vious that the slope of this curve plotted against temperature would be, like figures 1 and 2, of the "lambda" type.

We will await the results of the precision study now under way before attempting to reduce the qualitative picture we have given of the balance between energy and entropy to quantitative terms.

* This research is supported by the Atomic Energy Commission.

[1] *J. Colloid Science*, **7**, 551 (1952).

[2] Hildebrand, J. H., and Scott, R. L., *Solubility of Nonelectrolytes*, Reinhold Pub. Corp., New York, 1950. See Chap. VIII.

[3] Zimm, B. H., *J. Phys. Coll. Chem.*, **54**, 1306 (1950).

[4] Murray, F. E. and Mason, S. G., *Canad. J. Chem.*, **30**, 550 (1952).

[5] Hildebrand, J. H., Fisher, B. B., and Benesi, H. A., *J. Am. Chem. Soc.*, **72**, 4348 (1950).

[6] Kohler, F., *Monatsh. f. Chemie*, **82**, 913 (1951).

CHROMATOGRAPHIC SEPARATION AND IDENTIFICATION OF SOME PEPTIDES IN PARTIAL HYDROLYSATES OF GELATIN

BY W. A. SCHROEDER, LEWIS HONNEN, AND F. CHARLOTTE GREEN

GATES AND CRELLIN LABORATORIES OF CHEMISTRY,* CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA

Communicated by Linus Pauling, November 4, 1952

Recently we have been engaged in a study of the chemical structure of collagen and gelatin with the object of determining the sequence of the amino acid residues in the polypeptide chains of these proteins. In the course of this study we have made considerable progress in the chromatographic analysis of complex mixtures of peptides and we have isolated and identified several simple peptides which occur in partial hydrolysates of gelatin. The initial separation of the mixture into zones of one or more peptides has been made on a column of ion exchange resin; further separation of the peptides in each zone has been achieved by chromatographing in the form of dinitrophenyl (DNP) peptides on columns of silicic acid-Celite. It is to be hoped that the particular combination of chromatographic methods which has been successfully used in the present study will be helpful in the resolution of the complex mixtures which result from the partial hydrolysis of other proteins.

Although the amino acid composition of collagen and of its derivative, gelatin, is fairly well established by good analyses[1] (which have shown that there is virtually no difference in the amino acid composition of the two), relatively little is known about the sequences of amino acid residues in these proteins. In 1936 Grassman and Riederle[2] succeeded in isolating the tripeptide lysylprolylglycine (lys-pro-gly)[3] from partial hydrolysates of

gelatin. Later Gordon, Martin, and Synge[4] obtained evidence for sequences which contained glycine-leucine, proline-alanine, proline-glycine, and proline-alanine-glycine but the exact sequences were not determined. Very recently, Heyns, Anders, and Becker[5] found the peptides glu-gly, ala-gly, gly-asp, ala-ala-gly, and ala-(gly, glu) in partial hydrolysates of gelatin.

The sequences of the amino acid residues in collagen and gelatin are especially pertinent to the evaluation of the structures which have been put forward on the basis of data from amino acid composition and x-ray diffraction patterns. Some years ago Astbury[6] suggested that in collagen the sequence will be —P—G—R—P—G—R—P—G—R— where P, with the exception of one residue in 18, stands for either proline or hydroxyproline, G for glycine, and R for one or other of the remaining residues. Recently, Pauling and Corey[7] have described a definite configuration for the polypeptide chains in collagen and gelatin which embodies the same sequence of residues. In both of these quite different structures the packing of the residues appears to require that proline (or hydroxyproline) and glycine be present in the sequence of a prolyl (or hydroxyprolyl) glycine peptide and not in the reverse order. Thus far, only lys-pro-gly, isolated by Grassman and Riederle, and gly-asp, isolated by Heyns, Anders, and Becker, are in definite accord with this requirement.

