem. Rev. *32:* 135-172.

Wu, H., 1931, Studies on denaturation of proteins XIII. A theory of denaturation. Chin. J. Physiol. *5:* 321-344.

Discussion

Haurowitz: The bond energy of a hydrogen bond is certainly not higher than that of a peptide bond. The value of 10,000 calories might be true for the complete separation of water from a certain molecular group. In an aqueous system, however, the separated partners of a hydrogen bond will combine again with water or with other polar groups, so that most of the energy required for the cleavage will be liberated again. The over-all energy difference will hardly exceed 1000 calories.

Levy: Even if the free energies of formation in aqueous media of the specific hydrogen bonds determining the final structure are of the order of −1000 calories the extents of the spontaneous formations are considerably greater than the extents of the spontaneous formation of peptide bonds by reversal of hydrolysis ($\Delta F = +2000$ to $+4000$ cal. per mole). It is this difference in spontaneity of reaction which leads to the proposition that the hydrogen bonds required for the final and specific folded structure of the native proteins are formed before the final peptide bonds and that the final structure is a metastable one dependent on the sluggish kinetics of peptide hydrolysis in the absence of catalysts.

STRUCTURE AND ENZYMATIC BREAK-DOWN OF PROTEINS

K. LINDERSTRØM-LANG
Carlsberg Laboratory, Copenhagen, Denmark

I

Since the beginning of quantitative research in the field of protein chemistry much weight has been laid upon investigations attempting to prove that proteins are attacked by the same enzymes that are able to split peptide bonds in the simpler compounds. Numerous more or less reliable demonstrations of this fact have been regarded as valuable evidence that proteins have a peptide chain structure. On various occasions (Linderstrøm-Lang, 1933, 1939a, b) the author has issued warnings against a too optimistic opinion about this line of argumentation, the main points being:

1. The difficulty of establishing the homogeneity of an enzyme even when isolated as a crystalline protein.

2. The possibility that one and the same enzyme has several catalytic functions associated with spatially separated groupings in the molecule.

3. The existence of a reversible denaturation giving rise to an equilibrium

$$G \rightleftarrows D$$

between native and denatured protein molecules (Anson and Mirsky). Under these circumstances the splitting of a protein, even by a pure enzyme with strict peptide bond specificity, can tell nothing *a priori* about the presence of peptide bonds in the native molecules.

In the present discussion the question will be left undecided as to whether a multiple action of an enzyme is better explained in terms of a lack of homogeneity or as a simultaneous or subsequent action of several catalytic centers on the same enzyme molecule. The main problem will be the following:

Does a proteinase, the ultimate action of which unquestionably is a release of amino and carboxylic groups, directly attack the native protein molecule by splitting "available" peptide bonds in this molecule, or is this "endopeptidase" action preceded by a denaturation? In the latter case, is the denaturation spontaneous or is it initiated by the enzyme which thereby exerts a double action, namely, that of a denaturase and that of a peptidase? It is hardly possible to give a definite answer to these questions, the main reason being that there is no general answer valid for all proteins and proteinases. Some of the experiments reported in the following may however throw some light on the tryptic break-down of one protein, namely β-lactoglobulin, with which our group has been mainly concerned.

II

It is quite natural that a process in which two proteins are involved, and in which one of these proteins is broken down to lower peptides, cannot be described in great detail. In recent years the investigations of Tiselius and Eriksson-Quensel (1939), Haugaard and Roberts (1942) and others (Beloff and Anfinsen, 1948) have indicated that in the break-down of some proteins by pepsin end-products accumulate without formation of appreciable quantities of intermediary products, in fact as if by the touch of the enzyme the molecules exploded, one by one, to form the peptides of the length and composition that is characteristic of the protein digest after exhaustive enzymatic hydrolysis. It is quite clear that under these conditions no detailed picture of the type and sequence of the breakages can be obtained. However, we know of several other processes in which proteins are attacked by enzymes and where this "one by one" principle is not followed, *e.g.* the activation processes of zymogens or the (recently studied) formation of plakalbumin from ovalbumin. Here well-defined intermediary products are formed (see Jacobsen, 1947; Linderstrøm-Lang and Ottesen, 1949; Eeg-Larsen, Linderstrøm-Lang and Ottesen, 1948). These substances have a certain, often a pronounced stability, but if we refrain from teleological considerations, we may say that in the long run they are only stepping stones in the total degradation of the protein molecule.

In order to understand the divergent behavior of different enzymes, it is advisable to drop the picture of an explosion of the protein molecule, since it may emphasize too much the idea that chain reactions of some kind take place involving a large number of almost simultaneous breakages of peptide bonds. This is by no means necessary or probable although it cannot be entirely excluded. The simplest assumption that may explain the phenomena observed and at the same time secure the continuity of our description of enzyme action is that the protein molecule in some cases, prior to its break-down by endopeptidase, must undergo an initial reaction by which it is prepared for the attack of this enzyme. This initial reaction may be a denaturation or it may involve the splitting of a few strategically important peptide bonds. What is essential in the present connection is whether its rate is smaller or higher than that of the endopeptidase action proper. If it is much smaller the "one by one" reaction will be realized, and if it is of the same

order of magnitude we shall expect formation of intermediary products in considerable quantities. The following simple calculation will show the essential points of this consideration.

The protein Qs, having the concentration S, (initial concentration S_0) is attacked by an enzyme Q_E, total concentration E_0, in an initial process leading to Q_D, (denatured protein). Q_D has n peptide bonds that may be attacked by the same enzyme Q_E. The affinity of Q_E to each of these peptide bonds and the rate of splitting are assumed to be the same, independent of the position of the linkage in the molecule Q_D. If these assumptions are valid also for the further split-product of Q_D, we may write

$$-\frac{dS}{dt} = k[ES] \tag{1}$$

$$-\frac{dP}{dt} = n\frac{dS}{dt} + k_1[EP] \tag{2}$$

where P is the concentration of "available" peptide bonds, and [ES] and [EP] are the concentrations of the enzyme-substrate compounds. Hence

$$-\frac{dP}{dS} = n - \frac{k_1}{k}\frac{[EP]}{[ES]} = n - \frac{k_1}{k}q\frac{P}{S} \tag{3}$$

where q is the distribution factor of the enzyme between Q_S and the available peptide bonds.

Integrating (3) we obtain

$$\frac{P}{S_0} = \frac{n}{1 - k_1q/k}\left[\left(\frac{S}{S_0}\right)^{k_1q/k} - \frac{S}{S_0}\right] \tag{4}$$

The number of peptide bonds split is given by

$$B = nS_0 - nS - P = \frac{nS_0}{1 - k_1q/k}\left[1 - \frac{k_1}{k}q\left(1 - \frac{S}{S_0}\right) - \left(\frac{S}{S_0}\right)^{k_1q/k}\right] \tag{5}$$

hence

$$\frac{dB}{dS} = \frac{-n}{1 - k/qk_1}\left(1 - \left(\frac{S}{S_0}\right)^{k_1q/k-1}\right) \tag{6}$$

For

$$\frac{k_1}{k}q = 1, \quad \frac{dB}{dS} = n\ln\frac{S}{S_0}$$ (appearance of intermediary products)

For

$$\frac{k_1}{k}q \gg 1, \quad \frac{dB}{dS} = -n$$ (one-by-one process)

Hence if k is small in comparison with k_1 or if q is large (enzyme preferentially bound by "peptide bonds") n peptide bonds will be split for every molecule of Qs that disappears. However, if k_1q/k is unity the number of peptide bonds split for each disappearing Qs-molecule will increase as the enzymatic process proceeds, which means that potential substrates for the endopeptidase are found in noticeable quantities.

Now if the course of the break-down is determined by such a competition between two or more processes (and beforehand one would be inclined to assume this) any change in the conditions (temperature, pH etc.) under which the enzymatic reaction takes place would be likely to alter this course. The same enzyme may therefore in one case give rise to intermediary products, in another case not. Reference may here be made to a recent investigation by Moring-Claesson (1948) in which it was found that the character of the split products formed from ovalbumin by the action of pepsin undergoes a definite change as the reaction proceeds. This result is apparently incompatible with the previous work by Tiselius and Eriksson-Quensel, but the difference may be explained by a difference in the enzyme preparations or in the conditions of the reaction.

In our preliminary study of the volume change accompanying the enzymatic break-down of β-lactoglobulin at 30° and pH about 8, Jacobsen and the author (1941) obtained well reproducible results that definitely excluded a one-by-one reaction (see below). When, toward the end of the war and in the period following, the experiments were resumed and extended to include measurements at different temperatures, results were obtained which did not agree too well with our previous ones and which further revealed the fact that the course of the enzymatic hydrolysis by trypsin was essentially different at different temperatures. The lack of reproducibility had to be ascribed to differences in substrate or enzyme without further proof, since none of our older preparations were now available; but with our new preparations consistent results were again obtained and the temperature effect was studied. Some of our results will be reported in the following. The experiments have been carried out in collaboration with Gordon Johansen, M. Rørvig, H. Levi, C. F. Jacobsen and Korsgaard Christensen.

III

The cleavage of peptide bonds is accompanied by a volume change that is essentially due to the electrostriction effect of the new electric charges formed:

$$R_1CONHR_2 \rightarrow R_1COO^- + NH_3^+R_2$$

This effect will depend to some extent upon the spatial arrangement of neighboring ionic charges (in R_1 and R_2) but will in general be about −18 ml/

mole (−15 to −20 ml/mole) if the distance to other charges is larger than 5 A and if both the carboxyl and the amino group are fully ionized. (Reference may here be made to papers by Weber and Nachmansohn, 1929; Weber, 1930; Cohn, McMeekin, Edsall and Blanchard, 1934; McMeekin, Cohn, and Weare, 1935; and for general references, Cohn and Edsall, 1943). Investigations of the electrostriction effect in the pH-region 6-10, namely, the range of ionization of the amino groups (Linderstrøm-Lang and Jacobsen, 1941) showed that the above is valid in the whole region provided no other buffers are added than NH_3 — NH_4^+. Our experiments were therefore carried out with solutions containing this buffer or no buffer.

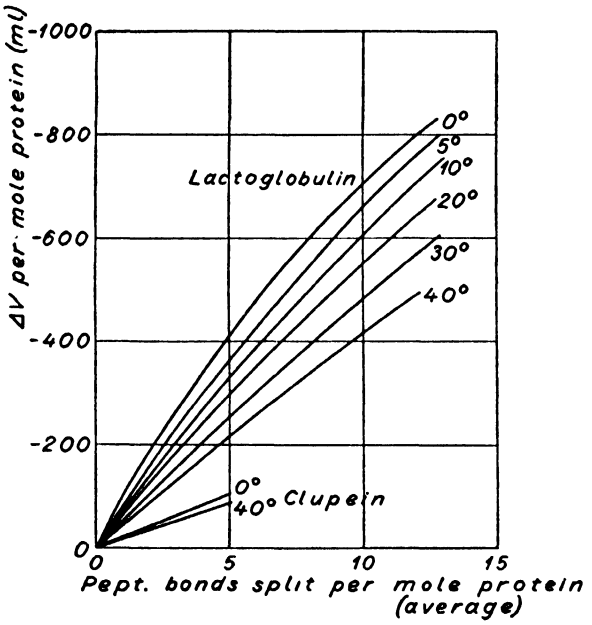


FIG. 1. Influence of temperature upon the volume change accompanying the tryptic break-down of β-lactoglobulin and clupein.

The dilatometric technique and the method for the determination of the number of peptide bonds split were described in the 1941 paper. It may here suffice to mention that the protein concentration varied between 0.5 and 2 percent, that the buffer was 0.1 molar and that the enzyme concentration was 0.05 percent (dry filter cake). The pH was 8.3 at 30° C.

The main results of our preliminary investigation were: 1) That the tryptic splitting of simple substrates like benzoylarginine amide (Jacobsen, 1943) and clupein gave values for the volume change that agreed with the theoretical figure (−15 to −20 ml/mole) so that the magnitude of other volume effects involved in the process of hydrolysis could not be appreciable (c. 2 ml/mole). 2) That the volume change accompanying the cleavage of peptide bonds in lactoglobulin was so high (about −40 ml per mole peptide bond) and changed so much with the progress of the cleavage that we had to assume the presence of other volume effects that were not directly connected with the electrostriction caused by the ions formed. Since the volume effect observed approached normal values with increasing length of the digestion time we felt justified in connecting its unexplained part with the initial stages of the break-down of the protein molecule.

The results of more recent experiments are shown in Figure 1, where the volume change per mole of protein is plotted against the number of peptide bonds opened. (Fig. 1 A shows the results of 4 experiments at 0° C. It should be pointed out that the solutions investigated were made up to have the same composition at all temperatures (2 percent protein, 0.05 percent trypsin and 0.1 molar ammonium chloride-ammonium buffer) and hence pH increased with decreasing temperature (about 1 pH-unit in the interval 30° — 0°). The pH value at 30° C was 8.3. This procedure has been deliberately chosen because the ionization equilibria in this way would be least disturbed by the change of temperature. However, if the heat of ionization of certain groups in the protein deviates much from that of others or of ammonia, certain changes in the state of ionization of the protein will occur also in this case (see later).

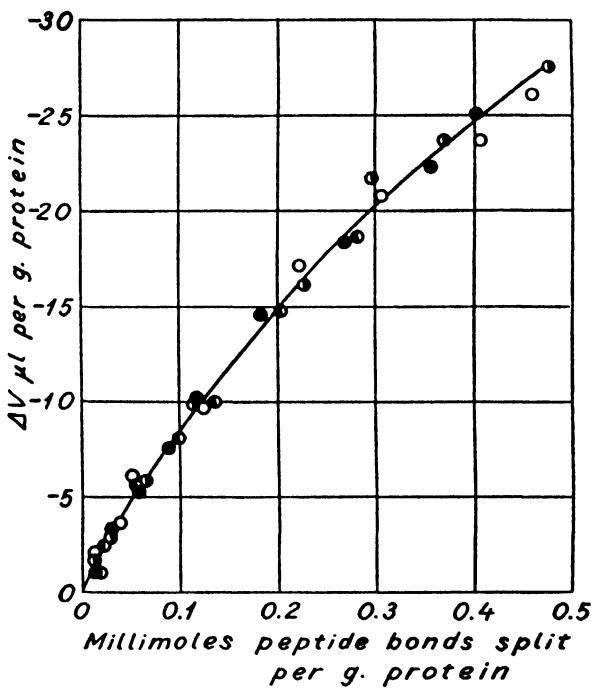


FIG 1 A. Reproducibility of the type of experiments leading to the curves in Fig. 1.

It will be observed from Figure 1 that the volume changes are higher than found before and mentioned above, that is, about −55 at 30° (slope of curve at the origin). However, the most striking feature of

this graph is the pronounced variation of the volume change with temperature found in the case of β-lactoglobulin as compared to the negligible one observed in the case of clupein. Since a one-by-one process would lead to a straight line in this graphical representation (compare Haugaard and Roberts, 1942) because the same process takes place throughout the whole course of the digestion, Figure 1 shows that the system approaches this type of breakdown as the temperature increases. Taking into consideration the rough calculation given above, the curves may be explained as follows:

The initial process, in which a cleavage of a small number of peptide bonds may be involved, is accompanied by a large volume change. The subsequent endopeptidatic process is accompanied by a normal electrostriction effect (c. −20 ml per mole peptide bond). The velocity of the former process

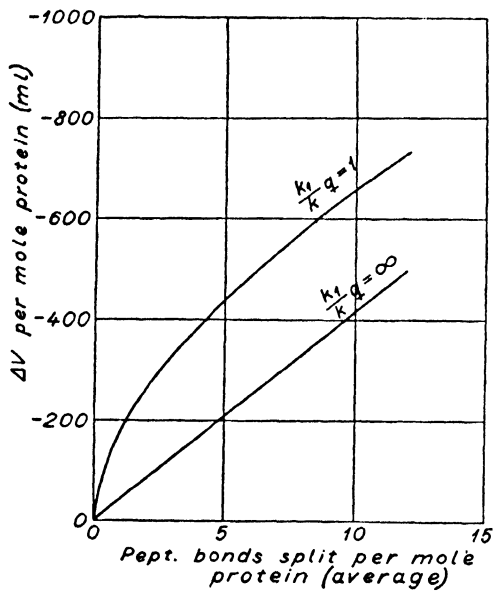


FIG. 2. Theoretical curves.

relative to that of the latter is much greater at 0° than at 40° where it is so small that the reaction approaches the one-by-one type, the curve at 40° being nearly a straight line. Assuming that this curve *is* a straight line and taking into account the value 35 for the number of peptide bonds split in the exhaustive break-down of lactoglobulin by trypsin (Miller, 1939) the total volume change due to electrostriction may be taken as −700 ml per mole protein. The slope of the curve at 40° C gives −1400 ml and hence the volume change in the initial process may be estimated at −700 ml per protein molecule, a value which is certainly a maximum one.

Figure 2 shows the results of a few calculations on the basis of equation (5). The numerical values introduced were:

$$n = 35 \qquad S_0 = 1$$
$$\Delta V = -700\,(S_0 - S) - 20\,B$$
$$\Delta \text{ pept} = B$$

It will be observed that there is a certain qualitative agreement between Figures 1 and 2 indicating that the above explanation may have some elements of truth in it. The question is only what kind of chemical reaction may be responsible for the initial contraction of about 700 ml/mole.

IV

In our preliminary paper (1941) Jacobsen and the author investigated the possibility that denaturation was the initial step of the proteolysis. However, we met with the difficulty that the heat denaturation to which we subjected the protein (9 minutes at 80° and pH 8) seemed to be semi-reversible in the sense that after cooling, a protein mixture was obtained which had lost the solubility at the isoelectric point, characteristic of the native protein, but in which a large portion of the molecules had regained (if they had ever lost it) the structure that was split with a high contraction. We therefore concluded that the experiments were "not in favor of the assumption that the ..... primary processes causing the high contraction are exclusively such as are involved in ordinary denaturation." In subsequent experiments, however, G. Johansen (unpublished data) found a volume change upon slow alkaline denaturation, at 30° C and pH 8.3 of −5 μl per g of lactoglobulin, *i.e.* −200 ml per mole. In view of the insufficiently defined criterion for denaturation, namely, the loss of solubility in a certain salt mixture at a certain pH (see below) the only conclusion which can be drawn from such experiments is that *one kind of denaturation* gives rise to a contraction of considerable magnitude. (Regarding the general problem, reference is made to the excellent review by Neurath, Greenstein, Putnam and Erickson, 1944). Further experiments however revealed that also in the denaturation with urea (Jacobsen and Linderstrøm-Lang, unpublished data) a substantial volume contraction is observed, about 250 ml/mole, and in addition Jacobsen and Korsgaard Christensen (1948) were able to show that the rate of this denaturation was much greater at 0° than at 30° (compare Hopkins, 1930) a phenomenon which fits very well into the picture given above of the enzymatic breakdown of lactoglobulin provided trypsin is able to exert a denaturing action on lactoglobulin similar to that of urea.

Figures 3 and 4 show the results of experiments by the above named authors on the denaturation of β-lactoglobulin by urea at pH 5.15.

The solutions were 2 percent and 38 percent with respect to β-lactoglobulin and urea respectively. The quantity of denatured protein was determined by adding aliquots to 10 volumes of a solution containing 0.8 mole of acetic acid, 0.4 mole of sodium acetate and 0.5 mole of magnesium sulphate per 1. Extensive comparative experiments by Mr. G. Johansen in this laboratory have shown that of all buffer-salt mixtures tried, this solution gives the highest yield of precipitated (denatured) protein, without precipitating undenatured protein.

Figure 3 needs no further explanation. Figure 4 shows the reversibility of the denaturation. The solid line curve represents an experiment where the denaturation at first was allowed to proceed at 0° C for 6 minutes. The solution was then heated to 37.4° C and was kept at this temperature for five hours, whereupon it was cooled to 0° C, kept at this temperature for one hour and again heated to 37.4° C.

A completely satisfactory explanation of these phenomena can hardly be given at the present moment, but the following qualitative assumptions may still be of value in the further development of the question:

β-lactoglobulin exists in two forms that are in mutual equilibrium. The rate of transformation from one form to the other is high. One of these forms has a tendency to predominate at low temperature and is characterized by its susceptibility to denaturation by urea. The denaturation proceeds in two steps, one rapid and reversible (I) the other slow, irreversible (II). The rates of both these steps increase with temperature but the rapid and reversible step runs to an equilibrium determined by the equilibrium ratio between the two forms of lactoglobulin under the prevailing conditions (urea concentration, temperature, pH). If we use this picture in the treatment of the tryptic hydrolysis of β-lactoglobulin we need only introduce into (5) the assumptions

$$Q'_s \rightleftarrows Q''_s \tag{7}$$

$$S' = k_s \cdot S'' \tag{8}$$

$$S = S' + S'' = S' (1 + 1/k_s) \tag{9}$$

$$S_0 = S_0' (1 + 1/k_s) \tag{10}$$

where Q_s' and Q_s'' are the two forms of β-lactoglobulin, S'' the concentration of the stable form, S' that of the unstable, and k_s a constant that increases with decreasing temperature. Under this condition (5) is transformed into

$$B = ns_0\left[1 - \frac{S}{S_0} - \frac{1}{\left(1 - \frac{k_1}{k}q\right)\left(1 + \frac{1}{k_s}\right)}\left(\left(\frac{S}{S_0}\right)^{k_1 q/k} - \frac{S}{S_0}\right)\right] \tag{11}$$

If x denotes the volume change involved in the process: $Q'_s \rightarrow Q_D$ we obtain from (11)

$$\left(\frac{\Delta V}{\Delta B}\right)_{t=0} = x\,\frac{1 + k_s}{n} - 20 \tag{12}$$

and hence the initial slope of the ΔV — Δ pept-curve (Fig. 1) will increase with decreasing temperature if we give k_s the above named properties.

Putting $n = 35$ and $x = -700$ gives $k_s = 5$ at 0° C and $k_s = 1.8$ at 30°; but these values would be entirely changed if a few peptide bonds were also split in the initial process.

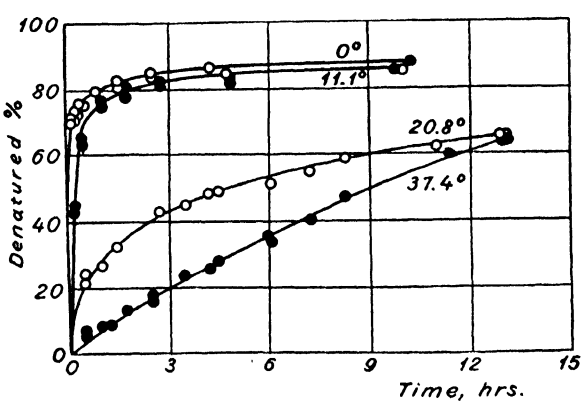


FIG. 3. Denaturation of β lactoglobulin by urea at different temperatures.

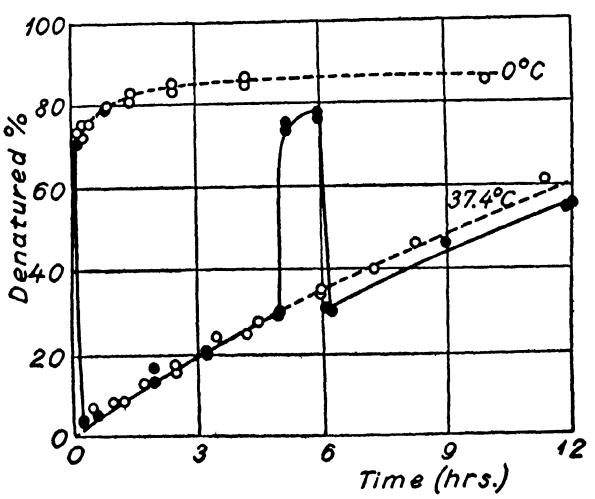


FIG. 4. Reversibility of rapid denaturation of β-lactoglobulin by urea.

Since (4) leads to an infinite initial slope in all cases except when $k_1q/k = \infty$ or when peptide bonds are opened in the initial process, (11) must be regarded as a definite improvement, but the assumptions made are so simple and gross that we shall refrain from giving elaborate calculations on this basis.

There is however one point that should be considered in this connection. As mentioned above the composition of the solutions investigated in the experiments shown in Figure 1 was the same at all temperatures. The possibility therefore exists that one or more groups with their buffer range around

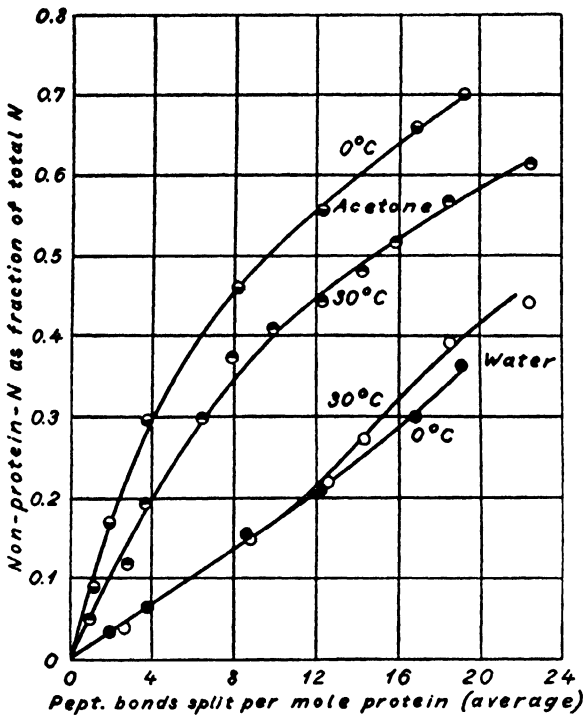


FIG. 5. Relation between cleavage of peptide bonds and formation of non-protein material in the tryptic breakdown of β-lactoglobulin.

the considered pH would have a much smaller (or higher) heat of ionization than that of the added buffer or of other groups determining the pH of the solution (NH_3^+ groups of the protein). A shift of temperature would therefore shift the equilibria of such groups (with low, or high, heat of ionization). It cannot be excluded that processes of this kind may be involved in the transformation

$$Q_s' \rightleftarrows Q_s''.$$

V

An interesting contribution has been given by Jacobsen (unpublished data) who investigated the tryptic break-down of β-lactoglobulin at 30° and 0° and compared the formation of non-protein nitrogen with the simultaneous cleavage of peptide bonds. The experiments were carried out at pH 8.3 (30°). Figure 5 shows some of the results. The lower curves represent experiments in which the protein was precipitated by 10 percent trichloroacetic acid (TClA) in water while in the case of the upper curves 10 percent TClA in 30 percent acetone was used (compare Jacobsen, 1947). It will be observed that the latter procedure is much more selective and that the ratio non-protein N/Δ pept is higher at 0° than at 30°, which is an indication of the fact that more intermediary compounds are formed at low temperature. The coincidence of the curves for aqueous TClA is only apparent, since measurements of the light absorption at 280 mμ (tyrosine, tryptophane) showed that the non-protein material had a quite different composition at the two temperatures (Fig. 6).

Experiments of this kind are valuable evidence that the separation of "attacked" and "non-attacked" molecules by TClA is highly arbitrary, and that great caution should be shown in interpreting

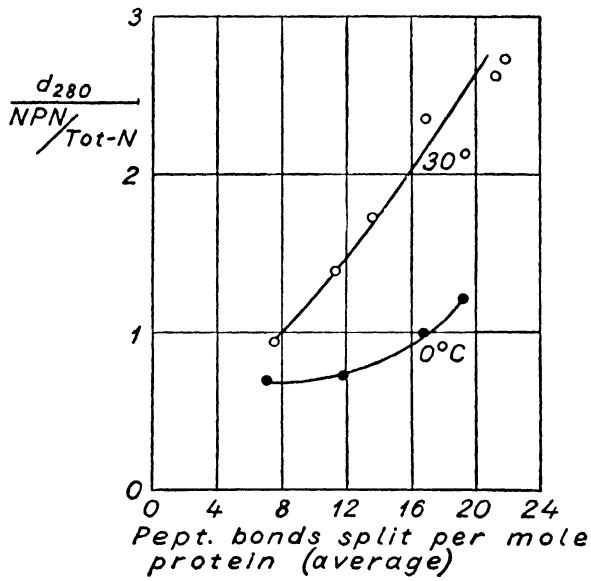


FIG. 6. Ultraviolet absorption of non-protein material formed in tryptic break-down of β-lactoglobulin.

the results of such precipitations (compare Jacobsen, 1947, and Linderstrøm-Lang and Ottesen, 1949).

VI

Other experiments showing the complexity of the break-down of lactoglobulin were carried out by Levi and the author who studied the temperature coefficient of the rate of hydrolysis as measured by the cleavage of peptide bonds. A comparison was made with clupein, L-benzoylarginine amide (BAA) and the corresponding racemic compound. The re-

sults are seen in Figure 7. k is the velocity constant for the monomolecular reaction in the case of BAA and clupein, while k for β-lactoglobulin is the initial slope of the pept/t curve. k is arbitrarily put equal to 1 at 0° C. ΔE is the "activation energy."

In the β-lactoglobulin experiments the protein concentration was 2.5 percent while that of the enzyme was 0.02 percent. The NH_4^+ — NH_3 buffer was 0.1 N; pH 8.3 (30°). The composition was constant at all temperatures.

Without going more fully into the explanation of the abnormal form of the β-lactoglobulin curve, it may just be mentioned that considerations of the type underlying (11) with a few but reasonable additional assumptions about the binding of the enzyme to substrate and reaction products are able to account for the positive curvature of the β-lactoglobulin curve.

VII

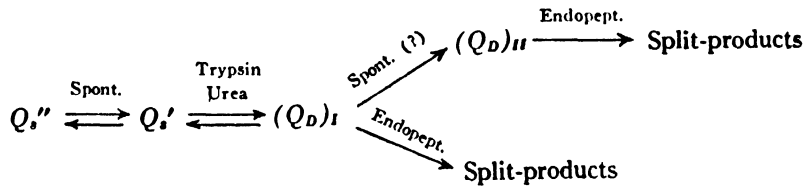
Experimental evidence that the initial reaction in the tryptic break-down is indeed a denaturation has been given by Korsgaard Christensen, (1949) who made the observation that denaturation by urea was accompanied by an increase in the optical rotation of β-lactoglobulin. The lower curve in Figure 8 demonstrates this fact for the following system: 2.17 percent β-lactoglobulin, 19 percent urea, 0.1N NH_4^+ — NH_3 buffer, pH 8.3, 30° C. Simultaneous determination of the proportion of protein insoluble in the above mentioned Johansen buffer showed the intimate connection between the rise in optical rotation and the formation of insoluble protein. The three other curves show the striking effect of the addition of trypsin. Denaturation is accelerated by the enzyme and only at higher concentration does another effect set in, namely, a decrease in laevorotation which in parallel experiments with initially denatured β-lactoglobulin was shown to be due to enzymatic break-down.

Almost identical results were found by Korsgaard Christensen in investigations of the denaturation of aqueous solutions at pH 9.3. Here too, there was a marked accelerating effect by trypsin (but not by heat denatured enzyme).

The simplest explanation of the above results is that trypsin first catalyses a process of the type

$$Q_n' \rightleftarrows (Q_D)_I$$

where $(Q_D)_I$ is a product of the type formed in the rapid and reversible denaturation with urea (see page 000). The fate of $(Q_D)_I$ is then determined by the following series of reactions.



Since Korsgaard Christensen has found that the fast and reversible denaturation at 0° is accompanied by a marked increase in optical rotation it must be assumed that the increase observed at 30° is essentially connected with the formation of $(Q_D)_I$ from Q_n' and that the transformation of $(Q_D)_I$ to $(Q_D)_{II}$, which is the stable form at 30°, is accompanied by no marked change in rotation.

It is apparent from the above experiments that by accelerating the initial process of the tryptic break-down it is possible to produce an accumula-

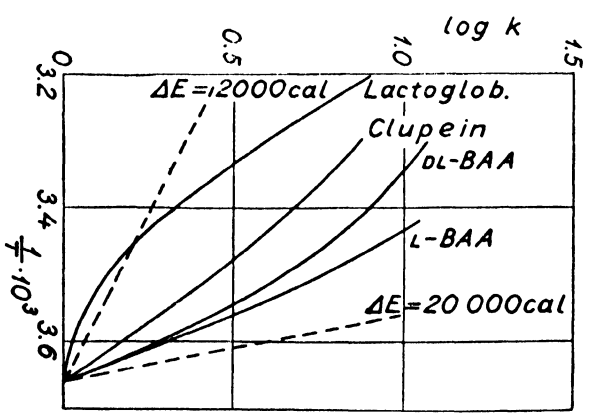


FIG. 7. Relation between velocity of tryptic catalysis and temperature in the case of benzylarginine amide, clupein and β-lactoglobulin.

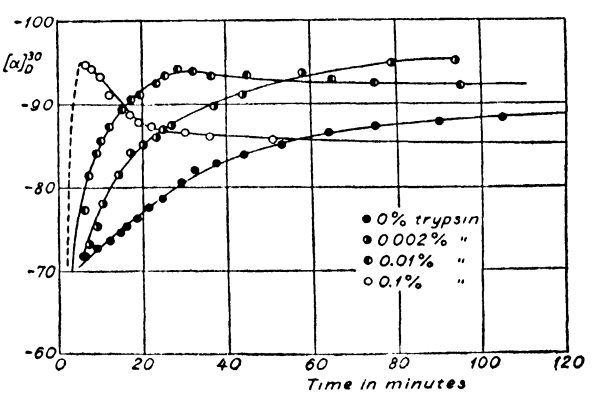


FIG. 8. Rate of denaturation by urea with and without addition of trypsin.

tion of denatured protein $(Q_D)_{II}$ at 30°. While in the urea experiments trypsin and urea may work hand in hand in both of the processes $Q_s' \rightarrow (Q_D)_I$ and $(Q_D)_I \rightarrow (Q_D)_{II}$ the influence of high alkalinity upon the rate of denaturation may be exerted on the process $Q_s'' \rightleftarrows Q_s'$. As mentioned above a shift in the ionization of a group in the lactoglobulin molecule may be involved here. If this group has a low heat of ionization an increase in pH by addition of alkali may give rise to a similar effect as the increase in pH caused by the lowering of temperature. Further speculations on this question must however be postponed until more detailed information is available.

The above scheme will complicate the simple equations (5) or (11) but will not cause any essential change in the qualitative picture except on one point. The side reaction $(Q_D)_I \rightarrow (Q_D)_{II} \rightarrow$ split products, which may well be imagined to have a high activation energy, is well suited to form the bridge between our present considerations and the experiments by Linderstrøm-Lang, Hotchkiss and Johansen (1938) which could not be fit into the simple theory leading to (11). The subject will be dealt with in a later publication.

The picture given above is practically identical with that given by Lundgren (1941) in his important paper on the splitting of thyroglobulin by papain. Indeed our opinion was originally formed on the basis of the findings of Lundgren. There is a pronounced similarity between his reaction chain $N \rightarrow \alpha \rightarrow D$ and the above $Q_s'' \rightarrow Q_s' \rightarrow Q_D$ only there is in the present case no evidence for or against the assumption that Q_s' is a "half unfolded" or "perturbed" form of the β-lactoglobulin molecule. However, our experiments are best explained assuming that the process $Q_s'' \rightleftarrows Q_s'$ is rapid and spontaneous.

VIII

Before concluding the present discussion it may be of value to form a physical picture of the processes involved in the break-down of the macro-molecule of β-lactoglobulin. The following suggestions have more the character of the "theses" in old doctors' dissertations than of well-founded scientific realities. The author believes however that they will form an excellent subject for discussion insofar as they contain anything new.

The lactoglobulin molecule is considered to be a small crystal, so small indeed that the major part of its ionizing groups may be regarded as being solvated and in equilibrium with the hydrogen ions of the surrounding medium. It has however a certain crystal lattice and the peptide chains or similar structures by which it is built up are knitted together in this lattice by forces that are well-known from macro crystals, Van der Waals forces and hydrogen bonds (see, for example, Mirsky and Pauling, 1936; Philippi, 1936; Huggins, 1942; Dervichian, 1943; Boyes-Watson, Davidson and Perutz, 1947; Bresler, 1944.

This micro-crystal is not "insoluble" but under normal conditions its "rate of solution" is small. Some catalysts added in small quantities (trypsin) may increase the rate of solution without influencing the solubility, other substances (urea) may do both. In this process of dissolution (formation of $(Q_D)_I$), a considerable volume contraction is involved. We shall postulate a ΔV of -500 ml per mole and we shall compare this phenomenon with the melting of ice and assume that the micro crystal lattice has some forced structure which makes it occupy more space than its liberated (or folded out) elements.

The "dissolved" micro crystals are unstable. They may either crystallize again or form "amorphous particles" that may gradually rearrange to crystal-like structures (completely or partly reversible denaturation). But they may also undergo chemical reactions and form new amorphous particles which cannot rearrange (irreversible denaturation, formation of $(Q_D)_{II}$). This back-reaction to stable particles may involve a volume increase so that the total volume change $Q_s' \rightarrow (Q_D)_{II}$ is smaller than -500 ml/mole, namely, the -200 to -250 found by Jacobsen and Linderstrøm-Lang and by Johansen. The stable particles might possess some elements of the original crystal lattice so that the volume change in the tryptic splitting is higher than with simple peptides.

According to this picture the resistance of many globular proteins to enzymatic digestion is a matter of steric hindrance (although other factors cannot be excluded). What happens when the trypsin molecule gets into contact with the substrate molecule cannot be discussed in detail at the present moment. A possible suggestion is that both molecules unfold or dissolve, from which process the trypsin molecule recovers more readily than that of the substrate and is able to split the peptide bonds laid open in the substrate.

IX

In connection with this discussion of the action of trypsin on β-lactoglobulin it may be mentioned that ovalbumin which is very slowly attacked by trypsin is not denatured by this enzyme (Korsgaard Christensen) even in the presence of urea. On the other hand chymotrypsin (Linderstrøm-Lang and Jacobsen, 1941; Jacobsen, unpublished data) seems from polarimetric data to act without denaturing β-lactoglobulin, from which may be concluded that specific peptide bonds are directly available on the surface of the micro crystal. Insulin, which is difficult to denature, is split by chymotrysin but not by trypsin (Butler, Dodds, Phillips and Stephen, 1948). Further investigation will be needed before a clear picture of the action of different enzymes on different substrates can be obtained. It should be emphasized, however, that the peculiar volume effect observed in case of the system trypsin-β-lactoglobulin has also been seen in other systems, namely, chymotrypsin-β-lactoglobulin, chymotrypsin-insulin, pepsin-ovalbulin, pepsin-β-lactoglobulin. In all these

enzymatic reactions the volume changes per peptide bond split are much higher than accounted for by the theory for the electrostriction effect of the appearing electric charges. Hence there seems to be a structural principle that is common in the molecules of these substrate proteins and that may be described in the above given way. However, the road to peptides and amino acids may be quite different depending upon the nature of substrate, enzyme, and external conditions (temperature, pH, etc.).

The author is indebted to Dr. C. F. Jacobsen and Dr. Korsgaard Christensen for valuable help and discussions.

References

Beloff, A., and Anfinsen, C. B., 1948, The products of proteolysis of some purified proteins. J. biol. Chem. *176:* 863-872.

Boyes-Watson, J., Davidson, E., and Perutz, M. F., 1947, An X-ray study of horse methaemoglobin. I. Proc. Roy. Soc., A *191:* 83-132.

Bresler, S. E., and Talmud, D. L., 1944, On the nature of globular proteins. I. Compt. Rend. (Dokl.) Acad. Sci. URSS *43:* 310-314. II. A few consequences of the new hypothesis. *Ibid. 43:* 349-350.

Butler, J. A. V., Dodds, E. C., Phillips, D. M. P., and Stephen, J. M. L., 1948, The action of chymotrypsin and trypsin on insulin. Biochem. J. *42:* 116-122.

Cohn, E. J., and Edsall, J. T., 1943, Proteins, Amino Acids and Peptides. New York: Reinhold Publishing Corporation.

Cohn, E. J., McMeekin, T. L., Edsall, J. T., and Blanchard, M. H., 1934, Studies in the physical chemistry of amino acids, peptides and related substances. I. The apparent molal volume and the electrostriction of the solvent. J. Amer. Chem. Soc. *56:* 784-794.

Dervichian, D. G., 1943, The structure of the protein molecule. J. Chem. Phys. *11:* 236-246.

Eeg-Larsen, N., Linderstrøm-Lang, K., and Ottesen, M., 1948, Transformation of ovalbumin into plakalbumin. Arch. Biochem. *19:* 340-344.

Haugaard, G., and Roberts, R. M., 1942, Heats of organic reactions. XIV. The digestion of beta-lactoglobulin by pepsin. J. Amer. Chem. Soc. *64:* 2664-2671.

Hopkins, F. G., 1930, Denaturation of proteins by urea and related substances. Nature, Lond. *126:* 328-330, 383-384.

Huggins, M. L., 1942, X-ray studies of the structure of compounds of biochemical interest. Ann. Rev. Biochem. *11:* 27-50.

Jacobsen, C. F., The contraction accompanying the splitting of peptide bonds in synthetic substrates for trypsin and chymotrypsin. C. R. Lab. Carlsberg. Sér. Chim. *24:* 281-287.

1947, The activation of chymotrypsinogen. C. R. Lab. Carlsberg. Sér. Chim. *25:* 325-437.

Jacobsen, C. F., and Korsgaard Christensen, L., 1948, Influence of temperature on the urea denaturation of β-lactoglobulin. Nature, Lond. *161:* 30.

Korsgaard Christensen, L., 1949, Trypsin splitting and denaturation of β-lactoglobulin. Nature, Lond. *163:* 1003.

Linderstrøm-Lang, K., 1933, Neuere Untersuchungen über die Anwendung von Enzymen zur Konstitutionsermittlung der Proteine. Ergebn. Physiol. exp. Pharmakol. *35:* 415-469.

1939a, Globular proteins and proteolytic enzymes. Proc. Roy. Soc. B *127:* 17-18.

1939b, Proteolytic enzymes. Ann. Rev. Biochem. *8:* 37-58.

Linderstrøm-Lang, K., Hotchkiss, R. D., and Johansen, G., 1938, Peptide bonds in globular proteins. Nature, Lond. *142:* 996.

Linderstrøm-Lang, K., and Jacobsen, C. F., 1941, The contraction accompanying the enzymatic break-down of proteins. C. R. Lab. Carlsberg. Sér. Chim. *24:* 1-48.

Linderstrøm-Lang, K., and Ottesen, M., 1949, Formation of plakalbumin from ovalbumin. C. R. Lab. Carlsberg. Sér. Chim. *26:* 403-441.

Lundgren, H. P., 1941, The catalytic effect of active crystalline papain on the denaturation of thyroglobulin. J. biol. Chem. *138:* 293-303.

McMeekin, T. L., Cohn, E. J., and Weare, J. H., 1935, Studies in the physical chemistry of amino acids, peptides and related substances. III. The solubility of derivatives of the amino acids in alcohol-water mixtures. J. Amer. Chem. Soc. *57:* 626-633.

Miller, L., 1939, Digestion of crystalline lactoglobulin by crystalline trypsin and chymotrypsin. C. R. Lab. Carlsberg. Sér. Chim. *23:* 45-53.

Mirsky, A. E., and Pauling, L., 1936, On the structure of native, denatured and coagulated proteins. Proc. Nat. Acad. Sci. Wash. *22:* 439-447.

Moring-Claesson, I., 1948, Frontal analysis studies on the absorption of crystalline egg albumin and its cleavage products. Biochem. Biophys. Acta *2:* 389-408.

Neurath, H., Greenstein, J. P., Putnam, F. W., and Erickson, J. O., 1944, The chemistry of protein denenaturation. Chem. Rev. *34:* 158-265.

Philippi, G. T., 1936, On the Nature of Proteins. Amsterdam. N.V. Noord-Hollandsche Uitgeversmaatschappij.

Tiselius, A., and Eriksson-Quensel, I. B., 1939, Some observations on peptic digestion of egg albumin. Biochem. J. *33:* 1752-1756.

Weber, H. H., 1930, Die Bjerrumsche Zwitterionentheorie und die Hydratation der Eiweisskörper. Biochem. Z. *218:* 1-35.

Weber, H. H., and Nachmansohn, D., 1929, Die Unabhängigkeit der Eiweisshydratation von der Eiweissionisation. Biochem. Z. *204:* 215-252.

Discussion

Butler: Dr. Linderstrøm-Lang has mentioned the action of trypsin on insulin. I think our experiments support the idea of a preliminary denaturation. We have never been able to be sure if the very slight splitting of peptide bonds is real, or due to impurities, but we have always found that insulin recovered after treatment with trypsin has only about two-thirds of the activity of the original insulin. Also, Dr. Pedersen did an ultracentrifuge run on insulin in the presence of trypsin and found that altho the mean sedimentation constant was about the same, the spread of the sedimentation curve was much greater. This might indicate a certain amount of denaturation.

Linderstrøm-Lang: Observations similar to those of Dr. Butler have been made in our laboratory but we have been unable to form a definite opinion about what happens to insulin when attacked by trypsin. It is not excluded, however, that

denaturation of some kind takes place.

PETERMANN: This very interesting hypothesis can be illustrated by practical experience in the enzymatic hydrolysis of the serum globulins. These proteins are an exaggerated example of the accumulation of intermediate split products—halves and quarters of the original globulin molecule. In the digestion of human gamma globulin, the yield of halves is a function of pH being maximal at pH 3.5. The yield of halves is also increased at lower temperatures. As for denaturation before splitting, however, the situation must be much more complicated. Altho serum globulins can exist in an opened-out or α-form, the split products are always less asymmetrical than the original globulin, and since the antibody activity of the original globulin is retained, denaturation cannot have disturbed the surfaces responsible for the antibody activity.

LINDERSTROM-LANG: I thank Dr. Petermann for calling my attention to this very interesting piece of work.

THE LIVER PROTEINS

J. MURRAY LUCK
Department of Chemistry, Stanford University, Stanford, California

Investigations on the proteins of liver fall into three principal categories: those concerned with the isolation and characterization of liver enzymes; those devoted to the fractionation of liver homogenates or the formed structural elements of the organ; and those concerned with problems of function in which attempts are made to determine quantitatively the changes in various protein fractions under different stimuli or stresses imposed upon the organ. This paper will be restricted in many respects but principally so by exclusion of almost all work on the liver enzymes except insofar as the studies are very relevant to the remaining subjects under discussion.

Introduction

To the protein chemist, the liver and indeed other solid organs and tissues, present formidable barriers. The separation of the formed elements from liver requires a number of manipulations which are unnecessary with blood. In fact the liver is a highly organized structure, consisting of cells which normally display a characteristic orientation toward the vascular trees, of sinusoids between the cords of parenchymal cells, of blood vessels, bile ducts, connective tissue, and of still other cells which are oriented toward the bile ducts. Each cell consists of a nucleus, cytoplasm, and the enveloping membranes. The nucleus contains the chromosomes, the nucleolus, and other particulate matter. The cytoplasm is heavily granulated and certain it is that the large granules and the small granules differ both in structure and function. While the enzymes responsible for much of the oxidation of fatty acids appear to be localized in the large granules (Kennedy and Lehninger, 1949), it is known that other enzymes are also present in this fraction. Some of the glycolytic enzymes are in either the "small granules" or the particle-free cytoplasm; and other enzymes, if present as particulate matter, will probably be found to run the entire gamut of particle size. The first problem, therefore, is philosophical. How much significance may be attached to analyses of liver homogenates, of extracts of the whole organ, or of pressed juice of liver? Should one focus his research only upon the isolated cellular elements? If so, in the case of cytoplasmic granules, how far down the size scale should he go? When does a granule cease to be a granule and where may the "cuts" in sedimentation separations logically be made? The answer to the first two questions, it seems to me, is simple: For some problems homogenates should be used, for others the isolated cellular elements.

From the practical point of view anyone who devotes his life to liver instead of to blood or urine will encounter many potential frustrations. The separation of the cellular elements is not yet satisfactorily worked out, even the preparation of clean undamaged nuclei, free of cytoplasm. The liver is rich in proteolytic activity—a constant source of annoyance in protein isolation. Depolymerases have to be circumvented, and, apparently, the oxygen tension should also be controlled for satisfactory fractionations (Sorof and Cohen, 1949).

Protein Storage

Fractionation of the liver proteins appears to have been initiated by Plósz (1873), a pupil of Kühne. Our own interest in the subject goes back to 1935 (Luck, 1935, 1936) when we became concerned with the problem of storage protein. Since this is still a live topic it may merit a brief discussion. The question to be settled is whether the increases in liver size and in total protein concentration that follow upon increased levels of protein intake reflect the deposition of a fraction that differs chemically from the proteins of the organ in the fasting state —the so-called "organized," "tissue," or "stable" proteins of the early investigators (*e.g.* Seitz, 1906; Tichmeneff, 1914). Do these terms have a metabolic significance only? Are there "reserve" proteins as such or is the deposition of liver protein under favorable metabolic conditions the result of a true hyperplasia or hypertrophy in which all of the more abundant proteins of the organ participate alike?

For the pursuit of this problem it would seem unnecessary to use isolated cellular elements in place of tissue homogenates, unless by chance a chemically distinct "storage" protein were to be found; in such a case its localization within the liver would be highly desirable. The system of separation we used gave us two principal fractions: those proteins soluble in 5 percent (0.86 M) sodium chloride at pH 5.0 and those extractable from the residue with 0.25 percent sodium hydroxide and quantitatively precipitable from such alkaline extracts by adjustment to pH 5.0. The former was readily subfractionated into albumin, pseudoglobulin and euglobulin by the use of ammonium sulfate and dialysis. These terms have little meaning without further definition of the conditions under which the fractions so named were isolated but it would lead us somewhat astray to go into this phase of the problem. A word must be said, however, about the alkali-

soluble fraction. We referred to it originally as globulin II but it is perfectly clear that it is a nucleoprotein fraction and may be considered as somewhat of a degradation product (only 0.6 percent phosphorus) or that it is grossly contaminated with a phosphorus-free protein. Almost 60 years ago it was discovered (Halliburton, 1892) that from homogenates of blood-free liver one may extract a fraction, then called a nucleoalbumin, which contained 1.45 percent P. This fraction, much like our globulin II in being quantitatively precipitable on mild acidification with acetic acid (or adjustment to pH 5) is apparently a mixture of nucleoprotein and phosphorus-free protein.

To return to the problem of storage proteins, we made use of a system of fractionation which accounted for 92 to 93 percent of the total protein of the liver. The proteins so extracted and accounted for were distributed over four fractions, each of which must be regarded as quite heterogeneous, but quantitatively reproducible from one run to another.

Table 1. Protein Content of Rat Liver*

Diet	Globulin II	Euglobulin	Pseudoglobulin	Albumin
Low protein†	5.07±0.70	4.58±0.93	1.06±0.07	0.86±0.04
High protein‡	7.39±0.25	6.84±0.44	1.65±0.13	1.39±0.15
Ratio (high/low)	1.46	1.50	1.56	1.61

Livers perfused in situ until blood free.
Fractionations carried out at 0°–3°.
* In gm. per 100 gm. perfused liver, fresh weight.
† Average values for 5 groups of rats.
‡ Average values for 4 groups of rats.

The results of this approach to the problem are set forth in Table 1. They show quite clearly that coincidentally with the increases in liver size and total protein that characterize the transition to a high protein diet, all four fractions of the liver proteins increase by the same proportion as conditions favorable for storage arise; no one fraction may be singled out as reserve protein. The liver, in respect to this function, behaves in a manner that is qualitatively unique. Not only does it increase in size but the concentration of protein per unit fresh weight also increases; the latter change is not evidenced by muscle, kidney, or intestine (Luck, 1936). Correspondingly, brief fasting causes the liver proteins to decrease markedly, with but small decreases in the protein content of other organs (Addis *et al.*, 1936). Thus the livers of albino rats lost 20 percent of their total protein in a two day fast but all other organs combined lost only 4 percent (Addis *et al.*, 1936b). Other observations (L. L. Miller, 1948) permit us to conclude that liver catalase, alkaline phosphatase, xanthine dehydrogenase, cathepsin, and arginine decrease on dietary protein depletion and increase on dietary protein repletion in parallel with total liver cell protein.

Within the last few years further work on the problem indicates that the liver nuclei may play a more passive role than the cytoplasm. Thus, in fasting, the cytoplasmic ribonucleic acid decreases appreciably, not in concentration but in total amount per liver, while the desoxyribonucleic acid remains unchanged (Davidson and Waymouth, 1944). Thus the ribonucleic acid P per liver, in the case of male rats, decreased from a normal value of about 35.7 mg. to 26.1 mg. after a 24 hour fast. The desoxyribonucleic acid P decreased, meanwhile, from a normal value of 8.37 mg. to 7.98 mg. So also the high turnover number of liver ribonucleic acid phosphorus as compared with the very low turnover rate for desoxyribonucleic acid phosphorus (Brues, Tracy, and Cohn, 1944; Hammarsten and Hevesy, 1946; Davidson, 1947) suggest that the proteins of the nucleus are metabolically more sluggish than the ribonucleoproteins of the cytoplasm. This would be in harmony with the tentative conclusion that the changes in liver proteins observed during fasting, and perhaps also at various levels of protein intake are primarily localized in the cytoplasm. This conclusion also follows from the work of Kosterlitz and associates (Kosterlitz, 1944, 1946; Kosterlitz and Cramb, 1943; Kosterlitz and Campbell, 1945, 1946; Campbell and Kosterlitz, 1946, 1947) on the effects of protein intake on constituents of the liver cell—especially phospholipid, nucleic acid, total protein, and number of liver nuclei. Although Kosterlitz concludes that gains or losses in liver protein in the dietary states under consideration are in whole cytoplasmic protein rather than in any special storage protein, he nonetheless distinguishes between "labile liver cytoplasm" and "remaining liver cytoplasm." It appears doubtful whether anything is gained by perpetuating such distinctions. The now classic studies of Schoenheimer and his associates (Schoenheimer, 1942; Shemin and Rittenberg, 1944) emphasize the fluidity of the body protein and almost necessitate the concept of a metabolic pool of proteins, the components of which are in a state of "dynamic equilibrium." This is emphasized by the observation (Shemin and Rittenberg, 1944) that one-half of the liver protein nitrogen is replaced in seven days by

nitrogen from food and tissue proteins. Insofar as the changes in tissue protein content that follow upon changing levels of protein intake, fasting, and tissue regeneration are concerned, it seems evident that all organs and possibly all tissue proteins participate. They do so, however, at quite different rates. The liver proteins, at least those of the cytoplasm, appear to have exceptionally high turnover rates: changes in the concentration of any one component lead rather rapidly to homeostatic readjustments in the concentrations of other proteins in the liver metabolic pool and are reflected in the apparent constancy in proportional amounts of the participating proteins. A very similar concept has been developed by Borsook (Borsook and Keighley, 1935; Borsook and Dubnoff, 1943).

