Dalgarno, L., and Gros, F. (1968), *Biochim. Biophys. Acta 157*, 52.
Gierer, L., and Gierer, A. (1968), *J. Mol. Biol. 34*, 294.
Gravilova, L. P., Ivanof, D. A., and Spirin, A. S. (1966), *J. Mol. Biol. 16*, 473.
Kaempfer, R., Meselson, M., and Raskas, H. (1968), *J. Mol. Biol. 31*, 277.
Kurland, C. G. (1966), *J. Mol. Biol. 18*, 90.
Mangiarotti, G., Apirion, D., Schlessinger, D., and Silengo, L. (1968), *Biochemistry 7*, 456.
Marchis-Mouren, G., Cozzone, A., and Marvaldi, J. (1969), *Biochim. Biophys. Acta 186*, 232.
Nomura, M., and Watson, J. D. (1959), *J. Mol. Biol. 1*, 204.
Osawa, S. (1968), *Ann. Rev. Biochem. 37*, 109.
Osawa, S., Otaka, E., Takuzi, I., and Fukui, T. (1969), *J. Mol. Biol. 40*, 321.
Santer, M., Ruebush, T. K., Brunt, J. V., Oldmixon, E., Hess, R., Primakoff, P., and Palade, P. (1968), *J. Bacteriol. 95*, 1355.
Schleif, R. (1967a), *J. Mol. Biol. 27*, 41.
Schleif, R. (1967b), *Mol. Gen. Genet. 100*, 252.
Spitnik-Elson, P. (1965), *Biochem. Biophys. Res. Commun. 18*, 557.
Spitnik-Elson, P., Atsmon, A., and Elson, D. (1969), *6th FEBS Meeting, Madrid, Symp. I.*
Sypherd, P. J. (1965), *J. Bacteriol. 90*, 403.
Tissières, A., Watson, J. D., Schlessinger, D., and Hollingworth, B. R. (1959), *J. Mol. Biol. 1*, 221.
Traub, P., and Nomura, M. (1969), *J. Mol. Biol. 60*, 391.
Traut, R. R. (1966), *J. Mol. Biol. 21*, 571.
Traut, R. R., Moore, P. B., Delius, H., Noller, H., and Tissières, A. (1967), *Proc. Natl. Acad. Sci. U. S. 57*, 1294.

An Investigation of the Products of an Enzymic Hydrolysis of Collagens*

Howard B. Bensusan

ABSTRACT: Ichthyocol and insoluble, acid-soluble, and salt-soluble bovine collagen were digested enzymically with collagenase, papain, and a mixture of prolidase and leucine aminopeptidase. Analysis of the protein-free product showed that 70–80% of the residues was present as free amino acids. The peptides present were isolated and their quantity and sequences were determined. Most of the peptides were tripeptides with the general sequence of Gly-Pro-X and X-Hyp-Gly as would be expected from the known specificities of the enzymes and the available sequence information concerning the collagen chains. The peptide, Gly-Pro-Hyp, was present to the extent of 40 and 32 moles per chain in the digests of bovine and ichthyocol collagen, respectively. Several peptides were found which have not been previously reported. A peptide having the tentatively determined sequence of Gly-(Gln, Hyp, ε-Lys) was also located. The quantity of each amino acid occurring in the peptides together with the quantity of each free amino acid accounted for essentially all of the residues when compared with acid hydrolysates of the enzymic hydrolysates. Except for two to three residues of glutamine in peptide sequences, all residues of glutamic acid and glutamine were liberated by the enzymes. Judging from the specificities of the pure enzymes, this should not occur if collagen contains γ-glutamyl linkages.

The difficulty faced in any attempt to obtain a complete enzymic hydrolysis of collagen stems from the fact that mammalian collagen contains about 24% proline plus hydroxyproline. A large portion of the proline occurs in apolar regions of the collagen chain involving sequences of the type, Gly-Pro-X, where X is most frequently an hydroxyproline or alanine residue. These sequences may be liberated through the activity of the collagenase from *Clostridium histolyticum* which hydrolyzes the bond between X and glycine in the sequence Pro-X-Gly-Pro-Y. Unfortunately, at the time this investigation was under way there was no known enzyme capable of hydrolyzing the tripeptide X-Pro-Y at a reasonable rate.

The first attempt at an extensive hydrolysis of collagen through the use of enzymes was conducted by Seifter *et al.* (1961). These investigators found that a combination of collagenase, prolidase, and leucine aminopeptidase resulted in the production of only a very limited amount of free amino acids. Their results showed that only 20% of the glycine, 10% of the proline, 20% of the hydroxyproline, and no glutamic or aspartic acid were liberated as free amino acids. It is not surprising that few free amino acids would be liberated by this enzyme mixture. The collagenase would liberate the tripeptides with Gly-Pro-X sequence leaving large peptides containing both proline and hydroxyproline. Since the action of leucine aminopeptidase essentially ceases when an

* From the Department of Biochemistry, Case Western Reserve University, Cleveland, Ohio. *Received June 19, 1969.* Please address all inquiries to University Hospitals, Cleveland, Ohio 44106. This research was supported by Grants AM-01825 and HD-00669 from the National Institutes of Health, U. S. Public Health Service.

imino acid occurs next to the *N*-terminal residue, it could not be expected to effect appreciable hydrolysis of polypeptides containing significant amounts of proline.

Hill and Schmidt (1962) reported a procedure for the complete enzymic digestion of proteins. Preliminary experiments demonstrated that papain would hydrolyze a third of the peptide bonds in denatured papain and was superior to other endopeptidases in bringing about extensive hydrolysis. When the treatment with papain was followed by a mixture of prolidase and leucine aminopeptidase, a complete hydrolysis of several proteins was realized. In view of their success, it seemed reasonable to investigate the use of their technique on collagenase digests of collagen although a complete hydrolysis of collagen could not be expected in view of products of collagenase digestion described above. This paper describes the results of the investigation.

