

152. A STUDY OF THE PARTIAL ACID HYDROLYSIS OF SOME PROTEINS, WITH SPECIAL REFERENCE TO THE MODE OF LINKAGE OF THE BASIC AMINO-ACIDS

BY A. H. GORDON, A. J. P. MARTIN AND R. L. M. SYNGE

From the Wool Industries Research Association, Torridon, Headingley, Leeds

(Received 19 November 1941)

THE classical Hofmeister-Fischer hypothesis of the peptide structure of proteins to-day finds almost universal acceptance as a basis of chemical study. This implies that proteins for the most part are built up of peptide chains of the type



where R , R' , R'' are the side-chains of the different amino-acids, of which only 20 odd molecular species have been isolated so far from the whole range of living organisms. It seems clear, therefore, that the distinctive properties of proteins reside in the arrangement rather than the nature of the amino-acid residues.

Bergmann & Niemann [1936; 1937; 1938] have put forward a theory of protein structure based on the idea that amino-acid residues are repeated at regular intervals along the peptide chain, and, further, that the 'frequency' of such recurrency must be factorizable entirely by 2 and/or 3, as must the total number of amino-acid residues in the molecule, and therefore the total number of residues of each amino-acid species in the molecule.

This hypothesis imposes considerable limitations on the astronomical possibilities of protein isomerism permitted by the Hofmeister-Fischer hypothesis. The main positive chemical evidence presented in its favour by Bergmann and others has been the results of amino-acid analysis of complete hydrolysates of proteins. It is clear that the only direct evidence in favour of a hypothesis concerning the order of amino-acid residues in a peptide chain can come from the isolation and identification of fragments of this chain containing at least two amino-acid residues.

Since n molecular species of simple amino-acids can only give rise to n^2 molecular species of dipeptides, an investigation of the dipeptides resulting from the partial hydrolysis of proteins should be possible technically, and, if we can be sure that the dipeptides studied do not arise by rearrangement, but are true fragments of the original peptide chain of the protein, the characterization of these compounds should yield information as to the structure of the original protein molecule.

A considerable body of data of this sort has been accumulated during the past 40 years. In the course of an examination of this literature undertaken by one of us (R. L. M. S.), special attention was paid to possible rearrangements occurring in the course of preparing and isolating the products of partial hydrolysis. No evidence could be found in the literature suggesting that partial

hydrolysis of proteins by strong mineral acid in the cold was likely to give rise to rearrangement or ring-formation in peptides such as might be expected to result from enzymic hydrolysis [Behrens & Bergmann, 1939], or from the action of hot dilute acids [Abderhalden & Funk, 1907; Abderhalden & Komm, 1924, 1, 2].

Cold concentrated acid has accordingly been used as the hydrolytic agent in the work described below. The possibility that the peptides studied result by rearrangement of previously existing structures has not been rigorously excluded: to devise a satisfactory control experiment for this purpose is not simple. The best control would seem to be to find a physically well-characterized and homogeneous protein containing only one residue of a particular amino-acid species per molecule. If this protein, as the physical and chemical evidence so far available strongly suggests, is a definite chemical compound in the same sense as are organic compounds of low molecular weight, then only two di-peptides of the single amino-acid residue in question should be isolable from a partial hydrolysate of the protein.

This would be the case if the protein had a definite chemical structure, irrespective of whether it had any particular structure, e.g. one conforming to the Bergmann-Niemann hypothesis. We have begun to undertake work in this direction, but as it is likely to be extremely protracted, it seems desirable to publish the evidence so far obtained on protein partial hydrolysis products, particularly since we have been unable to find any reference in the literature suggesting the rearrangement of peptide bonds under the hydrolytic conditions which we have employed.

The work here described falls into two sections:

I. A study of the course of hydrolysis of some proteins by cold concentrated hydrochloric acid, with special reference to the lability of the different types of bond present.

II. A study of the nature of the peptides containing basic amino-acids, which result from such hydrolysis.

Materials used in the investigation

In addition to wool, we have studied a preparation of edestin given to us by Prof. A. C. Chibnall, since for this protein his stoichiometric evidence seems better than any presented for other proteins in favour of the Bergmann-Niemann hypothesis. We have also studied gelatin (Coignet Gold Label) since current theories [cf. Bergmann, 1935; Bergmann & Stein, 1939; Astbury, 1940] assign to this substance a rather definite structure.

Analytical methods used

Estimation of cystine. Use has been made of the method of Sullivan [cf. Block, 1938, 2]. We have worked on a quarter the scale described in Block's book, and have employed direct visual comparison of the test with the standard solutions.

As cysteine residues involved in peptide linkage are not thought to give a colour in this reaction, it was important to show that the alkaline conditions employed did not give rise to significant hydrolysis of peptide bonds. For this purpose a mixed peptide solution of known amino-N content was kept with alkali of the same strength for the time usually allowed for the Sullivan reaction. A redetermination of amino-N showed no significant increase.

Estimation of arginine [cf. Kossel & Gross, 1924]. This has been done by the precipitation of arginine as its monoflavanate by addition of approx. 1.7 mol. of flavianic acid per mol. of arginine to the weakly acid unknown solution. The method was only applied to solutions known to be rich in arginine. The results are likely to be high owing to contamination with histidine, which, however, is a minor constituent of the proteins studied. The precipitation mixture was kept overnight at 0° and was then filtered through a weighed sintered glass funnel. The precipitate was washed with a little aqueous 0.5% flavianic acid and finally with alcohol, and then dried to constant weight *in vacuo* at room temperature over H₂SO₄. Control experiments with mixtures of arginine and alanine are described in § II below.

Estimation of higher monoamino-acids and proline. By acetylating a protein hydrolysate and extracting the acetylated mixture with chloroform under standard conditions it has proved possible [cf. Synge, 1939, 1; Martin & Synge, 1941, 1; Martin & Synge, accompanying paper] to obtain on the microscale a measure of the amount of these amino-acids present.

The work with the higher monoamino-acids described in the present paper was carried out before the liquid-liquid chromatogram [cf. Martin & Synge, accompanying paper] had been adapted for quantitative use. Accordingly use was made of total N (Kjeldahl) determinations on the chloroform extract prepared as described in the accompanying paper for obtaining an estimate of the aggregate of higher monoamino-acids present. The extract was also hydrolysed (by refluxing for 3 hr. in 2N HCl after it had been freed from alcohol by evaporation) and subjected to a Van Slyke amino-N determination (4 min. reaction time). The resulting figure gave an estimate of the higher monoamino-acids other than proline [cf. Van Slyke, 1911; Synge, 1939, 1]. The extract was subjected to qualitative investigation in the chromatogram, using B.P. chloroform saturated with water as the developing solvent.