In our own work, as has already been mentioned, we employed a system of fractionation that started from liver homogenates. It is now evident that this procedure would not permit one to single out and identify any proteins, present in comparatively small amounts, which possessed very low turnover rates. Because of their low concentration they would be "swallowed up" in the relative excess of other more active proteins in the fraction. For example, the nuclear proteins of liver constitute only 15 percent of the total liver protein (Price, Miller, and Miller, 1948) and their "passivity," if confirmed, might pass undetected by the methods we have used. However, the rate of participation of the nuclear proteins of the liver in fasting or in protein depletion or repletion is not necessarily the same as that of desoxyribonucleic acid; it is indeed conceivable that the nucleic acid may have a very low turnover number while the proteins associated with it, or present unassociated in the nucleoplasm, may be fully as active, metabolically, as those of the cytoplasm. Indeed, to pursue the argument further it does not follow that the turnover numbers for the phosphorus of the nucleic acids or nucleotides necessarily apply to the constituent nucleoside portions. From the high turnover number for the phosphorus of cytoplasmic ribonucleic acid we may not conclude that the constituent ribonucleosides turn over with equal rapidity.

Liver Protein Changes in Carcinogenesis

Another problem of current interest in which the liver proteins are involved concerns the changes induced in these proteins by carcinogens. Because of the great metabolic activity of the liver, it is conceivable that carcinogenic processes, seemingly remote from this organ, might induce profound hepatic effects. The work of Huggins is suggestive (Huggins *et al.*, 1949). In sera from many subjects with cancer and from many others with non-malignant pathologies a high correlation was found between the thermal stability of the whole serum protein and the incidence of cancer. If we make the plausible assumption that the serum proteins are of hepatic origin the hypothesis that the liver is involved in the carcinogenic process becomes more deserving of exploration. However I can recall no substantial piece of work in which a study has been made of possible hepatic changes in experimental animals with epithelial cancer, breast tumors, sarcomas, etc.

In our own work we have not yet skirted the periphery of the organism by the use of, say, carcinogenic hydrocarbons applied directly to the skin, transplantable epithelial tumors, or tumors of the mammary gland. The experiments are planned but not initiated. We have, however, been interested in the use of certain carcinogens which are known to induce hepatomas: *m'*-methyl-*p*-dimethylaminoazo-

Table 2. Effect of Carcinogenic Azo Dye on Liver Proteins

Purified basal diet +0.06 percent *m'*-methyl-*p*-dimethylaminoazobenzene

Component	Control basal diet	Time rats were fed diet containing azo dye				Hepatomas
		2 wks.	4 wks.	6 wks.	8 wks.	
Av. liver weight, gm.	10.7	7.8	9.7	9.6	10.9	
Nitrogen, mg. %	2,520	2,470	2,470	2,500	2,465	2,040
Phosphorus, mg. %	280	316	316	308	326	272
Desoxyribonucleic acid, mg. %	266	396	440	490	506	756
Ribonucleic acid, as ribose, mg. %	150	122	146	140	126	139
Riboflavin, γ %	2,466	1,967	1,740	1,395	1,278	
Albumin, %	1.3	1.1	1.2	1.1	1.1	
Globulin, %	1.3	2.1	2.0	2.2	2.2	
Non-protein nitrogen, mg. %	180	181	213	205	230	
Non-protein phosphorus, mg. %	81	98	104	108	108	

The values are expressed on a fresh liver basis and are averages of 3 or more groups per period, 4 to 6 rats per group. Cirrhosis was evident after 4 weeks and at 8 weeks most of the livers were extremely cirrhotic. Liver moisture, 75 to 78%. Food intake while on diet containing azo dye, 13.5 to 15 gm. per rat per day. Average liver weight, 9.7 gm. per rat (range 7.2 to 13.5). See J. Biol. Chem. *176*: 1228 (1948).

benzene and 2-acetylaminofluorene. The former is quite specific and is more active than its homologue, butter yellow (Miller and Baumann, 1945; Giese, Miller and Baumann, 1945). Thanks to the substantial investigations of the Wisconsin group the dietary factors that must be controlled are now so well known that hepatomas may be routinely and readily

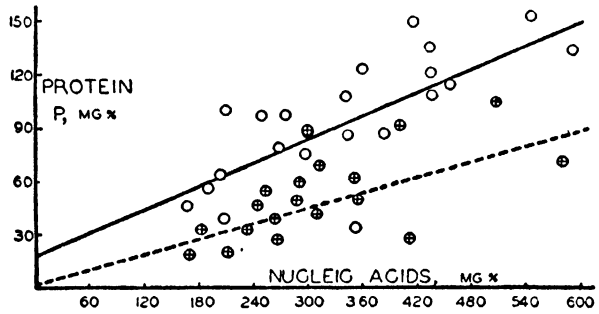


FIG. 1. The protein phosphorus content plotted against the total nucleic acid content of various NaCl extracts of liver, indicating that increases in the amounts of nucleic acids are accompanied by an increased phosphorus content. ○, corresponds to the 0.4 M NaCl extract (mostly ribonucleic acid); ⊗, corresponds to 1.0 M NaCl extracts (mostly desoxyribonucleic acid). The lines are calculated by the method of least squares.

produced in the white rat by both the azo dyes and acetylaminofluorene.

Our first studies (Griffin, Nye, Noda and Luck, 1948) were directed toward the liver protein changes that precede the appearance of discrete tumors. Use was made of *m'*-methyl-*p*-dimethylaminoazobenzene administered as 0.06 percent of the diet. The results are presented in Table 2. Attention should be drawn to the rather marked and early increase in desoxyribonucleic acid. This may merely reflect the increase in liver nuclei, known to be associated with azo dye carcinogenesis or it may mean this and more as shall presently be pointed out.

The observed increase in desoxyribonucleic acid is in harmony with the observations of others who employed butter yellow in similar experiments (Masayama and Yokoyama, 1940; Price, Miller and Miller, 1948; Dickens and Weil-Malherbe, 1943). Our evidence for the increase is based primarily upon use of the diphenylamine reaction of Dische as incorporated in the hot trichloroacetic acid method of Schneider (1945); confirmatory evidence may be gleaned from gross visual observation of the amount of the fraction precipitated, as well as from phosphorus determinations on aliquots of the desoxyribonucleoprotein fraction (Fig. 1). In the case of Ikubo's transplantable rat hepatoma no increase has been observed in protein phosphorus per unit wet weight of hepatoma over that of normal liver (Fujimara, Nakahara and Kishi, 1937). The total ribonucleic acid showed no change during carcinogenesis nor did the liver albumin. Liver globulin increased significantly. Riboflavin decreased sharply, in confirmation of the observations of others (Kensler, Sugiura and Rhoads, 1940; Griffin and Baumann, 1946; Miller, J. A., 1947). This probably indicates a corresponding decrease in flavoprotein as is suggested by the observations of Westphal (1944) and Lan (1944) on amino acid oxidase.

Control experiments in which the relatively noncarcinogenic compound, azobenzene, was fed revealed no significant changes in any of the constituents mentioned.

We next turned (Griffin, Cook and Cunningham, 1949) to 2-acetylaminofluorene incorporated at a

TABLE 3. EFFECT OF 2-ACETYLAMINOFLUORENE ON CERTAIN LIVER COMPONENTS
Purified basal diet+0.04% 2-acetylaminofluorene*

Component	Control basal diet	Time rats were fed diet containing carcinogen						24-week carcinogenic diet+4-week basal diet	
		2 wks.	4 wks.	6 wks.	14 wk.	20 wks.	24 wks.	Liver	Liver tumor
Av. liver weight, gm.	9.0	9.1	10.0		11.0	19.0	20	18	
Liver appearance	Normal	Normal	Normal	Mild cirrhosis	Mild cirrhosis	Moderate cirrhosis	†	Cirrhosis liver tumors	
Nitrogen, mg. %	2,523	2,590	2,560	2,380	2,370	2,410	2,210	2,150	1,965
Phosphorus, mg. %	291	339	297	301	270	292	290	287	228
Desoxyribonucleic acid, mg. %	273	261	320	326	244	205	218	256	316
Ribonucleic acid, as ribose, mg. %	140	141	102	120	101	124	96	161	119
Riboflavin, gamma %	2,540		2,100		1,670		1,630	1,430	320
Tissue moisture, %	75	73.5	73	74	75	73.4	70		

* The values are expressed on a fresh weight basis. Adult albino male rats. During experiment animals maintained weight or gained slightly.

† Livers enlarged, cirrhotic, small tumors evident.

level of 0.04 percent in the basal diet for a period of 28 weeks. The results are presented in Table 3. In contrast to the azo dye results, no increase in the concentration of desoxyribonucleic acid was noted. By the twentieth week, however, due to a sudden increase in liver size, the total desoxyribonucleic acid in the organ had increased considerably. We find a corresponding terminal increase in liver size (Cunningham, Griffin and Luck, in press) with the *m'*-methyl azo dye which manifests itself about the twelfth week (paired feeding experiments). During the first 8 weeks, however, as reported in Table 2, no increase in liver size is apparent though the concentration of desoxyribonucleic acid has markedly increased. Recently, nuclear counts have been made in normal livers and in the livers of animals given these two carcinogens, and the conclusion has been reached that the increase in desoxyribonucleic acid in both cases is due entirely to increase in numbers of nuclei per unit weight of tissue (Cunningham, Griffin and Luck, unpub.). Desoxyribonucleic acid determinations therefore appear to give a measure of the degree of precancerous hyperplasia in the liver.

The concentration of ribonucleic acid after acetylaminofluorene decreased slightly and that of riboflavin somewhat more appreciably, though less than in the azo dye experiments.

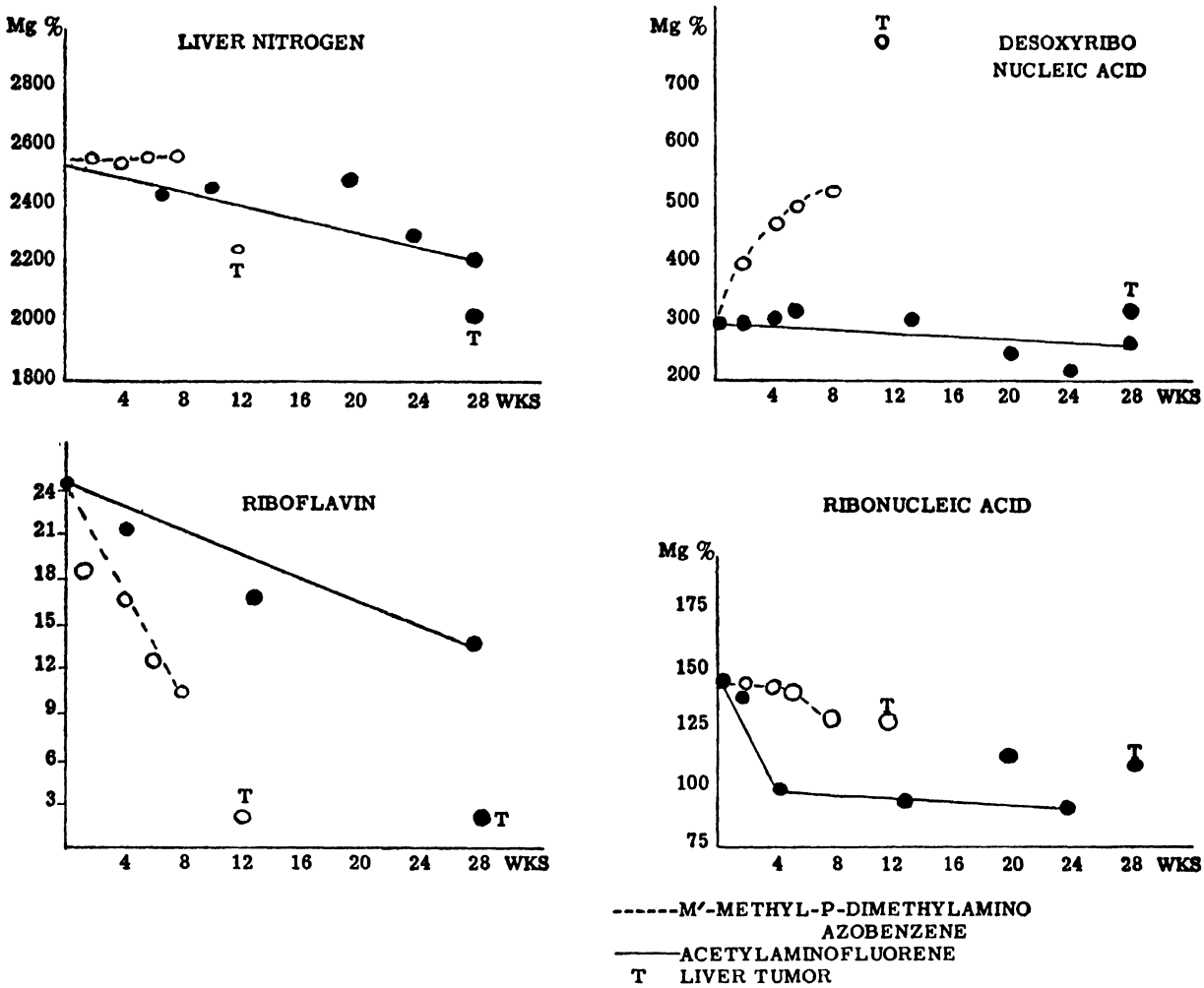


FIG. 2. Effect of two different carcinogens on certain liver components.

In Figure 2 the results of feeding these two carcinogens are presented graphically to facilitate comparison. The different responses elicited by the two carcinogens can be explained in part by the differences in the tumors that are formed. Tumors induced by the azo dye contain disoriented masses of closely packed cells with large, deeply-staining, nuclei. Those produced by acetylaminofluorene consist largely of small duct-like structures lined by flattened cells with small nuclei and containing rela-

tively much cytoplasm. We feel that those differences in tumor structure help to explain the different effects on desoxyribonucleic acid content.

Serum Protein Changes in Carcinogenesis

Since the liver is regarded by many as the principal site of formation of the serum proteins, it may not be inappropriate to discuss the serum protein changes during carcinogenesis. As will be shown presently they reflect some very fundamental changes in the liver. Table 4 gives the story quite completely for the m'-methyl dye (Cook, Griffin and Luck, 1949) with a typical electrophoretic pattern for the 6-week serum given in Figure 3. The most conspicuous change was a threefold increase in serum γ-globulin, evident after only two weeks' feeding of the dye, and a small but significant decrease (15 to 20 percent) in serum albumin. Concentration changes in α- and β-globulin were not observed. In control experiments with dimethylaminoazobenzene (butter yellow) similar changes but of lower magnitude were observed within a few weeks; likewise, and rather surprisingly, with the non-carcinogenic azobenzene.

The increase in γ-globulin is obviously not specifically associated with carcinogenesis, nor is the reduction in serum albumin. Many pathological conditions result in a reduction of serum albumin and an increase in globulin, frequently γ-globulin (Luetscher, 1947; Longsworth *et al.*, 1939; Moore *et al.*, 1943; Gutman *et al.*, 1941; and Dole, *et al.*, 1945). Various investigators have also shown, more specifically, that liver disease is frequently associated with decreased serum albumin and increased globulin (Luetscher, 1940, 1941; Gray and Barron, 1943; Moore *et al.*, 1945).

Table 4. Effect of *m'*-Methyl-*p*-Dimethylaminoazobenzene on Percentage Composition of Rat Sera by Electrophoretic Analysis*

No. of samples	Diet	Albumin		α-Globulin		β-Globulin		γ-Globulin	
		Average	Range	Average	Range	Average	Range	Average	Range
		percent†	percent	percent	percent	percent	percent	dercent	percent
6	Normal	66	60–74	14	11–26	13	10–14	6	2–10
1	1 week, m'Me-DAB‡	71		11		13		5	13–27
5	2 weeks, m'Me-DAB‡	53	44–61	15	12–16	14	12–17	19	16–19
3	4 weeks, m'Me-DAB‡	55	54–56	14	13–16	13	11–15	18	15–16
3	6 weeks, m'Me-DAB‡	58	55–60	14	13–16	13	12–14	16	16–18
3	8 weeks, m'Me-DAB‡	55	54–58	16	15–17	12	12–14	17	
1	8 weeks, m'ME-DAB‡+2 weeks, normal (hepatomas)	56		18		16		10	
1	10 weeks, m'Me-DAB+4 weeks, normal (hepatomas)	53		25		11		11	
1	2 weeks, m'Me-DAB+1 week (normal)	66		17		14		3	

* Buffer: sodium diethyl barbiturate; ionic strength: 0.1; pH 8.3 (8.21–8.40); potential gradient: 4 to 6 volts per cm.

† Percentages are the fractional areas of the electrophoretic diagrams due to each component, and represent the percentages of the total amount of protein in the serum present as these components.

‡ *m'*-Methyl-*p*-dimethylaminoazobenzene, 0.06 percent in the diet. See J. Biol. Chem. *177*: 376 (1949).

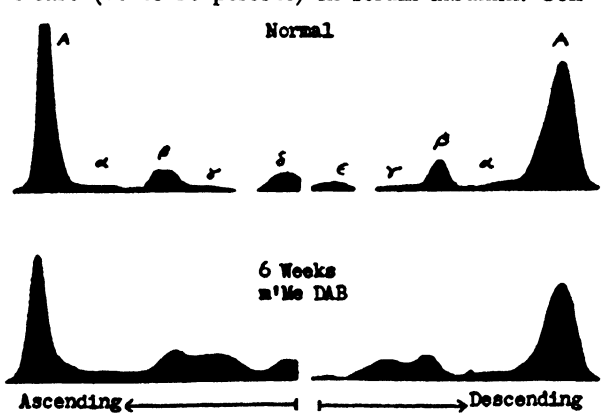


Fig. 3. Electrophoretic patterns of rat sera.

When we turn to the sera of rats that received acetylaminofluorene the picture is found to be quite different. As is shown in Table 5 (Griffin, Cook and Cunningham, 1949) serum γ-globulin shows little if any increase. Alpha and β-globulin, which are best considered together rather than separately, increase somewhat with the maximum effect in evidence after tumors appear. Serum albumin decreases slightly.

It was stated earlier that such serum protein changes are a reflection of more fundamental changes in liver protein distribution. Despite the seemingly rapid turnover of proteins in the normal liver and the equilibrium in protein distribution that is maintained at different levels of protein metabolism, quite a different situation prevails after the feeding of an appropriate carcinogen. *p*-Di-

methylaminoazobenzene is tightly bound by one or more of the liver proteins (Miller and Miller, 1947), giving rise to a complex that might well be expected to upset the equilibrium that normally prevails and to cause a somewhat different distribution pattern. Work is progressing in several laboratories, including our own, on the synthesis of a C^{14}-labelled azo dye carcinogen. Clearly, such a compound would be of invaluable aid in identifying the protein or proteins in the liver upon which the dye focuses its attack.

Electrophoresis of Liver Extracts

Another approach to the problem, so it seems to us, is to compare the electrophoretic patterns of appropriate extracts of normal and azo dye livers.

Table 5. Effect of 2-Acetylaminofluorene on Composition of Rat Sera*

Time rats were fed diet containing carcinogen	Abumin	α-Globulin	β-Globulin	γ-Globulin	α-+β-Globulin
	percent	percent	percent	percent	percent
Basal diet†					
Range‡	60–74	11–26	10–14	2–10	—
Average of 6 samples	66	14	13	6	27
2 weeks: acetylaminofluorene	61§	16	14	9	30
6 weeks: acetylaminofluorene	62	15	15	8	30
20 weeks: acetylaminofluorene	53	15	21	11	36
24 weeks: acetylaminofluorene	60	18	18	4	36
24 weeks: acetylaminofluorene plus 4 weeks: basal diet (tumors)	53	25	15	7	40

* Buffer: sodium diethylbarbiturate; ionic strength: 0.1; pH: 8.3±0.1; potential gradient: 4–6 volts/cm.

† Percentages are the fractional areas of the electrophoretic diagrams due to each component and represent the percentage of these components to the total amount of protein in the serum.

‡ Values for normal rat serum reported by Cook, Griffin, and Luck.

§ Percentages given at each interval on the acetylaminofluorene diet are for one pooled sample.

Such a comparison should reveal the dissimilarities believed to exist and should be a prelude to any endeavour to fractionate out the key proteins. If the azo dye-protein complex is colored as it may or may not be, depending upon the fate of the azo bond, fractionation would of course be facilitated.

This raises at once the question of what we mean by an "appropriate extract." The obvious but very general answer would be that extract which gives us a maximum yield of the proteins in which we are interested. The more practical definition of an "appropriate extract" must result from a good deal of searching about for the right procedure. At least five laboratories in this country alone are seriously interested in liver protein fractionation, but no extensive report has yet appeared which permits one to describe at length any but his own method.

Nucleoprotein-free Extract

The particular extract which we use for this specific purpose is prepared by homogenization of the blood-free liver for one or two minutes with a mixture of ice and water equal to three times the weight of the liver. The homogenate is then adjusted to pH 5.0 with acetic acid (Luck, 1937) and centrifuged for 45′ at 13,000 R.P.M. (top speed in Type SS1-Servall) under conditions which ensure a final temperature of the centrifugate of not more than 5°. The extract so obtained is free of nucleoproteins as determined by phosphorus analysis and direct application of the usual analytical methods for nucleic acid. The proteins contained in the extract are concentrated either by the freeze-concentration technique or by lyophilizing. Either method appears to be satisfactory and yields reproducible results. The lyophilized material is reconstituted by dissolving in enough water to give approximately a 10 percent solution, filtering off a very small amount of insoluble matter, and dialyzing against 150 volumes of veronal buffer (pH 8.6; ionic strength 0.1), or 150 volumes of phosphate buffer (pH 7.8; ionic strength 0.2). Most of our electrophoretic analyses of such extracts are now being made at pH 7.8 which, empirically, gives better patterns than at pH 8.6.

For electrophoresis this nucleoprotein-free extract is diluted with the buffer-NaCl solution to a suitable total protein concentration. In preliminary experiments electrophoretic patterns have been obtained by Mr. Noda from extracts of normal calf, rat, and rabbit liver. While rather marked species differences are in evidence there is a striking similarity, if not almost an identity, between the patterns for two of the three calf livers. Mrs. Eldridge, in continuing this work, has been able to compare the electrophoretic patterns obtained from extracts of normal rat liver and from the livers of animals in receipt of the azo dye. Again a close similarity in the electrophoretic patterns of the extracts prepared from normal rat livers is in evidence.

We propose to direct part of our fractionation program to isolation of several of the components

of special interest in the azo dye livers. There is no need to labor the point that the protein soluble at pH 5.0 under the condition indicated is only a small portion of the liver protein. There is much more that is soluble at low ionic strengths, extractable at pH 7 to 8, and precipitated out on acidification, still more that may be extracted with sodium chloride of high ionic strength at pH 7 to 8, and a considerable amount in the normal liver that may be extracted with dilute alkali (0.25 percent NaOH) and precipitated out on acidification to pH 5. These fractions may be of equal or greater interest and their study will not be overlooked.

Our preliminary experiments show marked differences between comparable electrophoretic patterns of extracts from the normal rat liver and of those from the livers of rats in receipt of the azo dye. The significance of these differences will not become clear without many more experiments. For example, even with the extracts of normal liver, the probability of protein-protein and protein-buffer interactions must be recognized (Scatchard and Black, 1949; Longsworth and Jacobsen, 1949). Even a neutral salt such as sodium chloride may be expected to make its own contribution, through protein interaction, to the complexity of an electrophoretic pattern. With the extracts of azo dye livers there is the added probability of interaction between the dye and certain of the liver proteins, resulting in addition products that almost certainly would have different mobilities than the uncombined proteins. It is with some apology that these rather elementary considerations are mentioned. It is because of such phenomena that great patience must be exercised in the electrophoretic analysis of liver extracts. The ever-present temptation to regard two components in seemingly comparable extracts as identical because of similar mobilities must obviously be resisted: to succumb to this temptation would be akin to a conclusion by the pilot of a plane that motor cars on the highway below him are identical if they are traveling at identical speeds.

Though these difficulties are very real in the case of liver extracts they are much less serious in the case of blood serum. The comparisons of azo dye and normal sera, referred to earlier, are probably completely permissible since no trace of azo dye has yet been reported in the systemic blood of animals receiving the dye.

Liver Desoxyribonucleoprotein

In working up the nucleoproteins we have thus far concentrated upon the desoxyribonucleoprotein fraction. For this purpose we have used whole liver homogenates, isolated nuclei (Hogeboom, Schneider and Pallade, 1948) and, more recently, chromosomes (Gopel-Ayengar and Cowdry, 1947). When using homogenates, at least, we have found it desirable to freeze the blood-free liver and store in liquid air or dry ice-acetone for 10 to 30 minutes—or to freeze and store at −15° for several weeks: the extractability of some of the proteins is thereby much increased. The liver is next extracted with 0.4 M NaCl at approximately pH 6.5 by homogenization with the salt solution (3 parts of 0.4 M NaCl to 1 part of liver) for two minutes in a Waring blendor. The mixture is centrifuged at 0° at 13,000 R.P.M. (SS1-Servall) for 20 minutes and the residue again extracted with the same volume of 0.4 M NaCl. The two extracts are combined. They are found to contain in typical cases 80 percent of the total nitrogen and phosphorus, 90 percent of the total riboflavin, and 90 percent of the total ribonucleic acid (as nucleoprotein). When aliquots of fresh liver were used, the freezing and storing being omitted, values of 55 percent, 40 percent, and 60 percent extraction were obtained for nitrogen and phosphorus, riboflavin, and ribonucleic acid, respectively. About 6 percent of the desoxyribonucleic acid was extracted. The use of 0.14 M NaCl (Mirsky and Pollister, 1946) was first attempted but found with rat liver to yield only a very partial (40 percent) extraction of the ribonucleic acid. In our experience 0.4 M NaCl is optimal; a high yield of ribonucleoprotein is obtained and very little desoxyribonucleoprotein is extracted (4 percent of the total). As the molarity of the sodium chloride is gradually increased, however, more and more desoxyribonucleoprotein is extracted, and the viscosity of the extract increases in parallel. The increase in extraction is sudden above 0.4 M NaCl. We also find that the nuclei are quite stable up to 0.4 M NaCl but at higher salt concentrations they disintegrate. It is this disintegration of the nuclei which obviously accounts for the increased extractability of desoxyribonucleoprotein. The residue is next extracted with 1.0 M NaCl (3 parts to one part of liver, original fresh weight). Sometimes the extraction is repeated and the extracts combined. The solutions so obtained are very viscous and contain 80 to 85 percent of the total desoxyribonucleic acid and only 3 percent or less of the ribonucleic acid.

For reasons that are still obscure the extractability of the desoxyribonucleoprotein by 1.0 M NaCl is greatly decreased in the case of livers in advanced carcinogenesis. Not uncommonly, as little as 30 percent of the desoxyribonucleoprotein may be extracted. An additional 5 percent may be extracted by turning next to 2 M NaCl, which Pollister and Mirsky (1946) found to be sufficient and necessary for some nucleoprotamines. Our own experiments have not gone far enough to permit a report on the non-extractable portion, approximately 65 percent of the total desoxyribonucleoprotein in the azo dye livers.

For preparative purposes we precipitate the nucleoprotein from the molar sodium chloride extracts by pouring with stirring into sufficient distilled water to bring about a solution 0.14 M with respect

to sodium chloride. The desoxyribonucleoprotein precipitates in long fibers which may be collected on a stirring rod and readily redissolved in 1 M NaCl. The protein is then reprecipitated in 0.14 M NaCl.

Material so prepared is being used in work on the characterization of the protein. It is also suitable for a study designed to answer the question of whether azo dye carcinogenesis leads to qualitative as well as quantitative changes in the desoxyribonucleoprotein. For this purpose we are using a product prepared from the pooled livers of normal rats maintained on the purified basal diet and a second preparation from the pooled livers of rats of the same age and sex after 6 to 10 weeks on the basal synthetic diet plus 0.06 percent *m'*-methyl-*p*-dimethylaminoazobenzene.

In addition to the lessened extractability of desoxyribonucleoprotein from the azo dye livers our preliminary findings indicate unexpected differences in the composition of the extractable portion. The investigations are now being repeated upon other samples of desoxyribonucleoprotein prepared from the whole liver, prepared from nuclei, and prepared from chromosomes.

Before going any further it would seem appropriate to ask whether qualitative changes in the liver cell are a reasonable expectation in azo dye carcinogenesis. This is partly a matter of opinion but the question is provocative. If, for example, it be true that a carcinogen operates by the induction of mutations (Tatum, 1949; Demerec, 1947) a change in the structural substance of the chromosome might be expected and might be of such a character and degree as to be readily detectable. If so, the desoxyribonucleoprotein should reflect such a change. The liver cell nucleus in azo dye carcinogenesis is by no means the same, in our experience, as the normal nucleus. Cytological studies (Cunningham, Griffin and Luck, in press) reveal gross differences which we hope will soon be amenable to closer analysis. Cerebrosides, normally absent from the liver cell nucleus, are reported to be present in the nucleus following butter yellow administration (Williams *et al.*, 1945). Stowell (1949) has also reported upon changes in the relative volumes of the nucleus, nucleolus, and cytoplasm in azo dye carcinogenesis that are in accord with the hypothesis that qualitative changes in the composition of the cell and its nucleus are to be expected. As for desoxyribonucleoprotein itself, its altered composition may be a result of any of the following changes:

(a) In the relative proportions of protein, nucleic acid, and other substances possibly associated (*e.g.* lipid).
(b) In the composition of the nucleic acid component.
(c) In the composition of one or more of the protein components or in the relative amounts of these components: one may increase at the expense of another.
(d) In the relative amounts of the several nucleic acids or nucleoproteins present in that heterogeneous product to which the generic name desoxyribonucleoprotein is applied.

In another investigation we have sought to purify and characterize the desoxyribonucleoprotein of normal liver. Our own studies constitute but a small fragment in a large mosaic to which so many are and have been contributing that it would be a formidable task merely to assemble the bibliography. In brief we have been carrying out electrophoretic and other related studies on liver desoxyribonucleoprotein prepared in various ways. Use has been made, for electrophoresis, of preparations such as those reported upon earlier. Miss Cook has also prepared it from nuclei (Hogeboom, Schneider and Pallade, 1947) and has studied electrophoretically aqueous solutions buffered at a very low ionic strength (Stern *et al.*, 1946, 1947) or in 0.6 to 0.8 M NaCl. The relatively salt-free preparations move quite quickly, while the solutions of high ionic strength require very long periods of electrophoresis.

Two typical electrophoretic diagrams have been obtained. The first shows the three components found when a water extract of isolated rat liver nuclei was electrophoresed in 0.02 μ phosphate buffer at pH 7.7. The mobility of the leading component, calculated from the descending side, was found to be 12.8×10^{-5} cm^2 $\times$ sec^{-1} $\times$ volt^{-1}, which agrees fairly well with that reported by Van Winkle and France (1948) and Frick (1949) although it is slightly lower than that found by Hall (1941). In general, anywhere from one to three components were found in different water preparations of low ionic strength.

The second pattern was obtained when the same nuclei were re-extracted with 1 M NaCl and electrophoresed in a 0.8 μ glycine-NaOH buffer at pH 9.0. This accords with the work of Frick (1949) who also reports two to three components in 1 M NaCl solutions. The mobility of the leading component was 7.5×10^{-5} cm^2 $\times$ sec^{-1} $\times$ volt^{-1} (descending boundary) which is slightly lower than values found by Van Winkle but agrees with that found by Frick for thymus nucleoprotein in molar salt solutions.

Since in nine electrophoreses of nucleoprotein in salt solutions of high ionic strength, two components were found, attempts were made to characterize these components further by preparative electrophoresis. The divided Tiselius cell was used in order to isolate the components. The cell compartment contents were then analyzed for desoxyribonucleic acid by the Dische reaction, and for protein by using the glyoxylic acid test for tryptophane, and the biuret or Sakaguchi reactions. These preliminary analyses indicate that the leading component is a nucleoprotein fraction containing a high ratio of desoxyribonucleic acid.

The slow component contains protein with a small amount of desoxyribonucleic acid. These results are very similar to those obtained by Stenhagen and Teorell (1939) on electrophoresis of artificial mixtures of nucleic acid and serum albumin.

The results seem to indicate that we have two nucleoprotein fractions: a fast component containing a high proportion of desoxyribonucleic acid and a slow protein component with a small amount of desoxyribonucleic acid. Either we have two different nucleoprotein fractions moving at different rates or two "partial dissociation" complexes of nucleohistone. Since nucleohistone is known to be partially dissociated in 1 M salt solutions (Mirsky and Pollister, 1946), this possibility is not excluded.

The direct precipitation of native protein (nucleic acid free) and of nucleic acid from solutions of high ionic strength is being attempted.

Liver Albumin

Earlier in this paper reference was made to the nucleoprotein-free fraction obtained by extraction with water and removal of all proteins insoluble at pH 5.0. We have become interested anew in this fraction since in it is to be found the liver albumin upon which we have reported before (Luck, 1936, 1937, 1938; Luck and Martin, 1937; Nimmo, 1942, Luck, Nimmo and Alvarez-Tostado, 1941). From this, or a comparable fraction, a small amount of a crystalline protein was prepared by Nimmo (1942) using dog liver. The yield, however, was small and the protein was found to be quite heterogeneous on electrophoresis. The electrophoretic migration of the principal component in phosphate-sodium chloride buffer at pH 7.4 was very little even after 10 hrs. The isoelectric point appeared to be pH 5.8 (Luck, Nimmo and Alvarez-Tostado, 1941). In more recent work Noda has concentrated upon a component of the nucleoprotein-free fraction of calf liver. From an extended study of the effects of ionic strength, pH, and alcohol concentration upon the solubility of this component it is found to be precipitated in fairly good yield at $-5°$ C at pH 8.6 (veronal buffer 0.12 M), ionic strength (sodium acetate) of 0.1, and an alcohol concentration of 15 percent. As determined by electrophoresis at pH 7.7 and 8.2 the product appears to be from 75 percent to 80 percent pure. It has been precipitated from solutions ranging from one percent to five percent total protein without any conclusive results as to the effects of initial protein concentration on the purity of the fraction. The component has an isoelectric point of about 5.7 (by extrapolation since it denatures at low pH, and a mobility of -3.4×10^{-5} $cm^2 \times volt^{-1} \times sec^{-1}$ in phosphate buffer of pH 7.7, ionic strength 0.1. Crystallization will be undertaken as more of the product is accumulated. It is a labile substance and lyophilization followed by reconstitution has always led to the separation of small amounts of insoluble protein.

Extensive studies on the fractionation of the liver proteins are underway in various laboratories. Thus far the reports that have appeared in the literature are fragmentary. This whole subject of investigation is new and several years may yet elapse before fractionation procedures as systematic and complete as those for blood serum will have been developed.

Grateful acknowledgment is made of the assistance and cordial collaboration in these present studies of A. Clark Griffin, Hyla Cook, Eleanore Frey, Donald Kupke, Alberta Rhein, Noreen Eldridge, Rene Mills, William Nye, and Lafayette Noda. Grants-in-aid from the Rockefeller Foundation, the American Cancer Society, Swift and Co., and the U. S. Public Health Service have made these studies possible.

References

Addis, T., Poo, L. J., Lew, W., and Yuen, D. W., 1936, Gravimetric methods for the determination of total body proteins and organ protein. J. biol. Chem. *113:* 497-504.

Addis, T., Poo, L. J., and Lew, W., 1936a, The quantities of protein lost by the various organs and tissues of the body during a fast. J. biol. Chem. *115:* 111-116.

1936b, Protein loss from liver during a two day fast. J. biol. Chem. *115:* 117-18.

1936c, The rate of protein formation in the organs and tissues of the body. I. After casein refeeding. J. biol. chem. *116:* 343-52.

Addis, T., Lee, D. D., Lew, W., and Poo, L. J., 1940, Protein content of the organs and tissues at different levels of protein consumption. J. Nutrition *19:* 199-205.

Borsook, H., and Dubnoff, J. W., 1943, The metabolism of proteins and amino acids. Ann. Rev. Biochem. *12:* 183-204.

Borsook, H., and Keighley, G. L., 1935, The "continuing" metabolism of nitrogen in animals. Proc. Roy. Soc. (London) B, *118:* 488-521.

Brues, A. M., Tracey, M. M., and Cohn, W. E., 1944, Nucleic acids of rat liver and hepatoma: their metabolic turnover in relation to growth. J. biol. Chem. *155:* 619-33.

Campbell, R. M., and Kosterlitz, H. W., 1946, The influence of sex on liver cytoplasm. J. Physiol. *105:* 33P-34P.

Cook, H. A., Griffin, A. C., and Luck, J. M., 1949, Tissue proteins and carcinogenesis. II. Electrophoretic studies on serum proteins during carcinogenesis due to azo dyes. J. biol. Chem. *177:* 373-81.

Cunningham, L., Griffin, A. C., and Luck, J. M., Effect of a carcinogenic azo dye on liver cell structure. Isolation of nucleic and cytoplasmic granules. Cancer Res. in press.

Davidson, J. N., 1947, Some factors influencing the nucleic acid content of cells and tissues. Cold Spr. Harb. Quant. Biol. *12:* 50-59.

Davidson, J. N., and Waymouth, C., 1944, The nucleic acid and nucleotide content of liver tisue. Biochem. J. *38:* 379-85.

Demerec, M., 1947, Production of mutations in Drosophila by treatment with some carcinogens. Science *105:* 634.

Dickens, F., and Weil-Malherbe, H., 1943, The metabolism of normal and tumor tissue. XX. A comparison of the metabolism of tumors of liver and skin with that of

the tissue of origin. Cancer Res. *3:* 73-87.

DOLE, V. P., WATSON, R. F., ROTHBARD, S., BRAUN, E., and WINFIELD, K., 1945, Electrophoretic changes in the serum-protein patterns of patients with scarlet fever and rheumatic fever. J. clin. Invest. *24:* 648-56.

FRICK, GÖSTA, 1949, Some physico-chemical properties of thymonucleoprotein prepared according to Mirsky and Pollister. Biochim. et Biophys. Acta, *3:* 103-116.

FUJIWARA, T., NAKAHARA, W., and KISHI, S., 1937, Comparison of chemical composition between hepatoma and normal liver tissues. III. Phosphorus compounds, cholesterol and fatty acids. Gann *31:* 51-62.

GIESE, J. E., MILLER, J. A., and BAUMANN, C. A., The carcinogenicity of m'-methyl-p-dimethylaminoazobenzene. Cancer Res. *5:* 337-40.

GOPEL-AYENGAR, A. R., and COWDRY, E. V., Desoxyribose nucleic acid from isolated chromosome threads in experimental epidermal methylcholanthrene carcinogenesis in mice. Cancer Res. *7:* 1-8.

GRAY, S. J., and BARRON, E. S. G., 1943, The electrophoretic analyses of the serum proteins in diseases of the liver. J. clin. Invest. *22:* 191-200.

GRIFFIN, A. C., and BAUMANN, C. A., 1946, The effect of certain azo dyes upon the storage of riboflavin in the liver. Arch. Biochem. *11:* 467-76.

GRIFFIN, A. C., COOK, H , and CUNNINGHAM, L., 1949, Tissue proteins and carcinogenesis. III. Precancerous changes in the liver and serum proteins of rats fed acetylaminofluorene. Arch. Biochem. *24:* 190-198.

GRIFFIN, A. C., NYE, W. N., NODA, L., and LUCK, J. M., 1948, Tissue proteins and carcinogenesis. I. The effect of carcinogenic azo dyes on liver proteins. J. biol. Chem. *176:* 1225-35.

GUTMAN, A. B., MOORE, D. H., GUTMAN, E. B., MCCLELLAN, B., and KABAT, E. A., 1941, Fractionation of serum proteins in hyperproteinemia with special reference to multiple myeloma. J. clin. Invest. *20:* 765-83.

HALL, J. L., 1941, The physical-chemical investigation of certain nucleoproteins. II. General electrophoretic behavior. J. Amer. Chem. Soc. *63:* 794-798.

HALLIBURTON, W. D., 1892, The proteids of kidney and liver cells. J. Physiol. *13:* 806-46.

HAMMARSTEN, E., and HEVESY, G., 1946, Rate of renewal of ribo- and desoxyribo nucleic acids. Acta Physiol. Scand. *11:* 335-43.

HOGEBOOM, G. H., SCHNEIDER, W. C., and PALLADE, G. E., 1948, Cytochemical studies of mammalian tissues. I. Isolation of intact mitochondria from rat liver; some biochemical properties of mitochondria and submicroscopic particulate material. J. biol. Chem. *172:* 619-35.

HUGGINS, C., MILLER, G. M., and JENSEN, E. V., 1949, Thermal coagulation of serum proteins. II. Deficient coagulation in cancer and the iodoacetate index. Cancer *9:* 177-82.

KENNEDY, E. B., and LEHNINGER, A. L., 1949, Oxidation of fatty acid and tricarboxylic acid cycle intermediates by isolated rat liver mitochondria. J. biol. Chem., in press.

KENSLER, C. J., and SUGIURA, K., and RHOADS, C. P., 1940, Coenzyme I and riboflavin content of livers of rats fed butter yellow. Science *91:* 623.

KOSTERLITZ, H. W., 1944a, Effect of dietary protein on liver cytoplasm. Biochem. J. *38:* 14P.

1944b, Effect of dietary protein on liver cytoplasm. Nature, Lond. *154:* 207-08.

1946, Dietary protein and histological structure of liver tissue. J. Physiol. *105:* 11P.

1947, Effect of changes in dietary protein on the composition and structure of the liver cell. J. Physiol. *106:* 194-210.

KOSTERLITZ, H. W., and CAMPBELL, R. M., 1945, Storage of protein in the adult animal. Nut. Abstr. Rev. *15:* 1-14.

1946, Assay of the biological value of a protein by its effect on liver cytoplasm. Nature, Lond. 157: 628.

KOSTERLITZ, H. W., and CRAMB, I. D., 1943, The effect of fasting on the protein content of the liver. J. Physiol. *102:* 18P.

LAN, T. H., 1944, The d-amino acid oxidase, uricase and choline oxidase in two transplanted rat tumors and in isolated nuclei of tumor cells. Cancer Res. *4:* 42-44.

LONGSWORTH, L. G., and JACOBSEN, C. F., 1949, An electrophoretic study of the binding of salt ions by β-lactoglobulin and bovine serum albumin. J. phys. colloid Chem. *53:* 126-35.

LONGSWORTH, L. G., SHEDLOVSKY, T., and MACINNES, D. A., 1939, Electrophoretic patterns of normal and pathological human blood serum and plasma. J. Exp. Med. *70:* 399-413.

LUCK, J. M., 1935, Fractionation of liver proteins. Abst. Communicat. XV. Int. Physiol. Cong., Leningrad and Moscow, 246.

1936, The question of protein storage. J. biol. Chem. *115:* 491-510.

1937, The Liver Proteins. "Perspectives in Biochemistry," 215-29 (Cambridge University Press).

1938, The extractability of the liver proteins. C. R. Lab. Carlsberg (Ser. Chim.) *22:* 321-28.

LUCK, J. M., EUDIN, J. G., and NIMMO, C. C., 1939, Autolytic changes in the protein and amino acid content of liver. J. biol. Chem. *131:* 201-09.

LUCK, J. M., and MARTIN, D., 1937, Liver albumin. Proc. Soc. exp. Biol., N.Y. *36:* 320-21.

LUCK, J. M., and NIMMO, C. C., 1937, Effect of hydrogen-ion concentration on the extractability of liver proteins. J. biol. Chem. *119:* lxv.

LUCK, J. M., NIMMO, C. C., and ALVAREZ-TOSTADO, C., 1941, Electrophoretic properties of liver proteins. J. biol. Chem. *140:* lxxxi.

LUETSCHER, J. A., JR., 1940, Electrophoretic analyses of plasma and urinary proteins. J. clin. Invest. *19:* 313-20.

1941, Electrophoretic analysis of the proteins of plasma and serous effusions. J. clin. Invest. *20:* 99-106.

1947, Biological and medical applications of electrophoresis. Physiol. Rev. *27:* 621-642.

MASAYAMA, T., and YOKOYAMA, T., 1940, Changes in the thymonucleic acid content of liver during cancer development. Gann *34:* 174-75. See also Chem. Abst. *35:* 3319 (1941).

MILLER, E. C., and MILLER, J. A., 1947, The presence and significance of bound aminoazo dyes in the liver of rats fed p-dimethylamino-azobenzene. Cancer Res. *7:* 468-80.

MILLER, J. A., 1947, Studies on the mechanism of the effects of fats and other dietary factors on carcinogenesis by the azo dyes. Ann. N. Y. Acad. Sci. *49:* 19-28.

MILLER, J. A., and BAUMANN, C. A., 1945, The carcinogenicity of certain azo dyes related to p-dimethylaminoazobenzene. Cancer. Res. *5:* 227-34.

MILLER, L. L., 1948, Loss and restoration of rat liver enzymes (activity) related to dietary changes in liver protein. Federation Proc. *7:* 174.

MIRSKY, A. E., and POLLISTER, A. W., 1946, Chromosin, a desoxyribose nucleoprotein complex of the cell nucleus. J. gen. Physiol. *30:* 117-147.

MOORE, D. H., KABAT, E. A., and GUTMAN, A. B., 1943,

Bence-Jones proteinemia in multiple myeloma. J. clin. Invest. *22:* 67-75.

MOORE, D. B., PIERSON, P. S., HANGER, F. M., and MOORE, D. H., 1945, Mechanism of the positive cephalin-cholesterol flocculation reaction in hepatitis. J. clin. Invest. *24:* 292-95.

NIMMO, C. C., Liver proteins. Fractionation, electrophoresis and enzyme studies. (Ph. D. dissertation, Stanford University).

PLÓSZ, P., 1873, Ueber die eiweissartige Substanen der Leberzelle. Arch. ges. Physiol. *7:* 371-90.

POLLISTER, A. W., and MIRSKY, A. E., 1946, The nucleoprotamine of trout sperm. J. gen. Physiol. *30:* 101-16.

PRICE, J. M., MILLER, E. C., and MILLER, J. A., 1948, The intracellular distribution of protein, nucleic acids, riboflavin, and protein-bound aminoazo dye in the livers of rats fed p-dimethylaminoazobenzene. J. biol. Chem. *173:* 345-53.

SCATCHARD, G., and BLACK, E. S., 1949, The effect of salts on the isoionic and isoelectric points of proteins. J. phys. colloid Chem. *53:* 88-99.

SCHNEIDER, W. C., 1945, Phosphorus compounds in animal tissues. I. Extraction and estimation of desoxypentose nucleic acid and of pentose nucleic acid. J. biol. Chem. *161:* 293-303.

SCHOENHEIMER, A., 1942, "The dynamic state of body constituents," (Harvard University Press).

SEITZ, W., 1906, Die Leber als Vorrathskammer für Eiweissstoffe. Arch. ges. Physiol. *111:* 309-34.

SHEMIN, D., and RITTENBERG, D., 1944, Some interrelationships in general nitrogen metabolism. J. biol. Chem. *153:* 401-21.

SOROF, S., and COHEN, P. P., 1949, Electrophoretic studies of the soluble proteins of rabbit liver. Federation Proc. *8:* 254.

STENHAGEN, E., and TEORELL, T., 1939, Electrophoretic properties of thymonucleic acid. Trans. Faraday Soc. *35:* 743-50.

STERN, K. G., and DAVIS, S., 1946, Studies on thymus nucleohistone. Federation Proc. *5:* 156-57.

STERN, K. G., GOLDSTEIN, G., WAGMAN, J., and SCHRYVER, J., 1947, Studies on desoxyribonucleoproteins: isolation and properties of genoprotein T. Federation Proc. *6:* 296.

STOWELL, R. E., 1949, Alterations in nucleic acids during hepatoma formation in rats fed p-dimethylaminoazobenzene. Cancer *2:* 121-31.

TICHMENEFF, N., 1914, Ueber Eiweissspeicherung in der Leber. Biochem. Z. *59:* 326-32.

WESTPHAL, U., 1944, d-Aminoacid oxidase of liver extracts from rats with chemically induced tumors. Z. physiol. Chem. *281:* 89-93. Chem. Abst. *37:* 3860 (1947).

WILLIAMS, H. H., KAUCHER, M., RICHARDS, A. J., MOYER, E. Z., and SHARPLESS, G. R., 1945, The lipid partition of isolated cell nuclei of dog and rat livers. J. biol. Chem. *160:* 227-32.

VAN WINKLE, Q., and FRANCE, W. G., 1948, Electrophoretic and ultracentrifugal investigation of rabbit liver nucleoprotein. J. phys. colloid. Chem. *52:* 207-216.

DISCUSSION

BENESCH: In connection with Dr. Luck's observations on the diminished extractability of desoxyribonucleic acid from hepatoma tissue, it seems interesting to recall Greenstein's results on the increased resistance to denaturation of such tissue, as revealed by the number of sulfhydryl groups released. With regard to the striking increase in the lipid content of hepatoma tissue, it might be mentioned that Goerner and Goerner also observed this phenomenon and found it to be correlated with a disappearance of vitamin A from these tissues.

SEIBERT: I was much interested in Dr. Luck's data showing, by means of electrophoresis, an increase in the gamma globulin in the sera of rats with liver tumors and in the alpha and beta globulins in those with tumors elsewhere. We have found essentially the same thing in cases of human carcinoma; namely, a rise in the alpha$_2$ globulin in all cases of carcinoma and, mainly in those where metastases to the liver, occurred, the gamma globulin was also high.

Your analyses, showing a significant amount of desoxyribonucleic acid in liver tumor tissue leads me to ask whether similar analyses were made on the sera. These would be most interesting in view of the fact that we have found in the sera of cases with carcinoma a significant increase in a substance which gives the color reaction with tryptophane and perchloric acid, a test claimed by Seymour Cohen to be specific for desoxyribonucleic acid.

MENKIN: Since in rats receiving 2-acetylaminofluorene, there is an increase in the α-globulins of their sera with the concomitant formation of liver tumors, and since it has been shown at the Rockefeller Institute that in various inflammatory conditions there is a rise in the α-globulins of the blood, I am wondering whether these observations can be correlated with changes in the leukocyte level of the circulation. I have shown that in exudates there is an alpha globulin (termed the leukocytosis-promoting factor) which offers an explanation for the mechanism of leukocytosis with inflammation. This globulin also causes a specific growth of granulocytes in the bone marrow. The formation of some tumors involves a certain degree of cellular injury, and therefore in a way acts as an inflammatory irritant. It may be that the formation of the neoplasm in the liver induces sufficient tissue injury, and thus releases the leukocytosis promoting factor. A correlation of the development of neoplasms with changes in the leukocyte level may give biological meaning to the increase in α-globulins in the sera of the rats utilized.

May I also inquire whether cirrhosis of the liver always precedes the formation of liver tumors? We have shown several years ago that cellular injury seems to liberate one or several growth-promoting factors. These can be recovered in exudates. It is just possible that the degree of cellular injury in the liver leading to cirrhosis by the carcinogen may also contribute to the eventual formation of the liver tumors through the release of growth-promoting factors.

LUCK: Dr. Seibert's suggestion interests me very much indeed. We have run no tryptophane-perchloric acid tests on sera but apparently we should do so.

We have not inquired into the possibility of inducing an inflammatory state by administration of acetylaminofluorene. Dr. Menkin's suggestion will encourage us to run some leukocyte counts. I am not sure, however, that one may generalize to the extent of regarding all tumors as inflammatory agents. At least, in our own experience, hepatoma formation by *m'*-methyl-*p*-dimethylaminobenzene is not accompanied by the α-globulin increase of which Dr. Menkin speaks. Furthermore it should be noted that with both carcinogens the characteristic serum protein changes ensue before discrete tumors become visible.

PETERMANN: Studies with adenine as a precursor have shown that in non-growing liver the purines of the desoxypentose nucleic acid are renewed only about one percent as rapidly as are the purines of the pentose nucleic acid.

LUCK: The work with adenine to which you refer had escaped my attention and I am grateful to you for reminding me of it. However, I would repeat the point that turnover numbers for the two nucleic acids do not necessarily throw any light on the turnover rates of the associated proteins. It is perfectly conceivable that the latter may be synthesized or broken down more or less rapidly than their prosthetic nucleic acids.

SEIBERT: In our analyses of carcinoma sera we based them on total protein values, and in spite of the fact that these were almost always very low in cases of carcinoma, the actual values for alpha$_2$ globulin were always high. Moreover, it was shown that it was this protein fraction that was accompanied by the polysaccharide giving the color reaction with tryptophane and perchloric acid, and also that a rise occurred in the sera of not only carcinoma but also tuberculosis and other diseases where there was tissue destruction. If the substance is of nucleic acid nature, it is probably a mixture of desoxyribonucleic and ribonucleic acid since the spectral absorption curves of the reacting mixture showed an absorption peak between those found for pure desoxyribonucleic acid (λ 5000 A) and for ribonucleic acid (λ 4500 A).