GENERAL EXPERIMENTAL PROCEDURE

The peptides which were present in the partial hydrolysates were resolved into a number of discrete zones by means of the ion exchanger Dowex-50. The procedure for this separation differed from that of Moore and Stein[8] for the analysis of amino acids on 100-cm. columns only in that a temperature of 37.5° was maintained throughout the chromatogram. The effectiveness of the separation was assessed by applying the ninhydrin procedure of Moore and Stein[9] to every other one of the approximately 1-ml. fractions of effluent. We have tried to improve the separation of the peptide zones by altering the conditions of development but none of our modifications seemed to offer any obvious advantage over the procedure which is indicated above.

Because the partial hydrolysate of a protein is so complex a mixture, one would not expect the relatively few peptide zones which can be isolated from the ion exchange column to be composed individually of a single compound. One is, therefore, faced with the choice of a method for the separation, identification, and estimation of the components of the mixtures. This choice is restricted because of the salt content of the buffers which are required by the ion exchange method. The quantity of salt is manyfold greater than the quantity of peptide, and hence the salt would interfere with attempts at further chromatographic separation, say by paper chromatography. We have resolved this difficulty by converting the materials

in the zones to the DNP-derivatives. The conversion to the DNP-peptides is not affected by the presence of the salts and the separation of salts and DNP-peptides occurs when the latter are extracted from the acidified reaction mixture with ether or ethyl acetate. The resolution of the DNP-peptides was then carried out on columns of silicic acid-Celite. The methodology was essentially that of Green and Kay[10] for the separation of DNP-amino acids except that stronger developers were required. Very satisfactory separations have been obtained by starting the development of the chromatogram with a solution which contains 8 volume per cent of acetic acid and 4 volume per cent of acetone in ligroin (abbreviated 8AA-4A-L)[10] and then changing to 3AA-15A-L, 4AA-20A-L, or 5AA-25A-L as may be necessary to develop strongly adsorbed DNP-peptides. If the zones isolated in this manner proved to be heterogeneous, 2AA-10A-B or 3AA-15A-B (B = benzene) were valuable in achieving further separations.

The constitution of the DNP-peptides so separated has been established by hydrolyzing them completely and determining the N-terminal amino acid by the method of Green and Kay.[10] The remaining amino acids of the peptide are present in the hydrolysate as the free amino acids; they were identified by dinitrophenylating them and identifying the DNP-amino acids so produced. The quantities were determined spectrophotometrically in glacial acetic acid.

RESULTS

Initial Separation of Peptides on Ion Exchange Chromatograms.—Figure 1 compares the results of identical chromatograms of a complete and of a partial hydrolysate of about 12 mg. of gelatin. In this figure, the optical densities of the various fractions after application of the ninhydrin procedure are plotted against the fraction number.

The positions of those zones which emerge during the development with pH-3.42 buffer are in good agreement with the positions of the amino acids in figure 1 of Moore and Stein's paper,[8] and very likely in this region the zones from the partial hydrolysate contain mainly free amino acids. That peptides may be present, however, is shown by the fact that the ninhydrin color from the proline zone of the partial hydrolysate definitely is not the usual brown-red color given by proline itself; the abnormal color indicates the presence of material with a free amino group. The difference in the chromatograms throughout the development with the other buffers clearly proves the presence of peptides in the zones of the partial hydrolysate.

The chromatograms which have been described were intended largely to explore the potentialities of the ion exchanger as a means of initial separation of the peptides in the partial hydrolysate: the results were encouraging. However, further investigation was necessary in order to determine whether or not other conditions of hydrolysis might yield more peptide zones or greater amounts of those already observed.