Estimation of total hydroxyamino-acids and threonine (also free ammonia). These have been determined as ammonia and acetaldehyde, respectively, liberated by periodic acid, as described by Martin & Synge [1941, 2]. After the commencement of the present work the superior techniques of Nicolet & Shinn [1941] for serine and of Shinn & Nicolet [1941] for threonine were published. For the sake of uniformity we have not modified our procedure in the course of the work, the yields of ammonia ('total hydroxyamino-acids') and acetaldehyde ('threonine') being recorded 'corrected' throughout this paper, on the assumption that they are 75 and 70% respectively [cf. Martin & Synge, 1941, 2]. Free ammonia was determined in the course of the total hydroxyamino-acid determination.

Estimation of amino-N. The manometric Van Slyke procedure (4 min. reaction time) was employed. Where the sample contained free mineral acid this was neutralized before analysis with excess of sodium acetate.

Estimation of total N. Semi-micro-Kjeldahl determinations were employed, using the incineration mixture of Campbell & Hanna [1937].

Estimation of free amino-acids. The ninhydrin procedure of Christensen *et al.* [1941] was employed.

Estimation of tyrosine. The method of Arnow [cf. Block, 1938, 2] was employed.

I. THE COURSE OF HYDROLYSIS OF SOME PROTEINS BY COLD CONCENTRATED HYDROCHLORIC ACID, WITH SPECIAL REFERENCE TO THE LABILITY OF THE DIFFERENT TYPES OF BOND PRESENT

Partial hydrolysis has been effected in all the work described in this paper by solution of the proteins in an excess of 10*N* HCl at 37°. Wool dissolves almost entirely in the course of the first 24 hr. if gently shaken at intervals, leaving a particulate residue which is mostly ether-soluble. Edestin and gelatin give an immediate clear solution; in the case of gelatin, a small amount of amorphous material is deposited during the first 2 days' hydrolysis. Wool and edestin hydrolysates show a brilliant purple colour which after 3-4 days darkens to a clear brown. Hydrolysis was effectively stopped at any desired stage by repeated evaporation of the mixture with water *in vacuo* below 40°, followed by dilution to large volume and storage at 0°.

The above-mentioned determinations were made on both the partial and the complete hydrolysates. The latter were prepared by refluxing the material in 6*N* HCl for 24 hr.

The results obtained are tabulated below (Tables 1-6). Fig. 1 shows the kinetics of each determination for the partial hydrolysate of Merino 64's, the results of each determination on the partial hydrolysate being expressed as 'B values' (see tables). Fig. 2 (same units) compares the kinetics of evolution of amino-N in the same proteins.

Table 1. *Evolution of amino-groups during partial hydrolysis*

Time of hydrolysis hr.	Wool (Merino 64's)		Wool (flannel)		Edestin		Gelatin	
	A	B	A	B	A	B	A	B
19	—	—	—	—	—	—	22.9	39.5
24	26.2	34.1	—	—	—	—	—	—
43	—	—	—	—	—	—	32.1	55.1
48	34.3	44.6	30.4	39.5	23.0	35.9	—	—
96	38.8	50.5	—	—	—	—	—	—
114	—	—	40.0	52.0	—	—	—	—
120	—	—	—	—	32.4	50.6	—	—
139	—	—	—	—	—	—	44.9	77.0
166	—	—	46.5	60.3	37.0	57.8	—	—
192	48.3	62.9	—	—	—	—	—	—
Complete	76.8	(100)	77.0	(100)	64.0	(100)	58.3	(100)

A values: Amino-N, as percentage of total N of protein.

B values: 'A value' as percentage of its value in complete hydrolysate.

Table 2. *Evolution of free amino-acids during partial hydrolysis*

Time of hydrolysis hr.	Wool (Merino 64's)		Wool (flannel)		Edestin		Gelatin	
	A	B	A	B	A	B	A	B
48	11.9	17.4	—	—	—	—	—	—
96	20.7	30.2	—	—	—	—	—	—
139	—	—	—	—	—	—	33.7	41.4
166	—	—	27.2	39.7	23.4	36.9	—	—
192	25.7	37.4	—	—	—	—	—	—
Complete	68.6	(100)	68.6	(100)	63.4	(100)	81.5	(100)

A values: CO₂ liberated by ninhydrin as N in percentage of total N of protein.

B values: 'A value' as percentage of its value in complete hydrolysate.

Table 3. *Evolution of amino-groups of hydroxyamino-acids during hydrolysis*

Time of hydrolysis hr.	Wool (Merino 64's)		Wool (flannel)		Edestin		Gelatin	
	A	B	A	B	A	B	A	B
24	6.4	42.6	—	—	—	—	—	—
48	9.3	62.0	—	—	—	—	—	—
96	11.0	73.0	—	—	—	—	—	—
139	—	—	—	—	—	—	4.9	94.0
166	—	—	10.2	76.5	7.3	102.0	—	—
192	12.1	80.6	—	—	—	—	—	—
Complete	15.0	(100)	13.2	(100)	7.1	(100)	5.2	(100)

A values: Ammonia-N liberated by periodic acid (corrected) as percentage of total N of protein.
B values: 'A value' as percentage of its value in complete hydrolysate.

Table 4. *Evolution of amino-groups of threonine during partial hydrolysis*

Time of hydrolysis hr.	Wool (Merino 64's)	
	A	B
24	1.9	50
48	2.9	76
96	4.0	105
192	4.3	113
Complete	3.8	(100)

A values: CH₃CHO as N (corrected) in percentage of total N of protein.
B values: 'A value' as percentage of its value in complete hydrolysate.

Table 5. *Evolution of ammonia during partial hydrolysis*

Time of hydrolysis hr.	Wool (Merino 64's)		Wool (flannel)		Edestin		Gelatin	
	A	B	A	B	A	B	A	B
24	5.6	66.0	—	—	—	—	—	—
48	6.5	75.5	—	—	—	—	—	—
96	7.0	81.5	—	—	—	—	—	—
139	—	—	—	—	—	—	0.5	55.0
166	—	—	7.1	86.0	9.4	92.0	—	—
192	7.2	84.0	—	—	—	—	—	—
Complete	8.6	(100)	8.3	(100)	10.2	(100)	0.9	(100)

A values: Ammonia-N as percentage of total N of protein.
B values: 'A value' as percentage of its value in complete hydrolysate.

