THE AMINO ACID COMPOSITION OF CRYSTALLINE PEPSIN*

By OLGA O. BLUMENFELD‡ and GERTRUDE E. PERLMANN

(From The Rockefeller Institute)

(Received for publication, July 10, 1958)

ABSTRACT

The amino acid composition of twice recrystallized pepsin (Worthington Biochemical Corporation) has been determined chromatographically on columns of Amberlite IR 120 resin. The results of the analyses obtained on four different preparations indicate a close agreement in their amino acid composition. Pepsin is unique in that it has a great predominance of acidic amino acids over basic ones. Moreover, all the preparations contain a small and constant amount of hydroxyproline, corresponding to about 0.1 residue per molecule.

Previous work from this laboratory has revealed that partial autodigestion of pepsin in urea gives rise to "modified" proteins which differ from the starting material by an enhanced specific activity (1). Inasmuch as a chemical characterization of these newly formed enzymically active components requires a detailed knowledge of the amino acid composition of the parent protein several crystalline pepsin preparations were analyzed for their amino acid content. As will be shown in this report, our results agree well, in part, with those of Brand obtained largely with microbiological methods (2). An entirely unexpected finding, however, is the occurrence in the pepsin preparations of a small quantity of hydroxyproline, an amino acid hitherto encountered only in collagen.

Experimental

Materials.—The twice recrystallized pepsin preparations used in this work were lots 611, 617, 622, and 623 of the Worthington Biochemical Corporation. These preparations were tested as routine for non-protein material which did not exceed 2 to 3 per cent. On free electrophoresis in monovalent buffers of pH 1.0 to 6.0, they migrate as a single component and are homogeneous in the ultracentrifuge. The sedimentation constant, $S_{20} = 2.9_6 \times 10^{-13}$ to 3.0×10^{-13}, corresponds to a molecular weight of 35,000 (3).

* This work was supported in part by Grant B-930 of the United States Public Health Service, National Institutes of Health.

‡ Holder of a Postdoctoral Fellowship in Medical Sciences, 1957–59, administered by the National Research Council and the National Science Foundation.

Methods.—For amino acid analyses 15 to 20 mg. samples of pepsin were hydrolyzed *in vacuo* with 3 ml. constant boiling hydrochloric acid at 110 ± 1°C. for 22 and 70 hours. The hydrolysates were taken to dryness in a rotating evaporator at 50°C. and the residue dissolved in 3 ml. distilled water. The washing procedure was repeated three times. 0.2 N citrate buffer of pH 2.2 was then added and the volume adjusted to 5.0 ml. The concentration of the hydrolysates has been determined from nitrogen analyses by the Pregl microKjeldahl method using a nitrogen factor for conversion to dry weight based on separate nitrogen, moisture, and ash determinations of the individual pepsin preparations.[1]

In the amino acid analyses the procedure of Moore, Spackman, and Stein was followed using a 153 cm. IR 120 column at 50°C. for the acidic and neutral acids and a 25 cm. column for the basic ones (4). Since in chromatography at 50°C. hydroxyproline does not separate from aspartic acid and its color value in the ninhydrin assay at 440 mμ is low, this amino acid was determined in separate experiments in which the column temperature was maintained at 30°C. (5). Here an aliquot of each fraction was tested with the modified ninhydrin reagent (6) and the remaining portion analyzed for hydroxyproline with the aid of the procedure of Martin and Axelrod (7). BRIJ, thiodiglycol, and phenol, usually added to the eluting buffer, interfere with the color reaction and were, therefore, eliminated.

Tryptophan analyses were carried out on the intact unhydrolyzed pepsin according to the method of Goodwin and Morton (8). As suggested, correction for extraneous absorption was applied and the tyrosine content derived from the chromatographic analyses substituted into the equation. The isobestic point was at 277 to 281 mμ (9).

RESULTS

Some of the striking features of the amino acid composition of pepsin are illustrated qualitatively with the aid of the elution patterns of Figs. 1 and 2. From these patterns it is apparent that pepsin has a high content of acidic, neutral, and hydroxy amino acids whereas the content of the basic ones is unusually low.