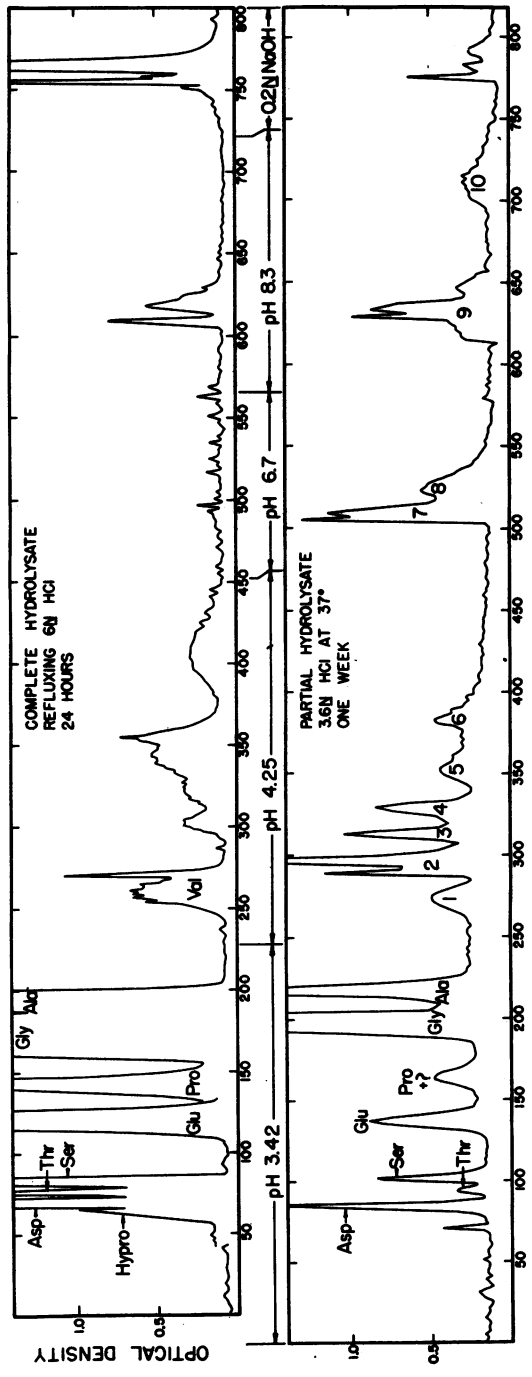

FIGURE 1

Comparison of chromatograms of partial and complete hydrolysates of gelatin on Dowex-50.

It was found that longer hydrolyses in 3.6 *N* hydrochloric acid at 37° exhibited no obvious advantages over the one-week hydrolysis as far as a source of peptides was concerned. A basic hydrolysate prepared essentially in the manner of Heyns, Anders, and Becker[5, 11] yielded a chromatogram which was very different from that of figure 1. During hydrolysis in dilute oxalic acid the preferential release of aspartic acid is a prominent feature, as reported by Partridge and Davis.[12]

Further Separation and Identification of Peptides.—In order to obtain sufficient material for a definitive study of the peptides in the various zones, the partial hydrolysate of about 250 mg. of gelatin which had been hydrolyzed in 3.6 *N* hydrochloric acid at 37° for one week was chromatographed on a large ion exchange column. The investigation of those zones of this large-scale chromatogram which are equivalent to zones 1, 2 and 3 in figure 1 is essentially complete.

Zone 1.—This zone consists almost entirely of thr-gly. The quantity is appreciable and as a minimum accounts for 40% of the threonine in gelatin on the basis of the analysis given by Tristram.[1] If corrections are applied, 55% of the threonine may be accounted for in this form. This correction which may or may not be applicable is largely a correction for incomplete dinitrophenylation of the peptides from the ion exchanger. In a series of experiments with known peptides a yield of more than 75% was not obtained despite wide variations in the conditions of dinitrophenylation; the presence of buffer salts in the reaction mixture was without effect on the yield. Whether such a correction is applicable to all peptides is open to question.

Zone 1a.—Due to incidental differences in the way in which the large- and small-scale ion exchange chromatograms were run, the large chromatogram exhibited a minor zone between zones 1 and 2. When the DNP-peptides which resulted from the dinitrophenylation of this minor zone were chromatographed, five zones separated, and two of these were present in sufficient amount to warrant further study.

The more strongly adsorbed zone contained a glycyl peptide or peptides of which the other constituent amino acids were glycine, alanine, proline, and hydroxyproline. Quantitatively, the N-terminal glycine, the other glycine, and the proline were present in equimolar proportions and in twice the molar amount of both the alanine and hydroxyproline. One might interpret these results to mean that a mixture of gly-(gly, pro, ala) and gly-(gly, pro, hypro) is present. However, on the basis of the chromatographic properties of known DNP-peptides such a conclusion is untenable: a more reasonable inference is that the mixture contains gly-(gly, pro, ala, gly) and gly-(pro, hypro).

The less strongly adsorbed zone was shown to be glu-gly. Some glu-gly is also present in zone 2.