Materials and Methods

Collagen Samples. The insoluble collagen, prepared from steer corium by the method of Veis *et al.* (1960), was a generous gift from Dr. A. Veis. Salt-soluble and acid-soluble collagen were obtained from sequential extracts of ground calfskin using 1 M NaCl and 0.1 N HOAc and were purified as described previously (Bensusan and Scanu, 1960). Ichthyocol was prepared according to Gallop (1955).

Enzymes. Collagenase, a twice-crystallized preparation, and papain, a twice-crystallized preparation from papaya latex, were purchased from Mann Research Laboratories. Prolidase was prepared by the method of Davis and Smith (1957) and had a C_1 activity of 11 after several months of storage. The leucine aminopeptidase was prepared according to Hill *et al.* (1958) except that the last step was omitted. The fresh preparation had a C_1 activity of 28. The carboxypeptidase B used in the sequence studies was obtained from Worthington Biochemical Corp.

Other Materials. The synthetic peptides used in this study were purchased from Cyclo Chemical Co. All other chemicals were the best quality obtainable commercially.

Enzymic Digestion. The soluble collagen solutions in 0.1 M HOAc were adjusted to pH 7.4–8.0 using 1 M K_3PO_4 and were incubated at 37° until fiber formation was complete. The reconstituted collagen was collected as small pellets by centrifugation at 60,000*g* for 30 min. The pellets were washed with several changes of water and suspended in a volume of 5×10^{-3} N Tris buffer, (pH 7.4) such that upon warming to 60° for 10 min a concentration of about 15 mg/ml was obtained. The gelatin solutions (2 ml) were used for the digestion. A larger quantity of insoluble collagen was used for the digestion. In the case of insoluble collagen, 1 g of the dry cubes of collagen was shaved into fine pieces with a razor blade, suspended in 30 ml of the Tris buffer, and heated to 80° to denature the collagen. A solution (2 mg/ml) of collagenase in Tris buffer was added to the gelatin solutions and the denatured suspension of insoluble collagen to give a collagen–enzyme concentration of 25:1. A crystal of thymol was added to each and the mixtures were incubated at 37° for 18 hr. The clear solutions were adjusted to pH 5.5 and the directions given by Hill and Schmidt (1962) were followed for the remainder of the digestion with papain and a mixture of prolidase and leucine aminopeptidase. The prolidase–leucine aminopeptidase mixture consisted of 29 mg of prolidase and 40 mg of leucine aminopeptidase in 5 ml of 0.005 M Tris buffer (pH 8.0), containing 0.005 M $MnCl_2$. A volume of the mixture was added to give approximately 4 μl/mg of original protein for the insoluble collagen, 8 μl/mg for the soluble calf collagens, and 12 μl/mg for ichthyocol. The protein digests were dialyzed and the free amino acids and peptides of the dialysate were concentrated by evaporation *in vacuo*.

Amino Acid Analysis. Acid hydrolysis, when used, was carried out on 2–3-mg samples in 3–4 ml of constant-boiling HCl. Prepurified nitrogen was bubbled through the solution which was then frozen in a Dry Ice–acetone mixture. The tube was then evacuated with a high-vacuum oil pump and sealed. Hydrolysis was carried out at 100° for 18–24 hr.

The analyses of enzyme digests and the control acid hydrolysates were performed on the Technicon amino acid analyzer using the 21-hr elution gradient and the 0.6 × 140 cm column. The resin column was equipped with a stream-splitting device to allow the collection of eluted fractions. The time of collection was correlated with the time of appearance of the ninhydrin peaks using dilute picric acid added to the sample. The picric acid is identified by its color in the collections and by a peak absorbing at 440 mμ on the analyzer record. Three-minute collections (1.5 ml) were used when isolating fractions for quantitative determinations.

Quantitative Determinations of Peptides. Peptides were collected from three column runs of aliquots of the enzyme hydrolysate of insoluble collagen by the technique described above. The collections of each peptide fraction were pooled, adjusted to pH 2.0–2.2, and desalted on a 0.8 × 12 cm column of Dowex 50-X8 resin in the hydrogen form, washing with water and eluting with 4 N NH_4OH. The eluate was freed of NH_3 by repeated flash evaporation. The peptides were then hydrolyzed in constant-boiling HCl and the amino acid contents were determined on the amino acid analyzer. The recovery obtained in this procedure was determined by isolating four different free amino acid fractions and putting them through the same procedure. Knowing the recovery to be 95 ± 2%, it was possible to determine the amount of each constituent amino acid in the peptide fractions of an aliquot of the original mixture. It was then possible to calculate the color yield of each peptide and, therefore, to determine the amount of each peptide in the enzyme hydrolysates of the other collagen samples without a repetition of this laborious procedure.

Except in the two cases discussed later, the isolated peptide fractions were found on acid hydrolysis to contain major amounts of either two or three amino acids in amounts related by whole numbers. The occurrence of other amino acids was limited to trace amounts. No further purification was attempted.