In Table I are presented the analytical data obtained from a 22 and 70 hour hydrolysate for the four preparations. It is of interest to note that the compositions of lots 611, 617, and 623 are identical within the experimental error of the method. Preparation 622, on the other hand, has less isoleucine. As is the case with most proteins, serine, threonine, and tyrosine are partly decomposed during acid hydrolysis. Thus, in the final calculations, a correction has been applied which was obtained by extrapolation of the 22 and 70 hour values to zero time.

A point of interest in these analyses, however, is that each pepsin preparation thus far examined contains a small amount of hydroxyproline. Although it has been shown by Spackman, Stein, and Moore (5) that on chromatography on

[1] We are indebted to Mr. T. Bella for these determinations.

IR 120 at 30°C. this amino acid emerges as a distinct peak ahead of the aspartic acid, the ninhydrin reaction was not sufficiently sensitive to detect any hydroxyproline in the corresponding region of our chromatograms. If, on the other hand, as also illustrated with the aid of Fig. 3, each fraction is assayed with the

TABLE I

Amino Acid Composition of Pepsin as Determined by Chromatography of Acid Hydrolysates on Columns of Amberlite IR 120

Amino acid (1)	Amino acid per 100 gm. protein						
	Lot 611		Lot 617		Lot 622	Lot 623	
	22 hrs. (2)	70 hrs. (3)	22 hrs. (4)	70 hrs. (5)	22 hrs. (6)	22 hrs. (7)	70 hrs. (8)
	gm.	*gm.*	*gm.*	*gm.*	*gm.*	*gm.*	*gm.*
Lysine	0.48		0.42	0.46	0.38	0.43	
Histidine	0.40		0.50	0.54	0.45	0.46	
Arginine	0.94		0.97	0.94	0.99	1.02	
Aspartic acid	16.70	17.32	16.40	17.39	16.38	16.31	16.64
Threonine	9.34	8.14	9.16	8.68	9.13	8.89	8.33
Serine	12.99	10.42	12.37	11.19	12.78	12.49	10.45
Glutamic acid	11.60	11.42	11.27	11.38	11.31	11.32	11.28
Proline	4.80	5.11	4.97	4.96	4.68	4.47	5.10
Glycine	8.19	8.05	8.04	8.25	8.13	7.97	8.05
Alanine	4.70	4.51	4.41	4.95	4.51	4.42	4.58
Half-cystine	1.92	1.29	1.50	0.35	1.40	1.64	2.11
Valine	7.38	7.13	6.85	7.70	7.05	7.11	7.34
Methionine	2.33	1.82	2.27	1.20	1.89	2.28	1.86
Isoleucine	9.79	10.29	10.07	10.21	8.65	10.74	9.75
Leucine	9.83	10.57	10.49	10.29	9.78	10.29	10.38
Tyrosine	8.80	8.02	9.00	7.12	8.94	8.19	8.33
Phenylalanine	6.99	6.75	6.74	6.55	6.78	6.61	6.56
Tryptophan	3.51		3.42		3.62	3.43	
Hydroxyproline	0.05		0.04		0.02	0.05	
Ammonia	1.91		2.02		2.05	1.98	
Nitrogen content, per cent	14.88		14.85		14.92	14.64	
Molar extinction coefficient, $\epsilon \times 10^{-3}$			52.4_7		53.1_4	51.3_9	

color test specific and more sensitive for hydroxyproline a small peak is present in the position in which this amino acid would be expected to occur. Separate experiments were performed in which a small amount of hydroxyproline was added to pepsin prior to the hydrolysis of the protein with hydrochloric acid. Here again, only one peak giving a positive isatin test was present in the chromatogram. Moreover, it should be noted that the added material was always

recovered quantitatively. Thus, in contrast to the other hydroxy amino acids, *e.g.* serine and threonine, hydroxyproline is stable in 6.0 N hydrochloric acid at 110°C.