Table 6. *Evolution of free cysteine (or '½ cystine') during partial hydrolysis*

Time of hydrolysis hr.	Wool (Merino 64's)		Wool (flannel)		Edestin	
	A	B	A	B	A	B
24	0.7	8.4	—	—	—	—
48	1.2	15.8	—	—	—	—
96	2.4	29.9	—	—	—	—
166	—	—	4.6	—	0.42	117.0
192	2.9	35.8	—	—	—	—
Complete	7.9	(100)	—	—	0.36	(100)

A values: Cystine-N (Sullivan) as percentage of total N of protein.
B values: 'A value' as percentage of its value in complete hydrolysate.

Size of peptides formed during partial hydrolysis

From Fig. 1 it is seen that free amino-acids are set free from the outset in the partial hydrolysis of wool, and that after about a week, 37% of the amino-acid residues are free, the rest being presumably in the form of peptides.

An estimate of the size of the peptides present may be made by comparing the number of peptide bonds split (change of amino-N value) with the number of amino-acid residues liberated (change of ninhydrin-CO₂ value) on complete hydrolysis.

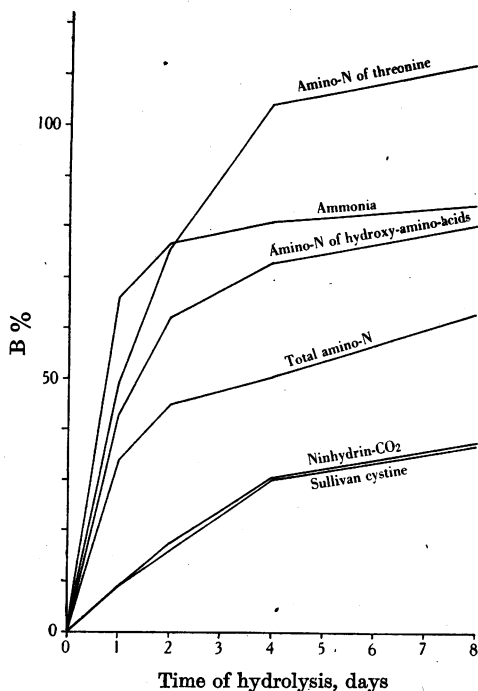


Fig. 1.

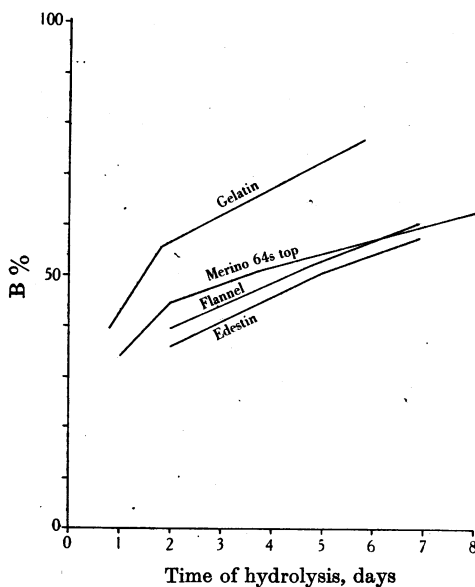


Fig. 2.

Fig. 1. Kinetics of liberation of various groups in partial acid hydrolysis of wool (Merino 64's top). The values are expressed as percentages of the value determined after complete hydrolysis.

Fig. 2. Kinetics of liberation of amino-N in partial acid hydrolysis of various proteins. The values are expressed as percentages of the value determined after complete hydrolysis.

Given for any partial hydrolysate,

$$r = \frac{\Delta (\text{'A value' for ninhydrin-CO}_2)}{\Delta (\text{'A value' for amino-N})} \text{ on completing the hydrolysis,}$$

we have for a dipeptide, which for the splitting of 1 bond liberates 2 amino-acid residues, $r = \frac{2}{1} = 2.0$, and for a tripeptide $r = \frac{3}{2} = 1.5$, and for a tetrapeptide $r = \frac{4}{3} = 1.33$.

Table 7 shows the derivation of r values for partial hydrolysates of the 3 proteins investigated.

Table 7

Protein	Time of hydrolysis hr.	Δ_A ninhydrin-CO ₂	Δ_A amino-N	<i>r</i>
Wool (Merino 64's)	192	42.9	28.5	1.50
Edestin	166	40.0	27.0	1.48
Gelatin	139	47.8	13.4	3.6

These figures suggest that the peptides present in the partial hydrolysates of wool and edestin are on the average tripeptides.

Since proline and hydroxyproline yield CO₂ with ninhydrin, but do not possess amino-N, the presence of peptide bonds involving the imino-groups of these residues will affect the *r* value obtained, as will deviations of amino-N and ninhydrin-CO₂ values from the 'theoretical' in the case of particular amino-acids and peptides. *r* values should therefore be interpreted with reserve.

The high *r* value, for gelatin may be explained by the high proline and hydroxyproline contents. In this substance (cf. Fig. 2) it seems that the bonds involving amino-groups are more readily split by cold concentrated acid than those of wool and edestin, or than the bonds in gelatin itself which involve imino groups.

'Amide-nitrogen'

In the partial hydrolysates the liberation of ammonia is found to proceed rapidly, and to reach a final value lower than that found for 'complete' hydrolysis. The lability of the ammonia agrees with the data of Vickery [1922; cf. Shore *et al.* 1936; Bailey, 1937].

It is suggested that the extra ammonia liberated on further refluxing for 24 hr. in 6*N* HCl is not true 'amide-nitrogen', but arises from the decomposition of hydroxyamino-acids such as serine. From Table 8 it will be seen that this 'extra' ammonia amounts consistently to about 10% of the hydroxyamino-acids present.

Table 8

Protein	Wool (Merino 64's)	Wool (flannel)	Edestin	Gelatin
'Extra' NH ₃ -N (see Table 5) as percentage of total N of protein	1.4	1.2	0.8	0.4
Hydroxyamino-acid-N (see Table 3) as percentage of total N of protein	15.0	13.2	7.1	5.2
NH ₃ as percentage of hydroxyamino-acids	10.0	9.0	11.0	8.0

The destruction of serine residues by hot mineral acid has been noted by Posternak [1927] and by Damodaran & Ramachandran [1941].