SIZE RELATIONSHIP AMONG SIMILAR PROTEINS

ASSOCIATION AND DISSOCIATION REACTIONS OF PROTEIN UNITS

KAI O. PEDERSEN
Institute of Physical Chemistry, University of Upsala, Upsala, Sweden

A quarter of a century ago it was generally assumed that the proteins consisted of particles varying considerably in size. It was therefore quite unexpected when Svedberg and Fåhraeus (1926) and Svedberg (1926) found that all the proteins first studied in the ultracentrifuge appeared to be monodisperse and to have definite molecular weights (M). This finding of Svedberg stimulated him and his pupils to investigate a large number of proteins in the ultracentrifuge. Since at that time the light absorption method was the most convenient to use, it was advantageous to study colored proteins and therefore a number of chromo-proteins were the first to be investigated. Soon it appeared to Svedberg that the M found often were multiples of 35,000. A lot of work was therefore done on the study of proteins from various sources in order to disprove or prove the validity of a multiple hypothesis for the molecular weights of proteins.

At the beginning of the thirties Svedberg and his collaborators started an extensive investigation of the respiratory proteins belonging to different classes of the animal kingdom (for details see Svedberg and Hedenius, 1934). It soon became evident that the respiratory proteins from species belonging to the same class generally had about the same sedimentation constants (s_{20}). In some instances two groups of s_{20} were found within the same class. In such cases the blood from some of the species often contained two components each representing one of the groups of s_{20} found within the class. By changing the pH of the protein solution it was often possible either to change the proportion between two such components or, in the proteins containing only one component, to produce another component having the other s_{20}. In order to find out whether the various respiratory proteins having the same s_{20} were the same proteins or were different, a number of them were studied in electrophoresis (Pedersen, 1933 and 1940). It was hereby found that even closely related species showed distinctly different electrophoretical behaviour. Later experiments by Eriksson-Quensel and Svedberg (1936) on the pH-stability regions of the hemocyanins have confirmed the difference between the various proteins.

Within each group a few of the more easily accessible were selected for more detailed studies and for determinations of M. The M found for different members of the same group generally agreed fairly well, although sometimes differences were observed which were larger than could be ascribed to differences in the amino acid composition. I think that the large deviations sometimes found must be due either to impure protein solutions being studied or to inaccuracy in the sedimentation and diffusion measurements. Thus a neglect of the variation of the apparent s_{20} with concentration may result in the use of too low a value for s_{20}, and consequently also in M. Some of the diffusion measurements may also be in need of revision. As to the sedimentation equilibrium method, it has recently been found that a systematic error in the determination of the speed of the electric ultracentrifuges may have resulted in about five percent too low values for M. Personally I feel quite convinced that a careful redetermination of M for respiratory proteins belonging to the same group will reduce the observed spread to an amount which may be explained mainly by differences in amino acid composition.

By studies of dissociation and association reactions on respiratory proteins it was also found that the molecules could be reversibly split or associated to molecules having the same M as those belonging to other groups of respiratory proteins.

A number of other types of proteins were studied in Upsala, and whenever their s_{20} indicated that a hitherto unknown M could be expected, a special determination of M was made. Generally, however, the s_{20} found coincided with one of the groups of s_{20} found earlier, and it was not considered worthwhile to spend time on further determination of their M. When, therefore, the tables giving the M of the proteins were published, all the large groups of proteins with the same M were only represented in the table of M by a very small fraction of their total, whereas those which did not belong to any of the large groups were *all* put into the table. By this selection the tables of M published from Upsala give a summary of all the different values of M found for different types of proteins. It is, however, quite evident that it gives no information whatever about the frequency with which a given M occurs among the known proteins.

The multiple hypothesis has been much discussed and severely criticized by, for instance, Bull (1941), Rothen (1942), Johnston, Longuet-Higgins and Ogston (1945), and Norris (1946). Bull made a logarithmic "spectrum" of the M values published in the tables and found that "there are no molecular weight classes. There is an apparent tendency for certain molecular weights to cluster around 17,000, but no one knows how many 'proteins' have smaller

weights than 17,000; the clustering may simply indicate that there are a large number of small molecular weight proteins in nature." Bull proposes a statistical analysis of the distribution of M in order to see whether or not the distribution departs significantly from a random one. Such analyses have been made both by Johnston *et al.* and by Norris. On the basis of the values given in the tables of M they find that the multiple hypothesis is disproved by the statistical analysis.

In 1940 I had found that Theorell's purest cytochrome c had a M of about 12,000 and at about the same time Rothen had found that ribonuclease had a M of about 13,000, a value which shortly afterwards was confirmed by me. There were thus at least two proteins with M considerably lower than the multiple unit of 17,600.

I suppose that after all this criticism many protein chemists have said: "Well, it has been proved that there cannot exist a multiple system for the proteins." Undoubtedly the hypothesis of the multiples is not of such general validity as first seemed to be, but still it seems quite possible that we have a limited number of parallel sets of multiple systems each characterized by the specific nature, function, and process of formation of the proteins in question. The M values of the respiratory proteins could probably be adequately described by three (or perhaps four) multiple systems with different basic units. As far as my knowledge of the proteins goes, there is considerable evidence against the statement that the M values are completely randomly distributed.

I would like now to show you some examples from investigations made in Upsala by Svedberg and his collaborators on two groups of proteins, namely, the respiratory proteins and the seed globulins.

TABLE 1. SEDIMENTATION CONSTANTS AND ISO-ELECTRIC POINTS FOR CO-HEMOGLOBIN FROM VERTEBRATES ACCORDING TO DETERMINATIONS MADE IN UPSALA

	s_{20}	I.P.	$\frac{du}{dpH}\cdot 10^5$		s_{20} I	s_{20} II	s_{20} III	I.P.	$\frac{du}{dpH}\cdot 10^5$
Mammalia:				Pisces:					
Erinaceus europaeus	4.5	7.49	4.6	*Raja clavata*	4.3				
Oryctolagus cuniculus	4.4	7.3	7.0	*Protopterus annectens*	4.3			7.21	7.5
Felis domestica	4.2	6.9	7.0	*Salmo irideus*	4.1				
Canis familiaris	4.2			*Cyprinus carassius*	4.4			6.45	4.0
**Equus caballus*	4.4	6.92	7.2	*Anguilla anguilla*	4.1				
Bos taurus	4.6	7.22	6.5	*Esox lucius*	4.2				
Ovis aries		7.09	7.0	*Gasterosteus pungilius*	4.5				
Cercopithecus sp.	4.0			*Lucioperca sandra*	4.5				
**Homo sapiens*	4.5	7.07	6.3	*Tautoga onitis*	4.2			7.45	2.9
				Pleuronectes platessa	4.3				
Aves:				*Prionotus carolinus*	4.4				
Gallus gallus	4.2	7.23	5.5	*Opsanus tau*	4.3			5.75	9.0
Columba livia	4.4	7.23	5.8					6.23	6.0
Anas platyrhyncha	4.4	7.51	5.4						
Syrnium aluco	4.3			Amphibia:					
Picus viridis	4.3			*Salamandra maculosa*	4.8	7.0			
Parus major	4.3			*Bufo viridis*	4.8	7.4			
Corvus cornix	4.4			*Bufo valliceps*	4.8	7.7	12.5		
				Rana temporaria	4.5				
				Reptilia:					
				Chrysemys picta	4.5	7.0			
				Anguis fragilis	4.8				
				Lacerta vivipara	4.6	7.1			
				Chamaeleon chamaeleon	4.6				
				Coluber longissimus	4.6			7.43	4.7

For details see Svedberg and Pedersen (1940). Molecular weight determinations were made only on the hemoglobins marked with *.

In Table 1 a summary of the results of the sedimentation measurements is given. No special efforts have been made for extrapolating the s_{20} to zero concentration, but all the measurements were carried out at a protein concentration below 0.5 percent. The s_{20}'s are all given in Svedberg units (1 S = $1\cdot 10^{-13}$ c.g.s. units). Direct determinations of M have been made only on the proteins marked with an asterisk.

Several of the hemoglobins were also studied in electrophoresis and the results obtained show clearly the electrochemical differences between these pro-

The s_{20} has been determined by means of the light absorption method for a large number of CO-hemoglobins from various vertebrates. In

teins. In one case, with the fish *Opsanus tau*, the electrophoresis diagram showed the presence of two components with different isoelectric points, whereas in all the other cases one component only could be observed.

In the sedimentation diagrams from the Amphibia and the Reptilia two components were often observed, in one case three. The normal components from these two classes did also show a somewhat higher value for s_{20} than those observed for the rest of the hemoglobins.

Red respiratory proteins, the so-called erythrocruorins, are also found in the blood of many invertebrates. In most cases they are directly dissolved in the plasma where they are present as giant molecules (Table 2). In some cases they are present in corpuscles, but their M are then lower than M found for the hemoglobins from the vertebrates. For the faster sedimenting proteins it was possible from a single run to determine whether two proteins had the same sedimentation constants within the limit of error ± 2 percent or whether they differed. Examples of two such test runs are given in Figure 1 and Figure 2. Such mixture tests were extensively used to decide whether two proteins belong to the same group or not. A higher accuracy for such tests may be obtained by the use of the Lamm scale method.

The most extended study on respiratory proteins have been made on the hemocyanins. Table 3 gives a summary of the results obtained for s_{20} and electrophoretic mobilities. The extended study of the dissociation of these proteins will be dealt with a little later.

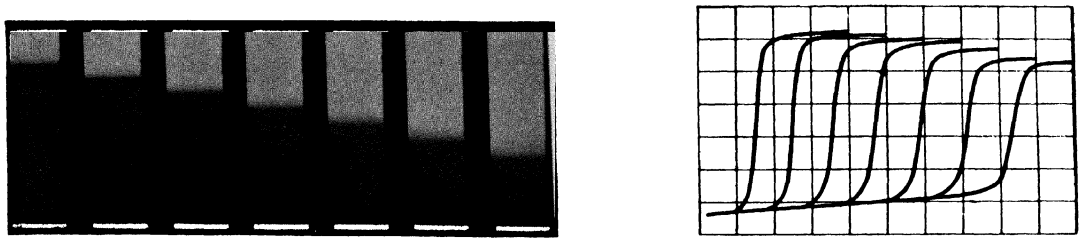
Various seed globulins have occasionally been studied for the last twenty years by Svedberg and collaborators. A more systematic investigation has been carried out for the last five years by Daniels-son (1949) working in Svedberg's laboratory. It was hereby found that the greater part of the isolated seed globulins have s_{20} of the order 12-13 S and M ≃ 300,000. The seed globulins from 10 different families were studied, and nine of these contained a component of this size (Table 4). The chemical and immunological properties are different for the various families so far investigated.

From the family Gramineae eight different species were investigated, and it was found that four different globulin components occurred in the various seed globulins. Three of these had lower s_{20} than the most common component found in the seed globulins from the other families, namely, with s_{20} ≃ 12-13 S. It is seen from Table 5 that barley, *Hordeum vulgare*, has a unique position in this family, since it contains all the four components, whereas in the rest of the Gramineae it was found that in a few cases only the α- or the γ-globulin was present. In most cases both of these two com-

TABLE 2. SEDIMENTATION CONSTANTS AND ISO-ELECTRIC POINTS FOR ERYTHROCRUORINS FROM INVERTEBRATES AND VERTEBRATES, CYCLOSTOMATA, ACCORDING TO DETERMINATIONS MADE IN UPSALA

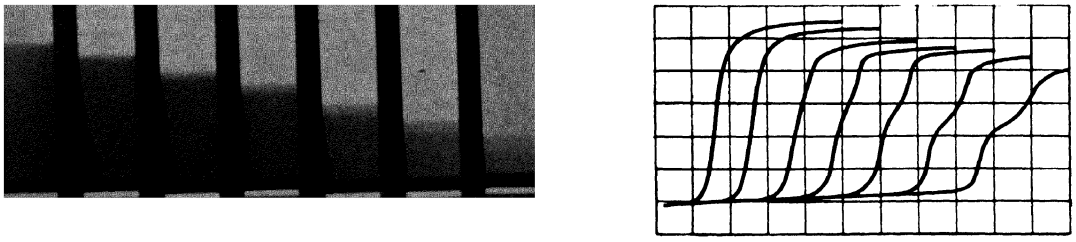
	Erythro-cruorin dissolved in	s_{20}	I.P.	$\frac{du}{dpH_0} \cdot 10^5$
Chaetopoda: Polychaeta				
Nereis virens	plasma	58.6	5.10	9.5
Lumbrinereis fragilis	plasma	59		
**Arenicola marina*	plasma	57.4	4.56	16
Eumenia crassa	plasma	59		
Pectinaria belgica	plasma	54 (12)		
Polymnia nebulosa	plasma	57		
Glycera rouxii	corpuscles	3.5		
Notomastus latericeus	corpuscles	2.1	6.0	4–5
Chaetopoda: Hirudinea				
Hirudo medicinalis	plasma	58	5.0	20
Haemopis sanguisuga	plasma	57		
Chaetopoda: Oligachaeta				
**Lumbricus terrestris*	plasma	60.9	5.28	12.6
Eisenia foetida	plasma	63		
Crustacea: Phyllopoda				
Daphnia pulex	plasma	16.3		
Athelges sp.	plasma	19.3		
Eutracheata: Diptera				
**Chironomus plumosus* (larvae)	plasma	2.0	5.40	3.6
Chironomus sp.	plasma	2.0		
Conchifera: Gastropoda				
**Planorbis corneus*	plasma	33.7	4.77	10.6
Planorbis umbilicatus	plasma	36		
Conchifera: Lamelli-branchiata				
**Arca pexata*	corpuscles	3.5	between 5 and 6	
Eleutherozoa: Holo-thurioidea				
Thyone briareus	corpuscles	2.6	about 5.8	
Cyclostomata: Hypero-artia, Hyperotreta				
**Lampetra fluviatilis*	corpuscles	1.87	5.60	3.2
Myxine glutinosa	corpuscles	2.3		

* Molecular weight determinations were made only on the erythrocruorins marked with *.



A B

FIG. 1. Sedimentation pictures (A) and photometer curves (B) of a mixture of Nereis and Sabella blood; centrifugal force 75,000 times gravity; time between exposures 5 minutes. The fact that only one boundary is visible shows that their respiratory proteins, erythrocruorin and chlorocruorin, are identical with regard to sedimentation constant. (Svedberg and Hedenius, 1934.)



A B

FIG. 2. Sedimentation pictures (A) and photometer curves (B) of a mixture of Sepia and Octopus blood; centrifugal force 78,000 times gravity; time between exposures 5 minutes. The fact that the boundary is double shows that their respiratory pigments, which are both of the hemocyanin type, have different sedimentation constants, *viz.* 51.1 and 57.1 S. (Svedberg and Hedenius, 1934.)

TABLE 3. SEDIMENTATION CONSTANTS AND ISO-ELECTRIC POINTS FOR HEMOCYANINS ACCORDING TO MEASUREMENTS MADE IN UPSALA

	s_{20}	I.P.	$\frac{du}{dpH} \cdot 10^5$		s_{20}	I.P.	$\frac{du}{dpH} \cdot 10^5$
Crustacea: Malacostrata				Conchifera: Gastropoda			
**Pandalus borealis*	17.4			**Paludina vivipara*	102.5	4.71	10.8
Palaemon fabrici	16			*Paludina contecta*	100	4.63	11.4
**Palinurus vulgaris*	16.4			*Littorina littorea*	99.7 (132)	4.34	12.8
Pagurus striatus	16			*Neptunea antiqua*	104.0	4.41	7
Eupagurus bernhardus	17 (22)			*Buccinum undatum*	102.1 (132)	4.61	13.7
Squilla mantis	24			**Busycon canaliculatum*	101.7 (130)	4.49	10.7
**Nephorps norvegius*	24.5	4.64	13.3	*Limnaea stagnalis*	98 (60.2)		
**Homarus vulgaris*	22.6	4.95	11.8	*Achatina fulva*	102 (64, 16)	5.03	7.8
Astacus fluviatilis	23.3	4.93	12.1	**Helix pomatia*	103.0	5.05	8.1
Hyas araneus	23			*Helix arbustorum*	91.2	5.50	7.6
Maja squinado	27			*Helix nemoralis*	101.0	4.63	11.4
Cancer pagurus	23.6 (16.4)	4.65	16	*Helix hortensis*	100.0	4.57	12.1
Carcinus maenas	23.3 (16.7)			*Agriolimax agrestis*	100.0		
Chiridothea entomon	23 (16)			*Limax maximus*	97.3		
**Calocaris macandrae*	34			*Arion ater*	99 (61)		
				Arion subfuscus	(64)		
Arachnomorpha: Xiphosura, Scorpionidea				Conchifera: Cephalopoda: Decapoda			
Limulus polyphemus	34.6 (56.6) (16.1) (5.9)	5.96	8.5	*Loligo vulgaris*	56.7		
				Sepiola oweniana	56		
Euscorpius carpaticus	34			**Rossia oweni*	56.2		
				Sepia officinalis	55.9		
Amphineura: Placophora							
Tonicella marmorea	61			Octopoda			
				**Octopus vulgaris*	49.3		
				**Eledone moschata*	49.1		
				Eledone cirrosa	48	4.6	14

* Molecular weight determinations were made only on the hemocyanins marked with *.

ponents could be observed although in different proportions as seen from Figure 3. The globulins in these sedimentation diagrams were obtained by extraction of the whole kernel. If only certain parts of the seeds are extracted different components predominate.

It is seen from Figures 4 and 5 that the α-globulin predominate in the flour from wheat, and the γ-component is present in larger amount than α in the brans. The γ-globulin was the only one present in the embryo. This is also in good agreement with the fact that in the seeds which contain little reserve proteins, the γ-component predominates in the whole kernel. Similar results were obtained in the corresponding fractions from barley seeds.

From the family Leguminosae, 34 different species were investigated by Danielsson, and the seed globulin vicilin could be isolated from all of them. Another seed globulin legumin was present in 29 of the species (Table 6). Some other globulin components could be isolated beside these, but they are probably either association or enzymatic degradation products of legumin. The legumin and vicilin on the other hand are distinctly different substances with entirely different chemical composition. Figures 6 and 7 give the sedimentation

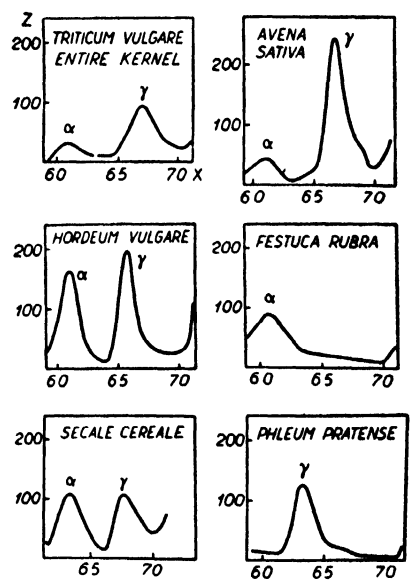


FIG. 3. Sedimentation diagrams showing the composition of globulin solutions prepared by extraction of seeds from different species of the family Gramineae. (C. E. Danielsson, 1949.)

TABLE 4. SEDIMENTATION CONSTANTS AND MOLECULAR WEIGHTS FOR VARIOUS SEED GLOBULINS ACCORDING TO DETERMINATIONS MADE IN UPSALA

Globulin	Species	Family	s_{20}	M
Exelsin	*Bertholletia exelsa*	Lecithydaceae	13.3	295,000
Edestin	*Cannabis sativa*	Moraceae	12.8	310,000
Amandin	*Prunus persica*	Rosaceae	12.5	330,000
Pomelin		Rutaceae	11.4	
Cocosin	*Cocos nucifera*	Palmae	12.0	
Hippocastanin		Sapindaceae	12.9	
Globulin	*Beta vulgaris*	Chenopodiaceae	12.7	
Globulin	*Senapis alba*	Cruciferae	12.7	
Legumin	*Pisum sativum*	Leguminosae	12.6	331,000
Vicilin	*Pisum sativum*	Leguminosae	8.1	186,000
Arachin	*Arachis hypogaea*	Leguminosae	13.1	
Conarachin	*Arachis hypoga*	Leguminosae	8.4	
Glycinin	*Glycine hispida*	Leguminosae	13.1	
Globulin	*Glycine hispida*	Leguminosae	8.0	
Phaseolin	*Phaseolus vulgaris*	Leguminosae	11.0	
Conphaseolin	*Phaseolus vulgaris*	Leguminosae	7.3	
Globulin α	*Hordeum vulgare*	Gramineae	2.5	29,000
Globulin β	*Hordeum vulgare*	Gramineae	6.2	
Globulin γ	*Hordeum vulgare*	Gramineae	8.3	210,000
Globulin δ	*Hordeum vulgare*	Gramineae	12.0	
Avenalin	*Avena sativa*	Gramineae	2.6	
Globulin	*Avena sativa*	Gramineae	8.1	
Globulin α	*Triticum vulgare*	Gramineae	2.5	
Globulin γ	*Triticum vulgare*	Gramineae	8.2	
Globulin α	*Zea mays*	Gramineae	2.6	
Globulin γ	*Zea mays*	Gramineae	8.5	
Globulin α	*Secale cereale*	Gramineae	2.6	
Globulin γ	*Secale cereale*	Gramineae	8.2	

For details see Danielsson, 1949.

diagrams obtained from globulin solutions prepared by extraction of seeds from different species of the family Leguminosae.

We have now seen a number of examples on the grouping of s_{20} around certain values. By changes in the environment, for example, in the pH or in the salt concentration, proteins which under normal conditions belong to one group of s_{20} may be associated or dissociated and thereby come to belong to another group of s_{20}.

TABLE 5. COMPONENTS PRESENT IN VARIOUS SEED GLOBULINS FROM THE FAMILY GRAMINEAE

Species	Sedimentation constants of components			
	α	β	γ	δ
Hordeum vulgare	2.5	6.2	8.3	12.0
Secale cereale	2.6	—	8.2	—
Triticum vulgare	2.5	—	8.2	—
Avena sativa	2.6	—	8.1	—
Zea mays	2.6	—	8.5	—
Plenum pratense	—	—	8.2	—
Festuca pratense	—	—	8.3	—
Festuca rubra	2.4	—	—	—

Danielsson, 1949.

The most comprehensive study of such changes has been made on the respiratory proteins, especially the hemocyanins. Thus Eriksson-Quensel and Svedberg (1936) found that even closely related species belonging to the same family could be distinguished by means of their pH-stability diagram. Figure 8 shows the pH-stability diagrams for four kinds of Helix. It is seen that within a certain range of pH around the isoelectric point only one component with s_{20} around 100 S is visible in the sedimentation diagrams from *Helix pomatia* (H.p.) and *Helix nemoralis* (H.n.). In the case of the hemocyanins from *Helix arbostorum* (H.a.) and *Helix hortensis* (H.h.), two components with s_{20} about 100 S and a little above 60 S are present around the isoelectric points. These two components correspond to whole and half molecules. In the case of H.p. and H.n. whole and half molecules are also present simultaneously at certain pH values as seen from Figure 8. It has recently been found by Strömbäck and Brohult (1949) that the dissociation of the whole molecules of H.p. which usually starts at about pH 7.0, in a 0.01 molar solution of calcium chloride does not begin before pH 9.3 and undissociated whole molecules are still present at pH 11.0-11.5, where H.p. hemocyanin under normal condition is irreversibly split into ill defined low-molecular products. The calcium ion is effective, however, only

if the hemocyanin solution is completely fresh. Old hemocyanin solutions show the same pH-stability in the absence of calcium ions as in the presence.

Very detailed studies of dissociation and association reactions with hemocyanins are being made by Brohult and his collaborators. They have made a careful redetermination of the molecular constants for the hemocyanins from H.p. and *Paludina vivipara* (P.v.). It is seen from Table 7 that the molecular weights for the dissociation products from these two proteins, within the limit of experimental error corresponds to halves and eighths of the normal molecules. Calculation of the lengths of the three kinds of molecules from the values of f/f_0 or from stream double refraction, both indicates that it is the same, which means that the dissociation takes place lengthwise, as illustrated in the lower part of Figure 9. The dissociation is supposed to be preceded by a swelling of the molecules. The presence of the swelled whole molecules, No. 2 and No. 3 in Figure 9, is also manifested by the appearance of the "intermediate compounds" between the peaks for the whole and the half molecules (Fig. 9). Similar swelling effects may also be responsible for certain variation in the sedimentation of bovine CO-hemoglobin observed by me ten years ago and to be dealt with later.

Dissociation of the hemocyanin molecules may in some cases be brought about by the addition of electrolytes. From Table 8 it is seen that H.p. and H.h. hemocyanins are only partly dissociated by electrolytes, whereas H.a. hemocyanin is 100 percent dissociated and the hemocyanins from P.v., *Littorina littorea* and *Buccinum undulatum* show no dissociation at all. In the cases where the hemocyanins are only partly dissociated by salts, it has been shown by Brohult (1947) that two different kinds of molecules may be isolated. Figure 10 shows sedimentation diagrams from such experiments. In Figure 10A only whole molecules are present; by increasing the concentration of sodium chloride from 0.2 molar to 1 molar the maximum dissocia-

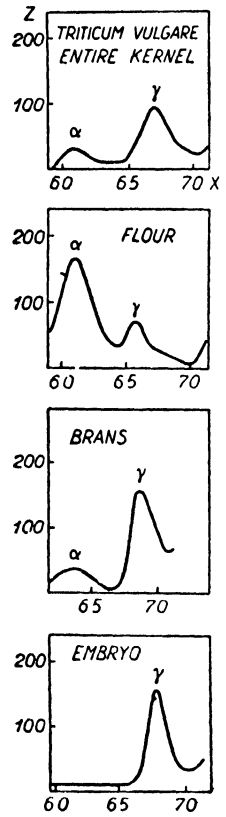


FIG. 4

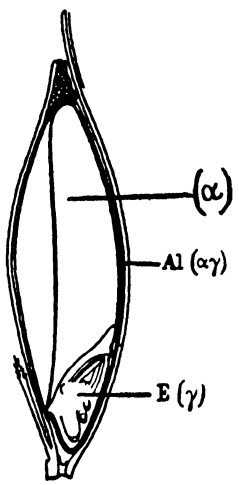


FIG. 5

FIG. 4. Sedimentation diagrams showing the composition of globulin solutions prepared by extraction of different parts of seeds from wheat. (Säverborn, Danielsson and Svedberg, 1944.)

FIG. 5. Diagram showing the localization of the α and the γ-components in seeds of wheat.—Al: aleurone grains. E: embryo.

TABLE 6. COMPONENTS PRESENT IN VARIOUS SEED GLOBULINS FROM THE FAMILY LEGUMINOSAE (DANIELSSON, 1949)

Species	Sedimentation constants in Svedberg units			
Acacia alata	1.25	7.90	11.6	
Acacia decipiens		8.02	12.7	
Acacia Farnesiana		8.04		
Acacia longifolia	2.72	7.59		18.0
Acacia penninervis	1.46	7.39		18.8
Acacia saligna		7.76	13.7	
Acacia verticillata		7.77		
Arachis hypogea	1.93	8.40	13.1	
Astragalus galegiformis		8.33	13.2	
Cytisus laburnum		8.08	14.0	
Cytisus supinus	1.84	8.03	13.4	
Dolichos Lablab		7.33	11.7	
Ervum Lens		7.25	13.2	
Genista tinctoria		8.54	13.3	20.3
Glycine Soja		7.97	13.1	
Lathyrus Clymenum		7.55	13.0	
Lathyrus odoratus		7.64	12.0	
Lathyrus sativus		7.46	13.0	
Lathyrus silvestris		7.48	13.0	
Lotus Tetragonolobus		8.32	13.1	
Lupinus albus		8.24	12.3	
Lupinus angustifolius		8.20	13.1	
Lupinus luteus		8.30	11.5	
Lupinus polyphyllus		8.69	12.2	
Medicago sativus		6.77	11.4	
Phaseolus coccineus	4.29	7.39	12.2	
Phaseolus nanus		6.55	10.1	
Phaseolus vulgaris	4.87	7.26	11.0	
Pisum sativum		8.10	12.6	
Trifolium hybridum		7.66	12.9	
Trifolium pratense		7.69	11.2	
Trifolium repens		7.27		18.2
Vicia faba		7.12	11.8	
Vicia sativa		7.09	11.9	

tion takes place as shown in Figure 10B. By spinning such a solution in one of the Beams type of preparative high speed centrifuges a supernatant solution, containing only half molecules, may be obtained (Fig. 10C). The deposit, when dissolved in the buffer containing 1 molar sodium chloride shows an increased amount of whole molecules. The half molecules corresponding to Figure 10C may be reassociated by decreasing the electrolyte concentration. Such a solution shows 100 percent dissociation by electrolytes and normal dissociation by change in pH.

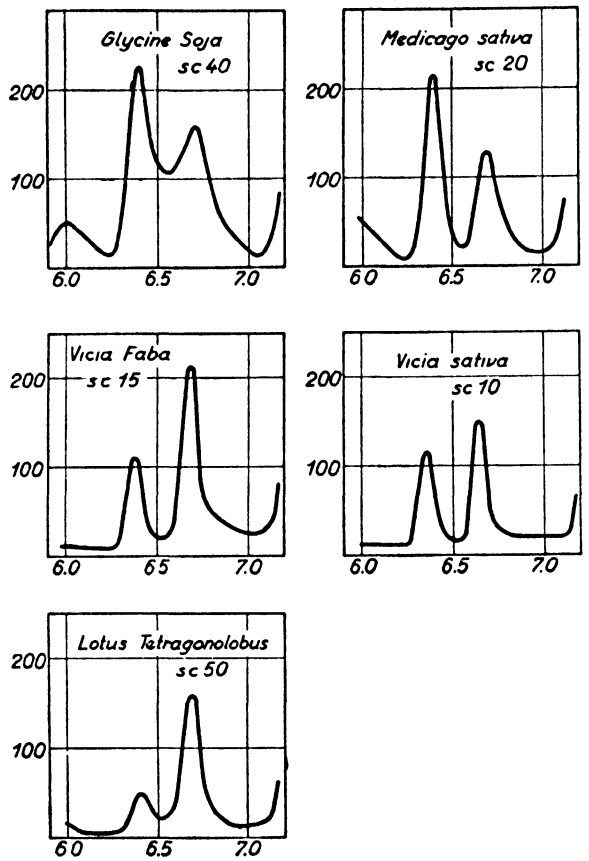


FIG. 6. Sedimentation diagrams showing the composition of globulin solutions prepared by extraction of seeds from different species of the family Leguminosae. (Danielsson, 1949.)

The study of association and dissociation reactions with more low molecular proteins (M below 200,000) has been taken up recently. Together with K. Olsson the author has been studying the behavior of insulin, especially at the acid border of the stability range. Insulin is not soluble in a pH region around its isoelectric point, but it may be dissolved, for example, in phosphate buffers above pH 7. We have made a series of determinations at different concentrations of insulin in phosphate buffers of pH 7.6, and they gave the following results (Table 9). The values for s_{20} thus seems to vary very little with concentration. Calculation of the line of regression gives $s_{20} = 3.455 - 0.00078\ c$.

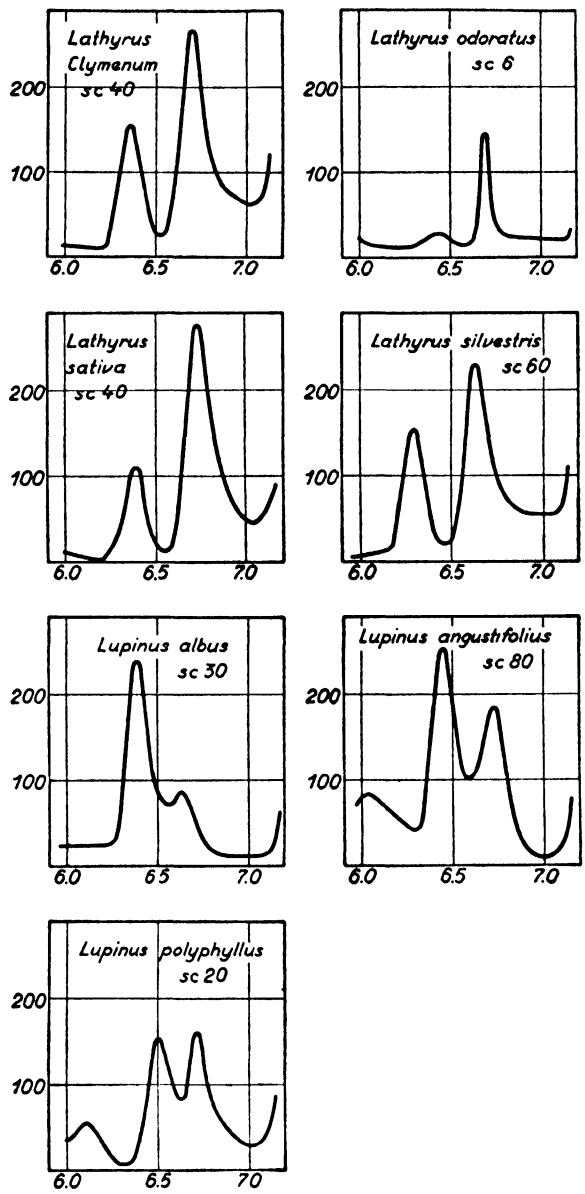


FIG. 7. Sedimentation diagrams showing the composition of globulin solutions prepared by extraction of seeds from different species of the family Leguminosae. (Danielsson, 1949.)

Differential diffusion experiments were carried out on the same sample of insulin with a concentration difference between the upper and lower solutions in the diffusion cell of about 0.25 percent. Three

TABLE 7. MOLECULAR CONSTANTS OF *Helix pomatia* AND *Paludina vivipara* HEMOCYANINS AND THEIR DISSOCIATION PRODUCTS (BROHULT, 1947)

Species	$(s_{20})_0$*	D_{20}	V_{20}	M	f/f_0	L_1†	L_2‡
		10^{-7} cm²/sec				A.	A.
Helix pomatia	103.0	1.07	0.738	8.91×10^6	1.45	1,130	890
	65.7	1.41	0.738	4.31×10^6	1.40	820	890
	19.7	1.77	0.738	1.03×10^6	1.79	820	960
Paludina vivipara	102.5	1.09	0.738	8.70×10^6	1.43	1,090	
	64.5		0.738				
	21.8	1.79	0.738	1.13×10^6	1.72	790	

* $(s_{20})_0$, the sedimentation constant at zero concentration, is expressed in Svedbergs (S).
† Length of the molecules from f/f_0.
‡ Length of the molecules from stream double refraction.

diffusion experiments were carried out; they gave for D_A 7.51, 7.61 and 7.64, or a mean value for D_{20} of $7.59\cdot10^{-7}$. Somewhat lower values were found for D_m, namely, 7.28, 7.31 and 7.28, or as a mean value: $D_{20} = 7.29\cdot10^{-7}$, indicating that the experiments were not as perfect as we would like to have them. For the calculation of the molecular weight, the D_A should be the right value to use.

The specific volume for insulin in the same solutions was kindly determined for us by Professor C. Drucker, who found a value of 0.735. From this value, s_{20} and D_{20} we calculate M = 42,000, which is lower than M = 46,000 found by Miller and Andersson (1942). Part of the discrepancy is due to their use of an old value for V, 0.749. They found s_{20} = 3.55 S and $D_{20} = 7.53\ 10^{-7}$. Gutfreund and Ogston (1946) found lower values for s_{20} as well as

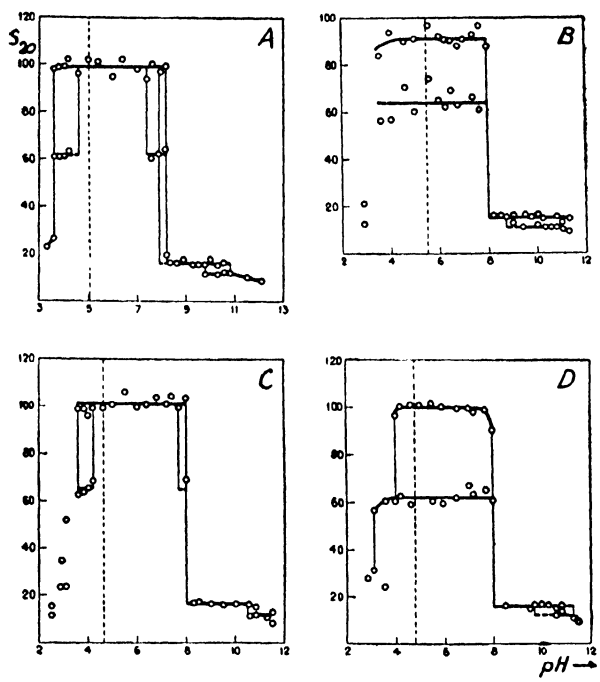


FIG. 8. pH-stability diagrams for hemocyanins of species listed below. Abscissae in all figures, pH; ordinates, s_{20}. The dotted lines indicate the position of the iso-electric points. *A. Helix pomatia; B. Helix arbustorum; C. Helix nemoralis; D. Helix hortensis.* (Ericksson-Quensel and Svedberg, 1936.)

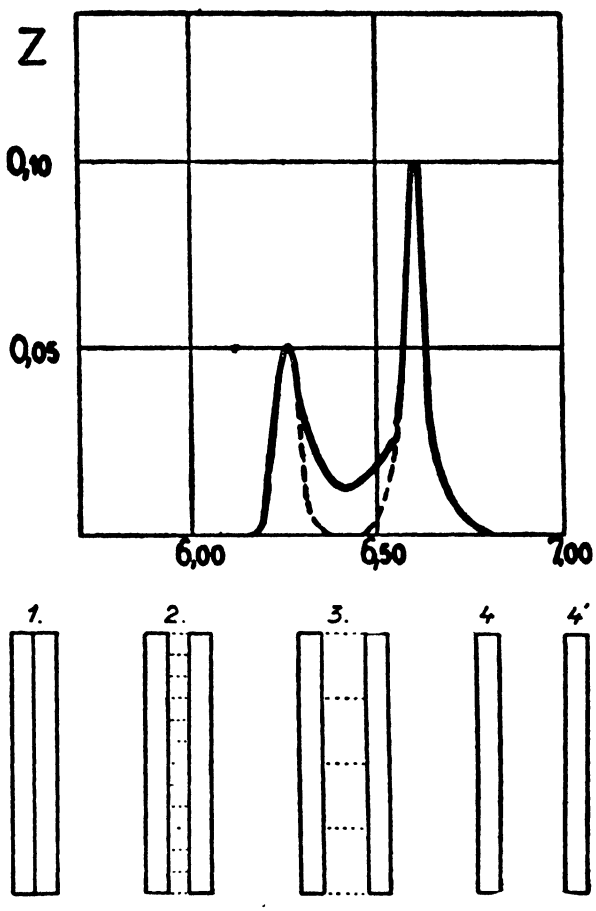


FIG. 9. *Upper part:* Sedimentation diagram by the refraction method, showing the "intermediate compounds." *Lower part:* Swelling of the hemocyanin molecule. *1.* whole molecule; *2* and *3* "intermediate compounds"; *4* and *4'* half molecules. (Brohult, 1947.)

for D_{20}, namely, $s_{20} = 3.35$ S and $D_{20} = 7.0 \cdot 10^{-7}$. Using $V_{20} = 0.749$ they found M = 46,600. Using the new value for V_{20}, we get 43,500 and 44,000, respectively. From osmotic pressure measurement at about the same pH Gutfreund (1948a) recently found M = 48,000 ± 3,000, or a much higher value than 42,000. On the other hand Crowfoot (1938) obtained a lower value for the dry unit cell weight, namely, 37,000 (or 36,000 if a more recent value is used for the water content of air-dried insulin crystals).

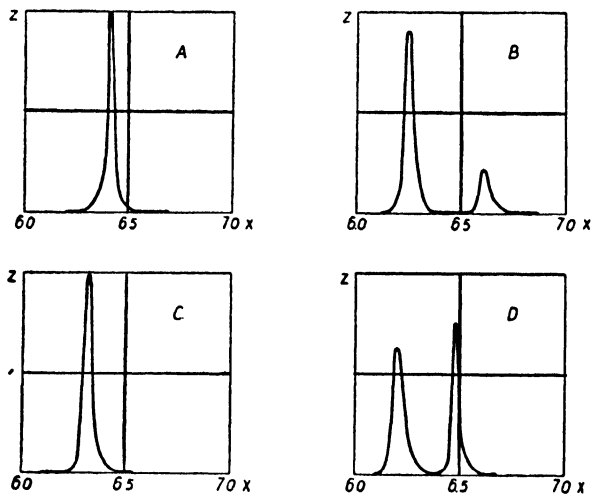


FIG. 10. Sedimentation diagrams showing the separation of the two kinds of hemocyanin molecules. *A:* original hemocyanin; whole molecules; buffer, 0.2 M sodium chloride + 0.08 M acetates; pH, 5.3. *B:* original hemocyanin; maximum dissociation by electrolytes; buffer, 1.00 M sodium chloride + 0.08 M acetates; pH, 5.2. *C:* separated hemocyanin; same buffer as in *B;* the solution contains only half molecules. *D:* solution of the redissolved centrifuge deposit; same buffer as in *B.* Concentration of the hemocyanin about 0.3 percent. (Brohult, 1947.)

The insulin used in our experiments was from a commercial preparation of Zn-insulin; it had kindly been placed at our disposal by Dr. J. Lens, N. V. Organon, Oss, Holland. It was recrystallized twice before being used in our experiments.

At a pH of 2.55 insulin, in a concentration of 0.5 to 1.0 percent, is almost immediately dissociated to half molecules, if the concentration of neutral salt is low. An increase in the concentration of sodium chloride to 0.2 molar results in a s_{20} only a little lower than the normal value $s_{20} = 3.46$.

In the pH region studied, the electrical charge on the insulin molecules greatly influences the movement of this protein. The sedimentation is decreased, and the diffusion increased. By using the experimentally found values for s_{20} and D_{20}, one would therefore find too low a value for M. This difficulty may be overcome, however, by plotting $1/\varkappa$ versus s_{20} or D_{20}, respectively, where $\varkappa$ is the conductivity of the insulin solution. Extrapolating to $1/\varkappa = 0$ from the values obtained for s_{20} or D_{20} results in the "true" value for s_{20} and D_{20}. In our experiments we found the following "true" values: $s_{20} = 1.95$ S and $D_{20} = 7.3 \cdot 10^{-7}$. Using $V_{20} = 0.735$ we obtain M = 24,500. We would probably have found M ≃ 21,000, if our accuracy had been greater.

Gutfreund (1948b) has recently, from osmotic pressure measurements, found that at an insulin concentration of 0.4 to 0.9 percent a maximum dissociation takes place between pH 2 and 3. His M versus c curve extrapolates for c = 0 to a value for M somewhat below 20,000. The experimental determinations made at concentrations below 0.4 percent actually extrapolate to M ≃ 12,000. If this is true, insulin should split to molecules having a quarter of the size of those existing around pH 7. On the other hand the X-ray investigation of Crowfoot showed that the unit cell of crystalline insulin has trigonal symmetry. One should therefore, from her values, expect a subunit of about 12,000. The English workers have explained this by assuming that insulin crystals contain three subunits, whereas insulin in solution consists of four subunits. I doubt whether this is true, and I think we have a discrepancy that must some day be solved.

At pH 3.0 and low salt concentration, the dissociation is not instantaneous, but takes a certain time. It seems, however, as if even in this case the dissociation results in the formation of half molecules. Toward the acid side the half molecules dominate at least down to a pH of 1 to 2 in the insulin solution, when the acidification has been made either with hydrochloric or phosphoric acids. If, however, acetic acid is present in not too low a concentration, the dissociation seems to go further

TABLE 8. DISSOCIATION BROUGHT ABOUT BY CHANGE IN THE pH AND BY ELECTROLYTES (BROHULT, 1947)

	Species	pH	Electrolytes
I	*Helix pomatia*	Dissociation	Dissociation to 75 percent
	Helix hortensis	Dissociation	Dissociation to 30 percent
	Helix arbustorum	Dissociation	Dissociation to 100 percent
II	*Paludina vivipara*	Dissociation	No dissociation
	Littorina littorea	Dissociation	No dissociation
	Buccinum undatum	Dissociation	No dissociation

TABLE 9. SEDIMENTATION VELOCITY FOR CRYSTALLINE INSULIN AT pH 7.6 AND DIFFERENT CONCENTRATIONS

c	2.368	1.186	0.795	0.586	0.393	(0)
s_{20}	3.46	3.44	3.49	3.42	3.47	(3.46)

and no measurable sedimentation can be observed in the centrifuge, neither by optical nor by analytical means. If such a solution is placed in a dialysis bag, the Kjeldahl nitrogen dialyses out rather quickly. The change produced by acetic acid seems to be reversible according to our experiments, at least as regards the size of the molecules. We dissolved insulin in glacial acetic acid and could not observe any sedimentation in the ultracentrifuge. The acetic acid was then removed by evaporation at room temperature, and the insulin was extracted with secondary sodium phosphate. The first insulin extract gave $s_{20} = 3.49$ S and the last $s_{20} = 3.42$ S, that is, within the normal range for insulin.

Similar experiments with bovine serum albumin showed that s_{20} (reduced to water at 20°C) was only one third of the normal, whereas D_{20} was about normal. The change is evidently less pronounced with the serum albumin than with the insulin. On the other hand the reversibility of the change as judged from the sedimentation diagrams, seems to be less good in the case of the serum albumin, where after the removal of the acetic acid, the albumin solution showed two distinct peaks in the sedimentation diagrams, corresponding to single and double molecules.

TABLE 10. DIFFERENTIAL DIFFUSION COEFFICIENTS (D_A) FOR A PATHOLOGICAL γ-GLOBULIN OBTAINED WITH A CONCENTRATION DIFFERENCE BETWEEN LOWER AND UPPER SOLUTION OF ABOUT 0.2%

$\Delta n \cdot 10^5$ for lower solution	54	695	1,054	1,172
D_{20}	3.58	2.27	1.45	1.17

In the presence of sodium thiocyanide in not too low a concentration insulin may be dissolved close to the isoelectric point. It was found, however, that the insulin in these solutions was more or less associated.

A dissociation and association reaction regulated by the temperature and protein concentration is at present being studied in Upsala. The reaction was observed with a pathological globulin from a myeloma serum placed at my disposal by Dr. L. W. Janssen, Amsterdam. On cooling, the serum separated into two layers, the lower one being an almost pure concentrated globulin solution. By repeated dissolution and cooling the globulin was further purified.

The globulin was investigated at different concentrations in order to get the value for s_{20} at zero concentration. Contrary to expectation I found that there was no straight line relationship beween s_{20} and concentration (Δn), but the curve showed a definite maximum (Fig. 11). Extrapolation of the curve from the lower concentration gave for $s_{20} \simeq$ 7 S, whereas extrapolation from the curve for the higher concentration gave $s_{20} \sim$ 11 S. The peaks on the sedimentation diagrams were most symmetrical at the very low and very high concentrations, and asymmetrical around the values where the s_{20} *versus* Δn curve showed the maximum (Fig. 12).

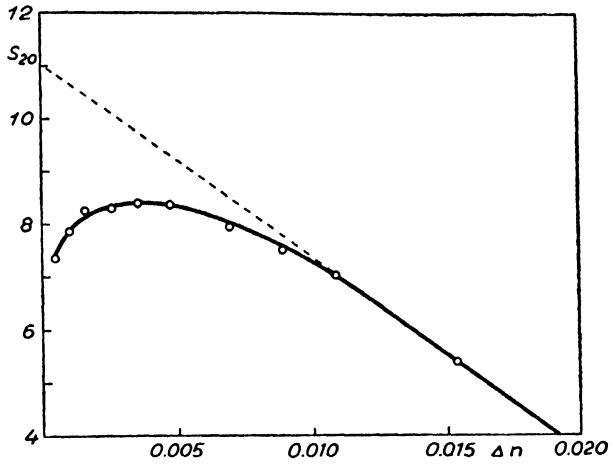


FIG. 11. Variation of the sedimentation velocity with protein concentration for a pathological serum γ-globulin. ($\Delta n = 0.002$ corresponds to a protein concentration of about 1%).

Differential diffusion experiments with a concentration difference of about 0.2 percent globulin between upper and lower compartment of the diffusion cell showed a very pronounced concentration dependence (Table 10).

Sedimentation runs have been carried out at dif-

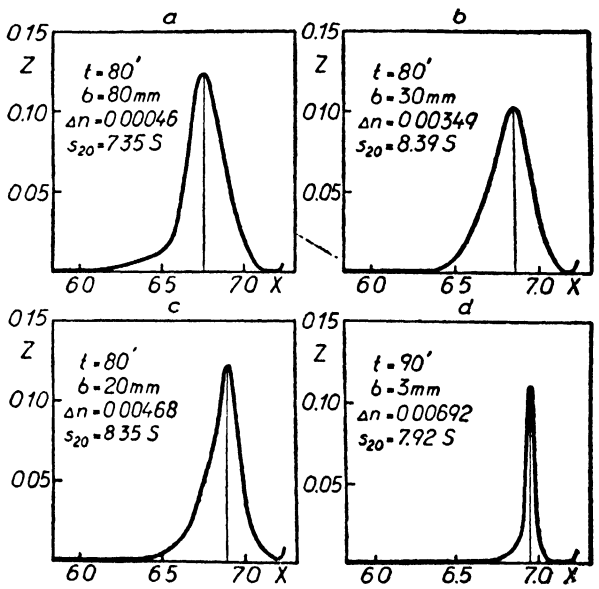


FIG. 12. Sedimentation diagrams from a pathological serum γ-globulin.

ferent temperatures, 10°, 20°, 30°, in the electrically driven ultracentrifuges. It was found that at 10°C the values down to at least 1 percent protein followed the extrapolated curve ($s^{\circ}_{20} \cong 11$ S), whereas at 30° the s_{20} values were situated considerably below the fully drawn curve.

Ten years ago the author started an investigation on the pH-stability range for bovine CO-hemoglo-

TABLE 11. SEDIMENTATION VELOCITY FOR VARIOUS PREPARATIONS OF BOVINE CO-HEMOGLOBIN

3% hemoglobin dissolved in {0.0666-M KH_2PO_4 / 0.2000-M NaCl

Prep. No.	Date of preparation	s_{20}
I	Apr. 1, 1948	4.31 (4.25–4.36)
II	Sept. 13, 1948	4.20 (4.20, 4.21)
III	Sept. 28, 1948	3.99
IV	Oct. 19, 1948	3.95 (3.93, 3.99)
V	Oct. 30, 1948	4.13
VI	Jan. 24, 1949	4.15
VII	Apr. 11, 1949	4.16

bin. Special attention should be put on the influence of protein and salt concentration and on the effect of temperature. The results obtained seemed at first to be very conflicting, but it was soon discovered that the irregular variation in s_{20} depended upon the temperature at which the hemoglobin solution had been kept immediately before being run in the ultracentrifuge. It was thus found that aliquot of the very same 3 percent CO-hemoglobin solution showed $s_{20} \sim 4.1$ S just after it was prepared. When it was left over night at 40°C and run immediately after, it showed $s_{20} \sim 4.7$ S, but if it was then kept for some days at 4 to 5°, s_{20} went down to $\sim$ 3.9 S. The change in s_{20} was completely reversible, and it could be changed up and down at will. As in the case of the dissociation of the hemocyanins, the nature of the neutral salts present played an important rôle, and the effect was most pronounced in about 0.2 molar sodium chloride.

Attempts were made to find similar effects by other proteins, but in vain. The problem was put aside during the war, until last year it was taken up again by one of my collaborators, H. G. Boman. In the first preparations we made, a small effect was observed, but it could not be reproduced in later preparations. Boman has made a total of seven different preparations, all in the same way. The solutions were all studied at a concentration of 3 percent hemoglobin in the same mixture of sodium chloride and primary potassium phosphate, but only the first preparation showed a definite temperature effect. On the other hand, it was found that the s_{20} observed for the various 3 percent solutions were not the same, but varied from preparation to preparation (Table 11).

For some of the later preparations, Boman determined the variation of s_{20} with the hemoglobin concentration (Fig. 13). For preparation VI it was found that the points in the diagram were scattered around a straight line starting from $s_{20} \sim 4.6$ at zero concentration. The same was the case with Preparation VII, where the results were not ready until after Figure 13 was made. Here the line of regression was found to be $s_{20} = 4.58 - 0.171$ c. For Preparation V the result was somewhat different, as s_{20} at concentrations below 2 percent were situated below the straight line through the points corresponding to the higher concentrations. The line of regression for the measurements from 2 percent and upwards was found to be $s_{20} = 4.86 - 0.235$ c, or rather different from the other two preparations. The variation in s_{20} for the different preparations is greater than could be expected from the experimental error. We are inclined therefore to think that it must be due to a difference in the hemoglobin molecule in the different preparations.

From time to time differences have been observed between various preparations of well-defined proteins. It has thus been found that fetuin, a globulin present as well in serum from cow's fetus as in serum from newly born calves, may differ quite

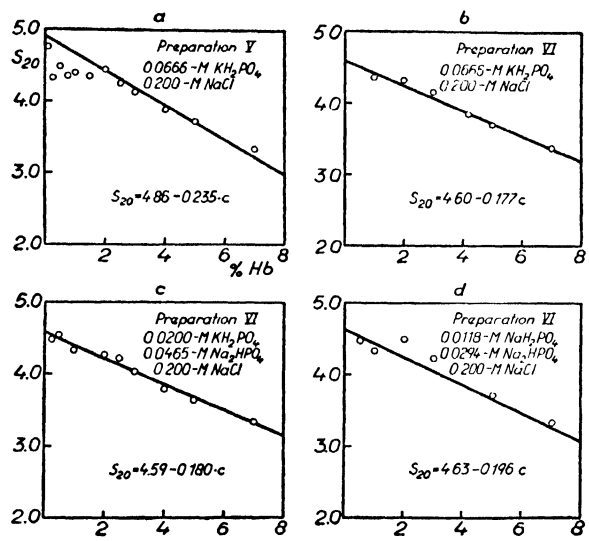


FIG. 13. Variation of the sedimentation velocity with protein concentration for two preparations of bovine CO-hemoglobin (H. G. Boman).

considerably (Pedersen, 1945, 1947). The different preparations all give about the same M, but s_{20} varied between 3.09 S and 3.28 S, D_{20} between 5.0 and $5.5 \cdot 10^{-7}$, V_{20} between 0.692 and 0.714, the nitrogen content between 12.3 g and 13.4 g per 100 g protein, and finally the phosphorus content between 188 mg and 101 mg per 100 g protein. In this case the variation in the protein may perhaps be due to a difference in the composition of the

fetuin present in the early fetus and in the fully developed fetus or the newly born calf.