The Determination of Serine, Threonine, and Amides. Since serine, threonine, glutamine, and asparagine in the enzyme digests elute together in unresolved fractions, it was impossible to determine their contents directly. To accomplish these determinations, two column runs of aliquots were used to collect this mixed fraction. The lysine peak was also collected and pooled with the collections of the mixed fraction to serve as an internal determination of final recoveries (95–98%). The pooled collections were desalted as before, and the amides were converted into the free dicarboxylic acids by acid hydrolysis. The four amino acids could then be quantitated by amino acid analysis using the lysine value to correct

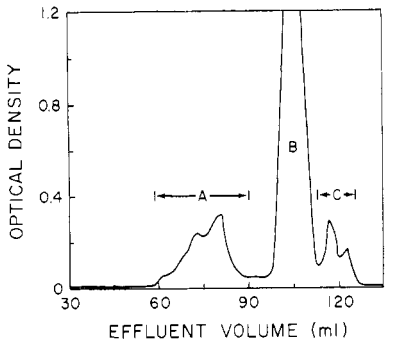

FIGURE 1: Elution pattern of enzymic digests of insoluble collagen from a 1.2 × 140 cm column of Bio-Gel P2. About 5 mg of the hydrolysate was added to the column. The eluting liquid was water flowing at the rate of 30 ml/hr. The ordinate is the optical density at 570 mμ of a small fraction of the effluent after reaction with ninhydrin.

for losses. These determinations were made on each of the collagen samples.

Glutamine determinations presented a special problem since the glutamine in the original digest could be converted in part into pyrrolidone-5-carboxylic acid on passage through the resin column and would not be part of the mixed collection. To determine the loss of glutamine, a standard glutamine solution was compared with a standard solution of norleucine for ninhydrin color yield by introducing each directly into the analytical portion of the analyzer without passage through the column. In this way, it was determined that glutamine had a color yield of 1.01 as compared with norleucine. Equal amounts of norleucine and glutamine were then mixed and analyzed by a column run on the amino acid analyzer. From the area of the peaks obtained from duplicate determinations it was found that only 25 ± 0.5% of the glutamine could be recovered in collections of the mixed fraction. This value was used to correct the results of the glutamine determinations.

The Determination of Pyrrolidone-5-carboxylic Acid. During the enzymic hydrolysis, some of the glutamine residues released cyclized to form pyrrolidone-5-carboxylic acid. The pyrrolidone-5-carboxylic acid cannot be determined when the enzymic hydrolysate is analyzed on the amino acid analyzer. In order to determine pyrrolidone-5-carboxylic acid, an aliquot of the hydrolysate was placed on a short column of Dowex 50-X8 in the hydrogen form and the pyrrolidone-5-carboxylic acid was eluted with water. The eluent was then placed on a short column packed with 2.5 ml of Dowex 1 in the acetate form. The column was washed with 0.5 M HOAc (30 ml) and the pyrrolidone-5-carboxylic acid was eluted with 2 N HCl (15 ml). The eluent was heated at 100° for 2.5 hr to convert the pyrrolidone-5-carboxylic acid into glutamic acid, which could then be determined on the amino acid analyzer after evaporation to remove the HCl. A standard solution of pyrrolidone-5-carboxylic acid was treated in parallel with samples from the different collagens in order to provide a correction for losses. A 91% recovery of the pyrrolidone-5-carboxylic acid as glutamic acid was obtained.

Large-Scale Isolation of Peptides. The dialysate from about 300 mg of digested insoluble collagen was concentrated to 4 ml and 0.5-ml aliquots were placed on a 1.2 × 140 cm column of Bio-Gel P2, 100–200 mesh, which had been equilibrated with water. The mixture was eluted with water at a rate of 30 ml/hr. The collected fractions were analyzed for ninhydrin color with the aid of a Technicon automatic sampling device. A typical elution diagram is given in Figure 1. Fractions labeled A and C in Figure 1 were collected and concentrated *in vacuo*. Each fraction was run on the amino acid analyzer using Dowex 50-X8, (minus 400 mesh) in a jacketed column (1.2 × 140 cm) at 60°. The peptide fractions, eluted with four times the normal gradient volume at four times the normal rate, were collected using a stream-splitting device. These fractions were desalted on short Dowex 50 columns, as described above, except that 4 M pyridine was used to elute the peptides.[1] The pyridine was removed by evaporation *in vacuo*.

Sequence Determinations. Amino acid analyses of acid hydrolysates of the isolated peptides afforded a preliminary identification and quantitation of the constituent amino acids. The amino acid sequences of most of the peptides could be determined by the Edman degradation, as described by Fraenkel-Conrat *et al.* (1955). Usually, all but the C-terminal amino acid were isolated as the phenylthiohydantoin derivatives. The C-terminal group was determined on the amino acid analyzer, using the 5.5-hr gradient system of Thomson and Miles (1964). When the Edman degradation failed to show the N-terminal residue, which occurred when arginine, serine, or glutamine occurred in that position, the residue was identified as the dansyl derivative as described by Gray and Hartley (1963). Dansyl derivatives of standard amino acids were prepared according to Boulton and Bush (1964). In one case (Gly-Pro-Pro-Gly), the hydrazinolysis technique (Akabori *et al.*, 1952) was used to determine the C-terminal glycine and in one case carboxypeptidase B and hydrazinolysis were used to confirm a C-terminal arginine.

Results and Discussion

Peptides Remaining in Enzymic Digests. The mixed fraction, labeled A in the elution diagram from the Bio-Gel P2 column (Figure 1), was found to contain the first 15 peptides shown in Figure 2, together with glutamic and aspartic acids and a small amount of proline, when rechromatographed on the amino acid analyzer. Fraction C from the Bio-Gel column contained tyrosine, phenylalanine, the basic amino acid residues, and peptides 18 and 19. Peptide 17 and fraction 16 had to be isolated by chromatography of the unresolved enzymic digest on Dowex 50. Fraction B was not analyzed.

The results of the sequence analyses of the various peptides are presented in Table I. As expected from the specificity of collagenase and the inability of either prolidase or leucine aminopeptidase to easily hydrolyze tripeptides containing either proline or hydroxyproline in the second position, it is not surprising that all peptides (with the possible exception of 14) contain either proline or hydroxyproline in the second position. All the peptides contain glycine, which is consistent with the established fact that glycine occurs at every third residue over a large portion of the collagen chains (see Harrington and von Hippel, 1961). The identification of the two main sequences, Gly-Pro-X and X-Hyp-Gly, is

[1] The most basic peptides had to be eluted with 4 N NH_4OH.