Free cysteine or '½-cystine'

In wool, this appears to be liberated during partial hydrolysis with acid at a rate not significantly different from the average rate of liberation of other kinds of amino-acids, as shown by the coincidence of the 'Sullivan' and 'ninhydrin-CO₂' graphs in Fig. 1. In edestin, under the conditions employed, all the cystine appears to be liberated as the free amino-acid.

Liberation of amino-groups of hydroxyamino-acids

The 'total hydroxyamino-acid' and 'threonine' determinations on the partial hydrolysates may be taken as a measure of hydroxyamino-acid residues having their α-NH₂ but not necessarily their —COOH groups free, since periodic acid

appears to react with compounds having —NH_2 and —OH groups bound to adjacent C atoms, but not to react when the —NH_2 group is acylated [cf. Nicolet & Shinn, 1939]. In the case of hydroxylysine the $\epsilon\text{—NH}_2$ group may be expected to react under the same conditions. These determinations may therefore be taken as a measure of the rupture of peptide bonds involving these —NH_2 groups. In the case of wool, all those bonds involving threonine residues (Table 4) appear to be readily broken, but not all those involving serine (Table 3). In gelatin and edestin all such bonds appear to be broken under the same conditions. This is also the case for all the derived fractions described in § II of this paper, consisting of basic peptides derived from wool, edestin and gelatin. It seems that, in these proteins at least, the bonds involving the amino-groups of serine and threonine are less than usually resistant to acid hydrolysis. In the field of simple compounds this is paralleled by Abderhalden & Bahn's [1932; 1935] study of the acid hydrolysis of acylated serine derivatives and of diketopiperazines containing serine.

The kinetics of hydrolysis of a peptide chain

It is useful in interpreting the above data to have some physical picture, however imperfect, of the mechanism of the hydrolysis of a protein, a chemical reaction in which one molecular species (the protein) is transformed into (say) 20 (the amino-acids) and in the course of which a number of intermediate substances transitorily appear and disappear. Both the relative amounts and the actual nature of these intermediates one may expect to be very considerably affected by quite small changes in the conditions of hydrolysis. The stability to hydrolysis of any particular peptide bond in a protein molecule will be determined by the nature of the amino-acids involved in it and by the nature of other neighbouring amino-acids, and these considerations may be expected to determine the point at which a long peptide chain containing a variety of amino-acid residues is first split. Once this split has occurred, a free positively charged —NH_3^+ or —NH_2^+ group is liberated, and this may be expected to repel H^+ ions and thus stabilize to acid hydrolysis the remaining peptide bond in which the amino-acid residue carrying the positively charged group is involved [cf. Moggridge & Neuberger, 1938; Syngé, 1939, 2]. Higher peptides will thus have a tendency to be shortened to dipeptides. An analogous mechanism should apply to the negatively charged —COO^- groups set free in alkaline hydrolysis, and it is reasonable to expect dipeptides to exhibit a greater stability than higher peptides both to acid and to alkaline hydrolytic agents. Those peptide bonds, then, in a protein that prove especially resistant to hydrolysis and are therefore recognizable as originally present in the protein through their isolation in simple peptides, may not initially have had any great intrinsic stability but may simply have been preserved through their proximity to a readily hydrolysable bond.

II. THE PEPTIDES CONTAINING BASIC AMINO-ACIDS RESULTING FROM HYDROLYSIS BY COLD CONCENTRATED HYDROCHLORIC ACID

In attempting to study the mixture of lower peptides resulting from the partial acid hydrolysis of proteins described in the previous section, it seemed to us desirable to begin with a broad separation based on the net electric charge on the molecule at pH 6, since such separation can be accomplished electrophoretically with some theoretical certainty as to the types of molecule that will be transported.

The naturally occurring amino-acid molecules can be classified as:

B (Basic). Arginine, histidine, lysine and hydroxylysine which carry a net positive charge at *pH* 6.

A (Acidic). Glutamic and aspartic acids, which carry a net negative charge at *pH* 6. (After partial hydrolysis as described above, all glutamine and asparagine residues will have been converted to glutamic and aspartic acid residues.)

N (Neutral). The other amino- and imino-acids, which carry no net charge at *pH* 6.

Considering now the six possible types of dipeptide resulting from the combination of these categories of amino-acids, on electrophoresis at *pH* 6:

BB and *BN* will migrate towards the cathode,

BA and *NN* will not migrate,

AA and *AN* will migrate towards the anode.

For a start, we have studied those components of a protein partial hydrolysate which migrate towards the cathode, since apparatus is already available [Albanese, 1940; Foster & Schmidt, 1923] for preparing such a fraction, and at least the free basic amino-acids are known not to be affected by secondary chemical action at the cathode.

Such a fraction is likely to prove of theoretical interest in revealing the mode of linkage of the basic amino-acids in the protein. Special importance has been attached to these in the 'kyrine' hypothesis of Siegfried [1903] and in the more recent 'anlage' concept of Block [1938, 1]. The results of work by Siegfried, Levene, Grassmann and others on the 'kyrine' question are of significance in connexion with the present work, as are also other investigations of basic peptides of protein origin, among which those of Kossel, Dirr, Felix and collaborators on partial hydrolysis products of clupein are outstanding.

All this earlier work on basic peptides made use of empirical precipitation methods; in the present work it will be seen that by use of a method based on theory we have succeeded in isolating from partial acid hydrolysates of proteins fractions consisting almost entirely of basic peptides and amino-acids, and containing a high proportion of the total basic amino-acid residues of the original proteins.

Experimental

The electro dialysis cell used in this work was similar to that described by Albanese [1940]. Some simplification was gained by making the cooling tubes reciprocate so as to stir all three compartments. It was found possible to use parchment paper for both the membranes as there was no very serious electromotive transfer of the compartment content. The drainage holes used by Albanese were omitted, the liquids being transferred by pipette. The various sections of the cell including the rubber gaskets were coated with paraffin wax to ensure that they were inert towards aqueous solutions. Either an ammeter or a milliammeter could be included in the circuit by means of a two-way switch. The source of current was a CuO rectifier giving 175 V. The circuit was stabilized by insertion of a 200 W. lamp. The cathode of the cell was made of platinum foil, and the anode was of gas carbon.