For many years, β-lactoglobulin was considered a prototype of a homogeneous well-defined protein. It was therefore quite unexpected when Li (1946) found that crystalline β-lactoglobulin consisted of two or three major components. Later on McMeekin *et al.* (1948) by fractionated crystallization succeeded in obtaining fractions with marked differences in solubilities. Recently Jacobsen (1949) has published a very careful investigation of various preparations of β-lactoglobulin made at different times at the Carlsberg laboratory. Two methods had mainly been used for the preparation, but Jacobsen found that when they were used on the same batch of milk, the two lactoglobulins were identical within the experimental error. When, however, β-lactoglobulin prepared in 1938 was compared with β-lactoglobulin prepared 1947, the two preparations were quite different. The solubility was much higher for the 1947 protein than for the 1938 one. The number of amide groups per 42,000 grams was 41 for the 1938 protein, but only 33 for the 1947 lactoglobulin. The number of basic groups was the same for the two preparations, but the carboxyl groups minus argine were 45 and 52, respectively. Jacobsen concluded that, unquestionably, β-lactoglobulin is not of constant composition.

If these findings can be verified and found in other proteins too, we will have to take them into serious consideration. It would then mean that we can only expect to find homogeneous proteins, if they have been prepared from a single individual. It may even be that such a protein may vary slightly when obtained at different occasions.

If we look at the problem as a whole it seems as if similar proteins, *i.e.* proteins that occur in the same way or have the same function, will either have about the same molecular weight or they will usually belong to the same multiple system.

If, however, we look into the problem in more detail, there seems to be some reason to believe that the same protein obtained from different individuals within the same species may vary within certain limits. When a given protein has been obtained from pooled material, as it has to be in most cases, we must expect that its properties will be statistically distributed around a certain average value. If the distribution is Gaussian and narrow, I think we may be content and for most purposes consider our protein to be "pure." On the other hand we cannot expect that the average values obtained on different preparations will be the same.

References

Brohult, S., 1947, Hemocyanins of the Gastropods. J. Phys. Colloid Chem. *51:* 206-217.

Bull, H. B., 1941, Protein structure. Advances in Enzymol. *1:* 1-42.

Crowfoot, D., 1938, The crystal structure of insulin. Proc. Royal Soc. (London) A *164:* 580-602.

Danielsson, C. E., 1949, Seed globulins of the Gramineae and Leguminosae. Biochem. J. *44:* 387-400.

Eriksson-Quensel, I.-B., and Svedberg, T., 1936, The molecular weights and pH-stability regions of the hemocyanins. Biol. Bull., Woods Hole, *71:* 498-547.

Gutfreund, H., 1948a, The osmotic pressure of insulin solutions. Biochem. J. *42:* 156-160.

1948b, The molecular weight of insulin and its dependence upon pH, concentration and temperature. Biochem. J. *42:* 554-48.

Gutfreund, H., and Ogston, A. G., 1946, The sedimentation constant of insulin. Biochem. J. *40:* 432-434.

Jacobsen, C. F., 1949, A comparison of different β-lactoglobulin samples. C. R. Lab. Carlsberg, Sér. chim. *26:* (no. 18): 455-75.

Johnston, J. P., Longuet-Higgins, H. C., and Ogston, A. G., 1945, On the distribution of the molecular weights of proteins. Trans. Faraday Soc. *41:* 588-593.

Li, C. H., 1946, Electrophoretic inhomogeneity of crystalline β-lactoglobulin. J. Amer. Chem. Soc. *69:* 2746-2747.

McMeekin, T. L., Polis, B. D., Della Monica, E. S., and Coster, J. H., 1948, Heterogeneity of crystalline beta-lactoglobulin. J. Amer. Chem. Soc. *70:* 881-882.

Miller, G. L., and Andersson, K. J. I., 1942, The molecular weight of insulin. J. biol. Chem. *144:* 459-464.

Norris, A. D., 1946, Molecular-weight distribution of proteins. Nature, Lond. *157:* 408.

Pedersen, K. O., 1933, Studien über isoelektrische Punkte von Eiweisskörpern, I. Kolloid-Z. *63:* 268-277.

1940, in Svedberg, T., and Pedersen, K. O., The Ultracentrifuge. Oxford: The Clarendon Press: 355-383.

1945, Ultracentrifugal studies on serum and serum fractions. Upsala: Almqvist and Wiksell.

1947, Ultracentrifugal and electrophoretical studies on fetuin. J. Phys. Colloid Chem. *51:* 164-171.

Rothen, A., 1942, The present status of molecular weights of proteins. Ann. N.Y. Acad. Sci. *43:* 229-241.

Strömbäck, O., and Brohult, S., 1949, Private communication.

Svedberg, T., 1926, Uber die Bestimmung von Molekulargewichten durch Zentrifugierung. Z. phys. Chem. *121:* 65-77.

Svedberg, T., and Ericksson-Quensel, I.-B., 1934, The molecular weight of erythrocruorin II. J. Amer. Chem. Soc. *56:* 1700-1706.

Svedberg, T., and Fåhraeus, R., 1926, A new method for the determination of the molecular weight of the proteins. J. Amer. Chem. Soc. *48:* 430-438.

Svedberg, T., and Hedenius, A., 1934, The sedimentation constants of the respiratory proteins. Biol. Bull., Woods Hole *66:* 191-223.

Säverborn, S., Danielsson, C. E., and Svedberg, T., 1944, On the globulins of the cereals and malt. Svensk Kem. Tid. *56:* 75-85.

Discussion

Anson: Dr. Cannan has asked for the latest results in the study of quantitative adsorption analysis of proteins whose adsorption is promoted by salt, as described by Tiselius. Both Dr. Pedersen and Dr. Stein have replied that such quantitive adsorption analysis of proteins has as yet been quite unsuccessful. I should like to point out that, altho the quantitative adsorption analysis of proteins is not yet perfected, nevertheless even a crude application of salt-

promoted adsorption can be very useful in the fractionation of proteins.

In the first step of the purification of spleen cathepsin, the spleen material is suspended in acid 0.3 saturated ammonium sulfate solution at 45° C (Anson, 1940, J. gen. Physiol. *23*:695). Vigorous autolysis takes place under these conditions, with digestion of much non-cathepsin protein. The cathepsin protein, however, which is adsorbed onto the solid material of the spleen survives. After the spleen solids with their adsorbed cathepsin have been separated from much soluble and solubilized material, the cathepsin is eluted by an amount of alkali which removes only a fraction of the total elutable adsorbed protein.

Eluted cathepsin, in the absence of spleen solids, is soluble in acid 0.3 saturated ammonium sulfate solution and is unstable in such solution at 45° C. Furthermore, altho cathepsin is completely adsorbed onto spleen solids in the 0.3 saturated ammonium sulfate solution, the adsorption is less complete when the concentration of ammonium sulfate is lower and cathepsin adsorbed onto spleen solids in 0.3 saturated ammonium sulfate solution can be washed off the spleen solids with water. Thus the very simple and very effective first step in the purification of cathepsin is based on salt-promoted adsorption, and it is to be expected that many other applications of salt-promoted adsorption to the practical fractionation of proteins can be found.

I would like to ask if Dr. Pedersen has calculated the heat of dissociation from the effect of temperature on the dissociation equilibrium.

PEDERSEN: As to the last question, I would like to say that we still know too little about the change in the dissociation equilibrium with temperature to make an exact calculation of the heat of dissociation. I wonder whether it is not a little when Gutfreund (1948a) calculates the energies involved in the dissociation of insulin from his osmotic pressure measurements.

(Note added in proof): During recent months there has been definite progress in the work on that kind of adsorption analysis of proteins, where the adsorption is promoted by salt. This progress has been reported in a paper in the Faraday Society Symposium on Chromatographic Analysis by Tiselius and Shepard, "Chromatography of Proteins. Effects of Salt-concentration and pH an Adsorption."

BUTLER: Is it possible that the large differences of specific volume which have been observed are due to differences in the degree of dryness of the insulin? Some of the water of insulin is very firmly held. We have found that when the breakdown products of insulin are thoroughly dried they become insoluble and will no longer dissolve in the buffer they were prepared in. This seems to be possibly due to the formation of intramolecular salt links between sidechains, which were previously separated by water molecules. The removal or partial removal of water from insulin may cause disorientations which are only slowly recovered from.

HUGHES: Oncley and Ellenbogen in our laboratory have recently carried out detailed physicochemical studies on insulin on both sides of the isoelectric point; I believe they obtained experimental results in accord with Pedersen's. However, their interpretation was quite different. They assumed the basic insulin unit of 12,000 to be in rapid equilibrium with its polymers, the degree of association depending on the net electrical charge and on the concentration of the insulin. This interpretation could rather satisfactorily explain the effects of pH, ionic strength and insulin concentration on the sedimentation diagram. However, it failed to predict the observed boundary spreading. (Eric Ellenbogen, 1949, "The determination of the physical-chemical properties of insulin and their application to the equilibrium between insulin of molecular weights 12,000 and 36,000," Ph.D. thesis, Harvard University.)

Experimental evidence that the monomer-polymer equilibria are rapid has been obtained from a limited number of light-scattering measurements by P. Doty and B. Rabinovitch in the Chemistry Department at Harvard University (and independently in our laboratory by H. Edelhoch). Measurements on the acid side of the iso-electric point have shown that the turbidity (a measure of the equilibrium distribution) changes to the new equilibrium value within one minute upon altering the insulin concentration. No further drift in turbidity was then observed during the next 24 hours.

PEDERSEN: The great differences found for the specific volume of insulin at Harvard (V 0.70) and at Upsala (V 0.735) has worried me greatly, but I cannot explain it. We usually agree fairly well in our specific volume determinations. As far as the nitrogen content of insulin is concerned, there seem, however, also to be very great differences between the results from different laboratories.

RANDALL: I should like to know to what extent the work of Wyckoff on hemocyanin in the electron microscope is in agreement with the ultracentrifuge results described in this paper.

PEDERSEN: As far as I remember there is qualitative, but I do not think there is quantitative agreement. Wyckoff has used the hemocyanin from *Busycon canaliculatum* for his investigations and Brohult has mainly used the hemocyanins from *Helix pomatia* for his detailed studies on the shape of these molecules.

TAYLOR: The partial specific volume of a protein can be calculated from the amino acid composition and from the partial specific volumes of the amino acid residues. Preliminary calculations for insulin at 25° yield 0.715 from the data of Chibnall and 0.725 from the data of Brand, without including a contribution for zinc.

SOME CHEMICAL INVESTIGATIONS ON THE STRUCTURE OF INSULIN

F. SANGER
Beit Memorial Fellow, Biochemical Laboratory, Cambridge, England

The problem of protein structure has been studied by two main types of methods. Specialized physicochemical techniques have given much information about the size and shape of molecules, and the newer analytical methods have made possible the accurate determination of the component amino-acids. Very little is known, however, about the order in which these amino-acids are arranged to form the whole protein molecule. This problem can probably best be attacked by the methods of organic chemistry, that is to say, by the gradual and systematic breakdown of these molecules and investigation of the structure of the products of partial degradation. Such an approach was largely hampered in the past by the extreme complexity of any mixture produced by the partial hydrolysis of a protein. The introduction of new chromatographic methods with a very high resolving power has made possible the reinvestigation of this problem, and in this paper an account is given of a study of the structure of insulin from the point of view of the organic chemist.

It is now well established that, except for a small amount of loosely-bound zinc, insulin is built up solely of amino-acids, and that it contains no special non-amino-acid prosthetic group which is responsible for its physiological activity. These amino-acid residues are joined together in peptide chains and there is no evidence to suggest that there are any other primary bonds between the amino-acids, except the -S-S- bridges of the cystine residues, of which there are more in insulin than in most other proteins.

The figure for the molecular weight of insulin that agrees best with the various published results is probably 12,000 (Gutfreund, 1948). In crystals of insulin three such molecules are combined to form the unit cell, and in relatively concentrated solutions aggregates are formed which probably contain four molecules. On dilution and at extreme pH values these aggregates dissociate. For the purposes of the following discussion the value of 12,000 will be assumed for the molecular weight.

Insulin contains all the commonly occurring amino-acids except methionine, tryptophan and hydroxyproline. Recently, two careful and complete analyses have been reported, one by Tristram (1946), Macpherson (1946) and Rees (1946) at Cambridge, and the other by Brand (1946). These workers have made use of several different methods for estimating each amino-acid and the figures should give an accurate picture of the over-all composition of the protein. The two sets of results are shown in Table 1, expressed as the number of residues of each amino-acid per molecule of molecular weight 12,000. For most amino-acids there is reasonable agreement, though it is difficult to understand the differences in the figures for threonine and amide-N, which should be accurate to at least five percent. These differences do suggest the possibility that all preparations may not, in fact, be completely identical.

TABLE 1. AMINO-ACID COMPOSITION OF INSULIN
(Results expressed as number of residues per molecule of insulin)

Amino-acid	Cambridge workers	Brand
Glycine	6.9	7.4
Alanine	6.8	—
Valine	8.0	9.1
Leucine	12.0	12.3
Isoleucine	2.6	2.7
Proline	2.7	3.1
Phenylalanine	5.9	5.7
Tyrosine	8.7	8.2
Glutamic acid	15.0	16.4
Aspartic acid	—	6.1
Serine	6.0	6.6
Threonine	2.1	3.2
Cystine/2	12.5	11.0
Arginine	2.1	2.4
Histidine	3.8	4.1
Lysine	2.1	2.1
Amide	11.6	15.1

TERMINAL RESIDUES OF INSULIN

Insulin differs from most other proteins in that it has a relatively high content of free α-amino groups (Chibnall, 1942). That at least one of these was located on a phenylalanine residue had been shown by Jensen and Evans (1935), who treated insulin with phenylisocyanate and isolated the phenylhydrantoin of phenylalanine from the hydrolysate of the treated insulin. In order to obtain more detailed quantitative information about the free amino groups of insulin and other proteins, a general method was worked out using the reagent 1:2:4-fluorodinitrobenzene. This reagent reacts quantitatively with the free amino groups of proteins to form a dinitrophenyl- (or DNP-) protein. The reaction is carried out under mild conditions,

which do not bring about any hydrolysis of peptide bonds. The DNP-protein is then hydrolysed with strong acid, which splits all the peptide bonds present and liberates the DNP-derivatives of the amino-acids which bear the free amino groups in the original protein. These DNP-amino-acids, which are relatively stable to acid, are bright yellow compounds and can be extracted into an organic solvent and successfully fractionated by chromatography on silica gel. They are identified by their behaviour on such columns, and can be estimated colorimetrically.

Besides reacting with the free α-amino groups, fluorodinitrobenzene reacts with the ε-amino groups of the lysine residues, giving rise to ε-DNP-lysine, which can be estimated in the hydrolysate, indicating whether all the ε-amino groups were reactive in the protein. It also reacts with the phenolic hydroxyl groups of the tyrosine residues and the imidazole groups of the histidine residues (Porter, unpublished). However, the products formed are colorless and thus do not interfere with the estimations.

The DNP method has been found to be generally applicable to proteins and peptides whatever the nature of the terminal residues, and to give reasonably accurate results. The chief limitation to the accuracy of the method is probably in the correction factor that has to be applied for the breakdown of the DNP-amino-acids on hydrolysis, since they are not completely stable to acid. This factor varies with the different amino-acids.

When insulin was investigated, three colored products were found in the hydrolysate of the DNP-derivative and these were identified chromatographically as DNP-phenylalanine, DNP-glycine and ε-DNP-lysine. On estimation it was found that there were two of each of these residues per insulin molecule, indicating that there are four free α-amino groups, two on glycyl residues and two on phenylalanyl residues. The presence of two molecules of ε-DNP-lysine demonstrated that both the ε-amino groups of the lysyl residues are free in insulin.

The number of free α-amino groups in a protein is equal to the number of open polypeptide chains present, that is to say chains with a free amino group at one end and a free carboxyl group at the other end. Insulin with four free α-amino groups has thus four open polypeptide chains. It is possible that there may be other types of chains, such as cyclic chains, which would not be detected by this method. However, since there was no evidence on this point, it was tentatively concluded that insulin was composed only of four chains and that these chains were joined together by the -S-S- bridges of the cystine residues, this being the only type of cross-linkage that is definitely known to exist in proteins.

The Splitting of Insulin by Oxidation

If this is indeed the structure of the insulin molecule it should be possible to split it into its separate peptide chains by breaking the -S-S- bridges. It was shown by Toennies and Homiller (1942) that cystine can be quantitatively converted to cysteic acid by oxidation with performic acid and that the only other commonly occurring amino-acids oxidized are tryptophan and methionine. Since insulin contains neither of these amino-acids treatment with performic acid should break the -S-S- bridges to give sulphonic acid groups without destroying any other part of the molecule. From insulin that had been oxidized in this way, two main fractions could be separated, an acid fraction A, and a basic fraction B. Fraction A contained only glycyl terminal residues and had no arginine, lysine, histidine, phenylalanine, threonine or proline. Determination of the amount of terminal glycyl residues by the DNP technique indicated an average molecular weight per chain of 2,900, which agreed well with the figure obtained by Gutfreund and Ogston (1949) using the ultracentrifuge. This corresponds to a weight of 2,700 for the individual chains in the intact insulin, the difference being due to the oxygen taken up during treatment with performic acid. The yield of fraction A was 30 to 40 percent of the original insulin, indicating that both the glycyl chains of the original insulin were present in this fraction. Many attempts, however, to separate two such chains were unsuccessful and it seemed probable that both the chains were in fact identical or almost identical.

Fraction B, the other main fraction of the oxidized insulin, contained all the amino-acids that occur in insulin and had phenylalanyl terminal residues. Estimation of these terminal residues indicated a molecular weight per chain of 3,800, whereas a value of 7,000 was obtained using the ultracentrifuge. This suggests that either the two phenylalanine peptides are joined together by some unoxidizable bond or that in solution aggregates are formed containing two molecules. In view of the known aggregation reactions of insulin and since no unoxidizable cross-linkage is known, the latter explanation seems the most probable. While the yield of pure fraction B was usually about 25 percent of the original insulin, no evidence of a second phenylalanyl chain could be obtained by fractionating the residues from the oxidized insulin, which appeared to consist of mixtures of the above two fractions and incompletely oxidized material.

Terminal Peptides of Insulin

By means of the DNP technique it was possible to determine the nature of the terminal residues by identification of the DNP-amino-acids produced on hydrolysis of the DNP-protein. If, however, the DNP-protein is only partially hydrolysed one obtains DNP-peptides, and an investigation of their structure reveals the nature of the amino-acids that occupy positions in the peptide chains near the terminal residues. The DNP-peptides can be largely separated from other unsubstituted peptides and

amino-acids by extraction into an organic solvent from acid solution and can be fractionated by chromatography on silica gel. The peptide mixtures produced by this method are very much less complex than those produced by the partial hydrolysis of an untreated protein, since only those peptides that contain the terminal residues are separated out. In order to simplify the peptide mixture still further, the purified fractions A and B of the oxidized insulin were used, since each of these contained only one terminal residue.

Fraction B

When the DNP derivative of fraction B (referred to as DNP-*B*) was hydrolysed for 8 days in 12N HCl at 37° four main bands were detected by chromatography of the ethyl acetate extract of the hydrolysate. Each of these was completely hydrolysed and gave rise to DNP-phenylalanine, which was identified by chromatography on silica gel, and the amino-acids shown in Table 2, which were

TABLE 2. AMINO-ACID COMPOSITION OF DNP-PHENYLALANYL PEPTIDES

Band	Amino-acids
B1 (DNP-phenylalanine)	None
B2	Valine
B3	Valine, aspartic acid
B4	Valine, aspartic acid, glutamic acid

identified by paper chromatography (Consden, Gordon and Martin, 1944). Quantitative paper chromatographic analysis (Polson, Mosley and Wyckoff, 1947) showed that for each peptide there was one residue of each amino-acid present per residue of DNP-phenylalanine. The increasing complexity of each of these peptides suggested that they were all breakdown products of the same peptide chain, containing the sequence DNP-phenylalanyl-valyl-aspartyl-glutamic acid, and this was confirmed by subjecting them to a further partial hydrolysis. Thus band B3 broke down to give band B2 and band B4 to give bands B3 and B2. A number of fainter bands were also present in the hydrolysate of DNP-*B*, but on further hydrolysis they broke down to the above three peptides, indicating that they were higher peptides from the same chain and no DNP-phenylalanyl peptides have been detected in fraction B that do not fit in with the above sequence. When a shorter time of hydrolysis was used a number of peptides were produced which contained the same amino-acids as B3 and B4 in the same order. These appeared to be amides, indicating that in the protein the aspartic and glutamic acid residues in these positions are in the form of asparagine and glutamine.

In the aqueous solution of the partial hydrolysate after extraction with ethyl acetate, a number of colored peptides containing ε-DNP-lysine were present. These could be fractionated very successfully from one another by chromatography but were usually contaminated with colorless peptides containing no DNP group. These, however, can be removed by making use of an adsorption chromatogram. The DNP compounds are strongly adsorbed on talc from acid solution whereas simple peptides and amino-acids pass rapidly through such a column. The composition of the four main bands are shown in Table 3. The terminal residues of the peptides were determined by the DNP method and by deamination with nitrosyl chloride (Consden, Gordon and Martin, 1947). On further hydrolysis of L4 the peptides L2 and L3 were both produced. From this and from the terminal residues only one structure for the peptide L4 was possible; that is threonyl-prolyl-(ε-DNP)-lysyl-alanine. No other peptides containing ε-DNP-lysine were detected which did not fit into the above sequence.

TABLE 3. AMINO ACID COMPOSITION OF PEPTIDES CONTAINING ε-DNP-LYSINE

Band	Amino-acids	Terminal residue
L1	ε-DNP-lysine	
L2	ε-DNP-lysine, alanine	ε-DNP-lysine
L3	ε-DNP-lysine, threonine, proline	Threonine
L4	ε-DNP-lysine, threonine, proline, alanine	Threonine

These results showed that the peptides phenylalanyl-valyl-aspartyl-glutamic acid and threonyl-prolyl-lysyl-alanine are present in insulin and the next problem was to estimate them, that is to say, to determine how many of the lysyl residues and the terminal phenylalanyl residues are present in the form of the above peptides. When DNP-insulin was hydrolysed for 8 days with 12N HCl at 37° only traces of DNP-glycyl peptides were present and the peptides containing DNP-phenylalanine and ε-DNP-lysine could be directly estimated colorimetrically. The results of such an analysis of the phenylalanyl peptides are shown in Table 4, expressed as moles of DNP-peptide as percent of the total terminal phenylalanyl residues present in the insulin. It can be seen that 92 percent of the total terminal phenylalanyl residues can be accounted for as compounds fitting in to the sequence phenylalanyl-valyl-aspartyl-glutamic acid and it would seem evident that in fact all the terminal phenylalanine is combined in this peptide sequence. It is however theoretically possible that the lower peptides (B2, B3) and DNP-phenylalanine are derived from a different sequence than that in B4 and in order to determine the actual amount of the B4 sequence present in DNP-insulin, it was necessary to study the breakdown of peptide B4 under the conditions of hydrolysis used. If the same products were produced from

TABLE 4. YIELDS OF DNP-PHENYLALANYL PEPTIDES FROM DNP-INSULIN

(Results expressed as percent of the total terminal phenylalanine)

Peptide	Yield from DNP-insulin	Yield from B4
B1	13	13
B2	16	14
B3	13	12
B4	30	55
Total known peptides	59	
Other bands giving B4 on hydrolysis	20	
Other uninvestigated bands	6	
Total	98	94

B4 in the same proportions in which they were produced from DNP-insulin, then there could be no doubt that the total amount of the B4 peptide present is equal to the total terminal phenylalanyl residues. In this way the amount of peptide B4 is estimated using a correction factor for the breakdown of the peptide itself on hydrolysis, and since the amounts of breakdown products (B1, B2, B3) produced from B4 are known three more separate estimations of the amount of B4 present can be obtained from the yields of these products from the DNP-insulin. This method of assay assumes that the rate of hydrolysis of each peptide bond is the same in the protein as in the free peptide. While this may not always be strictly true, it is probably a safe approximation, since several different bonds are involved. The yields of the various peptides and DNP-phenylalanine from a corresponding hydrolysate of peptide B4 are shown in the last column of Table 4. It can be seen that within the limits of the method they agree well with the yields obtained from DNP-insulin. The figure for unchanged B4 should be compared with the sum of the yields of B4 and higher peptides from DNP-insulin.

A similar analysis of the peptides containing

TABLE 5. YIELDS OF PEPTIDES CONTAINING ε-DNP-LYSINE FROM DNP-INSULIN

Peptide	Yield from DNP-insulin	Yield from L4
L1	14	14
L2	19	23
L3	32	32
L4	23	21
Total known peptides	74	
Other unidentified peptides	6	
Total	94	90

ε-DNP-lysine is shown in Table 5. The yields are compared with those obtained from a similar hydrolysate of peptide L4. The figures leave little doubt that all the lysine in insulin is present in the form of the peptide threonyl-prolyl-lysyl-alanine.

Fraction A

The investigation of the terminal glycyl peptides in fraction A of the oxidized insulin was rather more difficult due to the great lability of the glycyl peptide bond to concentrated acid. Thus partial hydrolysates of DNP-*A* contained larger amounts of DNP-glycine and very long peptides and only small amounts of short peptides that could readily be studied. The results of a preliminary investigation of the DNP-peptides obtained from DNP-*A* that had been hydrolysed for 24 hours with 12N HCl at 37° are shown in Table 6. The first three peptides fit into the sequence DNP-glycyl-isoleucyl-valyl-glutamic acid and that this was in fact the structure of A4 was shown by partial hydrolysis. Band Ax cannot however fit into this sequence and it was thought that it must represent a second glycyl chain (Sanger, 1948). However when the amino-acids present in this band were estimated it was found that there were about five to six residues of each amino-acid for each residue of DNP-glycine, indicating that the fraction contained a small amount of a DNP-glycyl peptide, which was responsible for the color of the band, and a large amount of a non-terminal peptide containing the four amino-acids. It had been extracted by the ethyl acetate from acid solution and moved on the chromatogram, in contrast to most amino-acids and peptides. The exact nature of this compound has not yet been determined. Thus the only peptides identified in this fraction were the di-, tri- and tetrapeptide corresponding to the sequence glycyl-isoleucyl-valyl-glutamic acid, though these peptides were obtained only in a very poor yield when 12N HCl was used for the hydrolysis. When, however, the hydrolysis was carried out in dilute (0.1N) HCl either at 100° or at 37° for several months, much better yields were obtained and another peptide (A5) was present in considerable proportions. This peptide contained the same amino-acids as A4 but on estimation it was found that there were about two residues of glutamic acid per molecule. On further hydrolysis it gave the other three peptides (A2, A3, A4) and

TABLE 6. AMINO-ACID COMPOSITION OF DNP-GLYCYL PEPTIDES

Band	Amino-acids
A2	Isoleucine
A3	Isoleucine, valine
A4	Isoleucine, valine, glutamic acid
Ax	Glutamic acid, leucine, serine, *O*-DNP-tyrosine

as it contained no amide group the only possible structure was DNP-glycyl-isoleucyl-valyl-glutamyl-glutamic acid. The yields of these peptides from DNP-A are shown in Table 7. Eighty-one percent of the terminal glycyl residues can be accounted for in terms of DNP-derivatives fitting the above sequence and it would seem likely that this is the only terminal sequence present in this fraction. It has not, however, been possible to estimate the amount of these peptides in the intact insulin since the DNP-insulin is too insoluble in the dilute acid used. Fraction A represents about 30 to 40 percent of the original insulin and there are some losses involved in the preparation of the DNP-derivative. While it seems probable that all the losses are incidental and do not involve the separation of a different fraction, it is nevertheless impossible to say whether one or two of the terminal glycyl residues are present in this peptide sequence in insulin.

Discussion and Conclusions

In this work we have demonstrated the presence of three peptide sequences in a hydrolysate of insulin and the question arises as to whether these sequences are actually present in the insulin molecule or whether they could have been synthesised during the degradation procedure. Behrens and Bergmann (1939) have shown that synthesis of peptide bonds can take place under the action of proteolytic enzymes and the theoretical possibility exists that it can also take place under the catalytic action of acids, though this has never been demonstrated experimentally. If any such synthesis has taken place in the present experiments it would have to be a completely specific synthesis, since for each DNP-amino-acid only one sequence of amino-acids has been detected and no peptides were found that would not fit these sequences. Since the equilibrium of the reaction involved is very much on the side of hydrolysis for every peptide bond and since the hydrolytic action of acid is extremely non-specific, all peptide bonds being split to some extent, it would seem that a completely specific synthesis of a peptide bond by acid would be impossible. The syntheses demonstrated by Behrens and Bergmann were brought about by the rapid removal of the product of synthesis either by crystallisation or by subsequent hydrolysis at a different bond. Since in the experiments described here the reaction mixtures were always in solution, only the second mechanism could play a part. Such a reaction may be illustrated by the following formula:

$$AB + CD \rightleftharpoons ABCD \rightarrow ABC + D$$

where A, B, C, D represent amino-acid residues. For synthesis to take place the bond C-D must be much more labile in ABCD than in CD. This might be expected in the case of acid hydrolysis, as the presence of a free α-amino group on the residue C in CD might render the bond more stable than in ABCD. However such an effect would be completely non-specific applying to any peptide and this could not be playing a part in the present results. In the case of hydrolysis with proteolytic enzymes, where the specificity may depend on the nature of the amino-acids not involved in the peptide bond a specific synthesis would seem more possible.

It would thus seem that the peptides detected in the present work are actually part of the insulin molecule and in fact the results would lend support to the assumption that peptide bond synthesis does not occur during the hydrolysis of proteins by strong acid. Such an assumption is the basis of all experiments involving partial hydrolysis and can probably only be proved by a sufficient amount of experimental evidence. Such a conclusion was also reached by Consden, Gordon, Martin and Synge (1947), who determined the sequence of amino-acids in "gramicidin S" by partial hydrolysis.

Table 7. Yields of DNP-Glycyl Peptides from DNP-*A*

Peptide	Yield	
A1 (DNP-glycine)		21
A2	1	
A3	5	
A4	20	
A5	34	
	—	
Total known peptides		60
Other unidentified peptides		12
		—
Total		93

It has been shown that all the terminal phenylalanyl residues in the insulin molecule are combined in a single tetrapeptide sequence and that all the lysine residues, which are in the same chains are present in another tetrapeptide. Thus the amino-acids occupying eight of the positions relative to the terminal amino groups and the lysine residues are identical in the two phenylalanyl chains and it would seem most probable from this that in fact the two chains are themselves identical. It is of course possible that they might differ only in one or in a few amino-acid residues in a part of the chain not yet investigated. While nothing is known about the principles that govern the arrangement of amino-acid residues in a protein, it would seem remarkable if the two chains were not identical, when eight out of thirty residues are the same.

The results with the glycyl chains are not so clear cut as with the phenylalanyl chains. We have demonstrated the presence of a single peptide sequence and could obtain no evidence of any peptides from another sequence. However, using a similar method Woolley (1948) has reported the identification of two other peptide sequences. It is difficult to understand why none of these peptides has

been detected in the present work, since for those sequences that have been detected, all possible intermediate peptides were present. We have also examined the terminal peptides present in the other smaller fractions of the oxidised insulin, but have found no peptides other than those already described. If there is indeed a second glycyl chain

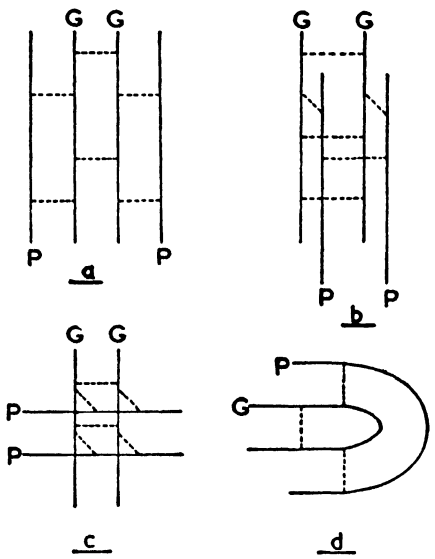


FIG. 1. Possible structures for the insulin molecule.

then the two chains must have the following properties in common. They must both contain no arginine, lysine, histidine, phenylalanine, threonine or proline, since fraction A which contains both chains, has none of these amino-acids, and they must have the same electrophoretic mobility and sedimentation constant. The difficulty of separating them is further evidence of their similarity.

The possibility was considered that besides the four open polypeptide chains, the insulin molecule might contain cyclic chains. Such chains have been shown to exist in tyrocidin (Christensen, 1945) and "gramicidin S" (Sanger, 1946). It would, however, be expected that if a cyclic chain were present it would have been detected as a separate fraction in the oxidation experiments, whereas there was in fact no evidence of any fraction that contained no terminal residue or less than would have been expected from a mixture of fraction A and fraction B. Thus the simple structure of four polypeptide chains seems the most plausible and this raises the question of how these chains are arranged in the molecule. From the sulphur contents of fractions A and B, which were kindly determined by Mr. M. W. Rees, a limited number of structures is possible. It was found that fraction A contained four atoms of sulphur per chain, representing four half cystine residues in the insulin and that fraction B contained two atoms of sulphur. The various structures that will satisfy these conditions are illustrated in Figure 1, where the full lines represent the polypeptide chains and the broken lines the -S-S- bridges. The chains marked G gave glycyl terminal residues and those marked P have phenylalanyl terminal residues. While it is impossible at present to choose between the different structures, those represented in *a* and *b* would seem rather more likely than the other two. Structure *d* would have a molecular weight of 6,000 which is not entirely impossible.

From time to time the question has arisen as to whether proteins can be regarded as real chemical entities or whether a single pure protein may be a statistical mixture of a number of different molecular species. X-ray studies on crystalline proteins have done much to show that all molecules of the same protein are very similar, though small differences in a few residues might not be detected. Thus, for instance, it might be supposed that a valine residue in one molecule might be replaceable in other molecules by a leucine residue or an aspartic acid residue by glutamic acid. In the present work it has been possible to identify the amino-acids that occupy certain positions in the molecule and in each case we have found only one amino-acid in each position. Thus, for instance, the second residue in the phenylalanyl chains is always valine and no trace of leucine or other amino-acid could be detected in this position. Similarly the third residue is always aspartic acid and so forth. These experiments do therefore offer support for the belief that proteins like other simpler compounds are chemical entities, each molecule being identical with other molecules of the same protein.

All the above work was carried out on crystalline bovine insulin. Crystalline insulins from a number of different animal species have been shown to have the same sulphur content, microscopic appearance and physiological activity (see Scott and Fisher, 1940) and no immunological differences could be found (Wasserman and Mirsky, 1942). It was thus suggested that insulins from all species were in fact chemically identical. The following experiment does, however, indicate that they do in fact differ in some details of their chemical structure. Samples of insulin from the cow, pig and sheep were oxidized and the A fractions prepared. These were hydrolysed and the amino-acids present identified and estimated semi-quantitatively by paper chromatography. It was

TABLE 8. SERINE, GLYCINE, THREONINE AND ALANINE CONTENT OF THE A FRACTIONS DERIVED FROM COW, PIG AND SHEEP INSULINS

Source of insulin	Serine	Glycine	Threonine	Alanine
Cow	++	+	−	++
Pig	++	+	+	+
Sheep	+	++	−	++

found that there were considerable differences in the content of the amino-acids glycine, alanine, serine and threonine but not of the other amino-acids. The most outstanding difference was the presence of threonine in the pig insulin, whereas it was absent from the other two. These results, which are summarized in Table 8, show that insulin in common with other proteins shows species differences.

Thus the conclusions that can be drawn from this work are that the bovine insulin molecule of molecular weight 12,000 is built up of four open polypeptide chains, that are joined together by -S-S- bridges. Two of these chains, which are identical contain all the amino-acids present in insulin and have a molecular weight per chain of 3,700. They contain the terminal peptide phenylalanyl-valyl-aspartyl-glutamic acid and the peptide sequence threonyl-prolyl-lysyl-alanine is present somewhere in the chain. The other chains, if not completely identical, are very similar. They have a molecular weight per chain of 2,700 and contain none of the basic amino-acids or threonine, phenylalanine or proline. Both have glycine terminal residues and one at least contains the terminal peptide glycyl-isoleucyl-valyl-glutamyl-glutamic acid.

References

Behrens, O. K., and Bergmann, M., 1939, Cosubstrates in proteolysis, J. biol. Chem. *129:* 587-602.

Brand, E., 1946, Amino-acid composition of simple proteins. Ann. N. Y. Acad. Sci. *47:* 187-228.

Chibnall, A. C., 1942, Amino-acid analysis and the structure of proteins. Proc. Roy. Soc. B. *131:* 136-160.

Christensen, H. N., 1945, The free chemical groups of tyrocidine. J. biol. Chem. *160:* 75-82.

Consden, R., Gordon, A. H, and Martin, A. J. P., 1944, Qualitative analysis of proteins: a partition chromatographic method using paper. Biochem. J. *38:* 224-232.

1947, The identification of lower peptides in complex mixtures. Biochem. J. *41:* 590-596.

Consden, R., Gordon, A. H., Martin, A. J. P., and Synge, R. L. M., 1947, Gramicidin S: the sequence of the amino-acid residues. Biochem. J. *41:* 596-602.

Gutfreund, H., 1948, The molecular weight of insulin and its dependence upon pH, concentration and temperature. Biochem. J. *42:* 544-548.

Gutfreund, H., and Ogston, A. G., 1949, A method for determining the sedimentation constant of material of low molecular weight: studies on oxidation products of insulin. Biochem. J. *44:* 163-166.

Jensen, H., and Evans, E. A., 1935, The nature of the free amino groups in insulin and the isolation of phenylalanine and proline from crystalline insulin. J. biol. Chem. *108:* 1-9.

Macpherson, H. T., 1946, The basic amino-acid content of proteins. Biochem. J. *40:* 470-481.

Polson, A., Mosley, V. M., and Wyckoff, R. W. G., 1947, The quantitative chromatography of silk hydrolysates. Science *105:* 603-604.

Porter, R. R., and Sanger, F., 1948, The free amino groups of haemoglobins. Biochem. J. *42:* 287-294.

Rees, M. W., 1946, The estimation of threonine and serine in proteins. Biochem. J. *40:* 632-640.

Sanger, F., 1945, The free amino groups of insulin. Biochem. J. *39:* 507-515.

1946, The free amino group of gramicidin S. Biochem. J. *40:* 261-262.

1947, Oxidation of insulin by performic acid. Nature, Lond. *160:* 295-296.

1948, Some peptides from insulin. Nature, Lond. 162: 491-492.

1949, Fractionation of oxidized insulin. Biochem. J. *44:* 126-128.

Scott, D. A., and Fisher, A. M., 1940, Bison zinc-insulin crystals and human zinc-insulin crystals. Trans. Roy. Soc. Canada *34(V):* 137-140.

Toennies, G., and Homiller, R. P., 1942, The oxidation of amino-acids by hydrogen peroxide in formic acid. J. Amer. Chem. Soc. *64:* 3054-3056.

Tristram, G. R., 1946, Observations upon the application of partition chromatography to the determination of the monoamino acids in proteins. Biochem. J. *40:* 721-733.

Wasserman, P., and Mirsky, L. A., 1942, Immunological identity of insulin from various species. Endocrinology *31:* 115-118.

Woolley, D. W., 1948, Some crystalline peptides isolated from enzymic digests of insulin and their relationship to strepogenin. Federation Proc. *7:* 200-201.

Discussion

Anson: It would be interesting to know whether amino group reagents can be made to react with certain amino groups and not with others, just as SH reagents can be made to react with certain SH groups and not with others. The SH groups of free cysteine, of cysteine peptides and of various cysteine proteins have very different reactivities. Suitable SH reagents under sufficiently favorable conditions react with all the different SH groups, despite the great differences in their reactivities. By varying the SH reagent and the pH, solvent, etc., however, one can find less favorable conditions under which the more reactive SH groups still do react and the less reactive do not react. Perhaps by varying the amino group reagent or by making the reaction conditions less favorable, one can similarly get an amino group reagent to react with an amino group or not, depending on the amino acid to which the amino group is attached and depending on the neighboring peptide or protein structure. Such selective reaction with amino groups, could it be achieved, would of course be useful for both identification and separation purposes.

It would also be interesting to know, again in analogy with SH experience, whether a combination of amino group reagent and reaction conditions can be found which would permit reaction of the reagent with all the free amino groups of a denatured protein but with none of the free amino groups of the same protein in its native form.

Sanger: No experiments have yet been done to determine the relative reactivity of the different amino groups in a protein. Our object has always been to obtain a complete reaction. R. R. Porter (1948, Biochimica et Biophysica Acta 2: 105) has

shown that in β-lactoglobulin and certain other proteins the ε-amino groups of the lysine residues are not all free to react with fluorodinitrobenzene when the protein is native, but become free after denaturation.

BUTLER: I congratulate Dr. Sanger on his results. I and my colleagues, D. M. P. Phillips and Miss Stephen, have been studying the degradation of insulin by enzymes and have obtained results which seem to offer an alternative path to the parts of the molecule more remote from the free amino ends. Trypsin has very little effect on insulin; pepsin breaks it down rather uniformly into products not precipitable by trichloroacetic acid; while chymotrypsin gives a rather massive fragment amounting to about half of the sub-molecule, together with a number, 14 or 15 in all, of small peptides.

The number of bonds broken by chymotrypsin is about the same as the total number of aromatic amino acids in the molecule. Since the specificity of chymotrypsin observed by Bergmann and Fruton is for bonds involving an aromatic group at the carboxyl side, this suggested that the bonds broken by the chymotrypsin in insulin might also involve the aromatic amino acids. The distribution of the aromatic acids between the large fragment (core) and the small peptides supports this. There are four or five aromatics in the core and nine or ten in the small peptides, as might be expected on this hypothesis. We have not made a complete separation of the small peptides, but they have been divided into groups and each group examined contains aromatic residues.

The core contains approximately two-thirds of the cystine of insulin, which presumably holds it together. On oxidation with hydrogen peroxide it gives a product which on electrophoretic examination now shows either three or four peaks. In this it resembles the whole insulin molecule. The amino-terminal groups of the core, as determined by Sanger's method, are about 3.7 percent of glycine and 0.8 percent of valine for the supposed molecular weight. Since the only glycines of Sanger's glycine terminal chains are the terminal ones, this indicates that the core is a residue from the amino ends of the original chains formed by the removal of the phenylalanine groups of the two chains.

The tentative picture of the action of chymotrypsin which we have formed, is that it works its way down from the carboxy ends of the chains, breaking a bond at each aromatic amino acid. These appear to be concentrated in the half of the molecule which is more remote from the amino ends, so that about half the original sub-molecule remains intact when all the aromatic bonds are broken. Altho the observations suggest this picture, it will be clear that a good deal more work will be required before it can be regarded as proven.

SOME ASPECTS OF THE BIOSYNTHESIS OF AMINO ACIDS[1]

DAVID SHEMIN
Department of Biochemistry, College of Physicians and Surgeons, Columbia University, New York

The biosynthesis of amino acids, for the purpose of this discussion, may be separated into two main processes, namely the synthesis and source of the carbon chain and the utilization and mode of participation of the nitrogen atom. The total synthesis of an amino acid may take place either from an α-keto acid formed from metabolites other than preformed amino acids, or from organic nitrogen compounds. If a keto acid is an intermediate in the synthesis of an amino acid it would have to be reductively aminated or be involved in transamination. If the intermediate is an organic nitrogen compound, which may or may not be an amino acid, it will have to be converted into an amino acid either by a condensation reaction or by a degradative step. In the latter case the amino acid will be formed utilizing the carbon atoms and the nitrogen of the precursor and the synthesis will not involve reductive amination or transamination.

The utilization of dietary nitrogen for amino acid synthesis has been demonstrated by the studies of Schoenheimer and his collaborators (1942). On feeding an animal an amino acid labeled with the stable isotope of nitrogen, N^{15}, and subsequently isolating the amino acids from the body proteins, it was found that most of the isolated amino acids contain N^{15}. This finding can be explained on the basis of the utilization of the dietary nitrogen by either of the processes previously mentioned above. The finding, however, of labeled or dietary nitrogen in the essential amino acids of the Mammal must be limited to reductive amination or transamination, since the carbon chain of this class of amino acids cannot be synthesized in the Mammal.

In discussing the synthesis of amino acids from keto acids, we shall attempt to show whether the last step is reductive amination or transamination. If reductive amination or transamination is generally involved in the synthesis of amino acids, α-keto acids must first be synthesized. The synthesis of α-keto acids is defined here as synthesis of these acids from compounds other than amino acids and not by oxidative deamination of preformed amino acids. As far as is known at present there are only three α-keto acids that are synthesized from metabolites other than amino acids, namely, α-keto glutaric acid, oxaloacetic acid and pyruvic acid, which all arise from the metabolism of glucose. Therefore the synthesis of amino acids through the intermediary formation of α-keto acids seems to be limited to glutamic acid, aspartic acid and alanine.

The conversion of α-keto glutaric, oxaloacetic and pyruvic acids and the keto acids, which may arise by oxidative deamination of performed amino acids, into amino acids may take place either by reductive amination or transamination. The evidence for reductive amination thus far obtained seems to be limited to one amino acid namely, glutamic acid (von Euler, Adler, Guenther and Das, 1938). There is as yet no experimental evidence for reductive amination playing a role in the formation of amino acids from their corresponding keto acids other than for the conversion of α-keto-glutaric acid to glutamic acid.

If reductive amination is limited to glutamic acid then any conversion of an α-keto acid to an amino acid must come about by transamination. However the view generally held today is that only glutamic acid, aspartic acid and alanine are concerned in transamination. The original conclusions of Braunstein and Kritzmann (1937, 1938; Braunstein, 1939) that practically all the monoamino monocarboxylic acids are involved in transamination have been challenged by Cohen (1939, 1940b) and the enzymes isolated thus far catalyze the reaction only between glutamic and aspartic acids and between glutamic acid and alanine (Cohen, 1940b; Green, Leloir, and Nocito, 1945). Indeed recently Braunstein (1947) has modified his views and states that of the monoamino monocarboxylic acids only valine, leucine and isoleucine transaminate.

Recently, utilizing another approach, we have begun an investigation to establish the range and mechanism of transamination (Tanenbaum and Shemin, unpub.). The mechanism of transamination as pictured by Braunstein and Kritzmann involves the intermediate formation of a Schiff base analogous to the non-biological reaction described by Herbst (Herbst, 1936; Herbst and Engel, 1934). If this mechanism is as described and not an oxidative deamination coupled with a reductive reamination, transamination will not involve the transient formation of free ammonia and the following experimental approach can be utilized. An amino acid can be synthesized with isotopic nitrogen and incubated together with an α-keto acid in an appropriate biological system known to carry out transamination. After a period of incubation a pure sample of the newly formed amino acid and a sample of ammonia

[1] This investigation was supported in part by a research grant from the Division of Research Grants and Fellowships of the National Institute of Health, U. S. Public Health Service.

are isolated and their respective N^{15} concentrations determined. If the formed amino acid contains a higher concentration of N^{15} than the ammonia, then the nitrogen of the formed amino acid did not exist as free ammonia at any time in the reaction and thus must have been transferred by transamination.

The first experiments were carried out with alanine and α-ketoglutaric acid which are known to transaminate. Alanine labeled with N^{15}, α-ketoglutaric acid and NH_4Cl were incubated with pig heart homogenate for one hour. The ammonium salt was added in order to isolate a sample of ammonia at the end of the incubation period. After the period of incubation, samples of alanine, glutamic acid and ammonia were isolated and their respective N^{15} concentrations determined. It can be seen from Table 1 that glutamic acid is formed with an appreciable N^{15} concentration and practically no N^{15} exists in the ammonia sample. This demonstrates that transamination occurred and is not an oxidative deamination coupled with a reductive amination. The alanine experiments can serve as a control for other amino acid systems.

TABLE 1. TRANSAMINATION

75 g. of pig heart homogenate suspended in 400 cc. of 0.125 percent $KHCO_3$ solution and incubated for 1 hour at 38° with 10 mM of the isotopic amino acid, 10 mM of non-isotopic NH_4Cl and with 10 mM of α-ketoglutaric acid.

Isotopic amino acid added		N^{15} concentration of compounds isolated. Atom percent excess			
Amino acid	N^{15} concentration. Atom percent excess	Alanine	Leucine	Glutamic acid	NH_3
DL-Alanine	0.50	—	—	0.167	0.027
DL-Alanine	0.50	0.214	—	0.176	0.038
L-Alanine	1.97	1.67	—	1.37	0.038
L-Alanine	1.97	—	—	1.45	0.008
L-Leucine	1.50	—	1.37	0.42	0.010
L-Leucine	4.19	—	—	1.15	0.40

Similar experiments may then be set up in order to study whether other amino acids can transaminate. Thus far only isotopic leucine has been studied. It can be seen from Table 1 that here again the glutamic acid formed contained an appreciable isotopic concentration, while the N^{15} concentration of the ammonia was lower. It would therefore appear that leucine nitrogen was transferred to α-keto glutaric acid without first going through free ammonia; leucine can transaminate with glutamic acid. Cohen (1939, 1940a) has shown for muscle, kidney and liver a slight transaminase activity for valine, leucine and isoleucine.

In the total synthesis of amino acids, amination or transamination seems only to be involved in the formation of the three amino acids, glutamic acid, aspartic acid and alanine, since only their respective α-keto acids are known to be formed from compounds other than amino acids. The formation of many of the remaining amino acids seems, from the data available, to take place rather by conversion reactions. In the Mammal we now know the source of all the non-essential amino acids but two.

The relationship of some of the non-essential amino acids in the Mammal can be seen in Figure 1. The biosynthesis of some of the amino acids essential to the Mammal has been studied in organisms capable of synthesizing these amino acids. It has been demonstrated that tryptophane is formed from serine and indole in Neurospora, the latter compound arising from anthranilic acid (Tatum, Bonner and Beadle, 1944; Tatum and Bonner, 1944). It has also been shown that methionine is formed from homoserine and cysteine via cystathionine and homocysteine (Horowitz, 1947; Teas, Horowitz and Fling, 1948) the reverse reaction of the formation of cysteine in the Mammal (Binkley and du Vigneaud, 1942; Stetten, 1942). There is also evidence that homoserine is converted to threonine in Neurospora (Teas, Horowitz and Fling, 1948). Although the mechanism of synthesis of the remaining

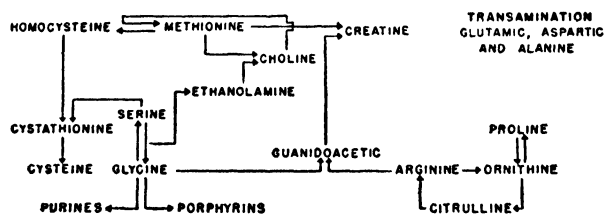


FIG. 1

amino acids is not as yet known it would appear that conversion reactions are involved rather than the formation first of the corresponding α-keto acid analogue.

It can be seen from Figure 1 that the carbon source of arginine or proline or ornithine, and the carbon source of serine or glycine are not as yet known. Since arginine, ornithine and proline are related to one another (Weil-Malherbe and Krebs, 1935; Stetten and Schoenheimer, 1944; Roloff, Ratner and Schoenheimer, 1940; Shemin and Rittenberg, 1945) and similarly since serine and glycine are interconvertible, one must seek for a source of one of the five carbon amino acids and the carbon source of either serine or glycine.

The remaining discussion will be concerned with serine and glycine, amino acids which enter into many biological syntheses. It has been demonstrated that serine is readily converted to glycine in the animal organism (Shemin, 1946). The manner in which these experiments were carried out will be given in some detail, for similar experiments, described below, were subsequently carried out to find the carbon source of glycine and serine. Isotopically labeled

compounds were administered to animals along with benzoic acid. The benzoic acid was included so that a sample of glycine could be readily obtained from the excreted hippuric acid.

In the original search for the carbon source of glycine, compounds were labeled with isotopic nitrogen administered to animals along with benzoic acid and the N^{15} concentration subsequently determined on the glycine of the excreted hippuric acid. Since the test compounds were labeled only with N^{15}, the appearance of excess isotopic nitrogen in the glycine of the isolated hippuric acid would only be indirect proof for the utilization of the carbon structure of the test compound, for the nitrogen might have been removed from it by deamination and then utilized for glycine formation by combination with some other compound. However, the determination of the ratio of the isotope concentration of the test compound, C_0, to that of the hippuric acid, C, (the dilution factor C_0/C) offers a method of differentiating between direct carbon utilization and mere nitrogen utilization. The dilution factor for glycine should be lower than for any other test compound. This is confirmed by the findings summarized in Table 2. The dilution factor for isotopic ammonia was also determined and found to be considerably higher than that of glycine.

The dilution factors of glycine and ammonia can be taken as the lower and upper limits respectively for comparison of "natural" amino acids as glycine precursors. One primary condition that a test compound must meet in order to be considered as a possible precursor of glycine is that its dilution factor must be lower than that found for ammonia. The nearer the dilution factor of the test compound to that found for glycine, the better the evidence for, and the more extensive is the conversion.