TABLE I: The Sequence and Content of Peptides Isolated from the Enzymic Hydrolysis of Various Collagens.

Peptide	Content (Moles/1000 Amino Acid Residues) Insoluble	Acid Soluble	Salt Soluble	Ichthyocol
1. Gln-Hyp-Gly	2.0	ND[a]	ND	ND
2. Gly-Pro-Hyp-Gly	8.8	3.1	3.6	0
3. Ala-Hyp-Gly	9.4	2.6	2.7	0
4. Ser-Hyp-Gly	2.9	1.1	0.8	0
5. Gly-Pro-Hyp	32.4	37.2	34.5	32.2
6. Gly-Pro-Ser	3.3	2.3	2.9	5.6
7. Leu-Hyp-Gly	1.3	Tr[b]	Tr	0
8. Ile-Hyp-Gly	0.4	ND	ND	0
9. Gly-Pro-Ala	10.5	8.4	7.0	14.3
10. Gly-Pro-Pro-Gly	2.1	ND	ND	ND
11. Gly-Pro-Val	1.5	ND	ND	ND
12. Gly-Pro-Gly	2.7	3.1	2.1	1.3
13. Gly-Pro-Ile	1.3	(1.4)[c]	(2.1)	(2.9)
14. Gly-(Gln, Hyp, ε-Lys)	0.6	(0.4)	(0.6)	(0.9)
15. Gly-Pro-Phe	2.1	Tr	Tr	Tr
16. Hyl,X	0.4	0.6	0.5	0.9
17. Gly-Pro-Lys	2.5	2.3	1.8	2.8
18. Gly-Pro-Arg	4.1	ND	ND	5.0
19. Arg-Hyp-Gly	1.0	ND	ND	ND

[a] ND represents not determined. [b] Tr represents trace. [c] Numbers in parentheses indicate estimates from unresolved peaks.

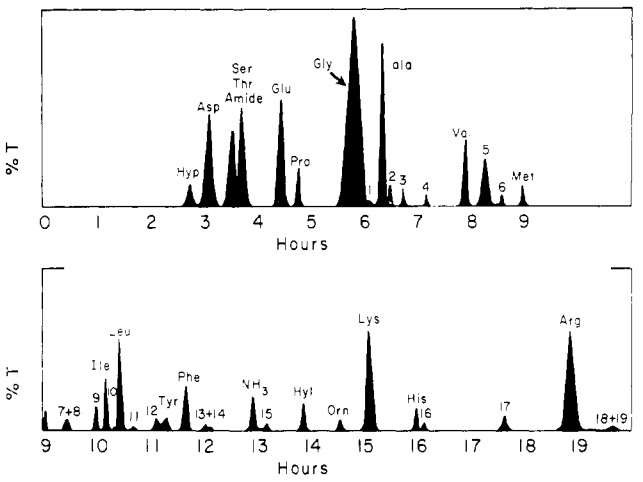

FIGURE 2: Elution pattern of the enzymic digest of insoluble collagen using the Technicon amino acid analyzer. The ordinate is given as transmittance in order to accentuate the size of the peptide peaks, which are numbered.

completely consistent with the demonstration that in collagen sequences, Gly-X-Y-Gly, proline occupies position X and hydroxyproline occupies position Y (Greenberg *et al.*, 1964). The single known exception occurs in earthworm cuticle collagen where hydroxyproline can occupy the second position (Goldstein and Adams, 1968).

Only in the case of the mixtures of peptides 7 and 8 and peptides 18 and 19 were sequences determined on unresolved peptides. The amino acid content of the mixture of peptides 7 and 8 was found to be Gly, Hyp, Leu, and Ile in a ratio of 1.0:1.0:0.8:0.2, respectively. Phenylthiohydantoin-leucine and phenylthiohydantoin-isoleucine were found in the first step of the Edman degradation and phenylthiohydantoin-hydroxyproline was found in the second step. Only glycine was found remaining after the second Edman cycle. These results indicated a mixture of Leu-Hyp-Gly and Ile-Hyp-Gly in a ratio of 4 to 1. The sequences given for the mixture of peptides 18 and 19 were deduced from the fact that the relative amounts of the amino acids in the mixture were 1.3:1.0:0.3:1.3 for glycine, proline, hydroxyproline, and arginine, respectively, and that both glycine and arginine were found in the first and third position, while proline and hydroxyproline occupied the second position. The proline-containing peptide was shown to contain the arginine in the C-terminal position since a majority of the arginine and only a small amount of glycine remained after hydrazinolysis.

Many of the peptides listed in Table I have previously been identified in partial digests of collagen. Gly-Pro-Hyp and Gly-Pro-Ala have been known for some time (Schroeder *et al.*, 1954). Peptides 3, 4, and 12 were isolated by Kroner *et al.* (1955), who also listed the peptide, Gly-(Hyp, Pro)-Gly, suggesting that they had obtained peptide 2. Grassmann *et al.* (1963) tentatively identified a peptide to be Gly-Pro-Ser (peptide 6). Ogle *et al.* (1961) reported peptide 18 in collagenase digests. Bornstein (1967) determined the sequence of the α1-CB2 polypeptide obtained from cyanogen bromide treatment of the α1 chain. This polypeptide contained peptides 2, 3, 6, 7, and 18 as well as peptide 10, which resulted from the incomplete hydroxylation of proline in position 3. In a personal communication, W. T. Butler (1969) showed that he had found peptides 17 and 19 in the sequence of the cyanogen bromide fractions α1-CB5 and α1-CB4, respectively. Thus, to the best of my knowledge, the peptides which have not been previously identified are peptides 1, 8, 11, 13, and 15.