The behaviour of the cell with mixtures of arginine and alanine was studied, since arginine could readily be determined by flavianic acid as described above,

while alanine could be estimated by a Van Slyke amino-N determination on the filtrate and non-alcoholic washings from the arginine flavianate. A control experiment with arginine alone showed that these carried away a negligible amount of arginine amino-N.

In the first experiment, the middle compartment of the cell contained initially 100 ml. of an aqueous solution containing 1 g. of *dl*-alanine and 1 g. of *l*-arginine hydrochloride. The anode and cathode compartments each contained 100 ml. of distilled water. The current rose and fell again in the course of 3 hr. No attempt was made to control the pH of the middle compartment. At the end of the run, no arginine and no Cl⁻ could be detected in the middle compartment; 101 % of arginine flavianate was recovered from the catholyte, and 39 % of the alanine used was found also to be present.

A second experiment was conducted using 0.5 g. of *dl*-alanine and 0.5 g. of *l*-arginine hydrochloride in the middle compartment, the pH of this mixture being first adjusted to 7 with NaOH. During the run the pH of the middle compartment was not controlled and at the end of the run the middle compartment was again found to be free from arginine and Cl⁻. This time 19 % of the alanine was transferred to the catholyte.

In a third experiment the same initial solution in the middle compartment was maintained throughout the run at pH 6 by addition of NH₃ (bromothymol blue used as external indicator). After 80 min. the current fell, and all the arginine and only 7.3 % of the alanine were found in the catholyte after removal of NH₃ by evaporation *in vacuo*.

Next the rate of diffusion between the compartments of the cell was studied. 100 ml. of water containing 0.5 g. of alanine were placed in the middle compartment, and 100 ml. of water in each of the end compartments. After 1.5 hr. stirring without passage of current 5 % of the alanine had entered the cathode compartment, and after 3 hr. 9.7 % of the alanine had entered.

The transfer of alanine in the third experiment to the cathode compartment is thus explicable in terms of diffusion, to which must be added transport in the electrosmotic stream (about 3 ml. per hr.) from the middle to the cathode compartment. It is probably this unidirectional flow of liquid through the membrane which prevents the electro-dialytic return of alanine from the catholyte to the middle compartment, for in the weakly alkaline catholyte alanine residues carry a net negative charge.

The electro-dialytic fractionation of partial protein hydrolysates

By thrice-repeated electro-dialysis, it has proved possible, in the light of the control experiments just described, to separate the basic peptides and amino-acids of partial protein hydrolysates from nearly all neutral and acid constituents.

As an example of the procedure which we have evolved, we describe in detail the fractionation of a partial hydrolysate of gelatin: 50 ml. of the solution to be fractionated, containing 197 mg. of N, were put in the middle compartment of the cell. This solution still contained a little of the HCl used to hydrolyse the gelatin, although most of this was already removed (see § I). All the compartments were made up to 100 ml. with distilled water, cooling and stirring were commenced and the current was switched on. It remained at a few mA. until after the addition of a few drops of NH₃ (sp. gr. 0.880) to the catholyte and of 10*N* HCl to the anolyte, thus raising the conductivity of the end compartments. The current rose sharply to about 80 mA., then steadily until after 2 $\frac{3}{4}$ hr. it had reached 450 mA. It then fell for 1 $\frac{1}{4}$ hr. to 200 mA.; previous

experiments having shown that no further fall was to be expected, the contents of the cathode compartment were transferred to a Claisen flask and concentrated *in vacuo* below 40°. When the volume was about 50 ml. the solution (which reacted in the alkaline range of thymol blue) was neutralized to bromothymol blue with 2*N* H₂SO₄. Great care was taken to add no more H₂SO₄ than was necessary just to shift the indicator from its alkaline colour. The solution was then put back in the middle compartment of the cell, and all compartments brought to 100 ml. with water as before. The run was conducted as before except that every 2–3 min. the *pH* of the middle compartment was inspected. It fell steadily during the run, but was all the time on the range of bromothymol blue. The current rose to 225 mA., and after 31 min. had fallen to 31 mA. The contents of the cathode compartment were evaporated, neutralized and returned to the cell exactly as before, and the next electro dialysis was conducted in the same way, except that this time one drop of conc. NH₃ was required to keep the middle compartment solution on the range of the indicator. The current rose to 270 mA. and then fell to 20 mA. after 31 min. The amino-N content of the middle compartment was then determined, and it was found to be 0.97% of the amino-N originally taken for the fractionation. As this was satisfactorily low, the contents of the cathode compartment were evaporated as usual to remove NH₃, then neutralized to *pH* 5 with HCl and made up to 50 ml. in a volumetric flask.

Electrodialysis of a complete protein hydrolysate

This served as a control of contamination of the 'basic peptide fraction' by neutral substances. Wool (Merino 64's) was hydrolysed completely by refluxing for 24 hr. in 6*N* HCl. The excess HCl was removed as usual by evaporation, and a solution containing 128 mg. N was taken for electro dialytic fractionation exactly as described above; 2% of the amino-N originally taken was found in the middle compartment at the end of the third run. The catholyte contained 25.4% of the N of the wool, of which 34.3% was amino-N. Hydroxyamino-acids, tyrosine and cystine were determined on the catholyte, as shown in Table 9. From this it can be seen that the electro dialysis procedure employed gave a satisfactory elimination of neutral contaminants from the basic fraction.

Table 9

Determination	N as % of N of catholyte	Total amount of given amino-acid in catholyte as % of amount originally present in unfractionated hydrolysate
Total hydroxyamino-acids	0.1	0.6
Tyrosine	0.04	0.5
Cystine (Sullivan)	0.08	0.25

The 'basic peptide fractions' from partial protein hydrolysates

The neutralized 'basic peptide' fraction, made up to 50 ml. as described above, was then investigated as follows.

Some of the solution was kept for determinations to be carried out on the unhydrolysed peptides (amino-N, ninhydrin-CO₂, Sullivan, etc.) and the rest was refluxed for 24 hr. in 6*N* HCl to effect complete hydrolysis, before carrying out further determinations.

Table 10 shows comparative data on the various 'basic peptide fractions' before and after hydrolysis.