It can be seen from Table 2 that, whereas the dilution factors for glycine in the rat and guinea pig are approximately the same, the dilution factors for ammonia in these animals are widely different; that is, 21 for the guinea pig and 415 for the rat. This indicates that the guinea pig can utilize ammonia more efficiently for the formation of glycine or for a glycine precursor than the rat. Since the dilution factor for ammonia in the rat is so high, the fact that the dilution factors for the other test substances are lower than that for ammonia is not a sufficient criterion for their conversion to glycine; the test substance may merely be a better nitrogen donor for glycine synthesis than ammonia. On the other hand it may be a relatively poor donor, as appears to be the case for L-leucine in the guinea pig. In the rat it is more desirable to consider how closely the dilution factor of the test substance approaches that of glycine.

It can be seen from Table 2 that the dilution factor of L-serine more closely approximates that of glycine than any other test compound in both guinea pig and rat, 5.5 in the rat and 3.9 in the guinea pig. This finding strongly indicates an extensive and rapid conversion of L-serine into glycine with utilization of the carbon chain and without loss of the α-amino group, unless a specific transamination is

TABLE 2. UTILIZATION OF NITROGEN-CONTAINING COMPOUNDS FOR GLYCINE FORMATION

0.35 mM of a compound labeled with N^{15} and 0.35 mM of benzoic acid per 100 gm. of body weight were injected intraperitoneally into fasting rats and guinea pigs.

Compound administered		N^{15} excess of isolated hippuric acid (C)		Dilution factor $\left(\frac{C_0}{C}\right)$	
Compound	N^{15} excess (C_0)	Rat	Guinea pig	Rat	Guinea pig
	atom percent	atom percent	atom percent		
Glycine	1.12	0.406	0.459	2.8	2.4
Ammonia*	2.35	0.006	0.115	390	20
Ammonia*	32.5	0.074	1.466	440	22
L-Serine	1.89	0.345	0.483	5.5	3.9
D-Serine	1.89	0.012	0.013	158	145
L-Glutamic acid	0.86	0.019	0.087	45	10
D-Glutamic acid	4.50	0.003	0.010	1,500	450
DL-Glutamic acid†	4.45	0.069	0.345	64	13
DL-Aspartic acid†	4.05	0.070	0.218	58	19
L-Alanine	1.97	0.021	0.092	94	21
DL-Proline	11.6	0.245	0.494	47	23
L-Leucine	4.67	0.039	0.087	120	54
Ethanolamine	2.00	0.006	0.035	334	57

* Administered as ammonium citrate.

† 0.35 mM of the L+0.35 mM of the D compound per 100 gm. of body weight.

involved. This conversion is specific for the L isomer, as can be seen from Table 2.

It would appear from these N^{15} experiments that serine is readily converted to glycine. Indeed this was proven with C^{13} and N^{15} labeled serine. It was found the serine was converted to glycine by the splitting of the serine between the α- and β-carbon atoms (Shemin, 1946). It was subsequently found that the conversion of serine to glycine is reversible. These reactions and the mechanism will also be discussed below.

Since the serine and glycine are non-essential to

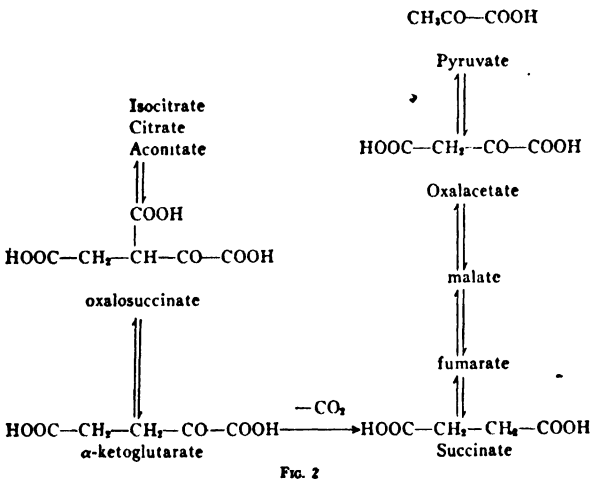


FIG. 2

the animal organism, at least one of these amino acids must have another carbon source. It would appear from Table 2 and from the experiments of Leuthardt (1941) that glutamic acid may be the carbon source for serine and glycine. The dilution factor found with glutamic acid is comparatively low. Either glutamic acid is utilized directly or it is merely an efficient nitrogen donor to the serine or glycine precursor. Leuthardt (1941) reported that he obtained increased amounts of hippuric acid on incubation of benzoic acid and glutamine with guinea pig liver slices.

Accordingly glutamic acid was synthesized with C^{13} in both the carboxyl group and α-carbon atom and with N^{15}, $HOOC\text{-}CH_2\text{-}CH_2\text{-}C^{13}HN^{15}H_2\text{-}C^{13}OOH$. This compound was administered to guinea pigs and rats along with benzoic acid in a manner similar to that described in the experiments on the conversion of serine to glycine (Shemin, Tanenbaum and Friedman, unpub.). If the glutamic acid were converted directly to serine or glycine the glycine isolated from the excreted hippuric acid should contain C^{13} in equal concentrations in both of its carbon atoms. On the other hand if some other member of the tricarboxylic acid cycle is more directly involved as a precursor of glycine, the glycine would only contain excess C^{13} in the carboxyl carbon atom. This would follow since the glutamic acid on being converted to succinic acid through α-keto glutaric acid would lose the C^{13} labeled carboxyl group. (See Fig. 2.)

The glycine isolated in these experiments of both the rat and guinea pig contained C^{13} only in the carboxyl-carbon atom (Table 3). This demonstrates that glutamic acid is not directly converted into serine or glycine. However from the results it would appear that some member of the tricarboxylic acid cycle is the source of glycine or serine. It would be very difficult to find the most immediate member since all the compounds in the cycle are readily convertible to one another. It is of interest to point out that the N^{15} dilution factor is much lower than the C^{13} dilution factor (Table 3). Apparently some of the glutamic acid is converted to the glycine or serine precursor which is then very efficiently transaminated with more of the N^{15} labeled glutamic acid. It may also be interpreted to mean that the tricarboxylic acid pool is several times greater than the nitrogen pool.

In an attempt to find a more immediate source of glycine or serine, alanine was synthesized with N^{15} and C^{13} in the carboxyl-carbon atom. This doubly labeled alanine was administered along with benzoic acid to rats and guinea pigs and the glycine isolated from the excreted hippuric acid (Shemin, Tanenbaum and Friedman, unpub.). It can be seen from Tables 3 and 4 that the C^{13} dilution factor was greater in the alanine experiment than in the glutamic acid experiment for the rat, while the C^{13} dilution factor was less in the alanine experiment than in the glutamic acid experiment for the guinea pig. It is difficult therefore to decide whether ketoglutaric acid or pyruvic acid is more closely associated with the immediate precursor of glycine or serine. Anker (1948) found that carbonyl labeled and carboxyl labeled pyruvic acids formed glycine

TABLE 3. UTILIZATION OF GLUTAMIC ACID FOR GLYCINE FORMATION

0.70 mM of DL-glutamic acid labeled with 28.2 atom percent N^{15} excess and with 25.6 atom percent C^{13} excess in both the carboxyl carbon atom and in the α-carbon atom ($HOOC\text{-}CH_2\text{-}CH_2\text{-}C^{13}HN^{15}H_2\text{-}C^{13}OOH$) and 0.35 mM of benzoic acid per 100 gm. of body weight were injected intraperitoneally into fasting rats and guinea pigs.

Animal	Isotope concentrations of the isolated glycine			Dilution factor	
	N^{15}	C^{13} Total carbon	C^{13} Carboxyl-carbon atom	N^{15}	C^{13}
	Atom percent excess	Atom percent excess	Atom percent excess		
Rat	0.76	0.040	0.085	37	300
Guinea pig	0.89	0.11	0.20	32	128
	—	0.07	0.18	—	140

labeled in the α-carbon atom and carboxyl-carbon atom respectively and postulated the following reaction to account for the results: pyruvic acid → serine → glycine. Elwyn and Sprinson (unpub.) found that β-carbon labeled serine gave rise to α-carbon labeled glycine. The latter results can be explained as follows (see Fig. 2). The β-carbon labeled serine gives rise to methyl labeled pyruvic acid which in turn may be converted to oxaloacetate, and the latter may form the symmetrical molecule, succinic acid, in which both methylene carbons will now be labeled. By the reverse process the succinic acid will give rise to α- and β-carbon labeled pyruvic acid, so that if pyruvic acid or a four carbon dicarboxylic acid were converted to serine the glycine formed would be labeled in the α-carbon atom. Therefore either pyruvic acid or a four carbon dicarboxylic acid may be the most immediate precursor of serine or glycine.

Proof of the conversion of serine to glycine was obtained in an experiment in which serine (CH_2OH-$CHN^{15}H_2$-$C^{13}OOH$ labeled with C^{13} in the carboxyl-carbon atom and with N^{15} was administered to rats and guinea pigs along with benzoic acid (Shemin 1946). The glycine isolated from the excreted hippuric acid contained both N^{15} and C^{13} in the same ratio as in the serine given (Table 5). This demonstrates that serine is converted into glycine by cleavage between the α- and β-carbon atoms of serine. The mechanism postulated for this reaction involved a dehydrogenation to formyl glycine which on hydrolytic cleavage yielded glycine and formic acid. The reverse process, the conversion of glycine to serine, was subsequently found by Ehrensvärd *et al.* (1947), Winnick *et al.* (1948) and Sakami (1948, 1949).

Sakami (1948, 1949) found that on feeding formic acid labeled with C^{14} and glycine labeled with C^{13} in the carboxyl group, the serine isolated contained C^{14} in the β-carbon atom and C^{13} in the carboxyl-carbon atom. Thus it would appear that formic acid condenses with glycine to yield an intermediate which is converted to serine. He subsequently found that α-carbon labeled glycine produced in the animal organism α- and β-labeled serine. He postulated therefore that the α-carbon atom of glycine was converted to formate which then condensed with another molecule of glycine. Independently Karlsson and Barker (1949) found that α-carbon labeled glycine produced uric acid labeled not only in the 5 position but also contained an appreciable isotopic carbon concentration in positions 2 and 8 which have been shown to originate from formic acid (Sonne, Buchanan and Delluva, 1948; Buchanan, Sonne and Delluva, 1948). The nitrogen of glycine as well as the carbon atoms are utilized for the biosynthesis of serine. On feeding N^{15} labeled glycine, both the cystine, which originates from serine carbon atoms, and the hydroxy amino acids contain a higher N^{15} concentration than the glutamic acid.

The mechanism of the conversion of glycine to serine can be further elucidated by studies on the biosynthesis of porphyrins. There appears to be a relationship between the mechanism of serine formation and the utilization of the α-carbon atom of glycine for porphyrin synthesis. From the evidence mentioned above it would appear that formic acid, as such, condenses with glycine to yield serine. The

TABLE 4. UTILIZATION OF ALANINE FOR GLYCINE FORMATION

0.7 mM of DL-Alanine labeled with 16.5 atom percent N^{15} excess and with 61.5 atom percent C^{13} excess in the carboxyl carbon atom (CH_3-$CHN^{15}H_2$-$C^{13}OOH$) and 0.35 mM of benzoic acid per 100 g. of body weight were injected intraperitoneally into fasting rats and guinea pigs.

Animal	Isotope concentrations of the isolated glycine				
	N^{15}	C^{13} Total carbon	C^{13} Carboxyl-carbon	Dilution factor	
				N^{15}	C^{13}
	Atom percent excess	Atom percent excess	Atom percent excess		
Rat	0.11	0.028	0.066	150	930
Guinea pig	1.15	0.58	1.18	14	52

TABLE 5. CONVERSION OF DL-SERINE, LABELED WITH C^{13} IN CARBOXYL GROUP AND WITH N^{15}, TO GLYCINE

0.35 mM of DL-serine (0.175 mM of the D+0.175 mM of the L compound) and 0.35 mM of benzoic acid per 100 gm. of body weight were injected intraperitoneally into fasting rats and guinea pigs. Serine administered contained 4.53 per cent $N^{15}(C_0)$ and 2.30 percent C^{13} in carboxyl group (C_0'); the N^{15} : C^{13} ratio $C_0/C_0' = 1.90$.

N^{15} excess of glycine of isolated hippuric acid (C)		C^{13} excess of carboxyl group of glycine of isolated hippuric acid (C′)		N^{15} dilution factor $\left(\frac{C_0}{C}\right)$		C^{13} dilution factor $\left(\frac{C_0}{C'}\right)$		N^{15}:C^{13} ratio of glycine of isolated hippuric acid $\left(\frac{C}{C'}\right)$	
Rat	Guinea pig	Rat	Guinea pig	Rat	Guinea pig	Rat	Guinea pig	Rat	Guinea pig
atom percent 0.614	atom percent 0.799	atom percent 0.346	atom percent 0.410	7.4	5.7	6.9	5.8	1.77	1.95

findings in the biosynthesis of porphyrins, however, throw some doubt on the utilization of formic acid in the formation of serine. It has been demonstrated that not only is glycine utilized in the synthesis of porphyrins but that all four nitrogen atoms of the porphyrin molecule come from glycine nitrogen (Shemin and Rittenberg, 1946a and b; Wittenberg and Shemin, 1948). It has also been demonstrated in an *in vitro* duck blood system (Shemin, London

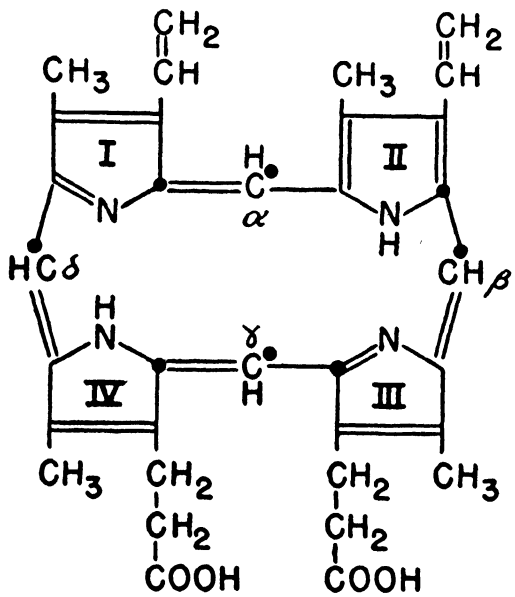


FIG. 3

and Rittenberg, 1948) that two α-carbon atoms of glycine are utilized for every nitrogen atom, or in the whole molecule of the porphyrin eight α-carbon atoms are utilized for the four nitrogen atoms of glycine (Radin, Rittenberg and Shemin, 1949). On the basis of degradation experiments, it was found that the α-positions, under the vinyl and propionic acid side chains, of each pyrrole and the methene carbon atoms of the porphyrin originate from the α-carbon atom of glycine (Fig. 3), (Wittenberg and Shemin, in press). Since in three pairs the α-carbon atoms of glycine are attached to each other and since α-carbon atom labeled glycine gives rise to α- and β-carbon atoms labeled serine (a ratio of two α-carbon atoms to one nitrogen atom) it might appear that serine is the intermediate utilized for porphyrin formation. However in experiments designed to test this point, it was found that serine is not utilized for heme synthesis (Shemin, London and Rittenberg, in press). It was also found in the *in vitro* system that formic acid is not involved in the porphyrin synthesis (Radin, Rittenberg and Shemin, 1949).

Although in porphyrin synthesis two α-carbon atoms of glycine are utilized for every nitrogen atom, serine and formic acid are not utilized. It is possible that a derivative of formic acid is the active compound which condenses with glycine to form an intermediate; the latter compound can then be converted to serine and also be utilized for porphyrin synthesis.

REFERENCES

ANKER, H. S., 1948, Some aspects of the metabolism of pyruvic acid in the intact animal. J. biol. Chem. *176:* 1337-1352.

BINKLEY, F., and DU VIGNEAUD, V., 1942, The formation of cysteine from homocysteine and serine by liver tissue of rats. J. biol. Chem. *144:* 507-511.

BRAUNSTEIN, A. E., 1939, Die enzymatische Umaminierung der Aminosäuren und ihre physiologische Bedeutung. Enzymologia *7:* 25-52.

BRAUNSTEIN, A. E., 1947, Transamination and the integrative functions of the dicarboxylic acids in nitrogen metabolism. Adv. in Protein Chem. *3:* 1-52.

BRAUNSTEIN, A. E., and KRITZMANN, M. G., 1937, Uber den Ab- und Aufbau von Aminosäuren durch Umaminierung. Enzymologia *2:* 129-146.

1938, Biochemia, U. S. S. R., *3:* 590.

BUCHANAN, J. M., SONNE, J. C., DELLUVA, A. M., 1948, Biological precursors of uric acid. II. The role of lactate, glycine, and carbon dioxide as precursors of the carbon chain and nitrogen atom 7 of uric acid. J. biol. Chem. *173:* 81-98.

COHEN, P. P., 1939, Transamination in pigeon breast muscle. Biochem. J. *33:* 1470-1487.

1940a, Transamination in liver and kidney. J. biol. Chem. *133:* XX-XXI.

1940b. Transamination with purified enzyme preparations. (Transaminase). J. biol. Chem. *136:* 565-584.

EHRENSVÄRD, G., SPERBER, E., SALUSTE, E., REIO, L., and STJERNHOLM, R., 1947, Metabolic connection between proline and glycine and in the amino acid utilization of *Torulopsis utilis.* J. biol. Chem. *169:* 759-760.

ELWYN, D. and SPRINSON, D. B., unpublished.

GREEN, D. E., LELOIR, L. F., and NOCITO, U., 1945, Transaminases. J. biol. Chem. *161:* 559-582.

HERBST, R. M., 1936, The reaction between α-ketonic acids and α-amino acids. J. Amer. chem. Soc. *58:* 2239-2243.

HERBST, R. M., and ENGEL, L. L. 1934, A reaction between α-ketonic acids and α-amino acids. J. biol. Chem. *107:* 505-512.

HOROWITZ, N. H., 1947, Methionine synthesis in Neurospora. The isolation of cystathionine. J. biol. Chem. *171:* 255-264.

KARLSSON, J. L., and BARKER, H. A., 1949, Biosynthesis of uric acid labeled with radioactive carbon. J. biol. Chem. *177:* 597-599.

LEUTHARDT, F., 1941, Über Hippursaurebildung aus Glutamin. Z. physiol. Chem. *270:* 113-124.

RADIN, N., RITTENBERG, D., and SHEMIN, D., 1949, Biosynthesis of heme. Federation Proc. *8:* 240.

ROLOFF, M., RATNER, S., and SCHOENHEIMER, R., 1940. The biological conversion of ornithine into proline and glutamic acid. J. biol. Chem. *136:* 561-562.

SAKAMI, W., 1948, The conversion of formate and glycine to serine and glycogen in the intact rat. J. biol. Chem. *176:* 995-996.

1949, The conversion of glycine into serine in the intact rat. J. biol. Chem. *178:* 519-520.

SCHOENHEIMER, R., 1942, The Dynamic State of Body Constituents, Harvard University Press, Cambridge, Mass.

SHEMIN, D., 1946, The biological conversion of L-serine to glycine. J. biol. Chem., *162:* 297-307.

SHEMIN, D., TANENBAUM, S. W., and FRIEDMAN, S., unpub.

SHEMIN, D., LONDON, I. M., and RITTENBERG, D., 1948, The *in vitro* synthesis of heme from glycine by the nucleated red blood cell. J. biol. Chem. *173:* 799-800.

SHEMIN, D., and RITTENBERG, D., 1945, On the mechanism of the conversion of ornithine to proline *in vivo;* intramolecular nitrogen shift. J. biol. Chem. *158:* 71-76.

1946a, The biological utilization of glycine for the synthesis of the protoporphyrin of hemoglobin. J. biol. Chem. *166:* 621-625.

1946b, The life span of the human red blood cell. J. biol. Chem. *166:* 627-636.

SONNE, J. C., BUCHANAN, J. M., and DELLUVA, A. M. 1948, Biological precursors of uric acid. I. The role of lactate, acetate and formate in the synthesis of the ureide groups of uric acid. J. biol. Chem. *173:* 69-79.

STETTEN, D., JR., 1942, The fate of dietary serine in the body of the rat. J. biol. Chem. *144:* 501-506.

STETTEN, M. R., and SCHOENHEIMER, R., 1944, The metabolism of L(—)-proline studied with the aid of deuterium and isotopic nitrogen. J. biol. Chem. *153:* 113-132.

TANENBAUM, S. W., and SHEMIN, D., unpublished.

TATUM, E. L., and BONNER, D., 1944, Indole and serine in the biosynthesis and breakdown of tryptophane. Proc. Nat. Acad. Sci. Wash. *30:* 30-37.

TATUM, E. L., BONNER, D., and BEADLE, G. W., 1944, Anthranilic acid and the biosynthesis of indole and tryptophane by neurospora. Arch. Biochem. *3:* 477-478.

TEAS, H. J., HOROWITZ, N. H., and FLING, M., 1948, Homoserine as a precursor of theonine and methionine in Neurospora. J. biol. Chem. *172:* 651-658.

VON EULER, H., ADLER, E., GUENTHER, G., and DAS, N. B., 1938, Über den enzymatischen Abbau und Aufbau des Glutaminsäure II. In tierischen Geweben. Z. physiol. Chem. *254:* 61-103.

WEIL-MALHERBE, H., and KREBS, H. A., 1935, Metabolism of amino acids. V. The conversion of proline into glutamic acid in kidney. Biochem. J. *29:* 2077-2081.

WINNICK, T., MORING-CLAESSON, I., and GREENBERG, D. M., 1948, Distribution of radioactive carbon among certain amino acids of liver homogenate protein, following uptake experiments with labeled glycine. J. biol. Chem. *175:* 127-132.

WITTENBERG, J., and SHEMIN, D., 1949, The utilization of glycine for the biosynthesis of both types of pyrroles in protoporphyrin. J. biol. Chem. *178:* 47-51.

WITTENBERG, J., and SHEMIN, D., in press.

SOME CONSIDERATION OF THE INTERACTION OF THE METAL PEPTIDASES WITH THEIR SUBSTRATES[1]

EMIL L. SMITH AND RUFUS LUMRY[2]
University of Utah, Salt Lake City, Utah

It is generally agreed that enzymes must form definite chemical compounds with their substrates. This thesis, first clearly stated by Michaelis and Menten (1913), has been tested by kinetic studies on a large number of enzymes. Direct physical evidence has also been obtained by Keilin and Mann (1937) and by Chance (1943, 1947) that peroxidase and catalase form complexes with their substrate, hydrogen peroxide. Although there is ample evidence that enzymes form compounds with their substrates, the nature of the interactions has remained largely unknown. From the viewpoint of the enzyme chemist, it is apparent that these interactions determine the specificity of the system. From the viewpoint of physical chemistry, we wish to know why these interactions lower the free energy of the activation and hence permit the reaction to proceed. It has now become possible to interpret some enzyme reactions both from the specificity and energetic viewpoints. In this paper, we shall deal mainly with a few peptidases although certain of the conclusions would appear to have wider significance. It is convenient to deal with these peptidases first mainly from the aspect of specificity.

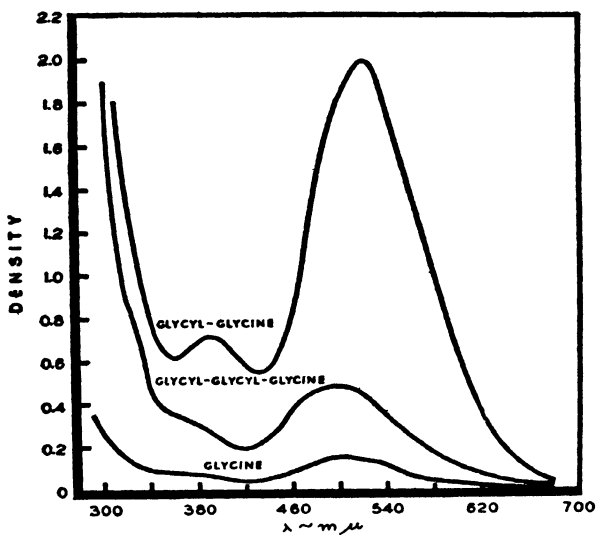


FIG. 1. Interaction of glycine, glycylglycine, and diglycylglycine (all at a concentration of 0.125 M) with 0.01 M $CoCl_2$ for 24 hours at room temperature. It is apparent that the dipeptide gives the greatest degree of interaction (Smith, 1948b).

In 1926, Euler and Josephson found that while glycylglycine was readily hydrolyzed by an extract of intestinal mucosa, the N-acylated peptide was resistant. They concluded that the enzyme combines with its substrate not only at the sensitive peptide linkage, but also with the free amino group; this was called the di-affinity concept. The later work of Waldschmidt-Leitz, Grassmann, and the Bergmann group has given us the classification of dipeptidases, aminopeptidases and carboxypeptidases, and the di-affinity (or polyaffinity) theory was extended to these enzymes. The nomenclature indicates that the free polar groups required in these substrates are essential for the action of the enzyme.

Since the discovery that peptidase activities are increased by certain metal ions (Johnson and Berger, 1942), it has been shown that many of the well-characterized peptidases are metal-proteins in which the metal is essential and reasonably specific for the enzymatic action (Smith, 1949b). Leucine aminopeptidase, carboxypeptidase and other peptidases are inhibited by the usual poisons of metal enzymes. With some enzymes like leucine aminopeptidase, purification or dialysis results in a decrease of activity unless metals are returned to the system. The recombination of metal and protein has been shown to follow the usual mass law equilibrium where one atom of the metal participates in each active group of the enzyme (Smith, 1946). Moreover, for many peptidases and certain other enzymes, the restoration of the enzymatic activity by the addition of metal ions is a slow reaction (Smith, 1949b).

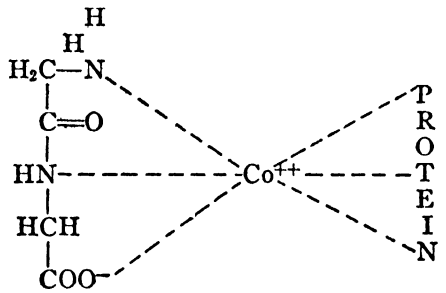
While the metal of most peptidases is rather easily separated from the protein, crystalline carboxypeptidase probably contains tightly bound magnesium. The enzyme is inhibited by such metal poisons as cyanide, sulfide, pyrophosphate, cysteine, citrate, oxalate, etc. (Smith and Hanson, 1949).

The requirement of the peptidases for substrates which contain free polar groups suggests that these polar groups combine, through complex formation, with the metal of the enzyme. With glycylglycine dipeptidase which requires Co^{++} for its activity, there is a definite parallelism between the ability of the enzyme to hydrolyze the substrate and the ability of Co^{++} to form a coordination compound with the substrate as judged by the change in the cobalt

[1] This work was aided by grants from the United States Public Health Service.

[2] Merck Fellow in the Natural Sciences.

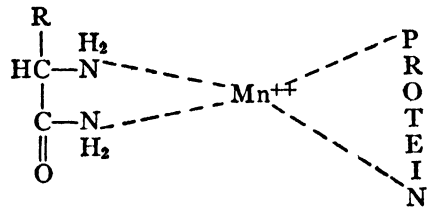
absorption spectrum (Smith, 1948b). This is shown in Figure 1. Co^{++} appears to combine with glycylglycine through the nitrogen of the free amino group, the nitrogen of the peptide bond and the free carboxyl group. While sarcosylglycine is split by the enzyme, dimethylglycylglycine, glycylsarcosine, and glycylglycinamide are resistant. Thus, the enzyme-substrate complex must be formed as shown in the following diagram:



It is well-known that Co^{++} tends to make chelate compounds, that is ring compounds in which there are only five members, and that is the type of complex found with glycylglycine and cobalt. It is interesting that β-alanylglycine does not coordinate with cobalt and is not split by the dipeptidase. Likewise, glycyl-β-alanine coordinates only weakly with Co^{++}, but there may be an extremely weak hydrolysis by the enzyme.

It appears from these data that for a compound to be susceptible to hydrolysis by glycylglycine dipeptidase, it must be able to form a complex with cobalt. Obviously, this is a necessary condition but not a sufficient condition since many peptides, for example, L-leucylglycine, form complexes with cobalt but these compounds are not hydrolyzed by this enzyme. On the other hand, glycyl-L-leucine, the isomeric peptide, does not form a strong complex with cobalt apparently because of the position of the isobutyl side chain (Fig. 2). Thus, peptides can exhibit considerable specificity in forming complexes with metal ions. While it is well-known that proteins are extremely specific in their combination with various substances, it is interesting to find such specificity at the dipeptide level (Smith, 1948b).

From the probable mode of combination of metal and substrate, we can begin to understand the differences between a dipeptidase and an aminopeptidase. The cobalt of glycylglycine dipeptidase almost certainly combines with its substrate at three points to form a bidentate chelate complex. With leucine aminopeptidase, it must be assumed that the metal, either manganese or magnesium, combines only with the free amino group and the nitrogen of the peptide bond. A free carboxyl group is not required since the enzyme can hydrolyze leucinamide just as rapidly as it can the dipeptide, leucylglycine. Here the complex may be indicated as follows (Smith, 1949a):



Unlike the substrates for aminopeptidases and dipeptidases, the substrates for carboxypeptidase do not contain free amino groups. The best substrates for carboxypeptidase are dipeptides in which the amino group is blocked by an acylating group (Hofmann and Bergmann, 1940). A typical substrate is carbobenzoxyglycyl-L-phenylalanine. The terminal carboxyl group is essential for the action of the

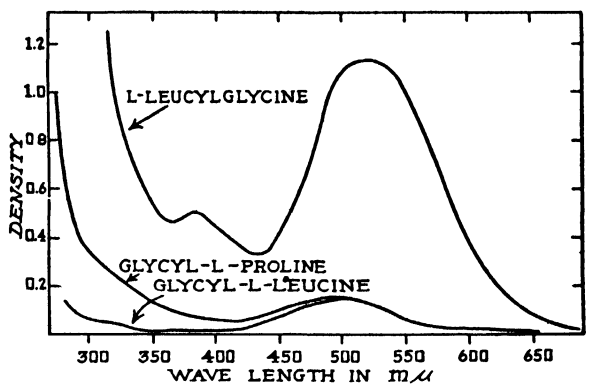


FIG. 2. Interaction of leucylglycine, glycylleucine and glycylproline (all at 0.125 M) with 0.01 M $CoCl_2$ for 24 hours. Glycylleucine gives about the same degree of interaction as glycylproline which does not possess a peptide hydrogen (Smith, 1948b).

enzyme. The peptide hydrogen is not essential for the action of this enzyme as indicated by the hydrolysis of such a compound as carbobenzoxytryptophylproline (Smith, 1948a). Since proline is an imino acid, there is no hydrogen at the peptide bond. The unessential character of the peptide nitrogen is shown very strikingly by the discovery of Snoke, Schwert and Neurath (1948) that carboxypeptidase can hydrolyze esters of the correct configuration; such a compound as hippurylphenyllactic acid is hydrolyzed by carboxypeptidase. Therefore, in carboxypeptidase the metal must bind with the free carboxyl group and with the carbonyl group at the peptide bond and not with the nitrogen at all (Smith, 1949a). Thus, for the three different types of exopeptidases, each of which is a metal-protein, the nature of the chelation is the determining factor in the mode of combination. The important and critical consideration seems to be that in each instance, the metal forms at least one bond on each side of the sensitive linkage.

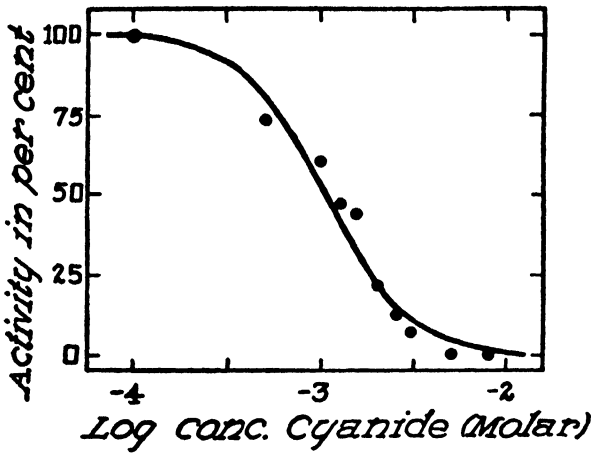


FIG. 3. Inhibition of carboxypeptidase by cyanide. The curve drawn through the data is the theoretical one derived from the mass law equation for the combination of two moles of inhibitor with each active group of the enzyme (Smith and Hanson, 1949).

Since the metal in carboxypeptidase is tightly bound, quantitative inhibition studies were performed in order to determine how many valences of the metal are available for combination with the substrate or, alternatively, for combination with an inhibitor which binds with the metal. It was found that each active group of carboxypeptidase is capable of binding two cyanide atoms (Fig. 3) or one orthophosphate or one sulfide. From these data it would appear that there are two free valences available for combination with the substrate or with an inhibitor (Smith and Hanson, 1949). This provides additional evidence for the di-affinity of the enzyme, but it may now be ascribed to the characteristic binding with the metal of the enzyme.

In addition to the binding of the substrate by the metal through the polar groups, there are additional binding forces exerted by the protein moiety of the enzyme through the side chains of the substrate. For example, carbobenzoxyglycylglycine is a poor substrate for carboxypeptidase, but carbobenzoxyglycylphenylalanine is an excellent substrate. This indicates that the protein of carboxypeptidase binds very strongly with the terminal phenyl group. We may speak of this as the R-group binding by the protein.

The polyaffinity concept first suggested by Bergmann and his coworkers (1935) indicates that there must be a correct steric relationship between the polar groups of the substrate and the side chain branches of the substrate. Hence, carboxypeptidase hydrolyzes only peptides which contain L-amino acids and not those with D-amino acids. A simple diagram (Fig. 4) may be used to indicate this polyaffinity relationship for carboxypeptidase. The polar

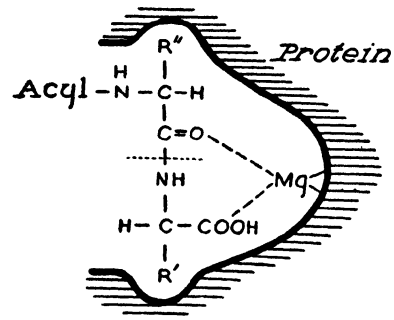
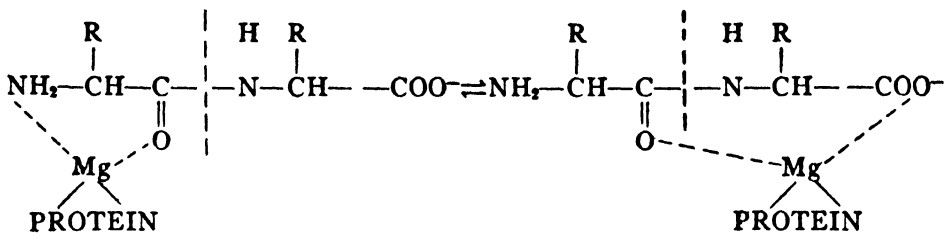


FIG. 4. The binding of the various groups of an acylated dipeptide by carboxypeptidase is shown diagrammatically.

groups are bound by the metal and the R groups are held by the protein. An explanation of the substrate specificities and inhibition properties of this enzyme may be made utilizing the polyaffinity concept and the knowledge that the metal binds the polar groups of the substrate. It is known that compounds which have free amino groups are either not hydrolyzed at all or are acted upon extremely slowly by this enzyme. It may be supposed that the metal has a high affinity for the amino group and a different type of complex is then formed than in the usual enzyme-substrate complex. Such a combination with a free dipeptide is not a fruitful combination in that the substrate is not hydrolyzed. An explanation for the lack of action is that the linkages are not on opposite sides of the peptide bond but are on the same side (Smith, 1949a).



It should also be noted that the ineffective combination through the amino group involves the formation of a five-membered ring as compared to the seven-membered ring of the fruitful combination.

Since the chelation in the smaller ring structure would be expected to be stronger, we can understand why the equilibrium is in favor of the five-membered ring. Similarly, it would be expected that free dipeptides might act as inhibitors of carboxypeptidase and we have found that this is indeed true (unpublished observations).

The validity of our concept of the mode of combination of carboxypeptidase and its substrates can be tested by inhibition studies of various kinds. For example, the enzyme may be inhibited by blocking the metal portion of the enzyme (as with cyanide) or by blocking the R-group position or both. It might be expected that compounds which can bind both with the metal and with the R group position of the protein would be strong inhibitors,

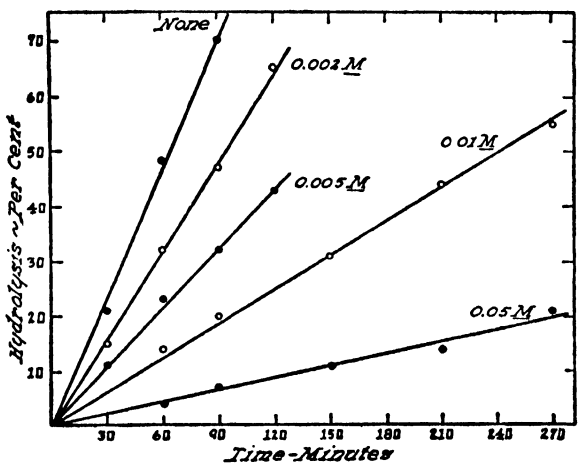


FIG. 5. The zero order hydrolysis of carbobenzoxyglycyl-L-tryptophan by carboxypeptidase in the presence of different concentrations of benzylmalonic acid.

and we can predict the type of substances which would be expected to inhibit this enzyme (Smith, 1949c). The sensitivity of the known substrates indicates that the strongest binding at the R-position of the protein is given by those compounds that contain aromatic residues. It may also be anticipated that those compounds which contain carboxyl groups would bind with the metal. It was found that benzylmalonic acid is such a compound. It has two carboxyl groups and it has a side chain much like that of phenylalanine, and it is indeed a strong inhibitor of carboxypeptidase (Smith, 1949c; and unpublished). In Figure 5 the hydrolysis of carbobenzoxyglycyl-L-tryptophan is shown in the presence of different concentrations of the inhibitor. The hydrolysis of this substrate at 0.05 M follows zero order kinetics to at least 80 percent of completion. Figure 6 indicates that one mole of benzylmalonic acid combines with each active group of the enzyme. The inhibition is "un-competitive" in character. It is of interest that the unsubstituted

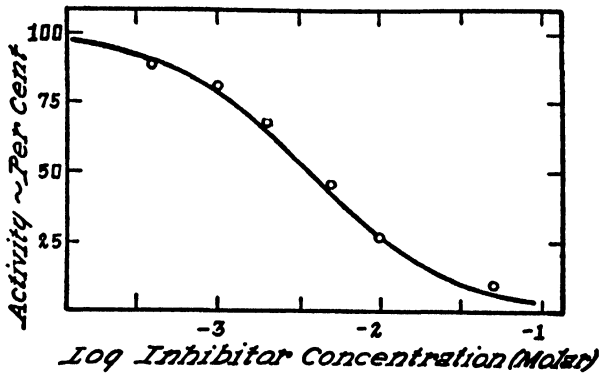


FIG. 6. Inhibition of carboxypeptidase by benzylmalonic acid. The theoretical curve is from the mass law equation for the combination of one mole of inhibitor with each active group of enzyme.

dicarboxylic acid, malonic acid, is an extremely poor and doubtful inhibitor of carboxypeptidase. Malonic acid at 0.5 M gives only 35 percent inhibition in contrast to the 50 percent inhibition produced by benzylmalonic acid at 0.003 M. The aromatic compound is about 250 times more effective. Similarly, other compounds which can block the R-group position and which bind with the metal also proved to be strong inhibitors.

From a series of inhibitors of different configuration, we can begin to evaluate and understand the precise limits of the configurations required for binding with the R-group position of the protein and with the metal simultaneously. Such a series is shown in Figure 7 for indole acetic, indole propionic and indole butyric acids. These plant growth hormones are all strong inhibitors of carboxypeptidase action. The effectiveness decreases with increasing distance between the carboxyl group and the indole ring. It should be noted that the curve drawn for the three sets of data is the theoretical one derived from the mass law equation for the combination of one mole of inhibitor per active group of enzyme. Another series of compounds which

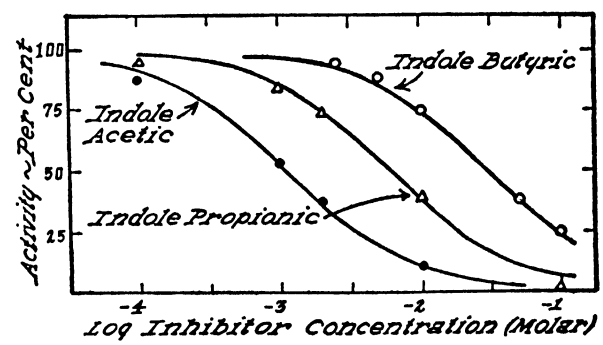


FIG. 7. Inhibition of carboxypeptidase by indole acetic, indole butyric and indole propionic acids. The same theoretical curve is drawn through the three sets of data as in Fig. 6.

we have studied is the phenyl substituted aliphatic acids: benzoic, phenylacetic, phenylpropionic and γ-phenylbutyric. The most effective inhibitor is phenylpropionic with a diminished effectiveness for the greater or smaller distance between the carboxyl and the phenyl residue.

It should be noted that Elkins-Kaufman and Neurath (1949) have recently shown that certain carboxylic acids are good inhibitors of carboxypeptidase. Independently, they have found the inhibition by benzoic, phenylacetic and some other acids.

As already mentioned, the normal enzyme-substrate combination occurs through chelate binding of the metal and the polar groups of the substrate. It

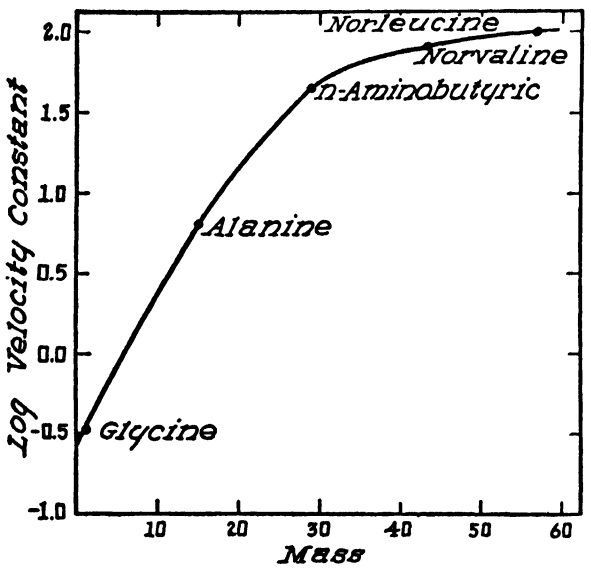


FIG. 8. The hydrolysis of straight chain aliphatic amino acid amides by leucine aminopeptidase. The logarithm of the relative first order velocity constant is plotted as a function of the molecular weight of the side chain bound to the α carbon (Smith and Polglase, 1949b).

is known from many studies of inorganic and organic compounds that where chelation occurs an extremely strong compound may be formed; where ring formation is not possible, only weak combination occurs. The effective chelation between metals and amino acids or peptides may be inferred from the striking experiments of Neuberg and Mandl (1948). They found that the metal ions known to be important in biological systems could not be precipitated as metal sulfides when amino acids or peptides were present.

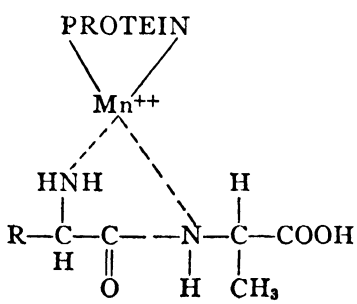
When a peptide is bound in chelate linkage with a metal ion, there is a definite shifting of electronic forces in order to form the ring structure. This is apparent from the enormously increased color given by Co^{++} in its combination with glycylglycine, or the alteration in absorption spectrum given by Mn^{++} with leucylglycine. It should be noted that the sensitive bond is in the ring structure since the necessary polar groups of the substrate are on the two sides of this linkage. It has been suggested that when such a combination occurs in the enzyme-substrate complex, a strong electronic pull is exerted by the metal through the polar groups of the substrate to produce a rearrangement of the electrons at the sensitive peptide linkage (Smith, 1949a). In other words, a large part of the role of the enzyme is to weaken the C-N bond or to increase the lability of the surrounding structure by the electronic pull through the metal. If this hypothesis is true, the actual hydrolysis of the peptide bond probably takes place through a conventional acidic or basic catalytic mechanism. Hydrogen and hydroxyl ions can catalyze the hydrolysis of peptide bonds. Ordinarily, this hydrolysis occurs only in strongly acid or strongly alkaline solutions, but with a lower free energy of activation through the mediation of the enzyme, the reaction involving the addition of OH^- and H^+ may cease to be rate-determining.

If this is true, then the enzymatic hydrolysis of the peptides should follow many of the relationships which govern acid and base catalysis. It is well-known from physical organic chemistry that with a series of related esters or amides the rate of the hydrolysis is a function of the strength of the acid which is linked in the ester or amide bond (Hammett, 1940). Comparison of the rate of hydrolysis of known substrates for carboxypeptidase shows that this is indeed correct. Chloroacetylphenylalanine is hydrolyzed much more rapidly than acetylphenylalanine. Obviously, chloroacetic acid is a far stronger acid than acetic acid. Similarly, it was reported by Hanson and Smith (1948) that carbobenzoxy-β-alanylphenylalanine is hydrolyzed about a thousand times more slowly than carbenzoxyglycylphenylalanine. Again it has been found that carbobenzoxy-β-alanine is a much weaker acid than carbobenzoxyglycine (Lumry, Polglase and Smith, unpublished). For carbobenzoxyglycine, pK′ is 2.75; for carbobenzoxy-β-alanine, pK′ is 3.85. Thus, a portion of the alleged specificity of this enzyme is not true specificity at all but is a function of the free energy of the bond which is exposed to the action of the enzyme (Smith, 1949a).

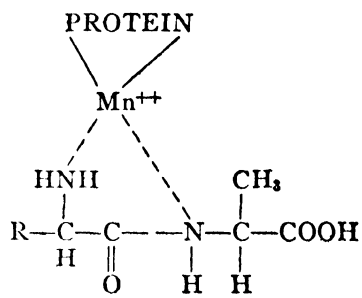
What may be termed a true specificity factor of the enzyme depends on the binding of the R groups by the protein itself and the contribution of this binding in lowering the free energy of activation. Such a study has been made for leucine aminopeptidase (Smith and Polglase, 1949a, b). In contrast to earlier assumptions that only leucine compounds are attacked, it has been found that the enzyme can hydrolyze a variety of peptides and amides containing different aliphatic residues. The simplest series is the homologous aliphatic amino acid amides. The data in Figure 8 show that the

rate of hydrolysis of the amide is a function of the molecular weight of the side chain. The initial slope indicates a linear proportionality of binding force and the mass of the residue attached to the alpha carbon. The significance of this will be discussed further.

Another factor which influences peptidase specificity appears to depend on a steric effect (Smith and Polglase, 1949b). With leucine aminopeptidase, L-leucinamide is hydrolyzed very rapidly. However, D-leucinamide is not hydrolyzed at all, at least within a factor of about five or ten thousand times the rate of splitting of the L compound. This difference between the L and D amides is obviously due to the polyaffinity relationship; namely, the isobutyl side chain of the D-amino acid amide lies in such a position that it cannot combine with the protein at the same time as there is combination of the amino group and the peptide nitrogen with the metal. When the relative rate of hydrolysis of L-leucyl-L-alanine is compared to that of L-leucyl-D-alanine a different optical effect is observed. Here the L-compound is hydrolyzed about 25 times more rapidly than its diastereoisomer, L-leucyl-D-alanine. The residue attached to the peptide nitrogen does not seem to be bound by the protein because the rates of hydrolysis for a variety of compounds are almost identical; L-leucinamide, L-leucyl-L-alanine, and L-leucyl-β-alanine are all hydrolyzed at about the same rate so that the presence or absence of a carboxyl group or the change in position from an α-amino group to a β-amino group has little or no effect. Since the residue attached to the nitrogen of the sensitive peptide bond does not appear to have any affinity for the protein, this must be interpreted as a steric effect in which the methyl group of the D-alanine lies in such a position as to interfere with the combination of the metal with the peptide nitrogen as suggested in the diagrams.



L-Leucyl-L-alanine

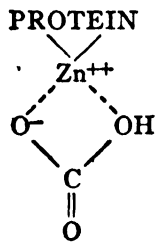


L-Leucyl-D-alanine

It would appear to be desirable at this point to summarize the previous discussion. For these metal-peptidases it has already been possible to distinguish five different factors which are concerned in the specificity of these enzymes. 1. The di-affinity concept of Euler and Josephson which we may relate to the ability of the metal to form a chelate complex involving the polar groups of the substrate. This is an absolute specificity factor since there can be no action whatsoever unless the chelate ring is formed. 2. There is the polyaffinity concept of Bergmann and his collaborators. The R group must lie in a definite spatial position in relation to the polar groups. In most cases, this would also appear to be an absolute specificity factor. 3. The free energy of the sensitive bond determines the *relative* sensitivity of the compound. This will depend on the acid strength of the carboxyl group, and on the energetic contribution of the amino group, imino group or ester group which is linked in the sensitive bond. It is noteworthy that the hydrolysis of an ester by carboxypeptidase discovered by Snoke, Schwert and Neurath (1948) is faster than the peptide hydrolysis. This is precisely what may be expected from the relative strengths of ester and peptide bonds. (It should be noted that the data on the hydrolysis of homologous peptides or amides and of esters by trypsin and chymotrypsin show this same relationship (Schwert, Neurath, Kaufman, and Snoke, 1948). This is suggestive that even in these non-metal enzymes similar forces are operative in the catalytic mechanism.) 4. The binding of the amino acid side chains is a relative specificity factor. 5. The steric effect mentioned above for the action of leucine aminopeptidase on the two forms of leucylalanine would also appear to be a relative specificity factor. Thus, we can already recognize at least five different effects in determining enzyme specificity. Undoubtedly, many more factors remain to be discovered.

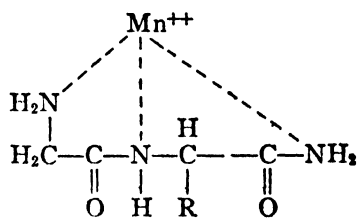
It is apparent that the above considerations may be used to interpret the action of other metal-enzymes. The notion of a minimal two-point combination to produce a ring structure in the enzyme-substrate complex seems to be capable of general application. However, certain interesting problems arise; one of these in the size of the ring structure. This might appear to offer certain difficulties since not too much is known about the binding characteristics of many of these metals, and it is ob-

vious that considerations which apply to the carbon ring compounds need not apply to the rings involving metal ions. Five or six membered rings offer no theoretical difficulties. However, carbonic anhydrase which is a zinc-enzyme (Keilin and Mann, 1940) must combine with bicarbonate ion to form a four-membered ring structure, uncommon though this interpretation appears to be.



As already indicated, the ring structure postulated for carboxypeptidase involves seven members. Fortunately, it is known that four and seven-membered rings can be formed when ionic bonds are involved (Johnson, 1943).

It was found by Smith and Slonim (1948) that Leucine aminopeptidase can hydrolyze compounds like glycyl-L-leucinamide to produce glycyl-L-leucine and ammonia. Here the essential free amino group is one peptide bond removed from the sensitive linkage, and we are presented with several possibilities. Either the metal combines with the free amino group and the amide nitrogen to make an eight membered ring, or the presence of the amino group changes the properties of the peptide nitrogen sufficiently to permit combination at this point forming a single five-membered ring; both of these possibilities appear to be rather unlikely. It appears more probable that a double chelate is formed in which all three nitrogens participate.



Ordinarily, one would expect the peptide nitrogen to give a weak binding but where bidentate ring structures are possible, the bonds are much stronger as beautifully demonstrated by Schwarzenbach and his co-workers (1947, 1948) in their recent studies of complex formation. It is interesting that the postulated coordination of Mn^{++} with glycylleucinamide is similar to the complex formed by glycylglycinamide and Co^{++}. In the latter instance, glycylglycine dipeptidase has no action on the dipeptide amide (Smith, 1948b) whereas the aminopeptidase hydrolyzes the terminal amide linkage of glycylleucinamide.

It is clear from all of the foregoing that we believe *primary* interaction between the enzyme and its substrate involves strong chemical bonds with the polar groups of the substrate. Additional examples are provided by the decarboxylation of the β-keto carboxylic acids studied by Kornberg, Ochoa and Mehler (1948) where the chelate complexes with metal ions were observed spectroscopically. It should be emphasized that with many enzymes, the chelate ring structure involving the metal and the polar groups may be the only interaction between the enzyme and substrate; this appears to be the case with catalase, carbonic anhydrase and others. It also should be noted that in many instances, labilization of the substrate may occur as the result of the complex formation with the metal in the complete absence of the specific protein. This occurs with some of the β-keto carboxylic acids and with the well-known metal-catalyzed oxidations of cysteine, ascorbic acid, and numerous other substances. In these instances, the electronic rearrangement produced by the combination with the metal alone can lower the free energy of activation sufficiently to permit the reaction to proceed. Many of these reactions appear to be ultimately caused by a basic catalysis.

As already mentioned, part of the role of the protein must be to increase the electronic effect of the metal. That binding to a protein may change the electronic character of a prosthetic group has been noted on many occasions as in the alteration of the redox potential of riboflavin phosphate as compared to that in the corresponding yellow enzyme. This is also the situation with the many heme or hematin proteins such as hemoglobin, catalase, and peroxidase, where the redox potentials vary considerably. Obviously, this depends largely on the protein since the prosthetic groups appear to be identical or very similar.