Zone 2.—The initial separation of the DNP-peptides from zone 2 yielded five definite bands on the chromatogram. For simplicity they will be designated as zones 2a, 2b, etc., in the order of decreasing adsorption affinity.

Zone 2a is present in relatively minor amount and appears to contain gly-(hypro, gly) as the main constituent. No definite conclusions can be made about the other peptides which are present in still smaller amounts.

Zone 2b is the main peptide in zone 2 and contains hypro-gly in amount equal to 13 to 19% of the hydroxyproline in gelatin. Although hydrolysis alters or destroys DNP-proline and DNP-hydroxyproline more easily than other DNP-amino acids, the detection of DNP-prolyl or DNP-hydroxyprolyl peptides is very simple because the spectra of these *imino* peptides differ from those of the *amino* peptides. Whereas the spectra of the DNP-amino acids in glacial acetic acid have the main maximum at 338 to 342 mμ, the spectrum of DNP-proline has a maximum at 360 mμ and that of DNP-hydroxyproline at 355 mμ and the latter two, in addition, have no inflection in the region from 380 to 420 mμ.[13] The spectra of DNP-prolyl and DNP-hydroxyprolyl peptides show the characteristics of DNP-proline and DNP-hydroxyproline.

Zone 2c was composed of a mixture of three peptides. Rechromatographing with 3AA-15A-B separated ser-gly from a mixture of gly-glu and glu-gly. The serine in ser-gly accounts for 23 to 33% of the serine in gelatin.

Zone 2d contains glu-ala and a small amount of ala-gly.

Zone 2e may contain thr-ala but the amount is small.

Zone 3.—Six definite zones separated when the DNP-peptides were chromatographed. So far only the main peptide, ala-gly, has been completely investigated.

DISCUSSION

The work of Grassman and Riederle,[2] of Heyns, Anders, and Becker[5] and of the present study has resulted in the identification of an appreciable number of amino acid sequences in gelatin: Grassman and Riederle isolated lys-pro-gly; Heyns, Anders, and Becker found glu-gly, ala-gly, gly-asp, ala-ala-gly, and ala-(gly, glu); and we have definitely identified thr-gly, glu-gly, glu-ala, gly-glu, ser-gly, hypro-gly, and ala-gly and possibly have also isolated thr-ala, gly-(pro, hypro), gly-(gly, gly, ala, pro) and gly-(gly, hypro). If we compare these peptides with the sequence —R—pro-(or hypro)-gly- which is suggested by the structures of Astbury[6] and of Pauling and Corey,[7] it will be observed that the lys-pro-gly, gly-asp, gly-glu, and hypro-gly fit the sequence whereas glu-gly, ala-gly, ala-ala-gly, ala-(gly, glu), thr-gly, glu-ala, and ser-gly do not. At first glance these data appear to cast doubt on the proposed structures. However, the most reliable analyses of gelatin[1] show that although one-third of the residues are

indeed glycine, as required by the structures, only about one-fourth instead of one-third are proline and hydroxyproline. Hence, at a maximum only three-fourths of the protein could have the sequence—R—pro (or hypro)-gly- and the remaining one-fourth would have no predictable sequence. Accordingly, we would expect to find appreciable amounts of peptides whose sequences do not agree with the over-all structures which have been proposed. The available data are still too meager to permit a definite decision as to the essential correctness or incorrectness of the suggested structures. It will be necessary to accumulate much more information, and we feel especially that quantitative information is essential. We believe that the isolation of trace amounts of various peptides from the partial hydrolysate of a protein is of dubious significance in elucidating the structure of a protein, but that a major step in the determination of structure will have been taken when the greater part of its content of amino acids is accounted for in the form of various peptides in a partial hydrolysate. To this end, our experiments have been designed to yield at least roughly quantitative data, and we have indicated the quantities which have been isolated in those instances in which the amount was appreciable.