Table 10. *Determinations on 'basic peptide fractions' before and after hydrolysis*

Determination	Wool		Edestin		Gelatin	
	I	II	I	II	I	II
N of basic peptide fraction as percentage of N of original protein	27.1	25.0	33.7	29.1	25.2	22.7
Amino-N as percentage of total N of basic peptide fraction before hydrolysis	23.7	25.6	—	25.1	28.5	28.6
Same after hydrolysis	43.3	45.8	41.0	40.0	47.0	—
Ninhydrin-CO ₂ as N in percentage of N of basic peptide fraction before hydrolysis	9.9	8.9	—	9.9	11.0	—
Same after hydrolysis	54.0	47.2	—	39.4	46.3	—
Δ ninhydrin-CO ₂	44.1	38.3	—	29.5	35.3	—
Δ amino-N	19.6	20.2	—	14.9	18.5	—
R	2.25	1.90	—	1.98	1.91	—
Cystine-N (Sullivan) as percentage of N of basic peptide fraction before hydrolysis	0.83	—	—	0.06	—	—
Same after hydrolysis	—	2.9	—	0.02	—	—
Total hydroxyamino-acid-N as percentage of N of basic peptide fraction before hydrolysis	—	7.7	—	3.4	3.3	—
Same after hydrolysis	5.5	5.1	2.7	2.0	3.1	—

The fractions described in Tables 10 and 11 were prepared by electro dialysis from partial hydrolysates of proteins in 20 parts of 10*N* HCl at 37° for the undermentioned lengths of time:

Wool I. Commercial Merino 64's top: 192 hr.

Wool II. A commercial loosely woven pure wool flannel: 166 hr.

Edestin I. A preparation of edestin (from Prof. A. C. Chibnall): 168 hr.

Edestin II. The same material hydrolysed on a different occasion for 167 hr.

Gelatin I. Coignet 'Gold Label' gelatin hydrolysed for 139 hr.

Gelatin II. The same hydrolysate subjected to a separate fractionation.

Table 11 shows the individual amino-acids determined in the hydrolysates of the basic peptide fractions. Corresponding determinations on unfractionated complete hydrolysates of the same proteins are shown in Table 12.

Table 11. *Amino-acid determination on hydrolysates of 'basic peptide fractions'*

Determination (N as % N of fraction)	Wool		Edestin		Gelatin I
	I	II	I	II	
Ammonia-N	1.2	2.3	0.7	0.1	0.3
Arginine-N	59.7	—	70.4	—	53.2
Cystine-N (Sullivan)	—	2.9	—	0.0	—
Tyrosine-N	1.7	1.6	1.0	1.1	—
Total hydroxyamino-acid-N	5.5	5.1	2.7	2.0	3.1
'Higher monoamino-acid-N'	10.6	10.6	11.1	9.0	5.8
'Proline-N'	5.1	3.8	0.7	3.1	4.2

Table 12. *Amino-acid determinations on complete protein hydrolysates*

Determination (N as % N of protein)	Wool	Edestin	Gelatin
	I		
Cystine-N (Sullivan)	7.9	0.36	—
Tyrosine-N	2.1	1.3	—
Total hydroxyamino-acid-N	12.2	7.0	5.2
'Higher monoamino-acid-N'	15.3	12.6	9.2
'Proline-N'	7.1	4.1	11.4

The mixture of acetamino-acids obtained in each case after acetylation and chloroform extraction as described above was investigated qualitatively by the chromatographic method already mentioned, using a column about 1 cm. in diameter. An amount of extract containing 1 mg. of N was put on the column, and the mixture of acetamino-acids obtained from the hydrolysate of the complete protein was run at the same time as the mixture obtained from the hydrolysate of the 'basic peptide fraction' of the same protein. In no case could any differences in the sets of bands present be observed. In the case of wool, the relative intensities of the bands were similar in the two columns. With edestin, the leucine-isoleucine complex band and the valine band were more intense for the 'basic peptide fraction' than for the whole protein, while the proline band was slightly diminished. With gelatin, the proline and alanine bands of the material from the 'basic peptide fraction' appeared enriched, and the leucine and phenylalanine bands diminished.

DISCUSSION

Yield of basic amino-acids in 'basic peptide fractions'

Table 13 illustrates the proportion of the total arginine residues of the proteins recovered in the 'basic peptide fractions'.

Table 13. *Recovery of arginine in 'basic peptide fractions'*

	Wool	Edestin	Gelatin
Arginine-N as percentage of N of 'basic peptide fraction' (cf. Table 12) = A	59.7	70.4	53.2
N of 'basic peptide fraction' as percentage of N of protein (cf. Table 10) = B	27.1	33.7	25.2
N of arginine in 'basic peptide fraction' as percentage of protein: $(A \times B)/100 = C$	16.2	23.7	13.4
N of total arginine of protein as percentage of N of protein (cf. Table 15) = D	20.2	28.1	14.6
Percentage of arginine residues of protein present in 'basic peptide fraction': $(C \times 100)/D$	80.0	84.0	92.0

It is seen that this recovery ranges from 80 to 92%. The unrecovered arginine sets an upper limit to the proportion of arginine bound in neutral and/or acidic peptides such as arginylglutamic acid [cf. Levene & Birchard, 1912]. Although the recovery of lysine and histidine has not been determined, the total N of the basic peptide fractions is compatible with a similar recovery of these residues, and it is apparent that in the present work we are dealing with the major proportion of the basic amino-acid residues of the proteins studied.

Character of the peptides present in the 'basic peptide fractions'

The use of the ratio r as a measure of the peptide size in a mixture of amino-acids and peptides has been explained in § I of this paper. Table 10 shows the value of this ratio for the 'basic peptide fractions' obtained from the proteins studied, and it suggests that the peptides present are almost entirely dipeptides. As already mentioned, this ratio is also affected by the number of proline and hydroxyproline residues involved in peptide linkage through their imino groups, since they evolve no amino-N on liberation, while becoming susceptible to decarboxylation by ninhydrin. Thus the presence of lysylprolylglycine ($r=3$) in the 'basic peptide fraction' from gelatin (cf. Grassmann & Riederle's [1936] isolation of this from a partial acid hydrolysate of gelatin)

would permit the presence of a proportion of tripeptides containing amino- rather than imino-groups in peptide linkage, while r remains approximately 2. Thus for gelatin, on account of its high content of proline and hydroxyproline, it is less certain than for wool and edestin that a majority of the peptides of the basic fraction are dipeptides.