For the peptidases, there is no definite information as to the actual magnitude of the energetic contribution of the protein nor is there any information as to just how much of the protein molecule is involved. However, it is possible to make some reasonable inferences from present information by utilizing our knowledge concerning leucine aminopeptidase. Our highly purified but still inhomogeneous leucine aminopeptidase reacts with dilute solutions of Mn^{++} to produce a strong pink color. The same enzyme preparations show marked changes in the electrophoretic mobilities of the protein components in the presence of Mn^{++} (Smith, 1946). The changes in mobility indicate a suppression of ionic groups with the formation of at least one primary bond, and the color intensification indicates a coordinate structure. This reaction is slow judged both by the color development and the formation of active enzyme. Hence the free

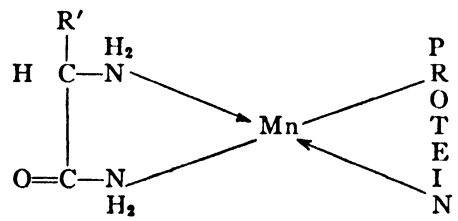
energy of activation must have a significant magnitude. If we exclude major changes in protein structure preceding the combination, the metal-protein bonds cannot be entirely electrostatic since such processes require no activation energy.

The standard free energy change can be calculated from the equilibrium constant of the Mn^{++} protein combination. This constant was found to be 2.5×10^4 per mole at 40° C (Smith, 1946). The resulting value for the free energy change is −6300 calories per mole which is of the expected order of magnitude. This value depends, of course, on the pH since this will influence the equilibrium constant.

The five factors which have already been indicated as influencing peptidase specificity can be regarded as of two essential types: those which are steric and influence the correct fit of the substrate with the various combination points, and those which are energetic in character. The combinations between protein and substrate appear to involve both primary and secondary bonds. First, let us see what can be determined concerning the nature of the primary bonds. A detailed study will not be feasible until the enzyme is obtained in homogeneous form. However, it is possible to study the interaction of the metal (Mn^{++}) and the substrate. In the absence of protein, Mn^{++} and leucylglycine interact to form a complex (Fig. 9) with an intense pink color (Smith, 1948b). Our observations, as yet unpublished, show that at pH 7.9, two molecules of the peptide react with a single ion of Mn^{++}. The reaction is exoergic and the product is stable. Indeed, the half-product composed of one molecule of peptide and one ion of Mn^{++} is just as stable. The unit of reaction colorimetrically is the half unit so that energetically the complete complex must be approximately equivalent to two half complexes.

Within the precision of our measurements, leucylglycine or leucinamide is not split in the presence of metal alone. It is as yet somewhat difficult to see how the properties of the metal ion which acts as a bridge between protein and substrate may be so altered that hydrolysis can easily occur. This is especially true in view of the low free energy change in the metal protein interaction. Rather, one would suspect that the protein forms bonds with the metal similar to those which exist between metal and substrate and that the enhanced reactivity of the intermediate enzyme-substrate complex is a matter of degree. This is certainly the case with the β-ketocarboxylic acids already mentioned, where the reaction of metal and substrate produces a complex in which the free energy of activation is lowered sufficiently for the reaction to proceed. A lower free energy of activation also occurs in the Mn^{++}-peptide complex since the spectral changes in the visible and ultraviolet indicate a rearrangement of electrons. It seems quite plausible that the complexes formed by the substrates of leucine aminopeptidase simply have a lower degree of reactivity which is improved by complexing with the protein. This idea is reinforced by considering the interaction of the protein with the side chains of the substrate.

From a number of considerations, the linking of Mn^{++} with protein and substrate appears to involve two covalent and two coordinate bonds with the most probable arrangement as shown in the diagram.



Considering only the hydrolysis of the aliphatic amino acid amides, we must conclude that the only differences between the metal-substrate complex and the enzyme-substrate complex is first the binding

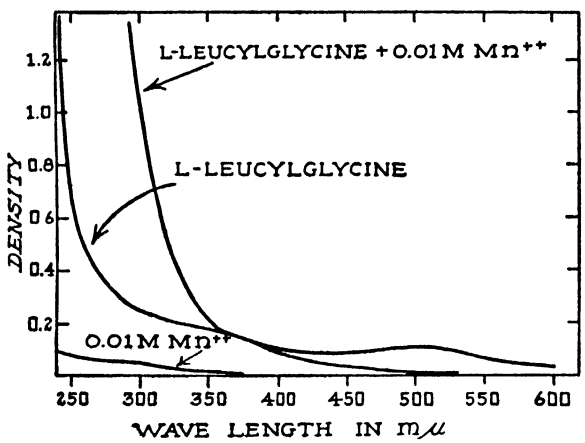


FIG. 9. Interaction of leucylglycine (0.125 M) with 0.01 M $MnCl_2$. The increase in the visible absorption spectrum is reflected in the pink color of the metal-peptide complex (Smith, 1948b).

or attachment of R′ by the protein and second, the possibility of secondary interaction by the protein with either the carbonyl oxygen or the amide hydrogen or both. There is a definite and fairly rapid splitting of glycinamide by the enzyme; consequently, the main protein effect cannot be due to the binding of R′ since in glycinamide only a hydrogen atom is available on the α carbon. Therefore, *we are forced to the conclusion that the protein has a definite interaction with one or both groups at the peptide bond.* At present, it would be fruitless to speculate concerning the nature of these forces. Nevertheless, this bonding at the peptide linkage by the protein must produce an additional contribution

towards the lowering of the free energy of activation. However, we must emphasize that this supplements the primary contribution derived from the bonding of the metal and the substrate.

Let us now consider what determines the change in rate which results from the modification of the aliphatic side chain; these data have already been presented (Fig. 8). It is well-known that rate and free energy of activation are related in a logarithmic manner, $\ln C \propto -\Delta F^*/RT$ where C is the proteolytic coefficient and ΔF^* is the change in free energy of activation. The linear portion of the data in Figure 8 where log C is proportional to molecular weight can be considered as a change in ΔF^* caused by the change in the size of the side chain.

From the classical Michaelis-Menten considerations of enzyme kinetics we know that two rate determining steps are possible: (1) where k_1 is the velocity constant which measures the rate of formation of the intermediate compound, and (2) where k_3 is the effective value of C since it measures the actual disappearance of the substrate. $K = k_1/k_2$

TABLE 1. BINDING ENERGY FOR ALIPHATIC SIDE CHAINS

The approximate side chain binding energy is taken as 2 times (latent heat—RT). The latent heats were obtained from the Landolt-Bornstein Tables (1923).

Substrate	Hydrocarbon analogue	Latent heat of vaporization	Approximated side chain binding energy
		Cals./mole	Cals./mole
Glycinamide	H_2	220	400
Alaninamide	CH_4	2,200	4,000
α-Aminobutyric acid amide	CH_3CH_3	3,800	6,850
Norvalinamide	$CH_3CH_2CH_3$	4,400	7,700
Norleucinamide	$CH_3CH_2CH_2CH_3$	5,300	9,500

$$\text{Enzyme} + \text{Substrate} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \text{Enzyme·substrate}$$

$$\text{Enzyme·substrate} \xrightarrow{k_3} \text{Products} + \text{Enzyme}$$

Since the acid strength of all of the aliphatic amino acids is practically identical, it appears quite certain that the bond strength of the various amides is identical. Therefore, it is likely that there are no significant variations in k_3 for these amides, since k_3 must be primarily determined by the strength of the bonds being hydrolyzed; the converse of this has already been indicated for the substrates of carboxypeptidase. In the present instance, what is actually being measured in the different rates of hydrolysis for these amides would appear to be alterations in K, the equilibrium constant for the formation of the intermediate compound or that fraction of ΔF_1 which contributes to k_1. ($\ln K = -\Delta F_1/RT$; $K = k_1/k_2$, thus $\Delta F_1 = \Delta F_1^* - \Delta F_2^*$.) We can be measuring alterations in rate which reflect the full change in free energy of binding or part of it. This will depend on whether the rate is being primarily determined by changes in K at equilibrium or in changes in k_1 under non-equilibrium conditions. This is not the place to assess the detailed equations which are involved in the two different conditions.

Can these secondary interactions of R′ with the protein be explained in terms of familiar low energy forces? Saturated hydrocarbon side chains can hardly be expected to interact with primary valence forces. In fact, saturated hydrocarbon molecules are known to interact among themselves only through Van der Waals "dispersion" forces (Margenau, 1939). We know that such Van der Waals forces generally involve small energies. Are they of sufficient magnitude to explain the observed differences in rates with the various amides. A good approximation of the binding energies for the R′ side chains which are all saturated hydrocarbons may be secured by comparing the latent heats of vaporization for hydrogen, methane, ethane, propane and butane at their boiling points (Table 1).

(This approximation yields too large values in that it neglects entropy changes, and substrate-solvent and protein-solvent bonds which must be broken in forming the intermediate compound. The fact that removal of translational and rotational contributions to the entropy depends primarily on the metal substrate bonds and that solvent-solvent interaction is not included reduces the importance of these two contributions. The approximation is too low in that it neglects the enforced propinquity of side chain and protein surface which will appreciably increase the time average interaction between protein and substrate side chain. The approximation is incorrect in so far as the local surface of the protein does not resemble a saturated hydrocarbon. What is important is that it is probably too small rather than too large.)

The logarithm of the velocity constants versus these approximate binding energies shows a hyperbolic curve similar to that obtained when log rate is plotted against molecular size of the side chains (Fig. 10). The initial slope indicates a variation of log rate with free energy of binding of -0.8×10^{-4} moles per calorie. This should be compared with a theoretical value which can be derived from the Michaelis-Menten mechanism. The first order (in substrate) proteolytic coefficient, C, may be expressed in this manner.

$$C = k_3 \cdot k_1/k_2 \quad \text{or} \quad C = k_3K$$

Substituting the modified Boltzmann function for k_3 (Glasstone, Laidler, and Eyring, 1941) where $k_3 = RT/Nh\ e^{-\Delta F_3^*/RT}$ and the thermodynamic function $K = e^{-\Delta F_1/RT}$ for K gives

$$C = \frac{RT}{Nh} e^{-\Delta F_3^*/RT} e^{-\Delta F_1/RT}$$

R is the gas constant, h is Planck's constant, T is the absolute temperature, N is Avogadro's number, ΔF_3^* is the free energy of activation for reaction 3, ΔF_1 is the change in free energy. Taking logarithms, we obtain the change in forming the intermediate compound.

$$\ln C = \ln \frac{RT}{Nh} - \frac{\Delta F_3^*}{RT} - \frac{\Delta F_1}{RT}$$

Differentiating this equation leads to

$$\frac{d \ln C}{d\Delta F_1} = -\frac{1}{RT}\frac{d\Delta F_3^*}{d\Delta F_1} - \frac{1}{RT}$$

However, we must assume that $d\Delta F_3^*/d\Delta F_1$ is equal to zero since we have assumed that k_3 does not change with the different amide substrates. Hence

$$\frac{d \ln C}{d\Delta F_1} = -\frac{1}{RT} = -1.6 \times 10^{-4} \text{ moles per cal.}$$

This may be compared with the value of -0.8×10^{-4} obtained from the data of latent heats. Since it is the reciprocals of these values which determine the increase in reaction velocity which is to be expected with a given increase in ΔF, at least for the first three members of the series, physical adsorption forces can actually supply more energy than is needed to explain the observed changes in reaction rate. This is probably true regardless of the precise meaning of our velocity constants C providing we consider only that range of enzyme activity in which the reaction is first order in substrate concentration.

Assuming that the rapidly rising portion of the curve is due to an increased Van der Waals interaction, there appear to be a number of possible mechanisms by which the leveling off at higher chain lengths can be explained. If the side chain interacts with a group which extends from the surface of the protein, it is obvious that the ends of the side chains will have a diminishing effect as their distance from such a group is increased. Another explanation would be that the relative increase in thermal energy of the larger residues over the increase in Van der Waals forces would result in a larger time-average distance of separation of the residue from the protein surface. Present knowledge does not permit a choice of mechanisms. We mention these only to point out that simple explanations are possible.

It would appear that we can now begin to investigate the detailed picture of the mechanism of action of these enzymes. Considerable work will be necessary to test some of the ideas which have been presented. Nevertheless, it would seem that a reasonable approach to some of these problems is at hand. Perhaps our concepts of the action of certain metal binding enzymes can be summarized in the following way. First, in the absence of enzyme, metal interaction to form chelates can, in some cases, lower the free energy of activation sufficiently to permit the reaction to proceed. Examples are the decarboxylation of some β-keto carboxylic acids. Second, there appear to be reactions of metal-enzymes where the only role of the protein is to effect an alteration in the electronic properties of the metal. This is a matter of degree. Such enzymes as carbonic anhydrase, catalase, and β-keto decarboxylases probably work in this manner. Third, there are those enzymes, which appear to include the vast majority of cases, where the protein, in addition to its combination with the metal ion, has one or several points of interaction directly with the substrate. These may involve both primary and secondary forces.

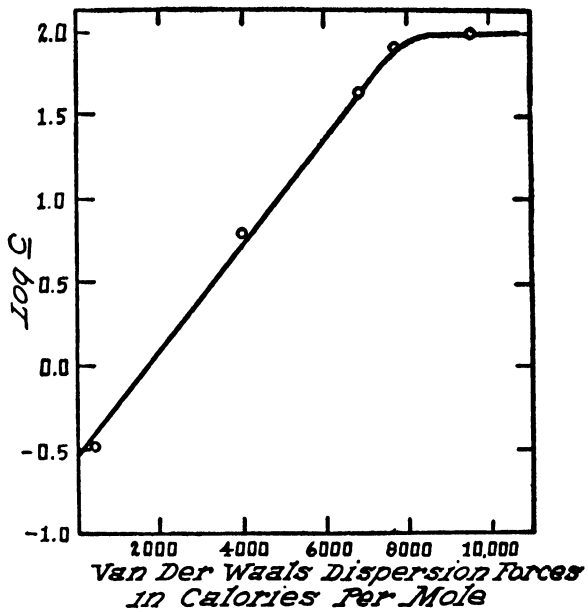


FIG. 10. The logarithm of the first order constants for the hydrolysis of the aliphatic amino acid amides by leucine aminopeptidase (same data as in Fig. 8) is plotted as a function of the approximated Van der Waals energy for the hydrocarbons which are analogous to the side chains of the amino acid amides (see Table 1).

According to our present picture, none of the bonds between substrate and enzyme, even when bridged by metal ions, appears to have special properties not found in the behavior of simpler chemical substances. No individual bond contributes all of the lowering of the free energy of activation while the others only determine specificity. Rather, it is likely that specificity is a measure of the number of structures on the protein which contribute to

the reduction in activation energy. Perhaps the only exception is steric hindrance where the contribution to the activation energy can be considered to be positive. Only a minor part of the internal energetic structure of the protein appears to be involved in any significant way in the enzymes that have been considered. If so, why is the protein of such high molecular weight? The large size of the protein molecule may be at least partly explained as a steric necessity in order to prevent two moles of the protein from combining with the metal ion; such a combination would not permit substrate interaction. This, of course, is in addition to permeability and solubility considerations.

In conclusion, the protein appears to do no more than several different kinds of smaller molecules can do in enhancing the rate of some reaction. The protein has the exceptional property of combining in a single entity all of the influences of a number of smaller molecules, and thus demonstrates a greater catalytic power than any of the smaller molecules. Hence, it becomes apparent that enzyme catalysis which has always been considered to be a special phenomenon, may be explained as a summation of several individual properties which are concentrated in a single large entity.

References

Bergmann, M., Zervas, L., Fruton, J. S., Schneider, F., and Schleich, H., 1935, On proteolytic enzymes V. On the specificity of dipeptidase. J. biol. Chem. *109:* 325-346.

Chance, B., 1943, The kinetics of the enzyme-substrate compound of peroxidase. J. biol. Chem. *151:* 553-577.

1947, An intermediate compound in the catalase-hydrogen peroxide reaction. Acta Chem. Scand. *1:* 236-267.

Elkins-Kaufman, E., and Neurath, H., 1949, Structural requirements for specific inhibitors of carboxypeptidase. J. biol. Chem. *178:* 645-654.

Euler, H. v., and Josephson, K., 1926, Enzymatische Spaltung von Dipeptiden. II. Z. physiol. Chem. *157:* 122-139.

Glasstone, S., Laidler, K. J., and Eyring, H., 1941, The theory of rate processes, New York: McGraw-Hill Book Co.

Hammett, L. P., 1940, Physical organic chemistry. New York: McGraw-Hill Book Co.

Hanson, H. T., and Smith, E. L., 1948, The application of peptides containing β-alanine to the study of the specificity of various peptidases. J. biol. Chem. *175:* 833-848.

Hofmann, K., and Bergmann, M., 1940, The specificity of carboxypeptidase. J. biol. Chem. *134:* 225-235.

Johnson, J. R., 1943, Modern electronic concepts of valence. Organic Chemistry an Advanced Treatise, Volume *2:* 1821-1942, 2nd ed. New York: John Wiley and Sons, Inc.

Johnson, M. J., and Berger, J., 1942, The enzymatic properties of peptidases. Advances in Enzymology and Related Subjects. *2:* 69-92.

Keilin, D., and Mann, T., 1937, On the haematin compound of peroxidase. Proc. Roy. Soc. London, Series B, *122:* 119-133.

1940, Carbonic anhydrase. Purification and nature of the enzyme. Biochem. J. *34:* 1163-1176.

Kornberg, A., Ochoa, S., and Mehler, A. H., 1938, Spectrophotometric studies on the decarboxylation of β-keto acids. J. biol. Chem. *174:* 159-172.

Landolt-Bornstein, 1923, Physikalisch-Chemische Tabellen, Berlin: Julius Springer.

Margenau, H., 1939, Van der Waals forces. Revs. Modern Phys. *11:* 1-35.

Michaelis, L., and Menten, M. L., 1913, Die Kinetik der Invertinwirkung. Biochem. Z. *49:* 333-369.

Neuberg, C., and Mandl, I. 1948, An unknown effect of amino acids. Arch. Biochem. *19:* 149-161.

Schwarzenbach, G., and Ackermann, H., 1948, Komplexone XII. Die Homologen der Äthylendiamintetraessigsäure und ihre Erdalkalikomplexe. Helv. Chim. Acta *31:* 1029-1048.

Schwarzenbach, G., Willi, A., and Bach, R. O., 1947, Komplexone IV. Die Acidität und die Erdalkalikomplexe der Anilin-diessigsäure und ihre Substitutionsprodukte. Helv. Chim. Acta *30:* 1303-1320.

Schwert, G. W., Neurath, H., Kaufman, S., and Snoke, J. E., 1948, The specific esterase activity of trypsin. J. biol. Chem. *172:* 221-239.

Smith, E. L., 1946, Manganese and L-leucine-aminoexopeptidase. J. biol. Chem. *163:* 15-27.

1948a, Action of carboxypeptidase on peptide derivatives of L-tryptophan. J. biol. Chem. *175:* 39-47.

1949a, The mode of action of the metal-peptidases. Proc. nat. Acad. Sci., Wash. *35:* 80-90.

1949b, Proteolytic enzymes. Ann. Rev. Biochem. *18:* 35-58.

1949c, The catalytic action of the metal-peptidases, Federation Proc. *8:* 581-588.

Smith, E. L., and Bergmann, M., 1944, The peptidases of intestinal mucosa. J. biol. Chem. *153:* 627-651.

Smith, E. L., and Hanson, H. T., 1949, The chemical nature and mode of action of pancreatic carboxypeptidase. J. biol. Chem. *179:* 802-813.

Smith, E. L., and Polglase, W. J., 1948a, The optical and side-chain specificity of leucine aminopeptidase. Federation Proc. *8:* 252.

1948b, The specificity of leucine aminopeptidase II. Optical and side chain specificity. J. biol. Chem. *180:* 1209-1223.

Smith, E. L., and Slonim, N. B., 1948, The specificity of leucine aminopeptidase. J. biol. Chem. *176:* 835-841.

Snoke, J. E., Schwert, G. W., and Neurath, H., 1948, The specific esterase activity of carboxypeptidase. J. biol. Chem. *175:* 7-13.

CHROMATOGRAPHIC DETERMINATION OF THE AMINO ACID COMPOSITION OF PROTEINS

WILLIAM H. STEIN AND STANFORD MOORE
The Rockefeller Institute for Medical Research, New York

There can be little doubt that the present extensive and fruitful use of chromatographic methods for the separation of water-soluble substances has been stimulated, to a very large degree, by the pioneer investigations of Martin and Synge and their co-workers. They developed the silica gel column for acetyl amino acids, the now famous paper chromatograms, and introduced the use of starch columns (Elsden and Synge, 1944; Synge, 1944). Our own work with chromatographic methods, like that of many others, has been a logical outgrowth of these earlier investigations.

In our studies, the first objective was to ascertain whether chromatographic methods could provide a convenient and accurate means to determine quantitatively the amino acid composition of proteins. For accurate work of this kind, column chromatography appeared to be more promising than paper chromatography. In addition, a column possesses the advantage that it can be operated on a scale sufficiently large to permit, where necessary, isolation of 20 mg. or more of a component. Accordingly, the starch column was chosen for development (Moore and Stein, 1948, 1949; Stein and Moore, 1948).

In Figure 1 is shown the result which may be obtained when a complex mixture of amino acids is fractionated on a starch column. In this case the mixture was made up to simulate an acid hydrolysate of bovine serum albumin and contained 18 components—17 amino acids and NH_4Cl. For this type of experiment, the effluent is divided into a regular series of very small fractions of known volume. These fractions are analyzed quantitatively. The data obtained permit the construction of effluent curves, such as the one shown, which reveal the detailed behavior and full resolving power of the column.

In the laboratory, the packed column is mounted over an automatic fraction collector which has been designed to collect continuously small fractions of specified volume (Stein and Moore, 1948). The basic parts of the machine are a turntable holding the receiving tubes and a photo-electric drop counting circuit which controls the amount of effluent delivered per tube. The curve shown in Figure 1 required the collection of about 400 0.5 cc. fractions, over a period of seven 24 hour days, and it is thus apparent that automatic machinery is essential before any considerable number of experiments of this type can be undertaken.

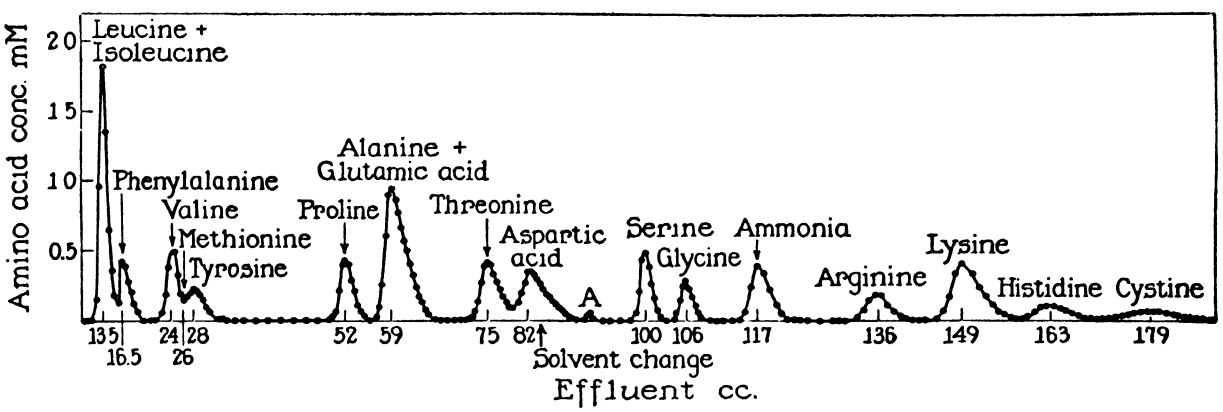


FIG. 1. Separation of amino acids from a synthetic mixture containing seventeen amino acids and ammonium chloride. Solvent, 1:2:1 *n*-butyl alcohol-*n*-propyl alcohol-0.1N HCl, followed by 2:1 *n*-propyl alcohol-0.5N HCl. Column, 13.4 gm. of starch (anhydrous); diameter, about 0.9 cm.; height, about 30 cm. Sample, about 3 mg. of amino acids. The effluent was collected in 0.5 cc. fractions. *A* is a small artifact peak.

At the beginning of the experiment the sample containing the amino acid mixture is added in a few tenths of a cc. to the top of the column. The sample is allowed to drain into the starch, and fresh solvent, containing no amino acids, is added and run through under slight pressure. The load on a column one cm. in diameter is about 2.5 mg. of an amino acid mixture. The apparatus and the procedures have been designed so that, when necessary,

the load can be increased about 100 fold on columns 8 cm. in diameter for experiments on a preparative scale.

The next step in the procedure is the analysis of the numerous fractions collected for their amino acid content. The amino acid concentration is determined by a simple photometric ninhydrin method which has been developed for this purpose (Moore and Stein, 1948). It is necessary for the analytical method to be highly sensitive, accurate, and readily adapted to the handling of hundreds of individual analyses. The ninhydrin method is sensitive to one part per million of amino acid in aqueous or alcoholic solution, and is accurate to ± 2 percent for two to three micrograms of amino acid nitrogen.

After the mechanics of the procedure were worked out, it became necessary to investigate the behavior of various solvent mixtures. The immediate objective in these studies was the development of quantitative methods for the determination of the amino acid composition of protein hydrolysates. For this purpose a great many combinations of solvents were investigated before the most convenient systems were established. It was found that a wide variety of neutral, acidic, and buffered solvent mixtures can be employed with starch columns. Dilute HCl mixtures, however, have proved to be the most useful for separating the common amino acids. In the curve shown in Figure 1, the chromatogram was started with a solvent composed of 1 part normal butyl alcohol, 2 parts normal propyl alcohol, and 1 part aqueous 0.1N HCl. After the emergence of aspartic acid, the solvent running through the column was changed to a mixture of 2 parts of propyl alcohol and 1 part 0.5N HCl, which accelerates the movement of the remaining amino acids. In this experiment, and all others in which known mixtures were employed, the positions of the various amino acid peaks were first established by a series of chromatograms performed with mixtures of increasing complexity.

It is important to note that in the experiment shown in Figure 1 the position on the effluent curve of a given amino acid is constant to ±5 to 10 percent, and is not influenced by the amino acid composition of the mixture fractionated. Proline, for example, emerges at the same place whether it is alone on the column, or whether it is originally admixed with 18 other components. Even more reproducible than the absolute position of a peak on the abscissa, is the position of a peak relative to its neighbors. This constancy of relative position leads to a pattern of peaks which is characteristic for each solvent, is readily recognized, and is, therefore, extremely useful when dealing with unknown mixtures.

The curve in Figure 1 illustrates the high resolving power of which the starch column is capable. The synthetic mixture of amino acids chromatographed in this instance contained 18 components and yielded 14 peaks. Nevertheless, the resolution is not satisfactory in all cases. Leucine, isoleucine,

TABLE 1. RECOVERY OF AMINO ACIDS FROM KNOWN MIXTURE CONTAINING 18 COMPONENTS

Solvents, 1:2:1 *n*-butyl alcohol-*n*-propyl alcohol-0.1N HCl followed, after the emergence of aspartic acid, by 2:1 *n*-propyl alcohol-0.5N HCl (*cf.* Fig. 1).

Constituent	Amount present	Percent recovery			
		Chromatogram 456	Chromatogram 457	Chromatogram 520	Average
	mg.				
Leucine-isoleucine	0.364	99.4	99.5	101.5	100.3
Phenylalanine	0.165	94.8	96.1	94.8	95.2
Valine-methionine-tyrosine	0.354	99.6	101.0	100.1	100.2
Proline	0.136	99.7	97.8	100.0	99.2
Glutamic acid*-alanine	0.515	95.2	94.6	96.8	95.5
Threonine	0.201	97.5	101.0	102.0	100.2
Aspartic acid*	0.267	93.5	94.1	94.7	94.1
Serine	0.118	100.0	99.8	101.2	100.3
Glycine	0.051	99.1	100.5	101.0	100.2
Ammonia	0.024	102.0	99.5	104.5	102.0
Arginine	0.143	97.7	102.8	105.0	101.8
Lysine	0.302	96.3	103.0	99.5	99.6
Histidine	0.094	99.7	104.6	97.4	100.6
Cystine	0.133	89.5	102.7	101.5	97.9
All constituents	2.867	97.3	99.3	99.6	98.7

* When the value for glutamic acid is corrected for the 7 percent low recovery due to esterification, the recoveries for glutamic acid plus alanine become 100.2, 99.7, and 101.7 percent. The aspartic acid recoveries, which run 6 percent low, may be similarly corrected to yield the figures 99.4, 99.9, and 100.8 percent. The total recoveries, on this basis, become 98.6, 100.8, and 101.0 percent

and phenylalanine are not well separated. The same is true of valine, methionine, and tyrosine. Glutamic acid and alanine emerge together. A satisfactory degree of resolution has been obtained for the remaining ten components—proline, threonine, aspartic acid, serine, glycine, NH_4Cl, arginine, lysine, histidine, and cystine. Integration of the curves gives quantitative recoveries of most of the components, as can be seen in Table 1. The overlapping amino acids are integrated as a group. For most of the single peaks, the average recoveries are accurate to ±3 percent. A low yield is obtained of aspartic acid and of glutamic acid + alanine. The low recoveries result from the loss of about 6 percent of the aspartic and 7 percent of the glutamic acid by esterification in the acidic alcoholic solvents. The degree of esterification is independent of the amino acid concentration, and hence, with unknowns, accurate aspartic acid figures may be obtained by incorporating a 6 percent correction factor into the final result. It may be noted that a single chromatogram requires about 3 mg. of total amino acid mixture, corresponding to between 5 and 30 hundredths of a mg. of each component. The recovery for the sum of the components in these experiments is 99 percent. For many investigations, this single type of chromatogram sometimes provides sufficient information in spite of the overlaps.

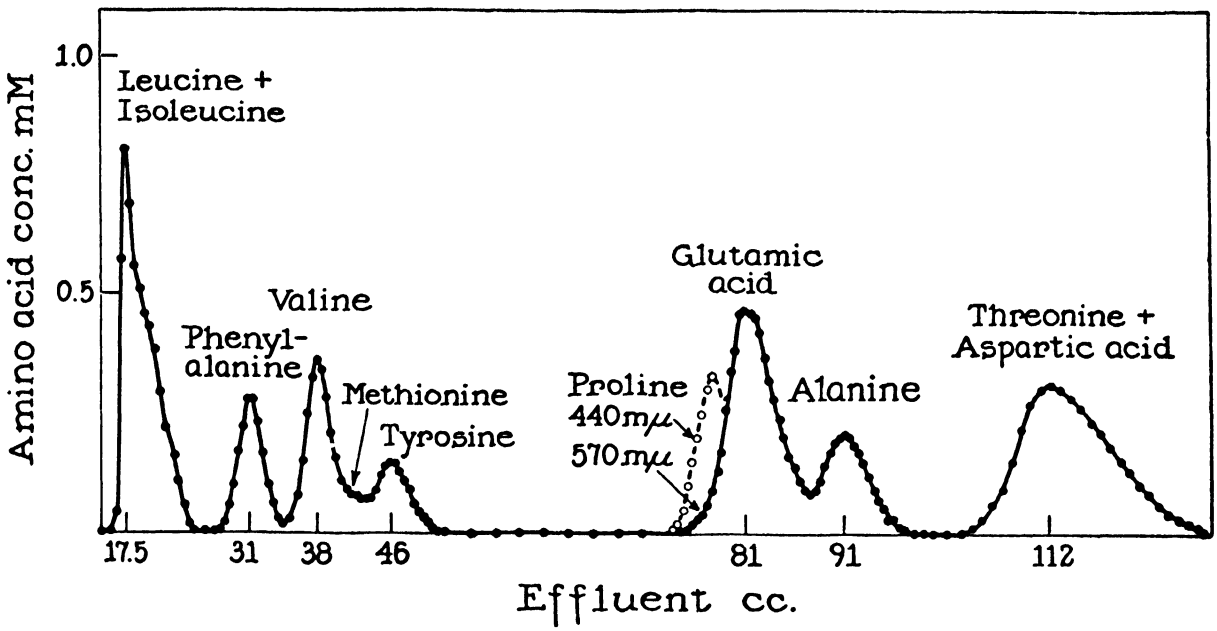


FIG. 2. Separation of glutamic acid, alanine, and other amino acids from a synthetic mixture containing 18 components. Solvent, 2 : 1 : 1 *tert*-butyl alcohol-*sec*-butyl alcohol-0.1N HCl. In other respects, the experimental conditions were the same as those described in Fig. 1.

It is possible to resolve the overlaps present in the curve shown in Figure 1 by the appropriate choice of solvents. Thus, glutamic acid and alanine, as may be seen in Figure 2, are separated by the use of a mixture of *tert*- butyl alcohol, *sec*- butyl alcohol and 0.1N HCl in the proportions of 2:1:1.

TABLE 2. RECOVERY OF GLUTAMIC ACID, ALANINE, AND OTHER AMINO ACIDS FROM SYNTHETIC MIXTURE
Solvent, 2:1:1 *tert*-butyl alcohol-*sec*-butyl alcohol-0.1N HCl (*cf.* Fig. 2). The mixture contained 18 components (*cf.* Table 1).

Constituent	Amount present	Percent recovery			
		Chromatogram 474	Chromatogram 543	Chromatogram 481	Average
Leucine-isoleucine	0.373	99.0	100.8		99.9
Phenylalanine	0.169	101.6	103.6		102.6
Valine-methionine-tyrosine	0.363	100.6	104.4		102.5
Glutamic acid	0.426	96.3	97.8	100.2	98.1
Alanine	0.102	97.3	101.3	97.5	98.7

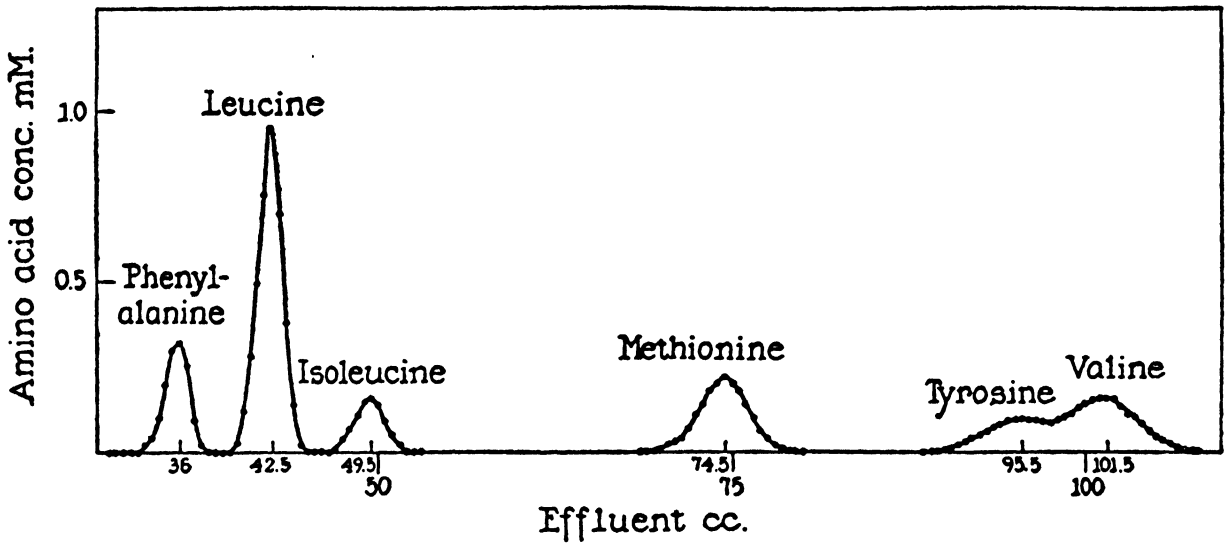


FIG. 3. Separation of amino acids from a synthetic mixture containing 18 components. Solvent, 1 : 1 : 0.288 *n*-butyl alcohol-benzyl alcohol-water, containing 0.5 percent thiodiglycol. In other respects, the experimental conditions were the same as those described in Fig. 1.

In addition to glutamic acid and alanine, this experiment yields a determination of phenylalanine. The quantitative data obtained from a chromatogram of this type are shown in Table 2. Glutamic acid is not esterified in this solvent, as evidenced by its quantitative recovery.

The overlaps still remaining from the chromatogram shown in Figure 1 can be resolved by the use of a third solvent composed of butyl and benzyl alcohol in equal proportions, 12 percent water and 0.5 percent thiodiglycol (Fig. 3). Thiodiglycol protects the sulfur of methionine from oxidation by traces of impurities present in the benzyl alcohol. It may be noted that in this solvent phenylalanine, leucine, and isoleucine emerge as well defined peaks. The separation of tryrosine and valine is not optimum, but is sufficiently good to permit calculation of the amount of each present. The quantitative data obtained from experiments such as the one shown in Figure 3 are given in Table 3.

Thus, by the use of three chromatograms, it is possible to obtain quantitative values for all of the 18 components of the synthetic mixture. The three columns are run with the butyl-propyl alcohol-0.1N HCl solvent, the butyl-benzyl alcohol mixture, and the tertiary butyl-secondary butyl alcohol solvent.

It should be mentioned that D, L, and DL amino acids have identical rates of travel on starch columns. This fact has been established in experiments with the L and DL forms of 12 of the amino acids discussed.

Attention can now be focussed on the application of the starch column technique to the determination of the quantitative amino acid composition of *purified* proteins. It should be noted that the word is "purified" not "pure." For analysis the protein

TABLE 3. RECOVERY OF AMINO ACIDS FROM KNOWN MIXTURE CONTAINING 18 COMPONENTS

Solvent, 1:1:0.288 *n*-butyl alcohol-benzyl alcohol-water (*cf.* Fig. 3). The composition of the mixture was similar to that given in Table 1.

Amino acid	Percent recovery				
	Chromatogram 270	Chromatogram 293	Chromatogram 338	Chromatogram 390	Average
Phenylalanine	103.6	101.8	98.4	99.4	100.8
Leucine	101.0	101.6	102.0	101.0	101.4
Isoleucine	102.0	103.6	100.0	102.8	102.1
Methionine			100.0*	100.6*	100.3
Tyrosine	106.4	100.6	101.1	102.7	102.7
Valine	96.7	99.0	100.1	100.8	99.2

* The solvent contained 0.5 percent thiodiglycol.

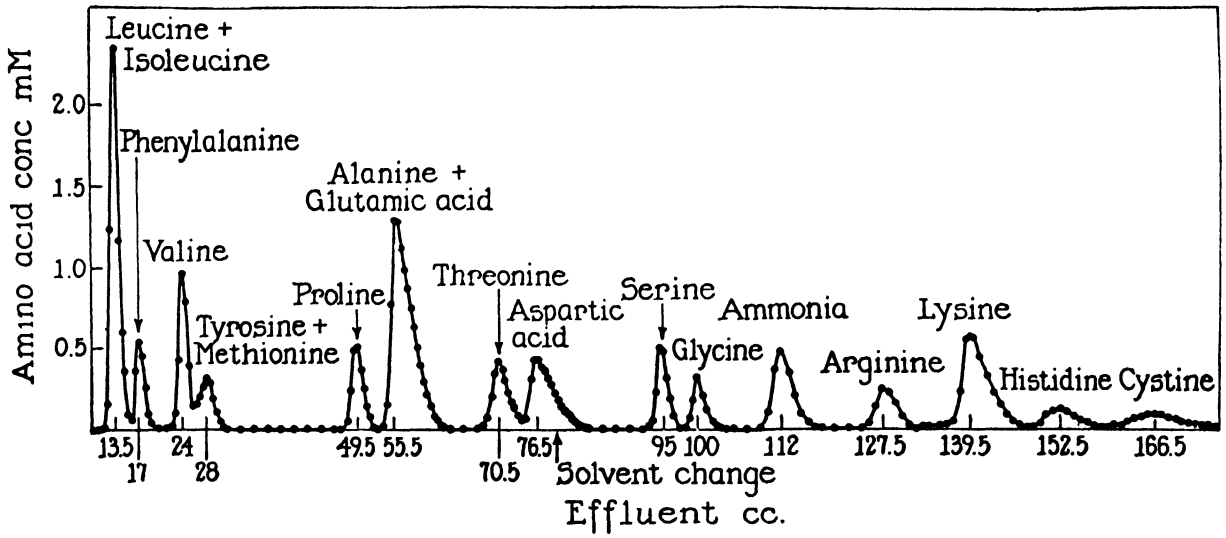


FIG. 4. Chromatographic fractionation of a hydrolysate of bovine serum albumin. Sample, an amount of hydrolysate corresponding to about 2.5 mg. of protein. The experimental conditions were the same as those described in Fig. 1.

preparations were hydrolyzed for 16 hrs. with 200 volumes of boiling 6N HCl, the excess HCl concentrated off, and the residue made up to a known volume. The chromatographic analyses were carried out on aliquots of the hydrolysate in exactly the manner described for the known amino acid mixtures. β-Lactoglobulin and bovine serum albumin have been analyzed by this procedure (Stein and Moore, 1949). The present discussion will be restricted to the latter. An effluent curve obtained with a hydrolysate of crystalline bovine serum albumin is given in Figure 4.

In the application of the chromatogram to unknown mixtures, the most important question concerns the correct identification of the peaks on the curve. In order to render the identification absolutely certain, it would be necessary to isolate the material responsible for each peak and to identify each substance by elementary analysis, optical rotation, or other means. The starch column can offer a means of accomplishing this end. But the isolation of all the components presents a formidable task under the best of circumstances. For analytical work it is worth while to consider whether a reason-

TABLE 4. CHROMATOGRAPHIC ANALYSES OF HYDROLYSATES OF BOVINE SERUM ALBUMIN

Solvents, 1:2:1 *n*-butyl alcohol-*n*-propyl alcohol-0.1N HCl and 2:1 *n*-propyl alcohol-0.5N HCl (*cf.* Fig. 4).

Constituent	Chromatogram 480		Chromatogram 481		Chromatogram 488	
	Gm. amino acid per 100 gm. protein	N as percent of protein N	Gm. amino acid per 100 gm. protein	N as percent of protein N	Gm. amino acid per 100 gm. protein	N as percent of protein N
Leucine-isoleucine	14.4	9.58	14.75	9.80	14.3	9.51
Phenylalanine	6.32	3.34	6.44	3.40	6.52	3.44
Valine-methionine-tyrosine		7.42 *ca.*		7.37 *ca.*		7.20 *ca.*
Proline	4.65	3.52	4.75	3.60	4.85	3.67
Glutamic acid-alanine		14.90 *ca.*		15.25 *ca.*		15.40 *ca.*
Threonine	5.44	3.98	5.72	4.19	5.50	4.05
Aspartic acid	10.91	7.15	10.86	7.12	10.96	7.19
Serine	3.91	3.24	3.76	3.12	3.77	3.13
Glycine	1.85	2.15	1.75	2.03	1.85	2.15
Ammonia	1.03	5.28	1.08	5.54	1.06	5.44
Arginine	5.96	11.92	6.03	12.07	5.72	11.45
Lysine	12.70	15.15	12.62	15.05	13.15	15.68
Histidine	4.29	7.24	3.67	6.18	4.04	6.82
Cystine	5.83	4.23	5.81	4.22	6.08	4.41
Total nitrogen recovery		99.1		98.9		99.5

ably sound answer can be arrived at without isolation. We believe that, in favorable cases, certain considerations and certain experiments make it possible to identify correctly the peaks on the curve with a very high degree of probability. The analysis of an acid hydrolysate of protein material is a favorable case. So much work has already been done with proteins, that it is possible to have a fair idea of what to expect. Moreover, the material analyzed has been subjected to exhaustive acid hydrolysis, thus virtually eliminating the possible presence of peptides.

TABLE 5. AMINO ACID COMPOSITION OF BOVINE SERUM ALBUMIN

The values for phenylalanine, leucine, isoleucine, tyrosine, and valine are from chromatograms carried out with 1:1:0.288 *n*-butyl alcohol-benzyl alcohol-water containing 0.5 percent thiodiglycol. Glutamic acid and alanine were determined with 2:1:1 *tert*-butyl alcohol-*sec*-butyl alcohol-0.1N HCl. The remaining chromatographic values are the average figures from Table 4. The nitrogen content of the protein was 16.07 percent, on an ash- and moisture-free basis.

Constituent	Gm. amino acid per 100 gm. protein	Gm. amino acid residue per 100 gm. protein	N as percent of protein N
Phenylalanine	6.59	5.87	3.48
Leucine	12.27	10.58	8.17
Isoleucine	2.61	2.25	1.74
Methionine	0.81*	0.71	0.47
Tyrosine	5.06	4.56	2.44
Valine	5.92	5.01	4.41
Proline	4.75	4.00	3.60
Glutamic acid	16.50	14.49	9.78
Aspartic acid	10.91	9.44	7.15
Alanine	6.25	4.99	6.12
Threonine	5.83†	4.95	4.27
Serine	4.23†	3.51	3.52
Glycine	1.82	1.38	2.11
Arginine	5.90	5.29	11.80
Lysine	12.82	11.25	15.30
Histidine	4.00	3.54	6.75
Cystine+cysteine	6.52‡	5.54	4.73
Tryptophan	0.58§	0.53	0.50
Amide-NH_3	0.95‖		4.87
Total		97.9	101.2

* The value for methionine is that determined by Brand *et al.* (1944, 1946). The chromatograms gave a figure of about 0.92.

† The average threonine and serine values of 5.55 and 3.81 from Table 4 have been divided by 0.95 and 0.90, respectively, in accordance with the estimates of Rees (1946) for decomposition of these amino acids during hydrolysis.

‡ The cystine+cysteine value is that determined by Brand *et al.* (1944, 1946). The chromatograms gave an average value of 5.91.

§ The tryptophan value is that determined by Brand *et al.* (1944, 1946).

‖ This figure is a maximum value for amide-NH_3 calculated from the total NH_3 of the hydrolysate corrected for the approximate amount of NH_3 formed on the decomposition of serine and threonine.

First of all, it will be noted that the curve (Fig. 4) obtained with the hydrolysate is similar to that obtained with a known mixture (Fig. 1). Careful inspection reveals that for both curves, the absolute positions of the peaks are the same (±5 percent) and the relative positions of the peaks are identical. The pattern is thus a familiar one. There are no new peaks on the curve. Hence there is no evidence for the presence in the hydrolysate of unsuspected amino-nitrogen containing substances.

Specifically, the proline peak is one which can be identified without question spectrophotometrically because of the reddish-yellow color which proline gives with ninhydrin. Other color tests can be used to help in the identification of amino acids such as tyrosine, histidine, arginine, and cystine. A more crucial experiment, however, is to add two or three amino acids to a sample of the hydrolysate and repeat the chromatogram. When threonine, serine, and histidine were added to this particular hydrolysate, the designated peaks rose without loss of symmetry and the added amount of amino acid was recovered quantitatively in each case. With these precautions, therefore, we feel that a curve, such as the one given in Figure 4, obtained on a hydrolysate of a purified protein can be interpreted with reasonable certainty.

The average values obtained by integration of the curve shown in Figure 4, and duplicates thereof, are given in Table 4. It can be noted that nearly all the nitrogen of the hydrolysate, 99 percent, has been accounted for. Individual determinations have been achieved for 10 of the components of the hydrolysate and the triplicate determinations are in good agreement. To complete the analysis for phenylalanine, leucine, isolueucine, methionine, tyrosine, and valine, butyl-benzyl alcohol columns were run. A third type of chromatogram run with the tertiary butyl alcohol-secondary butyl alcohol solvent yielded figures for glutamic acid and alanine. The amino acid composition of bovine serum albumin thus derived is given in Table 5.

The values in the first column are the average figures obtained chromatographically, with three exceptions. The values for tryptophan, methionine, and cystine are those of Brand and his associates (1944, 1946). Tryptophan is completely, and cystine partially destroyed under the hydrolytic conditions which we have used. Special hydrolytic methods are required before these amino acids can be determined, a problem with which we have not dealt. In the case of methionine, Brand's value is probably more accurate than the chromatographic figure, which was not determined under conditions optimum for a component present to less than one percent.

The figures at the bottom of the table are perhaps the most useful ones to emphasize at this time. In the third column it is seen that the total nitrogen

recovery is about 101 percent. In the second column, the total recovery in terms of amino acid residues, accounts for about 98 percent of the protein. Thus it is possible, primarily by chromatographic means, to obtain for a protein preparation a balance sheet which accounts for essentially all the amino acid components of the protein.

In the case of bovine serum albumin, Brand and his associates have previously estimated the amino acid composition, using microbiological assays for the largest part, and have determined all the amino acids except alanine. Other investigators have determined a good many of the components by various chemical, microbiological, and isotopic methods. A detailed comparison of these data with the results obtained by chromatography has already been given (Stein and Moore, 1949) and need not be repeated here. It is heartening to find that, for the most part, the analyses by the chromatographic and other methods are in good agreement. In analytical chemistry it is always advantageous and reassuring to have more than one way of getting the answer.

The experiments discussed thus far have dealt with purified protein preparations. Some times, however, it is desirable to determine the amino acid composition of the proteins of a whole tissue or a tissue extract. The application of the chromatographic techniques to the analysis of unfractionated biological material can be exemplified by the following experiments on tubercle bacilli. Through the generous cooperation of Dr. Gardner Middlebrook and Dr. Rene Dubos, it has been possible to chromatograph, for comparative purposes, hydrolysates of virulent and avirulent tubercle bacilli (Fig. 5). For these experiments the simplest possible procedure was followed, namely, the whole bacilli after being filtered and washed, were suspended in 6N HCl and refluxed for 18 hours. The humin and insoluble fat were filtered off and an aliquot of the complete hydrolysate chromatographed. The presence of non-amino nitrogen containing components offered no interference on the chromatogram. An alternative procedure, which has been employed on tissue proteins and bacterial cells, is to hydrolyze a trichloracetic acid precipitate.

The top curve in Figure 5 was obtained from the avirulent strain, the lower curve from the virulent strain. Within the accuracy of the method, no differences in the amino acid composition of the two strains are observed. The experiment is primarily an example of the applicability of the chromatographic technique to quantitative studies on the comparative chemistry of bacterial cells. It is sometimes useful to be able to look for differences by means of parallel experiments of this type.

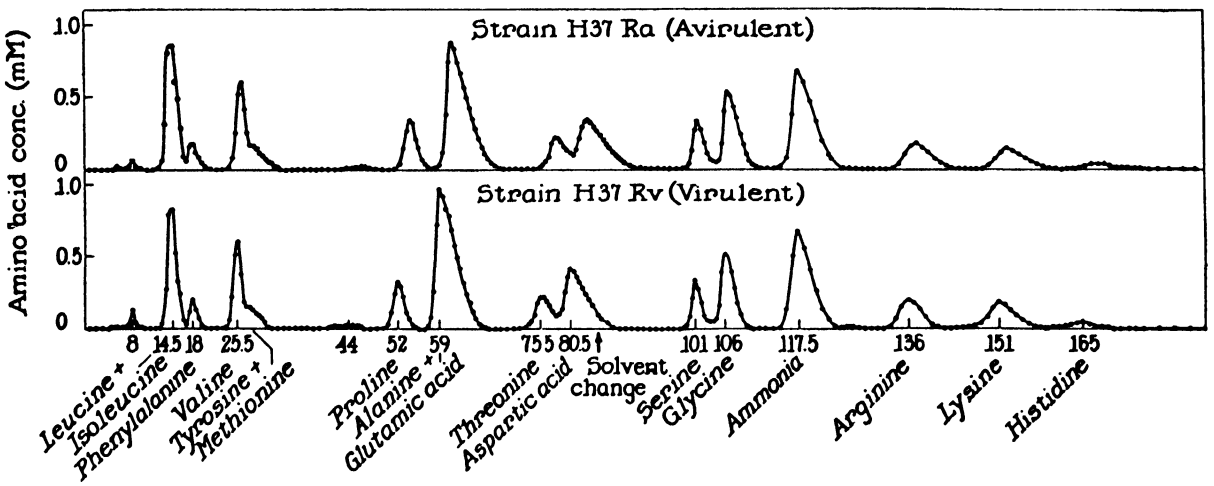


FIG. 5. Chromatographic fractionation of hydrolysates of virulent and avirulent strains of tubercle bacilli. The experimental conditions were the same as those described in Fig. 1.

Before proceeding to further applications of the starch column, it may be of interest to discuss some of the general implications of the work already outlined. The results have been achieved primarily because of two attributes possessed by the starch columns—their reproducibility and their high resolving power. These two attributes are related to one another and to certain mechanical considerations connected with the preparation of the column.

To the extent that the flow of solvent through a column is channeled, or non-uniform, the zones will be broadened and two adjacent peaks, instead of being separated by a clear valley, will merge into one another. For maximum efficiency and reproducibility, numerous experiments have shown that a chromatograph tube should be tightly packed with a material of small particle size, operated at a slow flow rate, and not overloaded with respect to the material being fractionated. It has been found that, with starch, the particles should be at least as small as 0.04 mm. in diameter, corresponding to about

300 mesh material. Potato starch granules average about this size. Canna starch, with larger granules, is less satisfactory, and rice starch, which is much finer than potato starch, gives flow rates that are inconveniently slow, even with automatic machinery.

The desired tightness of packing, which is important, can be achieved by utilizing the imbibition pressure of starch. If a slurry of starch in dry butanol is used in pouring the column, and a wet organic solvent is subsequently run through, the resultant swelling of the starch granules creates within the column an evenly distributed internal pressure capable of reducing the inter-granule interstices. It has thus been possible to devise a simple laboratory procedure which yields tightly packed, fully reproducible columns from experiment to experiment.

Starch columns are operated at the very slow flow rate of 2.0 to 3.0 cc. per hour per sq. cm. cross sectional area of the column. Under these conditions, it has been observed that the width of a slow moving amino acid peak is equivalent in effluent cc. to about 10 percent of the effluent volume at which the peak emerges. Thus, a peak emerging at 50 effluent cc. spreads over a narrow zone covering about 5 cc. (*cf.* Fig. 1, for example). Doubling the flow rate causes a slight but perceptible broadening and flattening of the peaks, which becomes intensified at markedly increased rates. High rates of flow result, therefore, in a loss of resolving power. Reducing the rate to one half of the recommended value does not alter the shape of the peak. The flow rate which has been found to be optimum for starch may not be so for other adsorbents. It seems clear, however, that chromatograms run at the excessively high flow rates frequently reported are, in many cases, operated at far from their maximum efficiency.