There are two unidentified fractions which are of interest. The material in peak 16 (Figure 2) chromatographed apart from hydroxylysine, yet hydroxylysine was the only ninhydrin-positive constituent seen on acid hydrolysis. It is possible that this fraction also contains part of the hexose which is bound to the δ carbon of hydroxylysine (Butler and Cunningham, 1966). Peptide 14 is of particular interest in view of the fact that Mechanic and Levy (1959) reported the isolation and identification of N^{ϵ}-(Gly-Glu)-Lys in partial acid hydrolysates. Peptide 14 yielded Gly, Glu, Hyp, Lys, and NH_3 in the ratio of 1.0:1.0:1.1:0.8:1.0. Treatment of this peptide with dansyl chloride, followed by acid hydrolysis, showed the presence of DNS-glycine and α-DNS-lysine when compared with authentic samples by thin-layer chromatography on silica gel G using a 1-butanol–HOAc–H_2O (40:6:15) solvent system in which α-DNS-lysine was well separated from ϵ-DNS-lysine and di-DNS-lysine. A single Edman degradation cycle was performed on a fresh aliquot and after the extraction

TABLE II: The Amino Acid Composition of Enzymic Hydrolysates of Insoluble Collagen.

	Amino Acid Content (Residues/1000 Total Residues)			
	Enzyme Hydrolysate			Original
	Free[a]	Free + Peptide	Acid Hydrolyzed	Acid Hydrolyzed
Hydroxyproline	41.2	98.4	93.8	93.1
Aspartic acid	46.9	46.9	48.1	47.2
Threonine	13.0	13.0	18.3	15.4
Serine	27.6	34.4	42.0	36.0
Glutamic acid	71.5	73.6	65.0	71.3
Proline	33.1	108	120	121
Glycine	221	323	329	325
Alanine	106	126	115	112
Valine	23.2	24.7	22.5	16.9
Methionine	5.9	5.9	7.2	7.6
Isoleucine	11.3	13.1	11.4	9.9
Leucine	24.3	25.6	23.8	28.4
Tyrosine	3.3	3.3	3.8	3.7
Phenylalanine	10.5	12.5	12.3	14.2
Lysine	22.8	26.0	27.3	30.2
Histidine	3.7	3.7	3.8	3.7
Arginine	48.4	53.6	49.4	52.3
Hydroxylysine	5.5	6.1	6.9	7.1
Ornithine	1.6	1.6	1.4	1.0
Total (μmoles)	8.60	11.95	12.15	

[a] The total number of micromoles of the free amino acids plus the number of micromoles of amino acids in the peptides (last entry in the second column) was used to calculate the residues of each free amino acid per 1000 total residues. The same calculation applies in Tables III–V.

TABLE III: The Amino Acid Composition of Enzymic Hydrolysates of Acid-Soluble Collagen.

	Amino Acid Content (Residues/1000 Total Residues)			
	Enzyme Hydrolysate			Original
	Free	Free + Peptide	Acid Hydrolyzed	Acid Hydrolyzed
Hydroxyproline	52.8	98.0	93.5	104
Aspartic acid	46.1	46.1	44.8	44.6
Threonine	17.4	17.4	18.5	14.6
Serine	35.0	38.4	39.2	38.1
Glutamic acid	72.4	73.2	74.3	76.0
Proline	53.3	111	116	118
Glycine	255	324	319	325
Alanine	112	123	118	120
Valine	21.0	21.0	22.9	15.0
Methionine	5.2	5.2	6.7	6.4
Isoleucine	9.5	11.2	10.9	6.2
Leucine	23.8	24.5	24.0	25.1
Tyrosine	3.4	3.4	3.6	3.7
Phenylalanine	11.6	11.6	12.3	12.0
Lysine	26.7	29.7	28.4	28.5
Histidine	4.3	4.3	4.3	4.2
Arginine	47.2	47.2	50.4	49.3
Hydroxylysine	5.8	6.4	4.6	6.6
Ornithine	2.2	2.2	2.2	0.9
Total (μmoles)	6.92	8.60	8.69	

of the phenylthiohydantoin derivatives, the remainder was treated with dansyl chloride. Acid hydrolysis and thin-layer chromatography of the dansyl derivative showed only one DNS derivative which was not identifiable. Further investigation of this fraction is being conducted and the results will be published subsequently.

The quantitative determination of the amounts of the various peptides in the collagen samples (Table I) should not be interpreted as representing the quantities of each in the original protein. It was determined that Gly-Pro-Ala could be hydrolyzed slowly by a mixture of prolidase and leucine aminopeptidase which is consistent with the observation by Hill and Schmidt (1962) that Gly-Pro-Leu is slowly hydrolyzed by these enzymes. However, experiments with Gly-Pro-Hyp demonstrated that this peptide is completely resistant to hydrolysis by the enzymes used in this investigation. Therefore, the quantities of this peptide together with the Gly-Pro-Hyp-Gly (peptide 2) and Gly-Pro-Pro-Gly (peptide 10) are a measure of the amount of the Gly-P-P′ sequences (where P′ is either proline or hydroxyproline) in the original protein. It can be seen that calfskin collagen contains about 40 such sequences per chain and ichthyocol contains about 32. These values are in good agreement with the values of 37 and 25, respectively, predicted by Josse and Harrington (1964) from their probability calculations based on the imino acid content in these collagens.