Relative stability to acid hydrolysis of bonds in proteins involving basic and other amino-acid residues

Siegfried [1903], in putting forward the 'kyrine' hypothesis of protein structure, suggested that the basic amino-acids were bound in a manner specially resistant to acid hydrolysis.

From the present work it is possible to estimate the ratio $\frac{\text{free basic residues}}{\text{total basic residues}}$ in the various partial protein hydrolysates, and to compare it with the corresponding figure for neutral and acidic residues taken together. This is done as follows.

The 'ninhydrin-CO₂' figure for the unhydrolysed 'basic peptide fraction' represents (1 amino-acid residue being equivalent to 1 mol. of ninhydrin-CO₂, expressed as 1 atom N in % of the total N of the protein) the free basic amino-acids in the partial hydrolysate, when allowance has been made (by subtracting the Sullivan value for the unhydrolysed 'basic peptide fraction') for 'free $\frac{1}{2}$ -cystine residues' (see following paragraph). This value is shown as A in Table 14. The total basic residues of the protein are known from already existing analyses, and are expressed as N in % of total N in the protein in Table 15. To render them comparable with A , the arginine figures are divided by 4, lysine by 2 and histidine by 3, and the total of basic residues thus arrived at is shown in Table 14 as B .

Table 14. *Relative stability of basic and other residues to acid hydrolysis*

Determination	Wool	Edestin	Gelatin
'Ninhydrin-CO ₂ ' of unhydrolysed 'basic peptide fraction' as N in percentage of total N of protein	2.48	2.89	2.77
'Sullivan N' of unhydrolysed 'basic peptide fraction' in percentage of total N of protein	0.22	0.00	—
A	2.26	2.89	2.77
B	6.9	9.3	7.2
'Ninhydrin-CO ₂ ' value for unfractionated partial hydrolysate (cf. Table 2) as N in percentage of total N of protein	26.4	23.4	33.7
C	24.1	20.5	30.9
'Ninhydrin-CO ₂ ' value for unfractionated complete hydrolysate (cf. Table 2) as N in percentage of total N of protein	68.6	63.4	81.5
D	61.7	54.1	74.3
A/B	0.33	0.31	0.38
C/D	0.39	0.38	0.41

The ninhydrin-CO₂ value (same units as A) for the unfractionated partial hydrolysate of the protein (cf. Table 2) *minus* A , gives the free neutral and acidic amino-acids in the partial hydrolysate (C), while the ninhydrin-CO₂ value for the complete hydrolysate of the protein (cf. Table 2) *minus* B , gives the total neutral and acidic amino-acid residues of the protein (D). The desired ratios are then expressed as A/B and C/D .

Neutral amino-acids of the basic peptides

Subject to the reservation mentioned in the introduction to this paper, that the peptides studied may conceivably be artefacts arising by rearrangement, the results of the present work are of some interest in considering the arrangement of amino-acid residues in proteins. Nearly all the neutral amino-acids found in a hydrolysate of the 'basic peptide fraction' derived from the proteins studied may be inferred to have been directly linked to a basic amino-acid, since the basic peptides present have been shown to be almost all dipeptides. No neutral amino-acid residue present in the original protein has been found to be absent, on searching for it, from the 'basic peptide fraction' except cystine, in the case of edestin, which has been shown to be set free completely during the hydrolysis. The control experiments already described make it extremely unlikely that neutral residues can enter this fraction in the form of free amino-acids. Nor are the proportions in which the various neutral amino-acid residues occur in the 'basic peptide fraction' very different from those in which they occur in the original protein. This is illustrated in Tables 15 and 16. In Table 16

Table 15. *Estimate of 'neutral amino-acid' content of proteins and derived basic peptide fractions*

N as % of total N of protein or fraction	Wool		Edestin		Gelatin	
	Whole protein	'Basic peptide fraction'	Whole protein	'Basic peptide fraction'	Whole protein	'Basic peptide fraction'
'True amide-N'	7.2	0.0	9.4	0.0	0.5	0.0
Arginine-N	20.2*	59.7	28.1†	70.4	14.6‡	53.2
Lysine-N	3.1§	9.2	2.4†	6.1	6.3‡	22.9
Histidine-N	1.2§	3.5	3.4†	8.5	1.3‡	4.7
Glutamic acid	6.0¶	—	10.4†	—	3.1‡	—
Aspartic acid	3.6¶	—	6.8†	—	2.0‡	—
Total of above	41.3	72.4	60.5	85.0	27.8	80.8
N of neutral amino-acids (by difference)	58.7	27.6	39.5	15.0	72.2	19.2

* Vickery [1940].

† Private communication from Prof.

A. C. Chibnall.

‡ Dakin [1920].

§ Block [1939].

|| Estimated.

¶ Private communication from Dr K. Bailey (minimum figure).

Table 16. *N of individual neutral amino-acids as % of total 'neutral amino-acid-N' of protein or derived fraction (cf. Table 15)*

Determination*	Wool		Edestin		Gelatin	
	Whole protein	'Basic peptide fraction'	Whole protein	'Basic peptide fraction'	Whole protein	'Basic peptide fraction'
Cystine-N	13.4	10.5	0.9	0.2	—	—
Tyrosine-N	3.6	6.0	3.3	7.0	—	—
Hydroxyamino-acid-N	20.8	19.2	17.7	15.7	†	†
'Higher monoamino-acid-N'	26.0	38.4	31.9	67.0	12.8	30.2
'Proline-N'	12.0	15.9	10.4	12.7	15.8	21.8

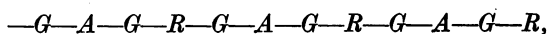
* The original determinations are shown in Tables 11 and 12. The figures taken from Table 11 are, where possible, the mean of two determinations.

† Not calculated, because of probable presence of hydroxylysine.

the N of the individual neutral amino-acid or group of neutral amino-acids is expressed as percentages of the total N of the neutral amino-acid residues present in the 'basic peptide fraction' and in the parent protein respectively.

This 'total N of neutral amino-acids' is arrived at (see Table 15) by subtracting from the total N of the protein, the N of arginine, histidine and lysine, the 'true amide-N' (cf. § I) and the N of glutamic and aspartic acids. For the 'basic peptide fractions' glutamic acid, aspartic acid and 'amide-N' are deemed absent. The lysine and histidine contents of the 'basic peptide fractions' have not been determined, but they are estimated as being present in the same ratio to the arginine as in the parent protein. This assumption introduces a source of second-order error only, and does not affect significantly the general conclusion.