The amount of material fractionated, that is the load, on a column also markedly influences the resolving power attainable. On starch, the optimum load has been found to be about 0.3 mg. of an individual amino acid per sq. cm. of cross sectional area of the column. At such a level sharp symmetrical peaks are obtained. If the load is increased markedly, say by a factor of 5 to 10, the peak broadens and skews so that there is observed a sharp fronted peak emerging ahead of its normal position, followed by a long tail. This behavior represents a loss in resolving power, since the single broad peak will engulf any peak which, at a more moderate load, would immediately precede it. The shape of the peak suggests that, at the higher amino acid concentration, the adsorption isotherm is no longer linear.

Another factor affecting the resolving power of the chromatogram is the degree of retardation of the substances being fractionated on the column. For example, at the point of emergence of the advancing front of solvent, which is about 6 cc. after the sample has been added to a column 0.9 × 30 cm., there will emerge all the substances which are not adsorbed at all by starch. A peak at this position, therefore, would, in all likelihood, be a mixture of substances. It is only when the starch has considerable affinity for the components, resulting in slow band rates, or R_F values, that the maximum resolving power of the column is attained. It can be shown, using the mathematical treatment of Mayer and Tompkins (1947), that for starch columns 30 cm. in height, maximum efficiency is attained after four or five column volumes of solvent have passed through the column, or after about 30 effluent cc. in the curve shown in Figure 1. The maximum plate-efficiency of the column is realized for the proline peak and those emerging thereafter. In order to obtain the same efficiency for peaks emerging at small effluent volumes, theoretically it would be necessary to add the sample in an extremely small volume and to cut much smaller fractions.

In order to get the best results from a chromatograph column, therefore, it is necessary to find a combination of adsorbent and solvent which gives sufficient retardation of the substances being fractionated. As larger effluent volumes are reached, however, the peaks become increasingly broad and eventually a point is reached where the concentration of solute in the effluent becomes too low to be measured accurately. After 10 to 20 column volumes of solvent on a 30 cm. starch column, it is usually necessary to change the solvent to one which gives faster rates of travel for the remaining zones on the column, as in the shift of solvent after aspartic acid in the experiment shown in Figure 1. The new solvent sharpens up the remaining peaks.

In the work discussed thus far, a satisfactory rate of travel on the column has been achieved for nearly all the amino acids by the use of appropriate solvent combinations. Many of the solvent mixtures employed are miscible with water. Indeed, the number of solvents potentially useful in chromatographic work with starch columns is increased greatly when it is realized that these columns are not liquid-liquid (partition) chromatograms. A detailed treatment of this theoretical question cannot be undertaken at this time. A study of distribution coefficients has revealed, however, that the new term "partition chromatography" is not applicable to the starch column. Starch appears to act as a solid, polar adsorbent, which can be used with a variety of solvents.

Amino acids are not the only substances for which suitable band rates on starch have been obtained. Reichard (1948) has fractionated the nucleosides, and Daly and Mirsky (1949), and Hammarsten and his co-workers (1949) have chromatographed successfully the purines and pyrimidines. There are, however, many mixtures of substances, such as the high molecular weight peptides, which we have been unable to fractionate efficiently on starch. Di-, tri-, and tetrapeptides can be handled, since their chro-

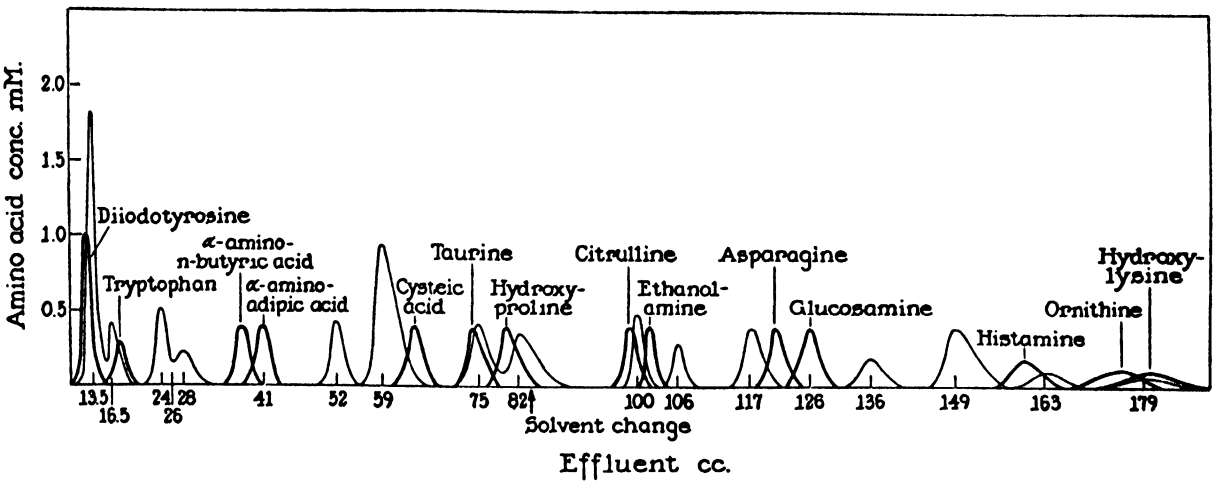


FIG. 6. Chromatographic behavior of a number of amino nitrogen containing substances. The background curve (thin line) is the same as the curve shown in Fig. 1.

matographic behavior appears to be similar to that of the amino acids. With polymyxin, or bacitracin, or other large peptides, however, we have not been able to obtain a suitable rate of travel. Craig, Gregory and Barry (1949) have brought peptides of this type within the range of the counter-current distribution method, but a satisfactory chromatographic system has not as yet been obtained. The peptides tend to run rapidly through the column and emerge with the solvent front, or else move scarcely at all on the column. At the present it would appear, therefore, that starch is not a good adsorbent for the higher peptides, and a search is currently being made for other adsorbent materials which might be applied effectively to this class of substances.

There remain two interesting applications of the starch column which warrant mention. They concern the chromatographic analysis of urine and of blood plasma. Obviously, urine may be expected to contain many ninhydrin positive compounds other than the common amino acids discussed thus far. Accordingly, the chromatographic behavior of a number of additional compounds has been investigated. The results are shown in Figure 6. The background curve (thin line) is the familiar amino acid

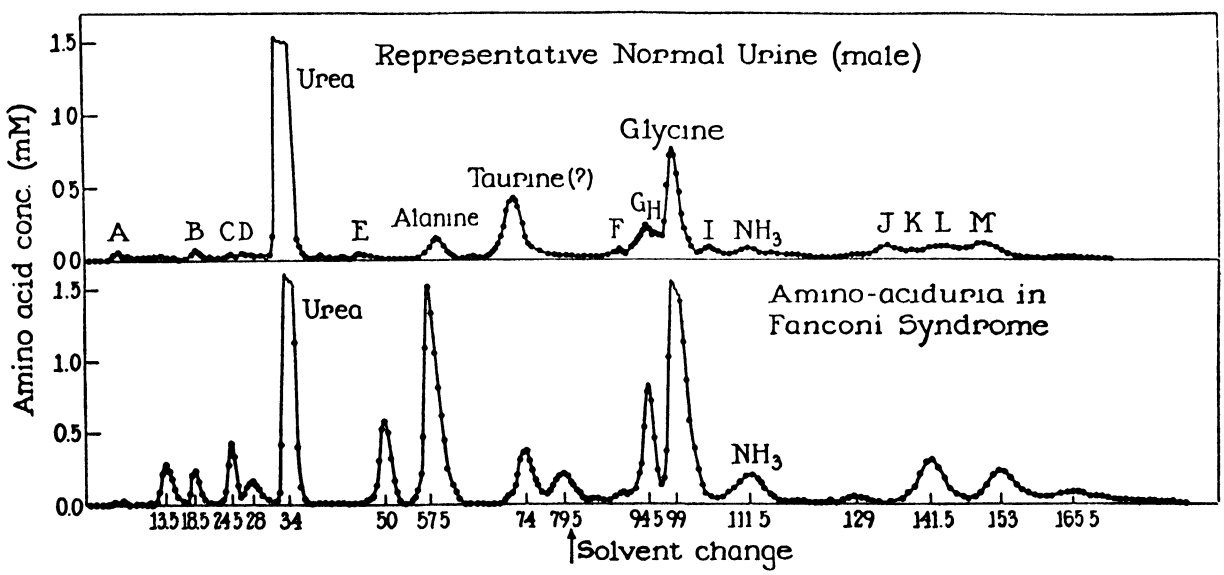


FIG. 7. Chromatographic fractionation of urine. The samples were withdrawn from a 24 hour specimen and desalted prior to chromatography. The amount of desalted urine chromatographed corresponded in each case to about 0.17 percent of the 24 hour specimen. The experimental conditions were the same as those described in Fig. 1.

pattern shown in Figure 1. Placed on this curve are a variety of other ninhydrin-positive substances. This picture indicates the necessity for the maximum resolving power attainable, and emphasizes the problems concerned with the correct identification of peaks when complex mixtures are chromatographed.

Preliminary curves obtained with samples of normal and pathological urines are illustrated in Figure 7. Most of the inorganic salts are removed prior to chromatography, using the electrolytic desalting device of Consden, Gordon and Martin (1947). The upper curve was obtained from a normal adult. There are at least 14 well defined, reproducible peaks on this curve. Ammonia, which was originally present in great amount, has largely been removed during the desalting process. The major amino acid peaks have been *tentatively* identified as alanine, taurine, and glycine. Many of the small peaks probably can be ascribed to other amino acids which are known to be excreted in small amounts. There appear to be at least four substances in the basic amino acid range, two of which are probably lysine and histidine. Glutamic and aspartic acids seem to be absent from urine, but appear after acid hydrolysis;—one would like to know from what source. The major peaks shown are resistant to acid hydrolysis.

The second curve was obtained from an infant exhibiting the Fanconi syndrome. We are greatly indebted to Dr. E. A. Park and Dr. Paul Hardy of the Johns Hopkins Hospital for the urine specimen. The curve is remarkably reminiscent of the ones obtained on protein hydrolysates. The starch column data are, therefore, in accord with the observations of others, notably Dent (1947), that the disease is associated with a marked amino aciduria. It should not be inferred, however, that all the peaks on this curve have been identified. Preliminary satisfactory curves have also been obtained with desalted dialysates of blood plasma.

The work with urine and plasma has been mentioned primarily to emphasize the fact that in chromatographing complex biological mixtures, such as these, particularly if they have not been subjected to acid hydrolysis, the identification of the peaks on the effluent curve becomes a major part of the research. Many substances other than the common amino acids are to be expected, including peptides, and in many cases even the most careful checks and controls cannot lead to more than tentative conclusions. It has seemed to us for some time that isolation of the components responsible for the peaks would be the only way to accumulate unequivocal data.

Work on convenient methods for the isolation of 25 to 50 mg. quantities of pure compounds from large scale columns is now in progress. Isolation experiments, however, present several major problems which are not encountered with chromatograms run on an analytical scale. As mentioned earlier, the capacity of a starch chromatogram is not great under conditions which maintain its high resolving power. To work up one gram of a protein hydrolysate on starch requires the use of at least two columns 8 cm. in diameter and about 25 liters of solvents. Furthermore, the effluent from a starch column is not carbohydrate-free, particularly when acidic solvents are used. Unless measures are taken to remove carbohydrate impurities, it is extremely difficult to isolate crystalline products from the effluent. Finally, in biochemical work there is frequently the problem of acquiring enough starting material to permit isolation. It is true that urine is not normally a scarce commodity, but only rarely would it be possible to obtain enough of a pathological plasma, for example, to warrant an attempt at isolation.

These considerations have led to several further avenues of experimentation. It has been found that the carbohydrate can be removed conveniently from the amino acids in the effluent by using a short column of an ion exchange resin. The fractions from a starch column comprising an amino acid peak can be pooled, concentrated to a small volume, and run through a resin bed. The carbohydrate passes through and the amino acids are retarded. In principle, the procedure seems to be sound. Analytically pure samples of amino acids have been obtained in mg. quantities from protein hydrolysates. The best conditions remain to be worked out, however.

A logical extension of this approach has been the examination in detail of the ability of some of the ion exchange resins themselves to fractionate mixtures of amino acids. Resins have been widely used as a means of separating the constituents of protein hydrolysates into groups of amino acids, and for the determination of a few of the individual components. Applying the principles and techniques which have been worked out to attain the best resolving power from starch, it has been found that some of the newer ion exchange resins now available commercially are capable of giving individual peaks for essentially all the common amino acids. This work is in a preliminary stage, but Dowex-50, for example, appears to possess fractionating power rivalling that of starch for some problems.

In Figure 8 is shown a result obtained recently upon passage of a complex mixture of amino acids through a column of Dowex-50, a nuclear sulfonic acid polymer. It will be noted that in a single chromatogram it has been possible to obtain individual peaks for all 18 components of the mixture, although the serine-threonine resolution is only partial. The column was 55 cm. in height. The first peaks were eluted with 1.5N HCl, followed by 2.5N and 4.0N acid. It is apparent from this curve that the resin is not functioning exclusively as an ion exchange medium. The separation of the monoamino acids, glycine, alanine, valine, etc., must result from non-ionic adsorption by the resin, which coupled with ion exchange, gives a very useful result.

The investigations with Dowex-50 are still in

their early stages. There is doubtless much yet to learn. For example, it is not known how reproducible the results will be with different batches of the resins. Nevertheless, some of the attributes of Dowex-50 columns warrant mention. In comparison with starch, the effluent from Dowex-50 is clean, the capacity is higher, and the performance of the column is not as sensitive to the presence of inorganic salts. The use of strong HCl (Fig. 8) as an eluent is disadvantageous. Experiments in progress indicate that it may be possible to employ Dowex-50 with a variety of buffered systems which would be much more convenient to use.

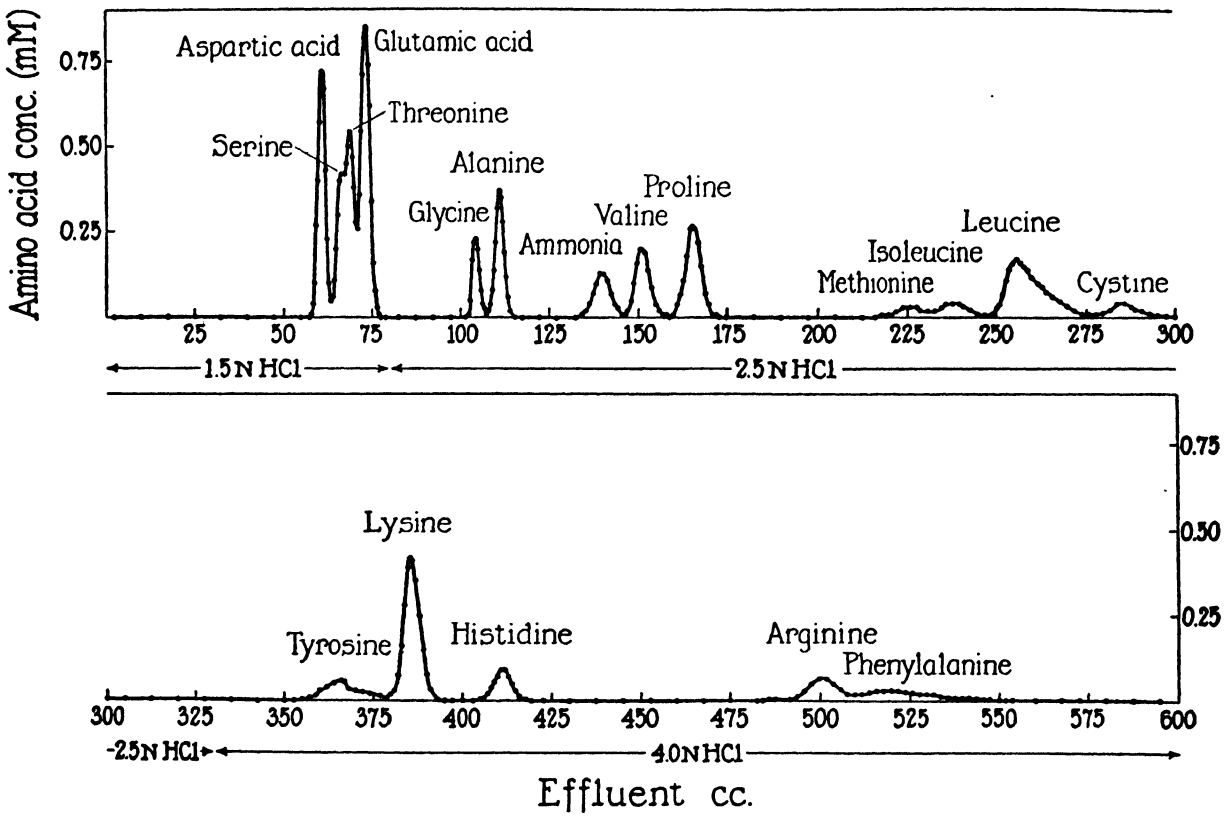


FIG. 8. Fractionation of a known mixture of amino acids on Dowex-50. Solvents, 1.5N, 2.5N, and 4.0N HCl. Column dimensions, 0.9 × 55 cm. Sample, about 3 mg. of amino acids.

One of the most interesting features of the curve in Figure 8, is the relative positions of the various amino acid peaks. The pattern on Dowex-50 is completely different from that obtained with starch. Thus in the aliphatic series of amino acids, glycine, alanine, valine, isoleucine, and leucine, the order of emergence is completely reversed. On starch glycine is last, whereas on Dowex-50, it is first. The other amino acids are also shifted to new relative positions. This reversal of the order arises in large part from the fact that starch is essentially a polar adsorbent, whereas Dowex-50 possesses non-ionic adsorptive properties characteristic of a non-polar adsorbent.

The sequential use of Dowex-50 and starch should yield a fractionating power of an exceptional order. It would be a rare coincidence indeed if two substances showed identical rates of travel on both types of columns. It also follows that the use of starch and Dowex-50 in conjunction should go far in identifying a peak in an effluent curve obtained from plasma, for example. Consider a given peak occurring in the glycine position on starch. If on rechromatographing on Dowex-50 the contents of this peak are recovered quantitatively again in the glycine position, and if, in addition, the peak is resistant to acid hydrolysis, and if, finally, glycine can be added and recovered quantitatively with no loss in symmetry of the peak, the identity of the substance in question would be well established. It is such a combination that may provide identification of components, without the need for isolation, when biochemical materials available in limited supply are being investigated.

Although starch and ion exchange columns used in conjunction appear to offer promise for the future, it is clear that much remains to be done. There is need for a variety of polar and non-polar adsorbents

that can be utilized in the form of columns of high resolving power. As more information accumulates following the application of these and related techniques to specific problems in the fractionation of mixtures of amino acids and peptides, it is to be hoped that further insight can be gained into some of the architectural mysteries of protein molecules.

References

BRAND, E., 1946, Amino acid composition of simple proteins. Ann. N. Y. Acad. Sci. *47:* 187-228.

BRAND, E., KASSELL, B., and SAIDEL, L. J., 1944, Chemical, clinical, and immunological studies on the products of human plasma fractionation. III. Amino acid composition of plasma proteins. J. clin. Invest. *23:* 437-444.

CONSDEN, R., GORDON, A. H., and MARTIN, A. J. P., 1947, The identification of lower peptides in complex mixtures. Biochem. J. *41:* 590-596.

CRAIG, L. C., GREGORY, J. D., and BARRY, G. T., 1949, Studies on polypeptides and amino acids by countercurrent distribution. Cold Spring Harbor Symposium. Quant. Biol. *14:* 24-31.

DALY, M. M., and MIRSKY, A. E., 1949, Chromatography of purines and pyrimidines on starch columns. J. biol. Chem. *179:* 981-982.

DENT, C. E., 1947, The amino-aciduria in Fanconi syndrome. A study making extensive use of techniques based on paper partition chromatography. Biochem. J. *41:* 240-253.

EDMAN, P., HAMMARSTEN, E., LÖW, B., and REICHARD, P., 1949, Partition chromatographic separation of adenine and guanine. J. biol. Chem. *178:* 395-398.

ELSDEN, S. R., and SYNGE, R. L. M., 1944, Starch as a medium for partition chromatography. Biochem. J. *38:* ix.

MAYER, S. W., and TOMPKINS, E. R., 1947, Ion exchange as a separation method. IV. A theoretical analysis of the column separation process. J. Amer. chem. Soc. *69:* 2866-2874.

MOORE, S., and STEIN, W. H., 1948, Partition chromatography of amino acids on starch. Ann. N.Y. Acad. Sci. *49:* 265-278.

1948, Photometric ninhydrin method for use in the chromatography of amino acids. J. biol. Chem., *176:* 367-388.

1949, Chromatography of amino acids on starch columns. Solvent mixtures for the fractionation of protein hydrolysates. J. biol. Chem. *178:* 53-77.

REES, M. W., 1946, Estimation of threonine and serine in proteins. Biochem. J. *40:* 632-640.

REICHARD, P., 1948, Partition chromatography of ribonucleosides on starch. Nature, Lond. *162:* 662.

STEIN, W. H., and MOORE, S., 1948, Chromatography of amino acids on starch columns. Separation of phenylalanine, leucine, isoleucine, methionine, tyrosine, and valine. J. biol. Chem. *176:* 337-365.

1949, Amino acid composition of β-lactoglobulin and bovine serum albumin. J. biol. Chem. *178:* 79-91.

SYNGE, R. L. M., 1944, Analysis of a partial hydrolysate of gramicidin by partition chromatography with starch. Biochem. J. *38:* 285-294.

Discussion

RANDALL: In view of the nature of the processes on which the work of this paper depends it is perhaps rather surprising that more attention has not so far been given to the effect of temperature on the degree of separation of different components of a mixture. This comment can legitimately be applied to starch-resin type columns. It is possible also that the overall efficiency of the method could be improved by a proper choice of temperature.

I understand that promising experiments have been carried out by the Oak Ridge group on the use of resin columns at 80° C. It is clearly a simple experimental matter to provide temperature control of columns and I should like to know the views of the authors on this problem.

STEIN: A few preliminary experiments have been carried out which bear upon the question raised by Dr. Randall. Mixtures of amino acids were passed through columns of Dowex-50 maintained at 70° C. Sodium citrate buffers were employed as eluting solvents. From these experiments it appeared that the order in which the amino acids emerged from the column was not markedly altered by the change in temperature from 25° to 70° C, although certain small differences were observed. At the elevated temperature the peaks were notably higher and sharper, however, a finding in accord with the observations of the Oak Ridge group. This definite advantage was offset by the fact that considerable decomposition of certain amino acids was encountered during the course of the experiments performed at 70° C. To minimize decomposition, we currently prefer to carry out chromatographic separations of amino acids at room temperature.

PHYSICAL AND CHEMICAL STUDIES OF GRAMICIDIN AND SOME IMPLICATIONS FOR THE STUDY OF PROTEINS

R. L. M. SYNGE
The Rowett Research Institute, Bucksburn, Aberdeenshire, Scotland

In the autumn of 1941, shortly after Dr. A. J. P. Martin and I had obtained our first successful results with partition chromatography at the laboratory of the Wool Industries Research Association in Leeds, we received a visit from Dr. B. C. J. G. Knight. He drew our attention to the then recently published studies of Dubos, Hotchkiss and others on gramicidin, and suggested that this might prove to be an admirable model substance, both for testing our new techniques of amino acid analysis and for assessing more closely the value of the isolation and recognition of products of partial hydrolysis as a means for the study of protein structure (Synge, 1943). Fortunately, we had no idea at the time what complexities lay hidden in gramicidin, and therefore, with Dr. A. H. Gordon, embarked on the study of it. Since that date, the problem of the structure of gramicidin has appeared, to all workers concerned with it, to become more and more involved, and at the present time has reached a sort of crisis, in that Gregory and Craig (1948) have established that gramicidin is not a single substance, but a mixture of substances. The present, therefore, seems a suitable occasion on which to recapitulate what is known, and to attempt a critical appraisal of the situation.

Although gramicidin did not serve our original purposes as a model substance, a suitable substance for those purposes soon came into our hands. This was gramicidin S ('Soviet gramicidin'), another antibacterial peptide also derived from strains of *Bacillus brevis*, which was discovered by Gauze and Brazhnikova (1943). Gramicidin S has turned out to be more closely related to tyrocidine than to the original gramicidin of Dubos. The specimens of gramicidin S which we have had behaved in a straight-forward way in all our experiments, and it has so far proved impossible to establish their heterogeneity. (Indeed one of them was recently examined by Dr. L. C. Craig by his counter-current distribution technique, and appeared substantially homogeneous under conditions in which a commercial specimen of gramicidin S had proved heterogeneous.) Furthermore, the data on amino acid composition and free functional groups (Synge, 1945a; Sanger, 1946) and on the products of partial hydrolysis by acid (Consden, Gordon, Martin and Synge, 1947; Synge, 1948) strongly suggest a definite cyclopeptide structure for gramicidin S. I will not discuss this subject further except to say that the results with gramicidin S gave us confidence in the technical methods used for the study of gramicidin. Without them, the results obtained with gramicidin would be even more difficult to assess than they are at present.

It is just the difficulty of interpreting the data obtained with gramicidin, a peptide whose composition is simple and molecular weight low, when compared with typical proteins, that makes gramicidin a very suitable subject for discussion at the present Symposium. I have the feeling that some of the difficulties encountered will reappear in future studies with proteins, and, to the extent that this is so, the study of gramicidin will turn out to be of value for the general progress of protein chemistry.

Origin and Homogeneity

Gramicidin has been obtained from a number of strains of aerobic spore-forming Bacillus spp., many of which have been identified as *Bacillus brevis*. They have been isolated from a wide variety of sources. Hotchkiss (1944) has given an admirable review of the various aspects of the study of gramicidin from its discovery by Dubos (1939) down to that date.

After growing the bacteria on fairly simple media, some antibacterial activity is observed in the culture medium. This increases on allowing the bacteria to autolyse, or on subjecting them to peptic or tryptic digestion, as though the active material were in some way associated with the bacterial protein. At this stage it can be extracted with acid alcohol, and, on diluting this alcoholic extract with aqueous salt solution, a precipitate forms. This material is known as "tyrothricin," and at this stage contains some lipid material that can be removed with fat solvents. In recent years, tyrothricin has become an article of commerce, being useful as an antibacterial agent applied topically in medical and veterinary practice. On account of its relative cheapness, and harmlessness *per os*, it probably has certain other commercial applications.

Tyrothricin consists largely of peptide material. By fractionation with organic solvents—particularly with acetone-ether mixtures—a soluble fraction and an insoluble fraction result. From the soluble fraction it is easy to isolate, by repeated crystallisation with acetone or other solvents, the material that has come to be known as gramicidin. The insoluble fraction, after repeated crystallisation with methanolic or ethanolic HCl yields basic peptide material (as the hydrochloride) which is known as tyrocidine.

Since all the many properties of these crystalline

fractions that have been studied remain unchanged on further crystallisation, it was natural to regard the isolated products as pure substances. However, it was clear from the outset that tyrothricin did not consist simply of tyrocidine and gramicidin. Most who have worked with tyrothricin would agree with the judgment of Hotchkiss (1944): "There is reason to believe that tyrothricin normally contains from 10 to 20 percent gramicidin and 40 to 60 percent tyrocidine hydrochloride, of which perhaps two-thirds can ordinarily be obtained in crystalline form. The unaccounted-for portion of tyrothricin appears to be largely made up of 'gramicidin-like' and 'tyrocidine-like' components, speaking in both a chemical and a biological sense."

It is the electrically charged, salt character of the tyrocidine-like components that readily permits their separation from the neutral uncharged gramicidin-like components by fractionation with organic solvents. No doubt electrophoretic and ion-exchange procedures would prove equally effective for the purpose. The mother-liquor materials from recrystallisation of gramicidin and of tyrocidine hydrochloride clearly differ in solubility, readiness to crystallise, and other properties from the respective crystalline products. In the case of gramicidin, detailed evidence was presented at an early stage that the material in the earlier mother-liquors shows small differences from the recrystallised product in amino-acid composition, optical rotation and chromatographic behaviour (Gordon, Martin, and Synge, 1943). There is also an early mention of gramicidin of divergent tryptophan content (White and Secor, 1946).

The highly heterogeneous nature of tyrothricin was thus apparent at an early stage. What was not fully appreciated was that crystalline gramicidin and tyrocidine hydrochloride, although they can be recrystallised with constant properties (see especially Tishler, Stokes, Trenner and Conn, 1941), are by no means single homogeneous substances. Gregory and Craig (1948) showed this to be the case with gramicidin by counter-current distribution between organic solvent phases. Diffusion (Pedersen and Synge, 1948) and adsorption (Synge and Tiselius, 1947) studies had failed to reveal heterogeneity. Gregory and Craig showed that at least five components, differing in amino-acid composition, were present in the usual recrystallised gramicidin. At the same time Synge and Tiselius (1947) had demonstrated by frontal analysis on charcoal that tryocidine consists of at least three different components differing in tryptophan content. Drs. Gregory and Craig (private communication) reached a similar conclusion about tyrocidine as a result of distribution studies.

I would like now to mention some experiences with gramicidin crystals. As ordinarily prepared, gramicidin gives very small crystals from acetone, that are unsuitable for X-ray studies. On one occasion, however, I obtained large chunky crystals from a specimen of gramicidin in alcoholic solution after very slow evaporation. In Dr. Crowfoot's hands these crystals gave very beautiful X-ray photographs. However, it proved impossible for a long time to repeat the preparation of the crystals. Then, some years later, I obtained from the Wallerstein Company a batch of gramicidin that regularly crystallised in this way. The crystals were subjected to detailed study by Dr. Crowfoot, and seemed to be the same as those obtained previously. Shortly after this came Gregory and Craig's (1948) announcement that gramicidin, as usually prepared, is heterogeneous. Although I had been unable to detect any other physical or chemical differences between the new batch of gramicidin and previous batches that would not crystallise in this way, nor between the new crystals and material in the mother liquors from them (Synge, 1949a), I thought it might be that the gramicidin was crystallising in this way because it was substantially homogeneous, and so submitted the material to Dr. Craig for examination. However, he found that all the usual components were present, although the proportion of the minor components was slightly lower than usual. Nevertheless, more than 20 percent of the crystalline material submitted to him consisted of components other than gramicidin A. Dr. Craig kindly returned some of the purified gramicidin A isolated by him from the crystalline material which he had examined. Dr. Crowfoot found that the X-ray photographs of this gramicidin A, which were very beautiful and detailed ones, appeared in no way to differ from those of the original material that had contained such a proportion of other components.

Dr. Craig, Dr. Crowfoot and Drs. Moore and Stein will perhaps be able, in discussion, to amplify the points just mentioned, and to tell you something about the properties of the gramicidin components A, B, *etc.* What I want to emphasize just now is that among peptides so far encountered in nature, the occurrence of families of substances is very frequently observed. Penicillins, polymyxins, the ergot alkaloids and clupein are among the better established examples, while the same is probably true for hypertensin, the toxic substances of *Amanita phalloides* and other peptide materials (*cf.* Synge, 1949b). The case of gramicidin is chiefly of interest in that the heterogeneity defied detection by a greater variety of techniques than in the case of these other families. The X-ray experiences just mentioned are particularly striking in this connection. One feels that proteins also probably occur as families of molecules related to one another in structure. One is forced further to conclude that if this is so, and the sort of differences existing are the same as with gramicidin components, the present-day methods used for the isolation of, and for testing the homogeneity of "pure" proteins are totally inadequate for revealing such differences. One also, after experience of attempts at stoichiometric interpretation of the amino-acid analysis of gramicidin, becomes highly

sceptical of all such interpretations so far applied to proteins.

In the remaining part of this paper, I can only for the most part refer to data obtained with heterogeneous gramicidin purified by simple recrystallisation. All this information will, at a subsequent date, require re-interpretation as a result of studies with individual gramicidins. The existing data seem, however, to establish facts of interest, and are therefore worth summarising and discussing at this stage.

Molecular Weight

Since stoichiometric data cannot yet be used, information from the properties of gramicidin in solution, and from X-ray crystallographic studies are the only available lines of evidence.

As would be expected from its neutral character and from its amino-acid composition, gramicidin is highly insoluble in water. Colloidal, slowly aggregating solutions result when alcoholic solutions of gramicidin are largely diluted with water. Electrolytes promote flocculation. Dispersion is better, and more permanent in the presence of proteins (as with serum), tyrocidine, urea, etc. However, solutions which appear to be stable indefinitely, and which are devoid of opalescence, can be made with a wide variety of organic solvents, and it is such solutions that have yielded data bearing on molecular weight.

Cryoscopic data in camphor give values 1250-1600 (Hotchkiss, 1941; Tishler *et al.*, 1941). The latter authors found values of 850-950 in phenol, and 600-1200 in cyclohexanol, in which solvent the apparent molecular weight showed a dependence on concentration that still demands physico-chemical interpretation.

The same authors give isopiestic data in methanolic solution which indicate a molecular weight of 3100, and mention a diffusion experiment with *n*-butanol as solvent giving a value in the range 2500-4500 (presumably calculated for spherical, non-solvated molecules). Pedersen and Synge (1948), for diffusion in 70 percent (*v/v*) aqueous ethanol, found values in the range 2800-5000 when certain corrections based on the behaviour of gramicidin S were applied.

Dr. Crowfoot's crystallographic data indicate that both gramicidin A and gramicidin B have molecular weights of the order of 3800. It would be just possible, in both cases, for the molecular weight to be double this figure, but this is not probable.

Thus at the present time nothing more conclusive can be said of the molecular weights of gramicidins than that they are probably of the order of a few thousand.

Elementary Composition and Component Residues

Gramicidin is composed entirely of the elements C, H, O and N (Hotchkiss and Dubos, 1941; Tishler *et al.*, 1941).

In acid hydrolysates of gramicidin there have been identified glycine, alanine, valine, leucine and tryptophan (see Hotchkiss, 1944) and ethanolamine (Synge, 1945b). Phenylalanine and tyrosine are also present in small amount, and appear to occur in the minor components (gramicidin B, C, *etc.*) (private communication from Dr. J. D. Gregory). The small amounts of ammonia and acetaldehyde present could arise by break-down of ethanolamine (see Hotchkiss, 1944; Synge, 1945b), and break-down products of tryptophan are present, although the decomposition is very slight when compared with that usually encountered in the acid hydrolysis of proteins. The content of lipid and phospholipid material is so small and variable that it must be a contaminant. Neither volatile alcohols, volatile or ether-extractable carboxylic acids, nor CO_2 are formed in appreciable amount during hydrolysis with acid (Hotchkiss, 1944; Synge, 1945b). The amounts found of the major components account for 96.9 percent of the N and give an atomic C/N ratio of 4.84. (See Table 1.) The ratio based on the elementary analyses of Hotchkiss and Dubos (1941) and of Tishler *et al.*, (1941) is 4.94. The unaccounted remainder is probably due to phenylalanine, tyrosine and to the margin of error in the determination of tryptophan and valine. There is no compelling reason for postulating the presence of yet other component residues.

Table 1. Component Residues of Gramicidin Estimated from Analyses of Acid Hydrolysates, etc.

Component	N as % of total N of gramicidin	Selected value based on data of:
Glycine	5.5	Synge (1949a)
Alanine	10.1	Synge (1949a)
Valine	18.3	Synge (1949a)
Leucine	20.3	Synge (1949a)
Tryptophan	36.9	Hotchkiss (1941); Edwards (1949)
Ethanolamine	5.8	Synge (1945b)
Total	96.9%	

Systematic quantitative analyses have shown that gramicidin requires unusually prolonged and vigorous hydrolysis with acid before all the valine is liberated. Valine figures after 24 hour hydrolysis with hot acid were consistently lower than those after 48 hours (Synge, 1949a). This is reasonably attributable to the occurrence of difficultly hydrolysable valylvaline, found by Christensen (1943, 1944) after rather vigorous partial hydrolysis with hot hydrochloric acid. The occurrence of valylvaline in acid hydrolysates of gramicidin is confirmed by Dr. J. D. Gregory (private communication), and it was probably responsible for one of two anomalous bands encountered on the chromatograms by Gordon, Martin, and Synge (1943) in their first analyses of acid hydrolysates as acetyl derivatives on silica gel columns. (The other anomalous band was presum-

ably due to phenylalanine.) The band corresponding to valylvaline diminished markedly on extending the hydrolysis from 24 hours to 48 hours (Synge, 1949a). Similar acid-resistant peptides have been reported by Barry, Gregory and Craig (1948) in hydrolysates of bacitracin. The occurrence of difficultly hydrolysable peptides is a source of error in the amino-acid analysis of proteins to which, perhaps, too little attention has been paid in the past.

Some data on the kinetics of liberation of various groupings and compounds during the acid hydrolysis of gramicidin, together with comparative data on the hydrolysis of some relevant dipeptides were reported by Christensen and Hegsted (1945) and by Synge (1945c).

Optical Configuration of the Component Residues

Amino-acids possessing the D-configuration have now been found in acid hydrolysates of a fair variety of bacterial and fungal products. It has usually been assumed that this implies the occurrence of the same residues with the D-configuration in the intact parent substance. In general, this is probably true, as far as we can judge from the behaviour on hydrolysis of synthetic compounds, but the possibility should be borne in mind that the same sort of process that effects racemisation of an amino-acid when it is free, or when its residue is linked in an otherwise symmetrical compound, may, when another centre of asymmetry exists in the molecule or in its environment, result not in racemisation but in a selective inversion or epimerisation. This possibility has been discussed in some detail by Neuberger (1948). In work with gramicidin it has been observed that the configuration of the valine isolated after acid hydrolysis depends on the conditions under which the hydrolysis was effected (Synge, 1944, 1949a). After acid hydrolysis for 48 hours at 110° in strong HCl-acetic acid, the valine isolated was racemic, whereas, when the same treatment in the hot followed ten days in the same mixture at 37°, there was a substantial predominance of the D-isomer. It seems unlikely that this difference is due simply to the liberation of different proportions of D-valine and L-valine residues from peptide linkage. It is more reasonable to suppose that the different conditions of treatment with acid differed in the extent of racemisation or inversion that they produced. Similar variations were observed in the optical form of lysine, isolated from acid hydrolysates of gelatin by Schryver and Buston (1927), although Thimann (1930) thought that this might have been due to racemisation by the alkali used for isolation.

The leucine isolated from acid hydrolysates of gramicidin is predominantly D-, while the alanine and tryptophan are predominantly L- (see Hotchkiss, 1944). The configuration of the valine residues in the intact gramicidin therefore appears to be the main problem. There is no inherent difficulty in postulating the co-existence of D-valine and L-valine residues either in the same molecule or in separate gramicidins. The simultaneous incorporation of D- and L-leucine residues into gramicidin and tyrocidine respectively by the same organism makes either view entirely plausible. The production of fumaryl- DL-alanine by *Penicillium resticulosum* should also be noted (Birkinshaw, Raistrick and Smith, 1942).

In studies of peptides obtained from gramicidin by partial hydrolysis with acid, the following peptides of valine have been identified: D-valyl-D-valine, and L-valyl-L-valine (as the racemic compound) by Christensen (1943, 1944), L-valylglycine and L-alanyl-D-valine by Synge (1944, 1949a). The valine obtained by acid hydrolysis of the last-mentioned peptide was considerably racemic, as was the leucine from some of the peptides of leucine isolated in the same work.

To sum up, it would be reasonable to suppose that both D- and L-valine residues occur in intact gramicidin, but this cannot be regarded as established until the possibility of epimerisation of valine residues during hydrolysis has been more closely investigated and excluded. The possible role of D-amino-acid residues in conferring toxicity or indigestibility on the products containing them has frequently been discussed (see especially Biochemical Society Symposia, 1948). Many of the compounds in which they have so far been found embody cyclic structures, and, according to Work (1948), Dr. Crowfoot has suggested that the presence of D- as well as L-residues in a peptide chain may produce an orientation favourable for cyclopeptide formation.

Peptides Resulting from Partial Hydrolysis of Gramicidin with Acid

Consideration of products obtained in the very early stages of acid hydrolysis of gramicidin will be deferred to the next section, and discussed in relation to the mode of linkage of the ethanolamine residues. Here I will deal only with peptides obtained by fairly far-reaching hydrolysis. Those so far

Table 2. Peptides Identified in Partial Hydrolysates of Gramicidin

Compound	Amount (N of peptide as % of total N of gramicidin)	Reference
L-valyl-L-valine	2.6	Christensen (1943, 1944)
D-valyl-D-valine	2.6	Christensen (1943, 1944)
L-valylglycine	7	Synge (1944, 1949a)
D-leucylglycine	1.0	Synge (1949a)
L-alanyl-D-valine	0.45	Synge (1949a)
L-alanyl-D-leucine	—	Synge (1949a)

recognised and fully identified are listed in Table 2. In addition, evidence of a less conclusive nature has been obtained for the occurrence of alanylvalylleucine or alanylleucylvaline or both (Synge, 1949a).

Most of the peptides in Table 2 were isolated by chromatography on starch or on paper, and the conditions of working were probably such that no dipeptides of glycine or of alanine present in the hydrolysate in substantial amount would have escaped recognition. The faster-moving fractions from the chromatograms, representing by far the greater part of the hydrolysate, were dipeptides and higher peptides incorporating a high proportion of leucine, valine and tryptophan residues. They proved difficult to fractionate by partition chromatography on starch or paper as "tailing" was considerable. Other methods will be required for further study of these fractions, and among possibly useful procedures may be mentioned counter-current distribution by the procedures of Craig and others, the use of kieselguhr partition chromatograms for the elimination of adsorption effects, the use of charcoal adsorption for fractionation according to the number of tryptophan residues in the molecule (Synge and Tiselius, 1949), and ionophoretic and electrophoretic methods taking advantage of differences in ionic mobilities, such as those of Martin and others.

While the peptides so far identified permit no far-reaching conclusions as to the detailed structure of the gramicidins, it is of interest that the greater part of the glycine was isolated as L-valylglycine. When it appeared that gramicidin might be a single substance, this implied that all the glycine residues must occur with their amino groups linked to valine. It was consequently very disconcerting to find leucylglycine present among the other products of partial hydrolysis. I was engaged in devising all kinds of hypothetical types of linkage to explain this phenomenon and, worse, beginning to doubt whether the products of partial hydrolysis with acid really do give reliable information about peptide structures, when the announcement by Gregory and Craig (1948) of the heterogeneity of gramicidin removed this conflict, at least temporarily. It will be interesting to see, with the so-called "pure" proteins, whether partial hydrolysis studies of the mode of linkage of those amino-acids that occur in small amount, or terminally, can prove useful for providing evidence as to heterogeneity.

Over-all Properties and Derivatives: The Problem of Chemical Structure

It seems to be generally agreed that gramicidin is a neutral substance without free amino or carboxyl groups. This has been established by electrometric titration (Christensen, Edwards and Piersma, 1941; James, 1949), other titrations with acid and alkali and Van Slyke amino-N determinations (Hotchkiss, 1941; James, 1949) and by its failure to form a stable yellow derivative on treatment with 1-fluoro-2:4-dinitrobenzene (James, 1949). The indole groups appear to occur free, as far as can be judged from spectrometric and other data (see especially Edwards, 1949), that is, they behave as in normal peptide and acyl derivatives of tryptophan. That -OH groups are present has been established by substitution with a variety of reagents, all of which give fairly concordant values, that are also approximately equivalent to the number of ethanolamine residues estimated to be present (see Table 3).

Table 3. Relation of Ethanolamine Content of Gramicidin to Capacity for Substitution by Various Radicals

Substituent grouping	Gm. res. per 10^4 gm. original gramicidin	Reference
Acetyl	8.2	Hotchkiss (1941)
	6	Fraenkel-Conrat, Brandon and Olcott (1947)
Succinyl	6.6	Olcott *et al.* (1946)
Sulphuric ester	5.9	Reitz *et al.* (1946a)
Phosphoric ester	6.1	Ferrel, Olcott and Fraenkel-Conrat (1948)
Methoxyl	5.4	James (1949)
Ethanolamine (after hydrolysis) (corr. for estimated destruction)	6.1	Synge (1945b)

Derivatives have also been prepared from gramicidin by treatment with formaldehyde, and some of these show diminished toxicity for animals compared with bacteria, as well as greater solubility in aqueous solvents (Lewis *et al.*, 1945). Indole groups appear to be chiefly concerned in the reaction with formaldehyde, methylol groups being substituted on the indole nuclei (Fraenkel-Conrat, Brandon and Olcott, 1947). The same authors showed that the newly resulting hydroxyl groups can be substituted with acetyl groups, or (Olcott *et al.*, 1946) with succinyl groups. Fraenkel-Conrat *et al.* (1946) studied the biological effects of a number of substitution products of gramicidin, including products formed by the action of chlorosulphonic acid, which appears likewise to attack the indole groups (Reitz

et al., 1946b). Schales and Mann (1947, 1948) studied biological effects of gramicidin treated with a variety of reagents, but here the products were not chemically characterised.

The chief problem in obtaining any general picture of the structure of gramicidin is to explain how, in a molecule apparently made up entirely from residues of monoamino-monocarboxylic acids and ethanolamine, there is a complete absence of basic or acidic properties. Basic groups, corresponding in number to those of the ethanolamine residues, must somehow be masked, and the molecule must likewise have a cyclopeptide type of structure.

Attempts to explain the masking of amino groups by fatty acids, volatile or otherwise, by other ether-extractable carboxylic acids, or by CO_2 (ureide structure) have all failed, and on hydrolysis only those products already mentioned have been detected. Nor do the data for elementary composition require the postulation of other residues.

These problems are at present being attacked by my former colleague, Dr. A. T. James, at the Lister Institute of Preventive Medicine, London. He has kindly permitted me to describe some of his results.

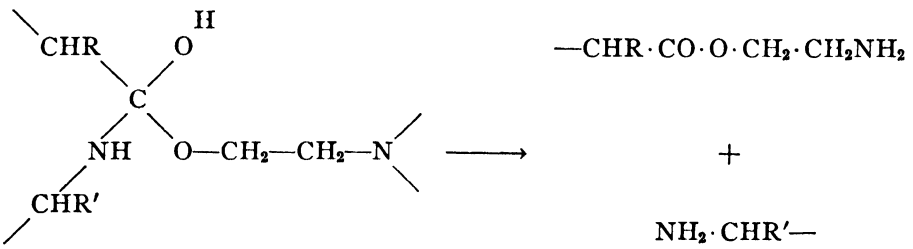
First of all, by electrometric titration he finds that in the very early stages of hydrolysis by acid, more basic than acidic groups are liberated.

At the same time, in agreement with earlier observations of mine (Synge, 1945b), there are produced basic materials other than ethanolamine. These migrate more slowly than ethanolamine towards the cathode on ionophoresis, and can also be separated from other hydrolysis products by chromatographic and counter-current distribution methods. The fraction is tryptophan-free, but yields other amino-acids and ethanolamine on further hydrolysis with acid. Its chromatographic behaviour shows that it is heterogeneous; however, in properties the components do not agree with *N*-aminoacyl- or *N*-peptidyl-ethanolamines. When the substances are treated with alkali, amino groups in them disappear, as can be demonstrated by electrometric forward and back-titration and by Van Slyke amino N determinations, both with the isolated basic materials and with unfractionated partial hydrolysates. It seems reasonable to postulate that the basic materials are *O*-aminoacyl- or *O*-peptidyl-ethanolamines, which undergo rearrangement to the corresponding *N*- compounds on making alkaline, a reaction studied in detail for ethanolamine derivatives by Phillips and Baltzly (1947) and for numerous other hydroxyamino compounds by the late Max Bergmann and colleagues more than twenty years ago:

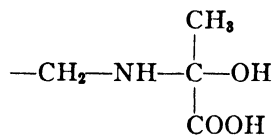
$$NH_2R \cdot CO \cdot O \cdot CH_2 \cdot CH_2NH_2 \rightarrow NH_2R \cdot CO \cdot NH \cdot CH_2 \cdot CH_2OH$$

Finally, on methylation of gramicidin with silver oxide and methyl iodide, Dr. James finds that OMe equivalent to the ethanolamine residues is introduced (see Table 3), but that on hydrolysis with acid under conditions not themselves destructive to *O*-methyl-ethanolamine, none of this compound is produced.

Taken together, these data seem to suggest that in gramicidin it is not the –OH group of ethanolamine that furnishes the free –OH group. It would seem reasonable, very tentatively, to postulate that the ethanolamine residue is linked as follows through its —OH group to part of a peptide chain, and that this structure is split on acid hydrolysis in the following way:



This proposed structure is similar to the intermediate postulated by Phillips and Baltzly (1947) in the *N*-acyl- to *O*-acyl-ethanolamine transformation. Otherwise, such a mixed ortho-peptide, ortho-ester linkage does not seem previously to have been postulated, although it bears similarity to part of the structure of lycomarasmin as formulated by Woolley (1948):



and a somewhat similar structure has been postulated in gliotoxin (for discussion and literature see Elvidge and Spring, 1949).

The amino group of the ethanolamine residue would be in acid-labile linkage of a kind still to be ascertained, possibly with an indole group or an enolised peptide bond.

These suggestions are admittedly very speculative, and possible alternative structures would be most welcome. There still seems no plausible explanation of how the extra basic group is masked, although it seems impossible to deny that it is masked. If linkages of this kind were to occur in proteins in small numbers, it would be very difficult

to detect them. I think also that the type of linkage of an –OH group to a peptide bond, postulated above, deserves very serious consideration in connection with the problem of how carbohydrate and hexosamine residues are incorporated in proteins.

I hope that this account of the present state of physical and chemical studies with gramicidin will focus attention on the numerous problems that require further investigation. Several of these look as if they will be just as perplexing with individual gramicidins as with the mixtures that have so far been studied. However, their elucidation will be watched with interest by workers concerned with many aspects of the study of proteins.

References

Barry, G. T., Gregory, J. D., and Craig, L. C., 1948, The nature of bacitracin. J. biol. Chem. *175:* 485-486.

Biochemical Society Symposia No. 1, 1948, The relation of optical form to biological activity in the amino-acid series. ed. R. T. Williams, Cambridge: University Press.

Birkinshaw, J. H., Raistrick, H., and Smith, G., 1942, Studies in the biochemistry of micro-organisms. 71. Fumaryl-DL-alanine (Fumaromono-DL-alanide), a metabolic product of *Penicillium resticulosum* sp. nov. Biochem. J. *36:* 829-835.

Christensen, H. N., 1943, The isolation of valylvaline from gramicidin hydrolysates. J. biol. Chem. *151:* 319-324.

1944, The configuration of valylvaline in gramicidin. J. biol. Chem. *154:* 427-436.

Christensen, H. N., Edwards, R. R., and Piersma, H. D., 1941, The composition of gramicidin and tyrocidine. J. biol. Chem. *141:* 187-195.

Christensen, H. N., and Hegsted, D. M., 1945, The course of the acid hydrolysis of gramicidin. J. biol. Chem. *158:* 593-600.

Consden, R., Gordon, A. H., Martin, A. J. P., and Synge, R. L. M., 1947, Gramicidin S.: the sequence of the amino-acid residues. Biochem. J. *41:* 596-602.

Dubos, R. J., 1939, Studies on a bactericidal agent extracted from a soil bacillus. 1. Preparation of the agent. Its activity in vitro. J. exp. Med. *70:* 1-10.

Edwards, B. G., 1949, Ultraviolet spectra of some indole derivatives, including tryptophan and gramicidin. Arch. Biochem. *21:* 103-108.

Elvidge, J. A., and Spring, F. S., 1949, Synthesis of a sulphur-containing degradation product from gliotoxin. Nature, Lond. *163:* 94.

Ferrel, R. E., Olcott, H. S., and Fraenkel-Conrat, H., 1948, Phosphorylation of proteins with phosphoric acid containing excess phosphorus pentoxide. J. Amer. chem. Soc. *70:* 2101-2107.

Fraenkel-Conrat, H., Brandon, B. A., and Olcott, H. S., 1947, The reaction of formaldehyde with proteins. 4. Participation of indole groups. Gramicidin. J. biol. Chem. *168:* 99-118.

Fraenkel-Conrat, H., Lewis, J. C., Dimick, K. P., Edwards, B., Reitz, H. C., Ferrel, R. E., Brandon, B. A., and Olcott, H. S., 1946, Chemical derivatives of gramicidin. Proc. Soc. exp. Biol., N.Y. *63:* 302-308.

Gauze, G. F., and Brazhnikova, M. G., 1943, Gramitsidin i yego svoistva. in Sovyetskiĭ gramitsidin i lecheniye ran, ed. Sergiyev, P. G., Moscow: Medgiz. pp. 5-36.

Gordon, A. H., Martin, A. J. P., and Synge, R. L. M., 1943, The amino-acid composition of gramicidin. Biochem. J. *37:* 86-92.

Gregory, J. D., and Craig, L. C., 1948, Counter-current distribution of gramicidin. J. biol. Chem. *172:* 839-840.

Hotchkiss, R. D., 1941, The chemical nature of gramicidin and tyrocidine. J. biol. Chem. *141,* 171-185.

1944, Gramicidin, tyrocidine, and tyrothricin. Advances Enzymol. *4:* 153-199.

Hotchkiss, R. D., and Dubos, R. J., 1941, The isolation of bactericidal substances from cultures of *Bacillus brevis.* J. biol. Chem. *141:* 155-162.

James, A. T., 1949, Private communication.

Lewis, J. C., Dimick, K. P., Feustel, I. C., Fevold, H. L., Olcott, H. S., and Fraenkel, Conrat, H., 1945, Modification of gramicidin through reaction with formaldehyde. Science *102:* 274-275.

Neuberger, A., 1948, Stereochemistry of amino acids. Advances Protein Chem. *4:* 297-383.