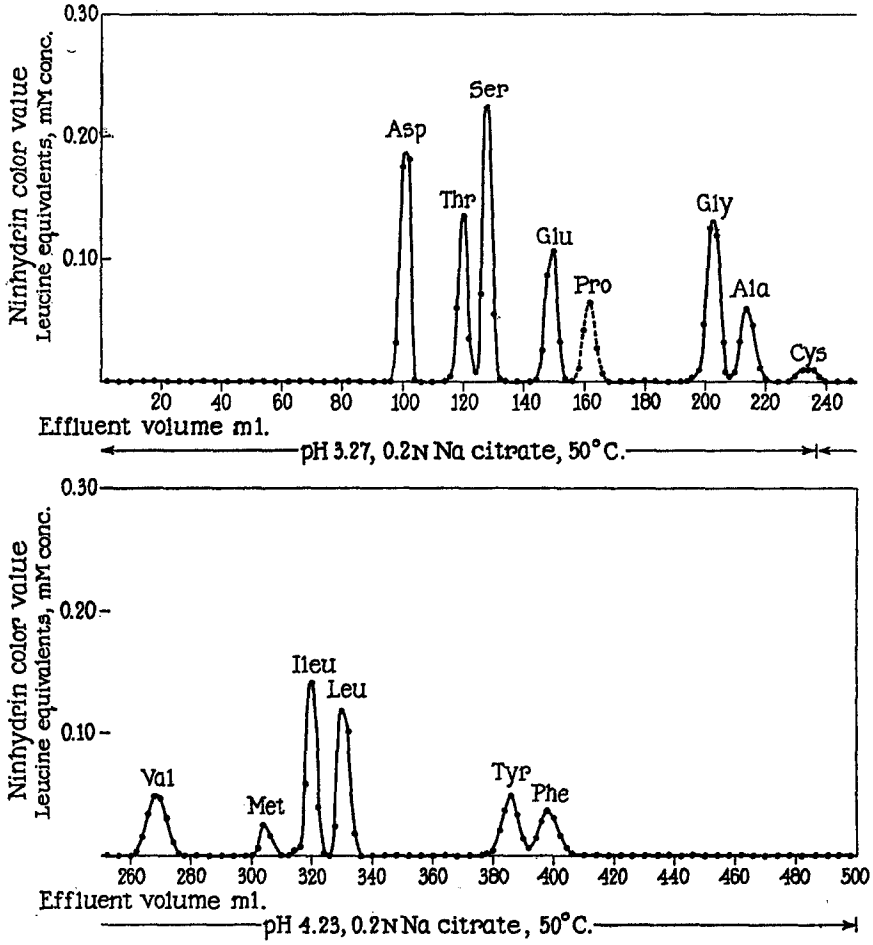

FIG. 1. Amino acids of a 22 hour hydrolysate of pepsin, lot 617, separated on a 0.9 cm. × 153 cm. column of IR 120. The dashed curve for proline gives optical density at 440 mμ. The amount of hydrolysate used in this analysis corresponded to 0.67 mg. pepsin.*

* The symbols used to denote the different amino acids follow those suggested by Sanger (*Advances Protein Chem.*, 1952, **7**, 1).

A summary of the analytical results is presented in Table II. Here columns 2, 3, and 4 list the average value for each amino acid derived from the data given in Table I. If these results are expressed as number of residues per mol-

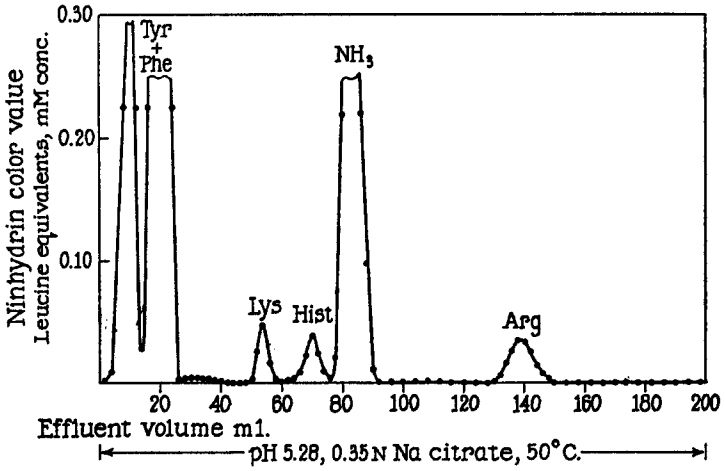

FIG. 2. Basic amino acids and ammonia of a 22 hour hydrolysate of pepsin, lot 617, separated on a 0.9 cm. × 25 cm. column of IR 120. The amount of hydrolysate used in this analysis corresponded to 6.12 mg. pepsin.*