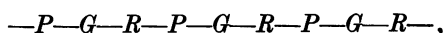
It is not easy to reconcile the picture of protein structure presented by the present work with the Bergmann-Niemann hypothesis, according to which in general the residues of any particular amino-acid will be directly linked to only a small number of different amino-acid species. Thus for silk [cf. Bergmann & Niemann, 1938], given the structure



where

G = glycine residue	}	in peptide chain,
A = alanine residue		
R = other residues		

a pentapeptide is the smallest molecule that can incorporate a residue of a basic amino-acid together with a residue of any amino-acid other than glycine or alanine. Similarly for gelatin [cf. Astbury, 1940]:



where

G = glycine residue	}	in peptide chain,
P = proline or hydroxyproline residue		
R = other residues		

a tetrapeptide is the smallest molecule that could contain a basic residue together with one of any species other than glycine, proline or hydroxyproline, unless the residue replacing one out of every 18 proline-hydroxyproline residues is involved.

While the present work does not by any means form a rigid refutation of the Bergmann-Niemann hypothesis, it seems to suggest for proteins a considerably more complicated structure. The hypothesis of random structure, with which it is scarcely incompatible, may be rejected on physical and biological grounds.

At present, the most hopeful line of extension of the work seems to be towards the isolation from a physically well-characterized and homogeneous protein of peptides containing a high proportion of the total residues present of one particular amino-acid. The smaller the number of such residues in the molecule the simpler should be the peptide mixture to be studied, and in the limiting case of one residue of the amino-acid in question per molecule, only two dipeptides should be obtainable containing this residue on any hypothesis of regular structure based on the peptide theory. Such a study, as mentioned above, would also be a valuable control on artefacts arising by rearrangement. Work in this direction is at present in progress.

SUMMARY

1. The characterization of the lower peptides resulting from the partial hydrolysis of proteins, and its significance for the elucidation of protein structure are discussed.

2. Concentrated hydrochloric acid, acting on wool, edestin and gelatin at 37° for a week, liberates about a third of the residues as free amino-acids. The rest are present in the form of lower peptides.

3. The general kinetics of this acid hydrolysis are described, with special reference to the splitting of bonds involving cystine, serine, threonine, proline, hydroxyproline and ammonia residues.

4. By electro dialysis in a diaphragm cell it has been found possible to isolate a large proportion of the basic amino-acid residues in the form of dipeptides.

5. A study has been made of the amino-acids directly linked to the basic residues. Its bearing on past and current theories of protein structure is discussed.

We are grateful to Prof. A. C. Chibnall for his valued criticisms and to Miss M. I. Whitehead for help in preparing this paper for the press.

We wish to thank the Director and Council of the Wool Industries Research Association and the International Wool Secretariat for permission to publish this work.

REFERENCES

- Aberhalden & Bahn (1932). *Hoppe-Seyl. Z.* **210**, 246.
 ——— (1935). *Hoppe-Seyl. Z.* **234**, 181.
 ——— & Funk (1907). *Hoppe-Seyl. Z.* **53**, 19.
 ——— & Komm (1924, 1). *Hoppe-Seyl. Z.* **134**, 121.
 ——— (1924, 2). *Hoppe-Seyl. Z.* **139**, 147.
 Albanese (1940). *J. biol. Chem.* **134**, 467.
 Astbury (1940). *J. int. Soc. Leath. Chem.* **24**, 69.
 Bailey (1937). *Biochem. J.* **31**, 1406.
 Behrens & Bergmann (1939). *J. biol. Chem.* **129**, 587.
 Bergmann (1935). *J. biol. Chem.* **110**, 471.
 ——— & Niemann (1936). *J. biol. Chem.* **115**, 77.
 ——— (1937). *J. biol. Chem.* **118**, 301.
 ——— (1938). *J. biol. Chem.* **122**, 577.
 ——— & Stein (1939). *J. biol. Chem.* **128**, 217.
 Block (1938, 1). *The Chemistry of the Amino-acids and Proteins*, p. 278 (ed. Schmidt).
 Springfield, Illinois.
 ——— (1938, 2). *The Determination of the Amino-acids*. Minneapolis.
 ——— (1939). *J. biol. Chem.* **128**, 181.
 Campbell & Hanna (1937). *J. biol. Chem.* **119**, 1.
 Christensen, West & Dimick (1941). *J. biol. Chem.* **137**, 735.
 Dakin (1920). *J. biol. Chem.* **44**, 499.
 Damodaran & Ramachandran (1941). *Biochem. J.* **35**, 122.
 Foster & Schmidt (1923). *J. biol. Chem.* **56**, 545.
 Grassmann & Riederle (1936). *Biochem. Z.* **284**, 177.
 Greenstein (1939). *J. biol. Chem.* **128**, 233.
 Kossel & Gross (1924). *Hoppe-Seyl. Z.* **135**, 167.
 Levene & Birchard (1912). *J. biol. Chem.* **13**, 277.

- Linderström-Lang & Jacobsen (1941). *J. biol. Chem.* **137**, 443.
Martin & Syngé (1941, 1). *Biochem. J.* **35**, 91.
——— (1941, 2). *Biochem. J.* **35**, 294.
Moggridge & Neuberger (1938). *J. chem. Soc.* p. 745.
Nicolet & Shinn (1939). *J. Amer. chem. Soc.* **61**, 1615.
——— (1941). *J. biol. Chem.* **139**, 687.
Posternak (1927). *C.R. Acad. Sci., Paris*, **184**, 306.
Shinn & Nicolet (1941). *J. biol. Chem.* **138**, 91.
Shore, Wilson & Stueck (1936). *J. biol. Chem.* **112**, 407.
Siegfried (1903). *Ber. sachs. Ges. (Akad.) Wiss.* **55**, 63.
Syngé (1939, 1). *Biochem. J.* **33**, 1913, 1918.
——— (1939, 2). *Biochem. J.* **33**, 1924.
Van Slyke (1911). *J. biol. Chem.* **9**, 205.
Vickery (1922). *J. biol. Chem.* **53**, 495.
——— (1940). *J. biol. Chem.* **132**, 325.