Olcott, H. S., Lewis, J. C., Dimick, K. P., Fevold, H. L., and Fraenkel-Conrat, H., 1946, Succinic acid derivatives of gramicidin and methylol gramicidin. Arch. Biochem. *10:* 553-555.

Pedersen, K. O., and Synge, R. L. M., 1948, Diffusion experiments on gramicidin S, tyrocidine, and gramicidin. Acta chem. Scand. *2:* 408-413.

Phillips, A. P., and Baltzly, R., 1947, Rearrangements betwen primary ethanolamides of carboxylic acids and the corresponding aminoethylesters J. Amer. Chem. Soc. *69:* 200-204.

Reitz, H. C., Ferrel, R. E., Fraenkel-Conrat, H., and Olcott, H. S., 1946a, Action of sulfating agents on proteins and model substances. 1. Concentrated sulfuric acid. J. Amer. chem. Soc. *68:* 1024-1031.

Reitz, H. C., Ferrel, R. E., Olcott, H. S, and Fraenkel-Conrat, H., 1946b, Action of sulfating agents on proteins and model substances. II. Pyridine-chlorosulfonic acid. J. Amer. chem. Soc. *68:* 1031-1035.

Sanger, F., 1946, The free amino group of gramicidin S. Biochem. J. *40:* 261-262.

Schales, O., and Mann, G. E., 1947, Gramicidin derivatives. I. Preparation; hemolytic and bacteriostatic properties. Arch. Biochem. *13:* 357-371.

1948, Gramicidin derivatives. II. Toxicity; effect of proteins on hemolytic and bacteriostatic activity; antibacterial effect *in vivo* Arch. Biochem. *18:* 217-228.

Schryver, S. B., and Buston, H. W., 1927, The isolation of some hitherto undescribed products of hydrolysis of proteins. Proc. Roy. Soc., Lond., *101B:* 519-527.

Synge, R. L. M., 1943, Partial hydrolysis products derived from proteins and their significance for protein structure. Chem. Revs. *32:* 135-172.

1944, Analysis of a partial hydrolysate of gramicidin by partition chromatography with starch. Biochem. J. *38:* 285-294.

1945a. "Gramicidin S": over-all chemical characteristics and amino-acid composition. Biochem. J. *39:* 355-362.

1945b, The hydroxyamino component of gramicidin hydrolysates. Biochem. J. *39:* 355-362.

1945c, The kinetics of low temperature acid hydrolysis of gramicidin and of some related dipeptides. Biochem. J. *39:* 351-355.

1948, The synthesis of some dipeptides related to gramicidin S. Biochem. J. *42:* 99-101.

1949a, A further study of hydrolysates of gramicidin. Biochem. J. *44:* 542-548.

1949b, Naturally occurring peptides. Quarterly Reviews *3:* 245-262.

SYNGE, R. L. M., and TISELIUS, A., 1947, Adsorption experiments with gramicidin and related substances: the heterogeneity of tyrocidine. Acta chem. Scand. *1:* 749-762.

1949, Some adsorption experiments with amino-acids and peptides, especially compounds of tryptophan. Acta chem. Scand. *3:* 231-246.

THIMANN, K. V., 1930, The precipitation of the basic amino-acids of proteins with phosphotungstic acid. Biochem. J. *24:* 368-378.

TISHLER, M., STOKES, J. L., TRENNER, N. R., and CONN, J. B., 1941, Some properties of gramicidin. J. biol. Chem. *141:*197-206.

WHITE, L. M., and SECOR, G. E., 1946, Micro-Kjeldahl determination of nitrogen in gramicidin and tryptophan. Comparison of the Gunning-Arnold-Dyer and Friedrich methods. Industr. engng. Chem. (Anal. Ed.) *18:* 457-458.

WOOLLEY, D. W., 1948, Studies on the structure of lycomarasmin. J. biol. Chem. *176:* 1291-1298.

WORK, T. S., 1948, D- and L-amino-acids in antibiotics. in Biochemical Society Symposia No. 1. The relation of optical form to biological activity in the amino-acid series. ed. R. T. Williams. Cambridge: University Press. pp. 61-77.

DISCUSSION

HAUROWITZ: The view that some proteins are not uniform, nor mixtures of two or three proteins, but families of a large number of similar compounds, is supported by the fact that all intermediates exist between normal serum globulins, over poorly adapted antibodies, to well adapted antibodies.

SANGER: The finding of groups or "families" of closely related polypeptides existing together in nature has suggested that proteins themselves may exist in nature, not as pure chemical entities but as "families" of very similar substances. It seems possible that the co-existence of these series of closely related polypeptides may be ascribed to species differences in the organisms producing them. Thus it may be that a single pure strain of bacteria produces a single unique polypeptide and that even if the initial strains used are pure, mutations during growth may lead to a heterogeneous family of bacteria producing a family of polypeptides. The types of differences found in these polypeptides are similar to those found between the same protein from different animal species. Thus, for instance, different insulins probably have the same general overall chemical structure but differ in the nature of a few amino acid residues. Similar differences are found in the polymixins from different strains. My results with insulin (p. 153-158) do suggest that each position in a protein chain is occupied by a single unique amino acid residue, at least for those positions that have been studied. Whether or not this will apply to all proteins or even to all positions in the insulin molecule cannot be certain, but it does suggest that proteins are real chemical entities with a unique structure, and this would seem to be the wisest working hypothesis on which to base a study of protein chemistry.

HAUROWITZ: My belief that antibodies are families of similar proteins is based on experiments where an immune serum produced by the injection of a single antigen is fractionally precipitated by several suitable test antigens. The higher the number of test antigens, the higher also is the number of antibody types found. By extrapolation one is led to the belief that a whole series of intermediates between normal serum globulin and well adapted antibodies is present in an immune serum.

PEPTIDE BOND SYNTHESIS IN NORMAL AND MALIGNANT TISSUE[1]

PAUL C. ZAMECNIK AND IVAN D. FRANTZ, JR.
The Medical Laboratories of the Collis P. Huntington Memorial Hospital of Harvard University, located at the Massachusetts General Hospital, Boston, Massachusetts

A general characteristic of malignant tissue is its capacity for sustained growth at a rate in excess of that of the tissue from which it originated. The word growth is, of course, a comprehensive term for the integrated result of many separate processes. Since protein synthesis is an indispensable part of the growth process, it may be stated as a general postulate that neoplastic tissue must have an increased net protein synthesis as compared with its normal counterpart. There is at the present time a very incomplete description of the nature of factors which regulate the protein metabolism of multicellular animals. It is a source of wonderment that the synthetic and degradative mechanisms are so carefully balanced in the adult animal that essentially no net change in protein mass occurs in the normal course of events.

There is little room for doubt that the potentiality for a sudden burst of cell growth is still present in most adult tissues, as experiments on regeneration of adult liver following partial hepatectomy testify. A problem more basic, and logically antecedent to that of trying to explain the increased growth rate of the neoplasm is that of uncovering the biochemical nature of growth regulators which keep the normal cell in balance.

The accompanying simplified diagram (Fig. 1), however, represents certain steps at which a comparison might, even at the present time, be made of the protein metabolism of a normal tissue and a neoplasm derived from it.

To begin with (1) it is possible that the neoplasm is favored by a rich blood supply, which makes available at the cell border a greater quantity of amino acids per unit time. Secondly, (2) it is conceivable that the neoplasm has a superior mechanism for transporting amino acids across the cell membrane. Recent evidence (Roberts and Tishkoff, 1949) indicates that the intracellular concentration of amino acids differs in certain tumors from that in normal control tissues. It is possible (3) that the neoplasm has a superior ability to synthesize amino acids within the cell, and is thus less dependent on the integrity of an outside supply. There may be (4) an increased rate of protein synthesis; due either to increased intracellular amino acid concentrations with unchanged proteosynthetic mechanisms, or to an increase in the latter. There may be (5) a decrease in the degradation rate of intracellular protein. Finally, the entire protein metabolic cycle may turn over at a greatly accelerated rate, with only a slight increase in net protein construction rate.

The initiation of this study was made possible by the synthesis of C^{14}-carboxyl labeled DL-alanine and glycine (Loftfield, 1947). The primary *p*-dimethylaminoazobenzene induced rat hepatoma was employed for tumor studies in order to provide a neoplasm which could be compared with non-

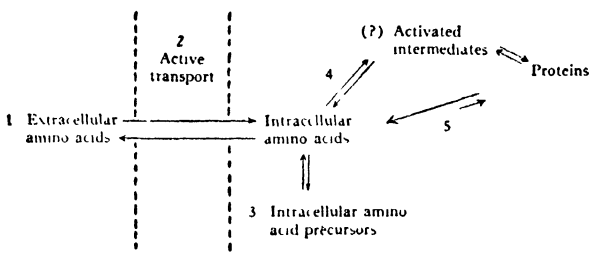


FIG. 1

malignant hepatic tissue in the same animal, as well as with embryonic, regenerating, and normal control liver in animals of the same stock. It is recognized that no statements on the behavior of neoplasms in general can be drawn from results obtained on one species of tumor. Isolated tissue slices have been studied for the most part, rather than whole animals, because the former provide the possibility of keeping more factors under experimental control. Until recently, the restriction in quantity of labeled amino acids available has been a practical consideration which prevented large scale whole animal studies.

Experimental procedure: This has been described in a previous publication (Zamecnik, Frantz, Loftfield, and Stephenson, 1948), but is briefly as follows: rats are killed by decapitation and liver slices approximately 8 x 10 mm in size and 0.5 mm thick are cut by means of a slicer (Stadie and Riggs, 1944). Slices are incubated in a Krebs-Ringer-phosphate medium for 2 to 3½ hours in the presence of the labeled amino acid, in Warburg vessels, with 100 percent oxygen in the gas phase. Ten percent KOH is placed in the center well. Slices are

[1] Aided by grants from the American Cancer Society and the Atomic Energy Commission. This is Publication No. 680 of the Harvard Cancer Commission.

then removed, washed 3 times with water, and homogenized in a solution containing inert alanine (or glycine), 10 mg/cc. The homogenate is precipitated with an equal volume of 20 percent trichloroacetic acid, then resuspended and recentrifuged in 10 percent trichloroacetic acid 3 times. At this point, in the earlier experiments, the precipitate was hydrolyzed in 6N HCl overnight in an autoclave. In later experiments, particularly with glycine, 2 additional washings in 2:2:1 alcohol, ether, chloroform have been carried out. These additional washings have not, however, influenced the slice results. The only recent change in the procedure for preparing protein hydrolysates for counting has been the trapping of $C^{14}O_2$ (either from the ninhydrin procedure or from combustions) in NaOH, then adding $BaCl_2$ to make the $BaCO_3$ precipitate. The $BaCO_3$ precipitate is finer and makes more uniform surfaces on the filtration discs used for counting than where the CO_2 is trapped directly in $Ba(OH)_2$.

TABLE 1. RATE OF INCORPORATION OF C^{14}-CARBOXYL-LABELED DL-ALANINE INTO VARIOUS TYPES OF HEPATIC SLICES FROM RAT

	n	c.p.m.
Normal control liver	24	38± 3.4
Control-hepatoma	7	91±11.4
Hepatoma	8	255±34.8
Regenerating liver	12	91± 6.9
Fetal livers	1	179

n = number of separate livers used in slice experiments. In the case of fetal livers, pooled slices from six 15 day old fetuses were used. The regenerating livers were produced by partial hepatectomies. (Brues, Drury, and Brues, 1949), and experiments on groups of 3 livers were run at 12, 24, 48, and 72 hours (Bucher, Loftfield, and Frantz, 1949). The "control-hepatoma" is the non-malignant part of a liver bearing a hepatoma. Standard errors of the mean are given for several groups. This table contains data on the hepatoma, previously published (Zamecnik, Frantz, Loftfield, and Stephenson, 1948). Experimental details in all experiments are as already described, 10,000 c.p.m. contained in 5μm of DL-alanine being used in all cases. The incorporation rates in the control-hepatoma, regenerating, and hepatoma groups are all statistically significantly different (Arkin and Colton, 1939) from those in the control group.

The carboxyl label makes it possible to use the ninhydrin procedure to liberate CO_2 from the hydrolysate. This is a valuable mechanism to identify the source of the C^{14} to be counted as the carboxyl group of an amino acid. There is, of course, the likelihood that the washed trichloroacetic acid precipitate which is hydrolyzed contains labeled organic material other than protein.

A summary of the results obtained when 5 micromoles of DL-alanine containing 10,000 c.p.m. (counts per minute) was added to slices of hepatomas and control livers is given in Table 1. An end window Geiger-Müller counter was used for measurements of C^{14}. The over-all efficiency of the counting arrangement was approximately 8 percent.

It appears that the hepatoma slice incorporates radioactivity into its proteins more rapidly than does either the normal control liver slice or the slice from the non-malignant part of the liver containing the hepatoma. The rate for the embryonic liver is similar to that for the hepatoma.

It was observed that although the initial pH of all flasks was 7.4, the final pH of the flasks in which normal liver slices were incubated (for 3½ hours) was consistently 6.9-6.7, whereas the pH of the incubation medium for the hepatoma slices was 7.4-7.3. In itself this is an interesting point which deserves further investigation. However, in order to expose both types of slice to the same incubation medium throughout the experiment, flasks were set up in which both normal and hepatoma slices were added to the same vessel. In these combination flasks the pH dropped terminally as low as 6.5, but the rate of incorporation of activity into hepatoma slice proteins still exceeded that into the normal slice proteins by as much as the difference found in the separate flask experiments.

In order to find out what fraction of the radioactivity derived from carboxyl-labeled alanine was incorporated in the form of carboxyl-labeled alanine, crystallizations were first carried out (Frantz, Loftfield, and Miller, 1947). These indicated that the great preponderance of the activity incorporated into the hepatoma was in the form of alanine, but left undecided the question as to whether a few percent might have been contained in glutamic and aspartic acids. In these experiments, KOH had been placed in the center well of the Warburg flasks purposely to serve as a trap for $C^{14}O_2$ liberated from alanine as a result of degradation. Thus it was hoped that the complicating effects of amino acid labeling as a result of $C^{14}O_2$ fixation might be minimized.

It was clear that a method for separating amino acids completely and quantitatively would be a valuable tool in this type of study. Doctors Stanford Moore and William Stein were kind enough to tutor us in the use of starch column chromatography. The first fractionations, carried out in their laboratories according to methods since published (Stein and Moore, 1948, 1949; Moore and Stein, 1948, 1949), demonstrated that protein hydrolysates could be used on the starch column, and that no unusual peaks or appreciable aberrations from the usual site of emergence of peaks occurred, as compared with chromatograms carried out on synthetic mixtures or hydrolysates of purified proteins.

Quite apart from localizing radioactivity among the amino acids of the hydrolysate, the starch column made possible a comparison of the amino acid composition of the proteins of normal liver and of hepatoma (Zamecnik, Frantz, and Stephenson,

1949). The data in Table 2 reveal no appreciable differences, with the exception that the leucine-isoleucine peak and probably the phenylalanine peak are consistently slightly smaller in the hepatoma. Similar comparisons have recently been made for other neoplasms (Dunn, Feaver, and Murphy, 1949; Sauberlich, Blades, and Baumann, 1949).

In the first starch column fractionations, effluent fractions were pooled for individual peaks, and ninhydrins (Van Slyke, MacFadyen, and Hamilton, 1941) were performed on aliquots, in order to prepare $BaCO_3$ precipitates for counting. The crystallization results were confirmed, and as shown in Table 3, 93 percent of the activity was accounted for in the glutamic-alanine fraction. Since these amino

TABLE 2. PERCENTAGE COMPOSITION OF AMINO ACID PEAKS IN LIVER HYDROLYSATES

Average of	6 control livers	5 hepatomas
Leucine, isoleucine	15.4±0.56‡	12.6±0.62*
Phenylalanine	4.0±0.19	3.3±0.16†
Valine, methionine, Tyrosine	10.6±0.60	10.1±0.48
Proline	5.3±0.37	5.4±0.32
Glutamic acid, alanine	17.6±0.77	19.2±1.0
Threonine	4.9±0.23	4.3±0.35
Aspartic acid	9.3±0.39	8.3±0.44
Serine	5.7±0.41	5.8±0.41
Glycine	9.1±0.53	11.0±0.72
Arginine	6.5±0.77	7.2±1.2
Lysine	8.1±0.89	9.1±1.3
Histidine	2.8±0.28	2.7±0.45
Cystine	0.7±0.25	1.0±0.6

* Difference between control and hepatoma significant ($p<.01$).

† Difference between control and hepatoma probably significant ($.05>p>.01$).

‡ Standard error of the mean.

TABLE 3. LOCATION OF RADIOACTIVITY IN AMINO ACIDS DERIVED FROM HEPATOMA SLICES WHICH HAD BEEN INCUBATED IN RADIOACTIVE ALANINE

Amino acid	Total counts c.p.m./0.1 mM $BaCO_3$	Percent of total activity incorporated
Alanine	1,346 ±75*	89.4
Glutamic acid	55 ±23	3.6
Aspartic acid	67.3 ± 2.6	4.5
Glycine	17.2 ± 1.3	1.1
Serine	11.4 ± 1.9	0.8
Leucine, isoleucine	0.6 ± 1.9	
Phenylalanine	0.4 ± 1.4	
Valine, methionine, tyrosine	2.6 ± 2.0	
Proline	2.1 ± 1.9	0.7
Threonine	1.4 ± 1.8	
Arginine	0.04± 1.9	
Lysine	2.8 ± 1.7	
Histidine	0.9 ± 1.8	

Fourteen Warburg flasks, each containing 3 hepatoma slices, 1 cc. Krebs-Ringer-phosphate and 0.1 cc. radioactive alanine (10,000 c.p.m. and 0.44 mg. alanine), plus 0.2 cc. 10% KOH in the center well, were incubated 3½ hours at 37.5° C in an atmosphere of O_2. After incubation all the slices were pooled, washed two times with 280 cc. portions of H_2O, homogenized in 70 cc. of inert DL-alanine solution (10 mg/cc.) and the proteins precipitated by addition of 70 cc. of 20% trichloroacetic acid. The precipitate was washed three times with 70 cc. portions of 10% trichloroacetic acid. It was then hydrolyzed in 14 cc. 6N HCl in an autoclave overnight.

10 cc. of the hydrolysate, containing 1.11 mM of amino acids and a total activity of 2,000 c.p.m. was placed on an 8 cm. starch column. 1,507 c.p.m. is accounted for in this table. Great loss in recovery occurred in concentration and transfer processes involved in preparing the effluents for counting, and in carrying out a second fractionation to separate the glutamic-alanine peak.

* Standard deviation due to statistical fluctuations in counting only.

acids emerged together in the starch column effluent, they were separated by putting them through a second fractionating column (Moore and Stein, 1949). There were definitely a few percent of activity in the carboxyl groups of aspartic and glutamic acid, 1 percent of the total activity in glycine, and recognizable minute amounts in serine. Chromatographic results obtained when carboxyl-labeled glycine was added to surviving slices in the usual way are given in Table 4.

Several points stand out in this single experiment, not yet repeated—of which the greater rate of incorporation of glycine into the hepatoma is the most striking. There is evidence of translocation of activity into a number of other amino acids. Because of the risk that "tails" of radioactivity may persist beyond the peaks of high activity, it appears conservative to look with suspicion on any orders of activity below 1 percent of the total.[2] It is interesting to compare these results with those of Ehrensvärd (1948).

Prior to our use of starch column chromatography, a number of comparisons were made of carbon dioxide fixation in the normal liver and in the hepatoma. The incubation medium for the slices was that of Anfinsen, Beloff, Hastings, and Solomon (1947). The results of the experiments (Table 5)

[2] In explanation of this term, once an amino acid has emerged from the starch column, we have found that a small tail of that amino acid continues to emerge thereafter, continuing indefinitely. The amount of amino acid which thus "tails" on after the emergence of an amino acid peak is so small as to be undetectable by ordinary chemical methods, and has only been found by the use of labeled amino acids. A curious feature of this "tail" is that it may be swept up into a tiny peak by the emergence of a second amino acid.

TABLE 4. LOCATION OF RADIOACTIVITY IN SLICES INCUBATED IN C^{14}-CARBOXYL-LABELED GLYCINE

	Normal		Hepatoma	
	c.p.m.	percent	c.p.m.	percent
Leucine, isoleucine	−7± 5†	0	28± 9	0.1
Phenylalanine	25± 5	0.5	25± 9	0.1
Valine, methionine, tyrosine	38± 5	0.8	381± 12	1.8
Proline	3± 5	0	42± 9	0.2
Glutamic acid, alanine	117*± 5	2.3	327± 16	1.5
Threonine	−5± 5	0	29± 9	0.1
Aspartic acid	17± 6	0.3	66± 13	0.3
Serine	2,165± 62	43.1	3,870±100	18.0
Glycine	2,550±325	50.7	16,580±300	77.2
Arginine	71± 5	1.4	91± 8	0.4
Lysine	10± 5	0.2	15± 9	0.1
Histidine	1± 1	0	12± 9	
Cystine	41± 7	0.8	29± 9	0.1

* This value represents the ninhydrin determination, all other values are from aliquots evaporated in dishes and counted directly.

† Standard deviation due to statistical fluctuation in counting only.

Four normal liver slices were incubated in 2 cc. Krebs-Ringer-phosphate solution containing 2.91 mg. glycine (4.42×10^6 c.p.m.) and 6 hepatoma slices were incubated in 2 cc. Krebs-Ringer-phosphate solution containing 1.32 mg. glycine (2.0×10^6 c.p.m.). 10% KOH was in the center wells. Incubation was carried out in O_2 at 37° for 3½ hours. Slices were washed and hydrolyzed in the usual manner. Each hydrolysate (1.5 cc.) contained about 0.2 mM amino acid. The normal contained 20,600 c.p.m. and the hepatoma 99,300 c.p.m. 0.65 cc. aliquots were put on 2.2 cm columns. Effluent in each amino acid peak was pooled and aliquots were taken for ninhydrin $BaCO_3$ c.p.m. and total c.p.m.

suggest that the surviving hepatoma slice has an increased ability to fix CO_2 into protein, as compared with a control liver slice. There is great variability of results in each series, but the difference between the normal and hepatoma series is statistically valid.

Starch column chromatography was now used to localize the activity among the amino acids in the CO_2 fixation experiments. It was found that if sufficient activity was added to the incubation medium, there was a high enough order of CO_2 fixation into protein to permit a more detailed analysis of the fractions for radioactivity. From effluent tubes, one 0.2 or 0.5 cc. aliquot could be used for colorimetric ninhydrin determination, and another 0.5 cc. aliquot for measurement of radioactivity. For this purpose, copper dishes were employed, and the 0.5 cc. aliquot of the effluent solvent was placed in the center of the dish and allowed to dry before counting. By this counting method, radioactivity present in any position of the amino acid is detected. Two continuous curves are obtained on the effluent fractions, and presumptive identification of the radioactivity as originating in an amino acid is made by coincidence of the peaks. If the activity originates in a certain amino acid in the effluent fractions containing that amino acid, there should be a constant specific activity as measured by the amount of activity per micromole of amino acid determined by the colorimetric ninhydrin. It is obviously possible for radioactivity to be present in a compound other than an amino acid, and to emerge from the column in the same fractions as does some amino acid. It is, therefore, desirable to apply additional criteria for positive identification. With the above reservation in mind, however, this procedure may be used as a convenient survey tool for studying interconversions of amino acids, and allied problems.

In the first experiment on $C^{14}O_2$ fixation into protein, pyruvate was omitted from the incubation medium by accident. In the succeeding two experiments, pyruvate was purposely omitted, in order to ascertain whether the activity distribution patterns obtained were constant, or variable from one slice experiment to the next. The patterns obtained were found to be in satisfactory agreement with each other. The latest chromatograms, obtained by the addition of 0.3 mc of $BaC^{14}O_3$ to each flask, are shown in Figures 2 and 3.

The total amount of CO_2 fixed in protein in the

TABLE 5. FIXATION OF $C^{14}O_2$ INTO PROTEINS OF SLICES OF NORMAL LIVERS AND HEPATOMAS

	n	c.p.m.
Controls	15	39± 6.3*
Hepatomas	15	92±14.6

n = number of separate animals used in experiments. At least two separate flasks were run on slices made from each liver. The differences between the two series are statistically significant ($p<.01$). Each Warburg flask contained 2 or 3 liver slices, 1 cc. Anfinsen's medium without alanine and 0.1 cc. of .05 M $NaHC^{14}O_3$ containing 90,000 c.p.m./0.1 cc. Experiments were run in 95 percent O_2, 5 percent CO_2 for 3½ hours at 37.5° C. Slices were treated the usual way.

* Standard error of the mean.

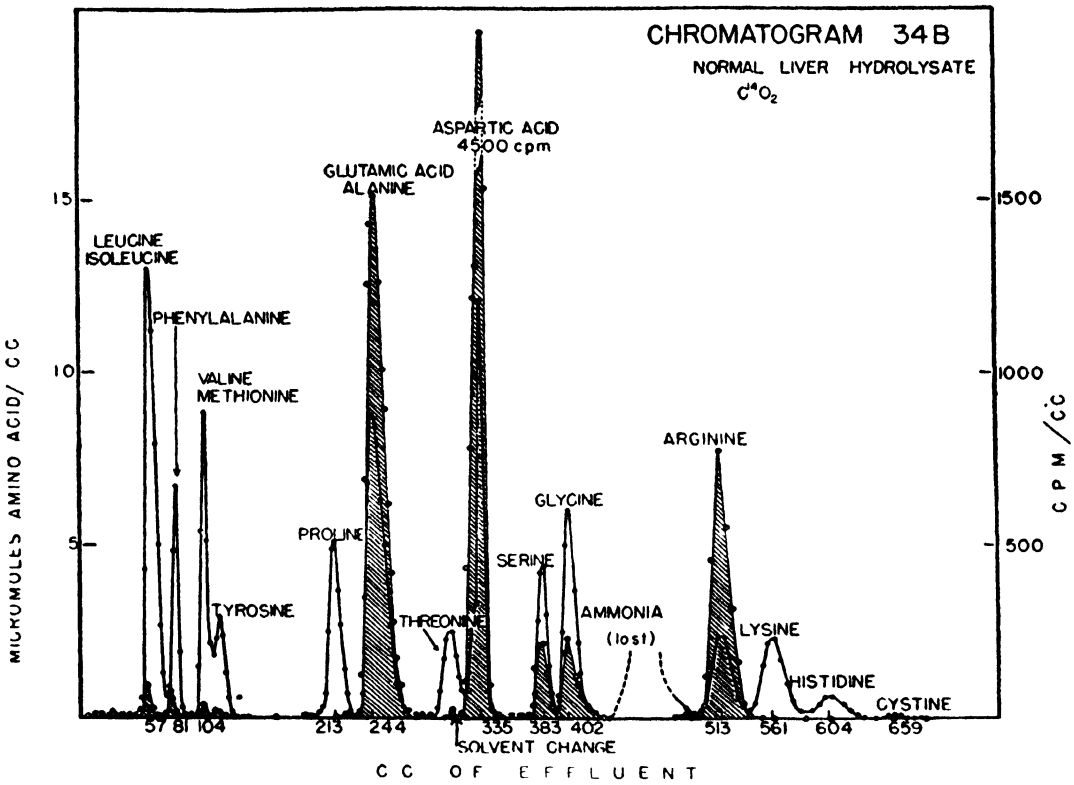


FIG. 2

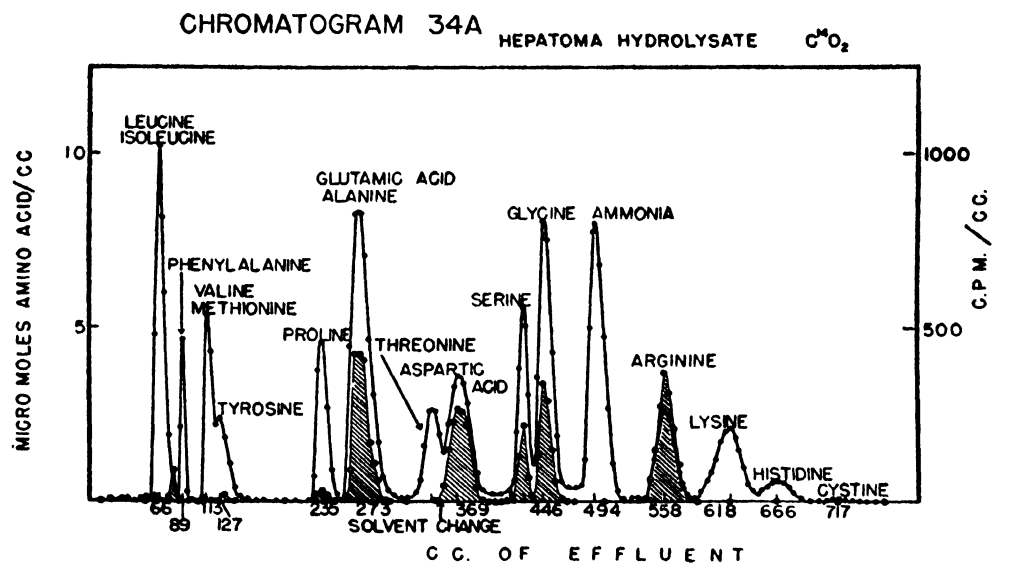


FIG. 3

FIGS. 2 and 3. Separation of amino acids from normal liver and hepatoma protein hydrolysates derived from liver slices which had been incubated in radioactive CO_2 Two double sidearm Warburg flasks, one containing 2 control liver slices and the other 3 hepatoma slices, were used. Each contained 1 cc Anfinsen's medium without pyruvate or alanine. Flasks were flushed with O_2 for 2 minutes, then the $C^{14}O_2$ was liberated from 15 mg $BaC^{14}O_3$ (0.3 mc) in the sidearms. Flasks were incubated 4 hours at 37°. After incubation, remaining CO_2 was trapped. Slices were treated in the usual way. The control hydrolysate (1 cc) contained 51,900 c.p.m. and 0.099 mM amino acid; and the hepatoma hydrolysate (1 cc) contained 18,200 c.p.m. and 0.115 mM amino acid. 0.9 cc of each hydrolysate was lyophilized, then taken up in solvent (1:2:1 butanol: propanol: 0.1N HCl). They were put on 1.9 cm. columns. The solvent was shifted to 2:1 propanol: 0.5N HCl at fraction 75. Aliquots for colorimetric ninhydrins were taken from every tube, and aliquots for counting in dishes from every tube containing amino acids, and every fifth tube where no amino acids were present.

hepatoma slices is less in these experiments than the fixation in normal liver slices, as determined by ninhydrin decarboxylation of aliquots of the protein hydrolysates. Thus the omission of pyruvate from the medium appears to decrease CO_2 fixation in proteins in the hepatoma more than it does in the normal liver slice.[8]

The chromatograms on the normal liver show clearly *that CO_2 is fixed into glycine and serine.* This fixation has not previously been described in animal tissue, although it has been observed to occur in simple plant forms (Ehrensvärd, 1948; Frantz and Feigelman, 1949). The small activity peaks in the leucine-isoleucine, valine-methionine, tyrosine, and threonine color peaks are suggestive of fixation in these amino acids, or in one member of the pair of amino acids emerging together. It should be pointed out that the curves as reproduced are very much compressed, and that the location and activity of these tiny peaks can be determined with some accuracy. The activity found under the phenylalanine peak does *not*, however, appear to be in phenylalanine. The activity peak and the colorimetric peak do not coincide. Furthermore, by spectrophotometry the absorption curve for phenylalanine in the ultraviolet differs from the radioactivity curve. Likewise, the radioactivity cannot be ascribed to tryptophane, which might conceivably emerge in this region, but which should be identifiable spectrophotometrically.

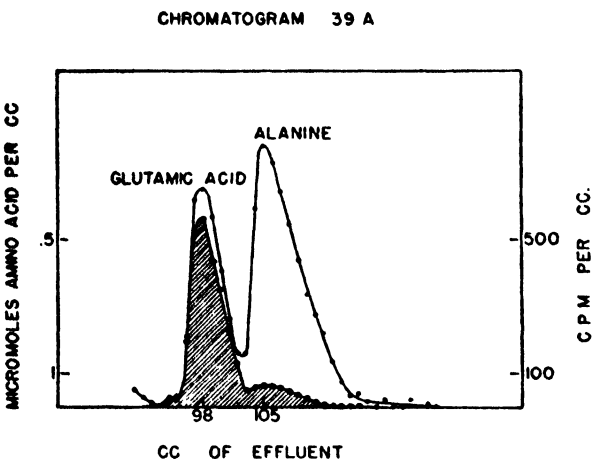


FIG. 4. Glutamic acid, alanine separation of chromatogram 34A. The glutamic acid, alanine fraction was distilled to dryness in vacuo at 50°C. The residue was taken up in 0.7 cc. of 2:1:1 tertiary butanol: secondary butanol: O.1N HCl (Moore and Stein, 1949). This was put on a 1.2 cm. column. 1 cc. fractions were collected. Colorimetric ninhydrin determinations and assays for radioactivity were done on every fraction.

The largest fraction of the CO_2 fixation occurs in the glutamic-alanine, aspartic acid and arginine peaks, in agreement with the results of Anfinsen, Beloff, Hastings, and Solomon (1947). If the glutamic-alanine fraction is put through a second column, a partial separation of these two amino acids is achieved (Fig. 4). Practically all of the activity is found in the glutamic acid fraction, but there is definitely a small amount of fixation in alanine also.

The hepatoma chromatogram (Fig. 3) reveals *less* activity in the glutamic-alanine, aspartic, and arginine peaks, *more* activity in the glycine peak, *and the presence of activity in proline.* No activity is found in the leucine-isoleucine, valine-methionine, and threonine color peaks, in contrast to the chromatogram for the normal liver.

The following data have been assembled on the location of radioactivity in the carbon chain of several amino acids (Table 6).

Combustion data on glutamic acid and alanine are not yet available. A comparison of ninhydrin results and copper dish counts suggests that the preponderance of activity in glutamic acid is located in the carboxyl group adjacent to the alpha position. The $C^{14}O_2$ fixation may be explained in part in terms of known reaction mechanisms of the tricarboxylic acid cycle, as shown in Figure 5. The

TABLE 6. LOCATION OF RADIOACTIVITY IN AMINO ACIDS IN CO_2 FIXATION EXPERIMENTS

	c.p.m./μM amino acid			
	Normal liver		Hepatoma	
	Ninhydrin	Total C	Ninhydrin	Total C
Glycine	300	352	305	354
Serine	306	317	288	255
Aspartic acid	3,380	3,480	546	435
Arginine	0	1,320	0	1,000

Aliquots were taken for ninhydrin $BaCO_3$ and total combustion determinations from pooled amino acid fractions from chromatograms 34A and 34B.

This type of comparison has been repeated several times on aliquots from amino acid peaks in other $C^{14}O_2$ fixation experiments. The results have been essentially the same as those listed above. Owing to the number of manipulations involved, however, we have been as yet unable to achieve a reproducibility of greater than ±10 percent in similar comparisons of the location of radioactivity carried out on amino acids with known positions labeled. The accuracy of the above data is therefore considered to be in the above range. One may conclude that essentially all the activity in aspartic acid and serine is accounted for in the carboxyl groups.

[8] We have lately completed chromatograms on $C^{14}O_2$ fixation when pyruvate is present in the incubation medium. Here the total fixation of $C^{14}O_2$ in hepatoma protein exceeds that in normal protein, as stated above, and the increased fixation occurs in the glutamic and aspartic acid peaks of the hepatoma chromatogram. Thus added pyruvate increased fixation of $C^{14}O_2$ into the glutamic and aspartic acid fractions of the hepatoma protein, for reasons as yet unexplained.

mechanism of fixation of $C^{14}O_2$ in glycine is unexplained. The interconvertibility of glycine to serine is well known (Shemin, 1946; Winnick, Moring-Claesson, and Greenberg, 1948; Sakami, 1948) and Ehrensvärd (1948) has found in yeast evidence that glycine may eventually be converted to proline. Three $C^{14}O_2$ fixation and separation experiments on normal liver and hepatoma slice protein hydrolysates have been done. More radioactivity was added each time the experiment was repeated, *and more small activity peaks rose distinguishably above the baseline as the activity was increased.* These results suggest the persistence in the highly specialized rat liver of vestiges of old CO_2 fixation mechanisms of more primitive living forms.

Thus far, only experiments carried out on isolated surviving liver slices have been discussed. In a recent series of experiments on whole animals, we have found *no* difference in the rate of incorporation of activity from C^{14}-carboxyl-labeled DL-alanine into the hepatoma as compared with the normal liver. These results are summarized in Table 7.

There is no adequate explanation for the discrepancy between these results and those found in the slice experiments.

The observations of Christensen (1949) suggest that growth acceleration may occur as a result of an increased concentration of intracellular amino acids. We have, therefore, considered (Cohen and Zamecnik, 1949) that the increased rate of incorporation of alanine and of glycine into hepatoma slices might be due, at least in part, to an augmentation in the active transport mechanism of amino acids across the cell membrane.

A comparison was made of the concentrating ability of normal liver slices and of hepatoma slices for radioactivity derived from carboxyl-labeled glycine.

Preliminary results imply that the hepatoma slice has a greater capacity to concentrate glycine across the cell membrane than does the normal liver.

To recapitulate these various results in terms of Fig. 1, there is evidence that the rat hepatoma slice as compared with a normal adult rat liver slice may have an increased concentrating mechanism for amino acids, an increased capacity to synthesize certain amino acids from precursors, and an increased ability to convert amino acids into proteins.

TABLE 7. COMPARISON OF RATES OF INCORPORATION OF DL-ALANINE INTO THE NORMAL LIVER AND THE HEPATOMA IN WHOLE FEMALE RATS

	n	c.p.m.
Normal control rats	5	56±4.5*
Control hepatomas	5	57±6.1
Hepatomas	5	50±4.2

The hepatomas were produced as previously described (Zamecnik, Frantz, Loftfield, and Stephenson, 1948). All rats were starved for 4 days, but were given water *ad lib.* Each rat was given a single subcutaneous injection of 1 cc. of solution consisting of 250,000 c.p.m. of C-14-carboxyl-labeled DL-alanine, contained in 0.63 mg. alanine. Three hours later, they were decapitated. One gram aliquots of hepatic tissue were homogenized in 5 cc. of solution containing 50 mg. of inert alanine, and the proteins were precipitated with 5 cc. of 20 percent trichloroacetic acid. The precipitate was washed 3 times with trichloroacetic acid and twice with 2:2:1 alcohol:either:chloroform. The washed precipitate was hydrolyzed overnight in 5 cc. of 6N HCl in an autoclave. Aliquots were taken for ninhydrin determinations.

* Standard error of the mean.

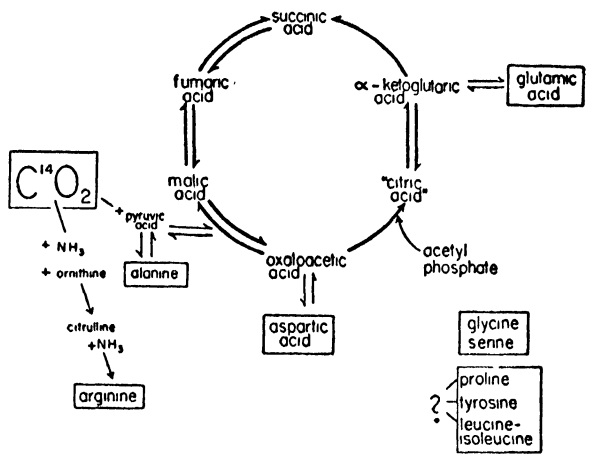


FIG. 5. Suggested mechanism of fixation of carbon from $C^{14}O_2$ into amino acids in rat liver slice proteins.

RELATED EXPERIMENTS

The preceding discussion has been centered on comparison of the behavior of the hepatoma and the normal liver in synthesizing certain peptide bonds. As mentioned previously, progress in understanding the protein metabolism of the neoplastic cell rests on development of greater insight into the general field of biological protein synthesis. It would be advantageous, for example, to have available a great many labeled amino acids, rather than a few, with which to study incorporation rates simultaneously. With this thought in mind, a biosynthesis of amino acids, labeled uniformly in all positions, was considered.

A chemotrophic organism, *Thiobacillus thioöxidans,* was chosen, which was capable of carrying out complete synthesis of organic compounds necessary for its existence and multiplication on a medium containing inorganic salts (including a nitrogen source), CO_2 as a sole carbon source, and inorganic sulfur as an energy source. Inside a closed vessel, $C^{14}O_2$ was generated from the $BaC^{14}O_3$, and the organism was allowed to grow. The bacteria were then harvested, and the proteins precipitated, washed, and hydrolyzed in the usual way. The amino acids were separated by starch column chro-

matography. Thus, a variety of amino acids of high specific activity (150 μc per millimole of carbon), with the carbon atoms uniformly labeled in all positions, can be made available in this way (Frantz and Feigelman, 1949).

In the study of rates of incorporation of amino acids into protein, one complicating factor is the complexity of the population mixture of proteins into which the incorporation takes place. It would be advantageous to be able to study incorporation of an amino acid into a well characterized protein species. With this thought in mind, C^{14}-carboxyl-labeled DL-alanine was injected into the giant silkworm, *Platysamia cecropia*. Twenty-four hours later, the silkworm initiated the spinning of a cocoon. The cocoon fabric exhibited radioactivity and a radioautograph was obtained. The silk was dissolved in a copper hydroxide-ethylene diamine solution, then precipitated out with acetone, and hydrolyzed. A ninhydrin decarboxylation revealed the presence of activity in the carboxyl position, and it is, therefore, presumed that the silk contains a radioactive carboxyl-labeled amino acid (Zamecnik, Loftfield, Stephenson, and Williams, 1949).

In all the preceding discussion, such terms as "synthesis," "turnover," and "exchange," have been avoided. The less committal term "incorporation" has been used instead, because the exact nature of the process under observation is not yet clear. The experimental observation is merely that the carboxyl-labeled amino acids become bound to the protein in such a way that they cannot be released except by hydrolysis of the protein. Until some alternative explanation capable of experimental test is proposed, it seems most profitable to proceed on the assumption that the labeled amino acids become part of the protein molecule by the only reaction now known which is capable of explaining the observed facts, that is, by peptide bond formation. Absolute proof of this interpretation is lacking.

Even granting that the incorporation into protein is by peptide linkage, many points remain obscure. There is no direct evidence to indicate whether the amino acids add to the end of peptide chains growing *de novo*, or whether they "exchange" with like or unlike amino acids in pre-existing protein molecules. If such exchange takes place, it is not known whether it can occur at any point in the molecule, or only at the end of a chain.

The tissue slice is a useful preparation for the study of the overall process, but so far it has provided little informaion concerning mechanisms. The only important observations appear to be the requirement for oxygen (Frantz, Loftfield, and Miller, 1947), the adverse effect of respiratory poisons (Winnick, Friedberg, and Greenberg, 1947, 1948), and the inhibition by an agent known to uncouple respiration and phosphorylation, dinitrophenol (Loomis and Lipmann, 1948; Frantz, Zamecnik, Reese, and Stephenson, 1948). Although it is tempting to interpret these findings as indicative of a coupling between peptide bond formation and known energy releasing mechanisms, proof of this viewpoint is by no means at hand. The relationship may be more remote. For example, the effects noted may result from interference with transport mechanisms. This latter explanation receives some support from experiments reported from other laboratories, in which incorporation of amino acids into the proteins of cell-free homogenates has been claimed. In these experiments, dependence on intact respiration and phosphorylation has not been consistently observed.

Of perhaps greater importance in furthering our understanding of protein synthesis is the more definitive knowledge of the manner of synthesis of several peptide-like compounds. Thus acetyl sulfanilamide (Lipmann, 1945), acetyl choline (Nachmansohn and Machado, 1943), hippuric acid (Borsook and Dubnoff, 1947), para-amino hippuric acid (Cohen and McGilvery, 1947), and glutamine (Speck, 1947), all appear to require energy-rich phosphate bonds for their formation. Recently, a true, naturally occurring peptide, glutathione, has been added to this list (Johnston and Bloch, 1949). Without more direct evidence, reasoning by analogy that a similar pathway is followed in the synthesis of the protein molecule is unwarranted. Nevertheless, a working hypothesis based on this idea appears to be the most attractive available at present.

Returning to the question of exchange, mention should be made of the observations of Axelrod on the phosphate esters (Axelrod, 1948). He showed that when an ester of phosphoric acid and an alcohol is allowed to split enzymatically in the presence of a second alcohol, formation of a second ester involving the second alcohol can be demonstrated. We have observed a similar effect, when glycylglycine is allowed to split in the presence of labeled glycine (Frantz, Loftfield, and Werner, 1949). The equilibrium concentration of the peptide to be expected can be calculated from thermal data (Borsook and Dubnoff, 1940), and has now been confirmed by direct analysis of equilibrium mixtures, with the aid of starch column chromatography. When glycylglycine is allowed to split in the presence of labeled glycine, and the reaction is stopped short of equilibrium, an amount of radioactivity considerably greater than that to be expected at equilibrium is observed in the peptide. This observation may be interpreted as a demonstration of the utilization of energy derived from the splitting of a peptide bond for the formation of a new, similar bond.

In summary of the data from experiments on partially purified systems, it is reasonably clear that labeled amino acids may be incorporated into peptides in at least three ways: (1) by reversal of proteolysis, (2) by exchange, and (3) by utilization of energy-rich phosphate bonds. The relationships of

these three processes to true protein synthesis, are problems for the future.

In conclusion, a number of incomplete studies have been discussed which have a bearing on the problem of peptide bond synthesis in biological systems, and which have some relationship to the problem of why a neoplasm grows at an accelerated rate. The present status of our knowledge may, however, be summarized in a two line poem by Robert Frost, entitled "The Secret":

"We dance round in a ring and suppose,
But the Secret sits in the middle and knows."

References

ANFINSEN, C. B., BELOFF, A., HASTINGS, A. B., and SOLOMON, A. K., 1947, The *in vitro* turnover of dicarboxylic amino acids in liver slice proteins. J. biol. Chem. *168:* 771-772.

ARKIN, H., and COLTON, R. R., 1939, An outline of statistical methods, New York, 4th edition, 126 pp.

AXELROD, B., 1948, A new mode of enzymatic phosphate transfer. J. biol. Chem. *172:* 1-13.

BORSOOK, H., and DUBNOFF, J. W., 1940, The biological synthesis of hippuric acid *in vitro*. J. biol. Chem., *132:* 307-324.

1947, Synthesis of hippuric acid in liver homogenate. J. biol. Chem. *168:* 397-398.

BRUES, A. M., DRURY, D. R., and BRUES, M. C., 1936, A quantitative study of cell growth in regenerating liver. Arch. Path. *22:*658-673.

BUCHER, N. L. R., LOFTFIELD, R. B., and FRANTZ, I. D., Jr., 1949, The effect of regeneration on the rate of protein synthesis and degradation in rat liver. Cancer Research *9:* 623.

CHRISTENSEN, H. N., 1949, Concentration of amino acids by the rat diaphragm suspended in artificial media. Federation Proc. *8:* 190.

COHEN, N. and ZAMECNIK, P. C., 1949, Unpublished data.

COHEN, P. P. and MCGILVERY, R. W., 1947, Peptide bond synthesis. III. On the mechanism of *p*-aminohippuric acid synthesis. J. biol. Chem. *171:* 121-133.

DUNN, M. S., FEAVER, E. R., and MURPHY, E. A., 1949, The amino acid composition of a fibrosarcoma and its normal homologous tissue in the rat, Cancer Research *9:* 306-313.

EHRENSVÄRD, G., 1948, Amino acid metabolism in *Torulopsis utilis*. Cold Spring Harbor Symposium Quant. Biol. *13:* 81-87.

FRANTZ. I. D., Jr., and FEIGELMAN, H., 1949, Biosynthesis of amino acids uniformly labeled with radioactive carbon, for use in the study of growth. Cancer Research *9:* 619.

FRANTZ, I. D., Jr., LOFTFIELD, R. B., and MILLER, W. W., 1947, Incorporation of C-14 from carboxyl-labeled DL-alanine into the proteins of liver slices. Science *106:* 544-545.

FRANTZ, I. D., Jr., LOFTFIELD, R. B., and WERNER, A. S., 1949, Observations on the equilibrium between glycine and glycylglycine in the presence of liver peptidase. Federation Proc. *8:* 199.

FRANTZ, I. D., Jr., ZAMECNIK, P. C., REESE, J. W., and STEPHENSON, M. L., 1948, The effect of dinitrophenol on the incorporation of alanine labeled with radioactive carbon into the proteins of slices of normal and malignant rat liver. J. biol. Chem. *174:* 773-774.

JOHNSTON, R. B., and BLOCH, K., 1949, The synthesis of glutathione in cell-free pigeon liver extracts. J. biol. Chem. *179:* 493-494.

LIPMANN, F., 1945, Acetylation of sulfanilamide by liver homogenates and extracts. J. biol. Chem. *160:* 173-190.

LOFTFIELD, R. B., 1947, Preparation of C^{14}-labeled hydrogen cyanide, alanine and glycine. Nucleonics *1:* 54-57.

LOOMIS, W. F., and LIPMANN, F., 1948, Reversible inhibition of the coupling between phosphorylation and oxidation. J. biol. Chem. *173:* 807-808.

MOORE, S., and STEIN, W., 1948, Photometric ninhydrin method for use in the chromatography of amino acids. J. biol. Chem. *176:* 367-388.

1949, Chromatography of amino acids on starch columns. Solvent mixtures for the fractionation of protein hydrolysates. J. biol. Chem. *178:* 53-77.

NACHMANSOHN, D., and MACHADO, A. L., 1943, The formation of acetylcholine. A new enzyme: "Choline Acetylase." J. Neurophysiol. *6:*397-403.

ROBERTS, E., and TISHKOFF, G. H., 1949, Distribution of free amino acids in mouse epidermis in various phases of growth as determined by paper partition chromatography. Science *109:* 14-16.

SAKAMI, W., 1948, Conversion of formate and glycine to serine and glycogen in the intact rat. J. biol. Chem. *176:* 995-996.

SAUBERLICH, H. E., BLADES, C. E., and BAUMANN, C. A., 1949, Amino Acids in certain normal and neoplastic tissues. Cancer Research. *9:* 613.

SHEMIN, D., 1946, The biological conversion of L-serine to glycine. J. biol. Chem. *162:* 297-307.

SPECK, J. F., 1947, The enzymatic synthesis of glutamine. J. biol. Chem. *168:* 403-404.

STADIE, W. C., and RIGGS, B. C., 1944, Microtome for the preparation of tissue slices for metabolic studies of surviving in vitro. J. biol. Chem. *154:* 687-690.

STEIN, W. H., and MOORE, S., 1948, Chromatography of amino acids on starch columns. Separation of phenylalanine, leucine, isoleucine, methionine, tyrosine, and valine, J. biol. Chem. *176:* 337-365.

1949, Amino acid composition of B-lactoglobin and bovine serum albumin. J. biol. Chem. *178:* 79-91

VAN SLYKE, D. D., MACFADYEN, D. A., and HAMILTON, P., 1941, Determination of free amino acids by titration of the carbon dioxide formed in the reaction with ninhydrin. J. biol. Chem. *141:* 671-680.

WINNICK, T., FRIEDBERG, F., and GREENBERG, D. M., 1947, Incorporation of C^{14}-labeled glycine into intestinal tissue and its inhibition by azide. Arch. Biochem. *15:* 160-161.

1948, The utilization of labeled glycine in the process of amino acid incorporation by the protein of liver homogenate. J. biol. Chem. *175:* 117-126.

WINNICK, T., MORING-CLAESSON, I., and GREENBERG, D. M., 1948, Distribution of radioactive carbon among certain amino acids of liver homogenate protein, following uptake experiments with labeled glycine. J. biol. Chem. *175:* 127-132.

ZAMECNIK, P. C., FRANTZ, I. D., Jr. and STEPHENSON, M. L., 1949, Use of starch column chromatography in study of amino acid composition and distribution of radioactivity in proteins of normal rat liver and hepatoma. Cancer Research *9:* 612-613.

ZAMECNIK, P. C., FRANTZ, I. D., Jr., LOFTFIELD, R. B., and STEPHENSON, M. L., 1948, Incorporation *in vitro* of radioactive carbon from carboxyl-labeled DL-alanine and

glycine into proteins of normal and malignant rat livers. J. biol. Chem. *175:* 299-314.

ZAMECNIK, P. C., LOFTFIELD, R. B., STEPHENSON, M. L., and WILLIAMS, C. M., 1949, Biological synthesis of radioactive silk. Science *109:* 624-626.

DISCUSSION

SCHULMAN: I would like to answer the objections of Dr. Keston regarding his proof of the non-existence of peptide bond formation in tissue homogenates. I am speaking on behalf of Dr. Greenberg in whose laboratory I participated in the protein synthesis problem *in vitro*.

It is difficult to conceive that the process of labeled amino acid incorporation is governed by adsorption or related non-enzymatic processes since it has been demonstrated that: a) the process of labeled amino acid uptake is inhibited by heating, lack of oxygen, or by the presence of small quantities of azide or cyanide; b) in experiments with $NH_2CH_2C^{14}OOH$ the radioactive protein material precipitated by trichloracetic acid on treatment with ninhydrin solution will not evolve isotopic CO_2, while complete acid or alkaline hydrolysis of the protein will release the C^{14} of the protein; c) similarly, enzymatic hydrolysis of the protein with pepsin or trypsin to the peptide stage liberated no $C^{14}O_2$ on ninhydrin treatment until hydrolysis to the amino acid stage was effected with a peptidase preparation; d) the many-fold increase in the uptake process by embryonic tissues and the response of regenerating liver to greater incorporation of label leaves little doubt that the radioactive protein formed in the incubations of tissue homogenates with labeled amino acids represents a true incorporation of the amino acid into the protein molecule.

INDEX TO VOLUME XIV