* The symbols used to denote the different amino acids follow those suggested by Sanger (*Advances Protein Chem.*, 1952, **7**, 1).

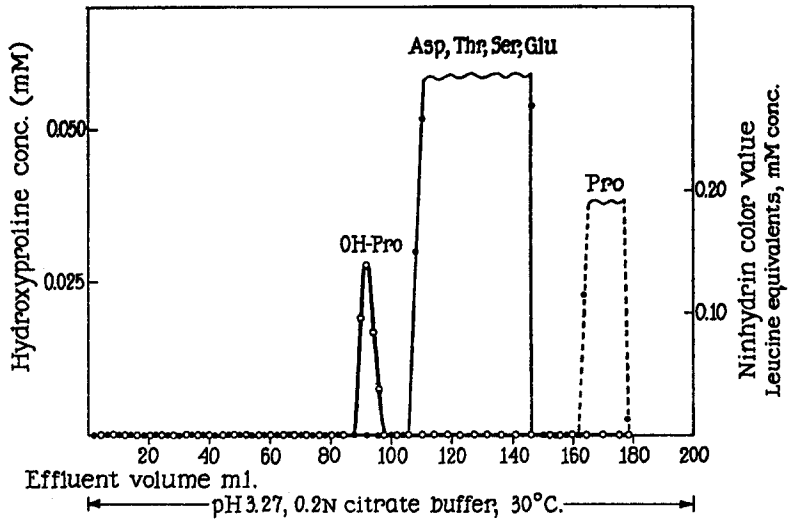

FIG. 3. Separation of hydroxyproline from the acidic and neutral amino acids of a 22 hour hydrolysate of pepsin, lot 611, on a 0.9 cm. × 153 cm. column of IR 120 at 30°C. The dashed curve for proline gives optical density at 440 mμ, (●) ninhydrin determination, (○) hydroxyproline assay. The amount of hydrolysate placed on the column corresponded to 43.1 mg. pepsin.*

* The symbols used to denote the different amino acids follow those suggested by Sanger (*Advances Protein Chem.*, 1952, **7**, 1).

TABLE II

*Amino Acid Composition of Crystalline Pepsin**

Amino acid	Amino acid per 100 gm. protein	Amino acid residues per 100 gm. protein	N as per cent of total N	Calculated No. of residues for mocleular weight 35,000	No. of residues per molecule	
					This work	Brand (2)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
	gm.	*gm.*				
Lysine	0.43	0.38	0.56	1.0_3	1	2
Histidine	0.47	0.42	0.86	1.0_6	1	2
Arginine	0.97	0.89	2.11	1.9_5	2	2
Aspartic acid	16.63	14.38	11.80	43.7_3	44	41
Threonine‡	9.50	8.06	7.54	27.9_2	28	28
Serine‡	13.20	10.94	11.86	43.9_6	44	40
Glutamic acid	11.34	9.95	7.28	26.9_8	27	28
Proline	4.90	4.13	4.02	14.9_0	15	15
Glycine	8.10	6.16	10.20	37.7_8	38	29
Alanine	4.51	3.60	4.78	17.7_1	18	—
Half-cystine	1.45	1.23	1.13	4.1_9	*Ca.* 4	4
Valine	7.09	6.00	5.72	21.1_8	21	21
Methionine§	2.07	1.82	1.31	4.8_6	*Ca.* 5	4
Isoleucine	10.03	8.65	7.22	26.7_6	27	28
Leucine	10.43	8.99	7.51	27.8_3	28	27
Tyrosine‡	9.40	8.47	4.90	18.1_6	18	16
Phenylalanine	6.73	5.99	3.85	14.2_6	14	13
Tryptophan	3.50	3.19	3.24	6.0	6	4
Hydroxyproline	0.05	0.04	0.04	0.1	0.1	—
Amide NH_3‖			9.72	36.0_9	(36)	32
Total		103.2	105.6		341	

* Average of ten analyses; only values which are within 5 per cent of the mean are included.