It is probable that the effectiveness of the combination of chromatographic methods which we have used rests in the differences in the mechanisms by which separations are effected by the ion exchanger and by the silicic acid. During the chromatography of the free peptides the anchoring groups, both ionic and non-ionic, which operate in the adsorption process on the ion exchanger, are different from those anchoring groups which are effective in the adsorption of the DNP-peptides on silicic acid. Thus, although a group of peptides may not be separable on the ion exchanger under the conditions which have been used, their conversion to the DNP-derivatives so alters their properties that they behave very dissimilarly when chromatographed on silicic acid.

Acknowledgment.—This investigation was supported in part by a grant-in-aid from E. I. du Pont de Nemours and Company.

* Contribution No. 1748.

[1] Tristram, G. R., *Adv. Prot. Chem.*, **5**, 143 (1949).

[2] Grassman, W., and Riederle, K., *Biochem. Z.*, **284**, 177 (1936).

[3] The abbreviations for the amino acids are those of E. Brand (*Ann. N. Y. Acad. Sci.*, **47**, 187 (1946)). A convention similar to that of F. Sanger and H. Tuppy (*Biochem. J.*, **49**, 463 and 481 (1951)) will be used to indicate known and unknown sequences; thus, hyphens as in gly-ala-gly show a known sequence whereas commas as in gly, ala, gly, denote an unknown sequence. Gly-(ala, gly) represents a glycyl peptide of partially known sequence.

[4] Gordon, A. H., Martin, A. J. P., and Synge, R. L. M., *Biochem. J.*, **37**, 92 (1943).

[5] Heyns, K., Anders, G., and Becker, E., *Z. physiol. Chem.*, **287**, 120 (1951).

[6] Astbury, W. T., *J. Int. Soc. Leather Trades' Chem.*, **24**, 69 (1940).

[7] Pauling, L., and Corey, R. B., these PROCEEDINGS, **37**, 272 (1951).

[8] Moore, S., and Stein, W. H., *J. Biol. Chem.*, **192**, 663 (1951).

[9] Moore, S., and Stein, W. H., *Ibid.*, **176**, 367 (1948).
[10] Green, F. C., and Kay, L. M., *Anal. Chem.*, **24**, 726 (1952).
[11] Heyns, K., Anders, G., and Becker, E., *Z. physiol. Chem.*, **286**, 118 (1950).
[12] Partridge, S. M., and Davis, H. F., *Nature*, **165**, 62 (1950).
[13] Sanger, F., *Biochem. J.*, **45**, 563 (1949).

VIABILITY OF HYBRIDS BETWEEN LOCAL POPULATIONS OF DROSOPHILA PSEUDOOBSCURA

BY M. VETUKHIV

COLUMBIA UNIVERSITY, NEW YORK

Communicated by Th. Dobzhansky, November 17, 1952

Artificial populations of *Drosophila pseudoobscura* have been made in laboratory experiments which involved two or more chromosomal types differing in the gene arrangements in the third chromosomes. Provided that the chromosomal types are derived from the population of the same geographic locality, the artificial populations eventually reach equilibria at which each type continues to occur with a certain characteristic frequency.[1, 2] The attainment of such equilibria means that the structural heterozygotes which carry two chromosomes with different gene arrangement are heterotic, i.e., possess higher adaptive values than do the corresponding homozygotes. If, however, the chromosomal types are derived from geographically remote populations, then some of the chromosomes may eventually be eliminated and others established.[3, 4] This has been interpreted to mean that the heterosis is produced by juxtaposition of chromosomes which carry different gene complexes coadapted, fitted together, by natural selection in the process of evolution. Chromosomes found in geographically remote populations need not be coadapted, and the structural heterozygotes obtained in interracial crosses need not show heterosis.

In the above experiments, the cytologically visible characteristics of the chromosomes were used as markers which permitted an observer to follow the fates of different gene complexes in populations. It seems, however, probable that heterosis is the normal state of natural populations of most sexually reproducing diploid organisms, regardless of whether their chromosomes are or are not visibly polymorphic. The experiments reported in the present article were designed to test this hypothesis.

Material and Technique.—The material used for the experiments described below consisted of four strains of *Drosophila pseudoobscura* from each of the following localities: Piñon Flats, Mount San Jacinto, California; Mather, Sierra Nevada, California; Bryce National Park, Utah; Ferron, Utah, and Black Canyon of the Gunnison National Monument,