Quantification of Enzymic Digests. A complete analysis of the enzymic digests of insoluble collagen is given in Table II. The amino acid content and color yield of each peptide were determined as described in the Methods section. Using these data the amount of each amino acid occurring in peptides could be determined from the chromatograms. When these values were added to the content of each of the free amino acids, the total content of each amino acid and the total amino acid content could be calculated. These values are given in the second column. The number of residues of each of the free amino acids could then be determined on the basis of their amount per total amino acid content (free plus peptide bound). These values are shown in the first column. The total amount of the free amino acids from the first column as compared with the total amino acid content from the second column shows that the enzymic hydrolysis was about 70% complete. Since the relative amount of each amino acid is not very sensitive to errors in the total amount, we determined the amino acid composition and total amount of the amino acids in acid hydrolysates of aliquot samples of the dialyzed enzymic hydrolysate. These values are given in the third

TABLE IV: The Amino Acid Composition of Enzymic Hydrolysates of Salt-Soluble Collagen.

	Amino Acid Content (Residues/1000 Total Residues)			
	Enzyme Hydrolysate			Original
	Free	Free + Peptide	Acid Hydrolyzed	Acid Hydrolyzed
Hydroxyproline	49.7	92.1	86.9	96.9
Aspartic acid	48.8	48.8	49.0	49.6
Threonine	20.5	20.5	22.6	18.2
Serine	37.5	41.2	44.8	42.3
Glutamic acid	77.4	78.0	76.2	78.8
Proline	61.1	115	111	119
Glycine	242	306	302	314
Alanine	105	115	113	114
Valine	25.4	25.4	26.2	14.6
Methionine	6.7	6.7	8.4	9.1
Isoleucine	13.5	15.6	14.1	8.0
Leucine	27.6	27.6	28.1	26.9
Tyrosine	4.5	4.5	6.7	5.1
Phenylalanine	11.7	11.7	13.0	12.3
Lysine	29.7	32.1	33.8	30.3
Histidine	6.1	6.1	5.8	5.9
Arginine	47.1	47.1	49.4	47.0
Hydroxylysine	5.1	5.6	5.7	6.6
Ornithine	1.8	1.8	2.0	0.9
Total (μmoles)	7.34	9.47	10.58	

TABLE V: The Amino Acid Composition of Enzymic Hydrolysates of Ichthyocol.

	Amino Acid Content (Residues/1000 Total Residues)			
	Enzyme Hydrolysate			Original
	Free	Free + Peptide	Acid Hydrolyzed	Acid Hydrolyzed
Hydroxyproline	48.1	81.2	80.8	77.8
Aspartic acid	43.8	43.8	46.3	47.0
Threonine	24.8	24.8	29.2	26.7
Serine	38.2	43.8	43.0	35.2
Glutamic acid	70.1	71.0	61.5	70.3
Proline	33.8	98.0	106	116
Glycine	259	326	329	334
Alanine	124	138	128	126
Valine	18.7	18.7	19.2	14.2
Methionine	12.4	12.4	16.0	15.2
Isoleucine	11.7	14.7	11.4	7.8
Leucine	21.9	21.9	23.3	21.4
Tyrosine	3.9	3.9	4.0	3.8
Phenylalanine	12.8	12.8	13.3	13.9
Lysine	24.2	27.9	27.0	27.1
Histidine	3.9	3.9	3.6	3.3
Arginine	43.1	48.2	50.3	50.2
Hydroxylysine	4.5	5.5	5.3	8.7
Ornithine	3.3	3.3	3.0	1.8
Total (μmoles)	6.00	7.49	7.23	

column. The data obtained from acid hydrolysates of insoluble collagen without enzymic treatment are given in the last column.

After the dialysis of the enzymic hydrolysates, that fraction of the material remaining in the dialysis sack which was soluble in 10% trichloroacetic acid was hydrolyzed with HCl and an aliquot was analyzed on the amino acid analyzer. The results showed that the total amount of trichloroacetic acid soluble material remaining after dialysis represented only 3% of the total collagen and contained only 200 residues of glycine, 64 residues of hydroxyproline, and 30 residues of arginine per 1000 total residues. These values indicate that only about two-thirds of the material was undialyzed collagen. Therefore, it appears that the analyses of the enzymic hydrolysates reported above represent the product of digestion of essentially all of the collagen.

The data obtained from samples of acid-soluble and salt-soluble calfskin collagen and from ichthyocol are presented in Tables III–V. The data in these tables are not as reliable as those in Table II because the amounts of the peptides which could not be resolved on the amino acid analyzer were not determined since such a determination involved a very long and laborious procedure. However, it can be seen that in all four collagen samples, the agreement is very good[2] between the calculated content of each amino acid in the enzymic hydrolysates (second column) and actual amino acid composition after acid hydrolysis (third and fourth columns) with two exceptions. The quantities of valine and isoleucine in the enzymic hydrolysis agree with those found after acid hydrolysis of the enzymic hydrolysates but were consistently greater than those in the acid hydrolysates of the original protein, although the latter were corrected using the values given by Piez *et al.* (1960). Thus, it is obvious that their correction factor was not sufficiently great to correct for the hydrolytic conditions in these determinations.

It has been a consistent observation that all of our acid-hydrolyzed bovine collagen samples contain a single residue of ornithine per chain and that ichthyocol has one or two residues. Ornithine appears as a resolved peak in a single-column amino acid analyzer but could be missed in the two-column system since it elutes with lysine. To my knowledge

[2] Although the standard deviation of the values for the amino acid analyses was determined to be 3.2% for all but the very minor components in collagen, additional errors in the values given in the second and third columns are possible. The values in the second column depend on the accuracy with which the color yields of the peptides were determined. In addition there is no certainty that all traces of peptides were recovered. This is obvious from Table I which shows that some peptides were not determined since they could not be resolved in the chromatograms. Errors could be introduced in the determinations of acid hydrolysates of the enzymic digests by the contamination of the enzymic digests with trace peptides resulting from the autolysis of the enzymes.

only Steven and Jackson (1967) have reported its existence in acid hydrolysates of collagen. The material was shown to be ornithine in my laboratory since it and the didinitrophenyl derivative cochromatographed with authentic samples on thin-layer plates using a 1-butanol–HOAc–H_2O (40:6:15) solvent system. Whether or not ornithine exists in the native protein is a question worthy of investigation.