‡ The values for threonine, serine, and tyrosine were obtained by extrapolation of the 22 and 70 hour hydrolysates to zero time.

§ Methionine corrected for 5 per cent loss during chromatography.

‖ NH_3 content of the hydrolysate corrected for decomposition of serine, threonine, and tyrosine, but considered as approximate.

ecule of 35,000 (column 5), they lie close to an integral value in each instance and agree well, in most part, with those of Brand (column 7) (2). Moreover, if the molecular weight of pepsin is computed from these data, the value of 36,212 is in close agreement with that of 35,000 based on physicochemical measurements (10, 11).

The values which are subject to greatest error are those for methionine and cystine.[2] As can be seen from Fig. 1, the peaks are the smallest ones on the

[2] As shown by Kern, pepsin has no free -SH groups (20).

effluent curve. Both amino acids are subject to losses during hydrolysis. Accurate determination of these two constituents would require analysis of performic acid–oxidized pepsin for cysteic acid and methionine sulfone.

Even though the basic amino acids are present in relatively small amounts, it was possible to determine them accurately. Three to 6 mg. samples were analyzed on a 25 cm. column of IR 120 rather than on a shorter column as recommended by Moore, Spackman, and Stein (4). Similarly, amounts of 35 to 40 mg. of pepsin had to be used in each determination of hydroxyproline whereas only 0.5 to 0.7 mg. was necessary for the analyses of the acidic and neutral amino acids.

DISCUSSION

The homogeneity of pepsin has been the subject of extensive investigations (12–14). Although the physicochemical properties and the analytical results point to a uniformity of our preparations, this does not preclude the fact that crystalline pepsin contains a few molecules with a slightly different amino acid composition.

A close examination of the amino acid distribution of pepsin reveals the presence of 44 aspartic, 27 glutamic acids, and only four basic residues; *i.e.*, one lysine, one histidine, and two arginines. This marked predominance of dicarboxylic acids with 35 free carboxyls[3] and the occurrence of one phosphate group (12, 15) explain that pepsin even in 0.1 N hydrochloric acid still moves electrophoretically as a negatively charged ion (15, 16).

The point of interest, however, is that pepsin has 44 serine, 28 threonine, and 15 proline residues, representing 26 per cent of the total amino acid content. It is, therefore, not unlikely that the sequential arrangement of these constituents, particularly that of proline, contributes to a specific configuration of a chain segment, or segments, essential to the catalytic activity of the enzyme.

The finding in this protein of small amounts of hydroxyproline was unexpected. Although it is possible that this amino acid originates from a contaminant, *e.g.* collagen or a collagen breakdown product which is tightly adsorbed to pepsin and which even on extensive purification cannot be removed, it is striking that in all the preparations thus far analyzed the amount of hydroxyproline is constant.[4] This constancy might be interpreted as an indication that we deal with a mixture of closely related proteins and that in every tenth molecule one proline residue is substituted by an hydroxyproline. We favor this latter explanation since preliminary experiments have indicated that the enzymically active pepsin modifications obtained on mild autodigestion contain

[3] Amide values computed from the chromatography results and not based on independent determinations.

[4] Analysis for hydroxyproline on crystalline pepsin, Pentex, lot 3713, indicated also the presence of 0.1 residue of this amino acid per mole of pepsin.

one to five residues of this amino acid (3, 17). The fact should be kept in mind that large amounts of a protein are necessary for the detection of traces of hydroxyproline. Thus this amino acid could occur in a variety of proteins of non-collagen origin and could have escaped the attention of investigators.

The authors are greatly indebted to Dr. Stanford Moore for generously supplying the Amberlite IR 120 resin used in this investigation and for his advice and interest in this work, and to Dr. R. Trautman for the ultracentrifuge measurements.

Our sincere thanks are also due to Miss Hessy Levinsons for her able assistance in the chromatography experiments.

Addendum

As stated earlier in this paper, one of the values subject to error in the analysis of amino acids is that of cystine. Since an accurate determination requires analysis of performic acid–oxidized pepsin for cysteic acid, performic acid oxidation was carried out according to the method of Schram, Moore, and Bigwood (18). The excess reagent was removed by freeze-drying (19) and the oxidized protein hydrolyzed at 155°C. under reflux with constant boiling hydrochloric acid for 18 hours. Aliquots containing 3 to 4 mg. of protein were then analyzed chromatographically on Amberlite IR 120 columns (4).

The results obtained with the twice recrystallized pepsin preparations, lots 611, 617, and 623, indicated the presence of 5.9_5, 5.8_7, and 5.7_4 residues of cysteic acid, respectively. It is thus clear that the discrepancy between the values given in Tables I and II and that found after performic acid oxidation is due to loss of cystine during the acid hydrolysis. Moreover, the integral value of six half-cystine agrees well with that reported by Kern (20). As in the work of Kern no free sulfhydryl groups are present in our preparations as shown with the aid of the amperometric titration of Benesch, Lardy, and Benesch (21).

BIBLIOGRAPHY

1. Perlmann, G. E., *Arch. Biochem. and Biophysics*, 1956, **65,** 210.
2. Brand, E., *Ann. New York Acad. Sc.*, 1946, **47,** 187.
3. Perlmann, G. E., and Mycek, M. J., Modification of pepsin by autodigestion, *in* Symposium on Protein Structure, (A. Neuberger, editor), New York, John Wiley and Sons, Inc., 1958, 179.
4. Moore, S., Spackman, D. H., and Stein, W. H., *Ind. and Eng. Chem., Analytical Edition*, 1958, **30,** 1185.
5. Spackman, D. H., Stein, W. H., and Moore, S., *Ind. and Eng. Chem., Analytical Edition*, 1958, **30,** 1190.
6. Moore, S., and Stein, W. H., *J. Biol. Chem.*, 1954, **211,** 907.
7. Martin, C. J., and Axelrod, A. E., *Proc. Soc. Exp. Biol. and Med.*, 1953, **83,** 461.
8. Goodwin, T. L., and Morton, R. A., *Biochem. J.*, 1946, **40,** 628.
9. Beaven, G. H., and Holiday, E. R., *Advances Protein Chem.*, 1952, **7,** 319.
10. Steinhardt, J., *J. Biol. Chem.*, 1938, **123,** 543.
11. Edelhoch, H., *J. Am. Chem. Soc.*, 1957, **79,** 6100.

12. Northrop, J. H., *J. Gen. Physiol.*, 1930, **13,** 739.
13. Herriott, R. M., Desreux, V., and Northrop, J. H., *J. Gen. Physiol.*, 1940, **24,** 213.
14. Hoch, H., *Nature*, 1950, **165,** 278.
15. Perlmann, G. E., *J. Gen. Physiol.*, 1958, **41,** 441.
16. Tiselius, A., Henschen, G. E., and Svensson, H., *Biochem. J.*, 1938, **32,** 1814.
17. Perlmann, G. E., Blumenfeld, O. O., and Mycek, M. J., *Rockefeller Inst. Bull., Ann. Rep.*, 1956–57, **2,** No. 2, 24.
18. Schram, E., Moore, S., and Bigwood, E. J., *Biochem. J.*, 1954, **57,** 33.
19. Thompson, E. A., *Biochim. et Biophysic. Acta*, 1954, **15,** 440.
20. Kern, H. L., cited in The Mechanism of Enzyme Action, (W. D. McElroy and B. Glass, editors), Baltimore, The Johns Hopkins University Press, 1954, 44.
21. Benesch, R. E., Lardy, H. A., and Benesch, R., *J. Biol. Chem.*, 1955, **216,** 663.