The values for the free serine, threonine, and aspartic, and glutamic acids in the enzymic hydrolysates could not be determined directly as discussed in the Methods section. The aspartic acid values in the first columns in Tables II–V are the sum of determinations of aspartic acid and asparagine while those for glutamic acid are the sum of values of glutamic acid, glutamine, and pyrrolidone-5-carboxylic acid. A breakdown of the individual values for all four collagen samples is given in Table VI.

TABLE VI: The Free Dicarboxylic Acids in Enzymic Hydrolysates of Various Collagens.

	Content (Residues/1000 Amino Acid Residues)				
	Enzyme Hydrolysate				Original[a]
	Glu	Gln	PCA[c]	Total[b]	Total
Insoluble	49.3	18.9	3.3	71.5	71.3
Acid soluble	42.0	22.0	6.4	72.4	76.0
Salt soluble	48.4	24.8	4.2	77.4	78.8
Ichthyocol	46.7	16.2	7.2	70.1	70.3
	Asp	Asn		Total	Total
Insoluble	33.2	12.6		45.8	47.2
Acid soluble	30.9	15.2		46.1	44.6
Salt soluble	32.5	16.3		48.8	49.6
Ichthyocol	27.5	16.3		43.8	47.0

[a] Values from acid hydrolysates. [b] Does not include the 0.4–2.6 residues of glutamine remaining in peptides 1 and 14. [c] PCA = pyrolidone-5-carboxylic acid.

In order to determine whether or not the values of pyrrolidone-5-carboxylic acid were reasonable, glutamine was incubated under the same conditions of time, temperature, and pH as the last step in the enzymic hydrolysis procedure. Under these conditions the quantity of ammonia liberated showed that about 10% of the glutamine was converted into pyrrolidone-5-carboxylic acid which would indicate that the values in Table VI are reasonable ones.

With the exception of the small amount of glutamine-containing peptides listed in Table I, all of the glutamic acid and glutamine appear to be accounted for as free residues in the enzymic hydrolysates. Since none of the enzymes used in the hydrolysis procedure are known to be capable of hydrolyzing γ-glutamyl bonds, these results are contrary to the investigations of Gallop *et al.* (1960) and Franzblau *et al.* (1963), who, on the basis of the results of a series of chemical reactions, suggested that there are a minimum of 20 γ-glutamyl linkages in collagen. This apparent discrepancy has been investigated and the results of the research are presented in the following paper (Bensusan, 1969).

Since the specificities of the enzymes used in this investigation are known, it is not surprising that a complete enzymic digestion was not obtained. However, the recovery of 70–80% of the residues as free amino acids represents the best enzymic hydrolysis of collagen yet attained. Since all of the remaining small peptides contain an imino acid residue and glycine, as expected, it now becomes possible to search for unusual linkages in collagen, such as in peptide 14, under circumstances which reduce the danger of destruction during an acid or alkaline hydrolysis.

After the completion of this work, Nordwig and Dehm (1968) published a preliminary report describing the preparation of two enzymes capable of hydrolyzing amino acids attached to proline. One, an aminopeptidase, releases the N-terminal glycine from Gly-Pro-Hyp, Gly-Pro, and Gly-Pro-Met-Gly-Pro-Ala. The other, a carboxypeptidase, releases the C-terminal group from Z-Gly-Pro-Ala and Z-Pro-Ala (Z = benzyloxycarbonyl). Theoretically, these enzymes should complete the hydrolysis of all the remaining peptide bonds except those involving Pro-Pro and Pro-Hyp. Work is now under way in this laboratory to investigate the use of these enzymes.

Acknowledgments

I gratefully acknowledge the capable technical assistance of Mr. Smith D. McKnight and Mrs. Audrey Scott.

References

Akabori, S., Ohno, K., and Narita, K. (1952), *Bull. Chem. Soc. Japan 25*, 214.

Bensusan, H. B. (1969), *Biochemistry 8*, 4273.

Bensusan, H. B., and Scanu, A. (1960), *J. Am. Chem. Soc. 82*, 4990.

Bornstein, P. (1967), *Biochemistry 6*, 3083.

Boulton, A. A., and Bush, I. E. (1964), *Biochem. J. 92*, 11P.

Butler, W. T., and Cunningham, L. W. (1966), *J. Biol. Chem. 241*, 3882.

Davis, N. C., and Smith, E. L. (1957), *J. Biol. Chem. 224*, 261.

Fraenkel-Conrat, H., Harris, J. I., and Levy, A. L. (1955), *Methods Biochem. Anal. 2*, 383.

Franzblau, C., Gallop, P. M., and Seifter, S. (1963), *Biopolymers 1*, 79.

Gallop, P. M. (1955), *Arch. Biochem. Biophys. 54*, 486.

Gallop, P. M., Siefter, S., Lukin, M., and Meilman, E. (1960), *J. Biol. Chem. 235*, 2619.

Goldstein, A., and Adams, E. (1968), *J. Biol. Chem. 243*, 3550.

Grassmann, W., Hannig, K., and Nordwig, A. (1963), *Z. Physiol. Chem. 333*, 154.

Gray, W. R., and Hartley, B. S. (1963), *Biochem. J. 89*, 59P.

Greenberg, J., Fishman, L., and Levy, M. (1964), *Biochemistry 3*, 1826.

Harrington, W. F., and von Hippel, P. H. (1961), *Advan. Protein Chem. 16*, 1.

Hill, R. L., and Schmidt, W. R. (1962), *J. Biol. Chem. 237*, 389.

Hill, R. L., Spackman, D. H., Brown, D. M., and Smith, E. L. (1958), *Biochem. Prepn. 6*, 35.

Josse, J., and Harrington, W. F. (1964), *J. Mol. Biol. 9*, 269.

Kroner, T. D., Tabroff, W., and McGarr, J. J. (1955), *J. Am. Chem. Soc. 77*, 3356.
Mechanic, G. L., and Levy, M. (1959), *J. Am. Chem. Soc. 81*, 1889.
Nordwig, A., and Dehm, P. (1968), *Biochim. Biophys. Acta 160*, 293.
Ogle, J. D., Arlinghaus, R. B., and Logan, M. A. (1961), *Arch. Biochem. Biophys. 94*, 85.
Piez, K. A., Weiss, E., and Lewis, M. S. (1960), *J. Biol. Chem. 235*, 1987.
Schroeder, W. A., Kay, L. M., LeGette, J., Honnen, L., and Green, F. C. (1954), *J. Am. Chem. Soc. 76*, 3556.
Seifter, S., Gallop, P. M. and Franzblau, C. (1961), *Trans. N. Y. Acad. Sci. 23*, 540.
Steven, F. S., and Jackson, D. S. (1967), *Biochem. J. 104*, 534.
Thomson, A. R., and Miles, B. J. (1964), *Nature 203*, 483.
Veis, A., Anesey, J., and Cohen, J. (1960) *J. Am. Leather Chemists' Assoc. 55*, 548.

The Glutamyl Linkages in Collagen*

Howard B. Bensusan

ABSTRACT: In the preceding paper (Bensusan, 1969) it was shown that the sequential digestion of collagen with collagenase, papain, and prolidase plus leucine aminopeptidase resulted in the release of essentially all of the glutamic acid residues as free glutamic acid. None of these enzymes is known to hydrolyze γ-glutamyl linkages. Since the presence of a relatively large number of such linkages has been reported previously (Franzblau *et al.*, 1963), a resolution of the apparent discrepancy was deemed necessary. Preliminary experiments with the enzymes showed them incapable of hydrolyzing the γ-glutamyl peptide linkage in γ-glutamylalanine, γ-glutamylglutamic acid, or glutathione. The free carboxyl groups of ichthyocol gelatin were modified by coupling them with L-alanine methyl ester. When this derivative was subjected to the enzymic hydrolysis, it was found that only 2.6 residues out of a possible 50 residues of glutamic acid appeared in the digest. This demonstrated not only that the γ-carboxyl groups are free in ichthyocol but that the enzymes do not hydrolyze γ-glutamyl linkage. The method used by Franzblau *et al.* (1963) to demonstrate the presence of γ-glutamyl linkages in ichthyocol was then applied to ichthyocol, oxidized ribonuclease, oxidized insulin A, and oxidized insulin B. The results with ribonuclease and insulin B chains indicated the presence of γ-glutamyl linkages. Insulin A chains gave no reaction. A similar method based on the report by Hoare *et al.* (1968) was investigated using ichthyocol and salt-soluble collagen. The results showed that all reacting free carboxyl groups of glutamyl residues were the γ-carboxyl groups. It was concluded that essentially all the glutamyl residues in collagen are involved in the normal α-glutamyl linkages.

Gallop *et al.* (1960) first reported the possible existence of γ-glutamyl linkages in collagen. Their technique involved the esterification of the free carboxyl groups and then the conversion of the esters into hydroxamates with neutral hydroxylamine. The hydroxamates were treated with 1-fluoro-2,4-dinitrobenzene (FDNB)[1] to form the dinitrophenylhydroxamates, which, when heated in strong alkali, undergo the Lossen rearrangement to yield an amine and carbon dioxide. An α-linked glutamyl residue would be expected to yield 2,4-diaminobutyric acid by this treatment, whereas a γ-linked glutamyl residue would yield succinic semialdehyde which could be identified as the 2,4-dinitrophenylhydrazone derivative. These investigators found that 35 residues of glutamic acid had reacted but that only 8.4 residues of 2,4-diaminobutyric acid had been produced. A positive identification of succinic semialdehyde was made. They cautiously recognized the possibility that the treatment of the esterified gelatin with hydroxylamine could result in an intramolecular carboxyl group exchange which would give false results. With this in mind, they formed the hydroxamates of the free carboxyl groups at approximately pH 4 by treating the gelatin with a water-soluble carbodiimide in the presence of hydroxylamine (Franzblau *et al.*, 1963). When the dinitrophenylation and Lossen rearrangement were carried out on these hydroxamate derivatives of ichthyocol, they found that 25 residues of glutamic acid were lost and that 4 residues of 2,4-diaminobutyric acid and 22 residues of succinic semialdehyde were formed on acid hydrolysis, indicating that a large majority of the glutamyl residues are involved in γ-glutamyl linkages.

Joseph and Bose (1960) used a sodium hypobromite oxidation of trypsin digests to investigate possible γ-glutamyl linkages. The succinic acid which would be produced from the oxidation of N-terminal glutamyl residues involved in γ-linkages was determined. These investigators found 1.3 and 0.14 residues of succinic acid in buffalo hide collagen and alkali-processed gelatin, respectively.

Bensusan (1969), using an enzymic digestion technique

* From the Department of Biochemistry, Case Western Reserve University, Cleveland, Ohio. *Received June 19, 1969.* Please address all inquiries to University Hospitals, Cleveland, Ohio 44106. This research was supported by Grants AM-01825 and HD-00669 from the National Institutes of Health, U. S. Public Health Service.

[1] Abbreviations used are: CMC, 1-cyclohexyl-3-(2-morpholinyl-(4)-ethyl)carbodiimide metho-*p*-toluenesulfonate; EDC, 1-ethyl-3-dimethylaminopropylcarbodiimide; FDNB, 1-fluoro-2,4-dinitrobenzene.