Among the products obtained upon the digestion of Achilles tendon collagen of cattle by the extracellular enzymes of Clostridium histolyticum are two peptides which give characteristic spots on paper chromatograms and seem to be present in significant quantities (1, 2). The isolation and identification of these peptides were undertaken in an effort to gain additional information regarding the amino acid sequence of collagen.

Much of the present knowledge concerning the arrangement of the amino acid residues in collagen is derived from the study of peptides obtained by partial acidic or partial basic hydrolysis of collagen or gelatin (3-9). These studies have resulted in the isolation and identification of a large number of di- and tripeptides. However, in most cases the yields were quite low. An investigation of a tryptic digest of procollagen has given some information as to the distribution of glycine, the imino acids, the acidic amino acids, and the basic amino acids (10). These studies, together with structural information obtained from x-ray diffraction studies (11), have shown that the structure of collagen is much more complex than the simple repeating sequence structures originally proposed by Bergmann (12) and Astbury (13).

EXPERIMENTAL

Preparation of Enzyme — The enzyme used in this study was prepared from C. histolyticum strain CHT, from the stock collection of this laboratory. The stock culture medium, medium for enzyme production, and filtration of the culture have been described previously (14, 15). The enzyme was prepared by methanol precipitation as described by Ogle and Tytell (16) with the following modifications. Before the precipitation step, the filtrate was dialyzed in a rocking dialyzer against running cold tap water for 24 hours. The water extract of the precipitate was lyophilized and stored in a desiccator over anhydrous calcium sulfate (Drierite).

Preparation of Collagen — The collagen was prepared from Achilles tendons of cattle as described by Neuman (17).

Preparation of Collagen Digests — Samples of collagen, 500 mg. each, in 20 ml. of distilled water were incubated at 37° for 16 hours with five times the quantity of enzyme necessary to give complete solution of the collagen at that time. This amount of enzyme and incubation time produced a high yield of the peptides under consideration as indicated by paper chromatographic estimation. The collagen nitrogen of the digests was determined by a micro-Kjeldahl procedure (18). Fig. 1 shows the ascending chromatogram of the digest using as the solvent ethanol-water-ammonia (80:20:1). The spots were located by spraying with 0.2 per cent ninhydrin in water-saturated n-butanol followed by drying in an oven at 100° for 10 minutes. The peptides with which this study is concerned give intense yellow spots at \( R_F \) 0.53 and 0.43. These will be referred to as Peptide A and Peptide B respectively.

Isolation of Peptides — The collagen digest was passed through a 2.5 x 35 cm. column of Amberlite IRC-50 (Rohm and Haas) in the hydrogen cycle and eluted with distilled water at the rate of 20 ml. per hour until the effluent was ninhydrin negative. The eluate was evaporated at 40° to a volume of 20 ml. It contained all of Peptides A and B and showed a chromatographic pattern similar to the original digest except for materials with \( R_F \) values lower than that of glycine.

The basic materials remaining on the column accounted for more than 50 per cent of the dry weight of the original collagen. The column was washed with 0.1 N HCl and water in order to recover the basic fraction and regenerate the column. Treatment of this fraction with enzyme did not produce additional amounts of Peptides A and B.

Columns (2.5 x 100 cm.) of Dowex 50-X4 (200 to 400 mesh) in the sodium cycle, prepared by a procedure similar to that of Moore and Stein (19), were equilibrated with sodium phosphate buffer, 0.2 N in respect to Na at pH 3.1, and operated at room temperature with phosphate buffer of constant Na concentration.

The use of phosphate buffer instead of the citrate-acetate buffer (19) permitted the progress of the fractionation to be followed by direct, paper chromatography of the fractions without prior desalting, using the ethanol, water, and ammonia solvent. The paper chromatograms obtained with fractions from the columns differed from those from desalted fractions only in that they showed a purple salt spot at the origin.

The water eluate from the Amberlite IRC-50 column was adjusted to pH 2 and passed on to the column at a rate of 20 ml. per hour. The column was then eluted with pH 3.1 buffer at the same rate and 10 ml. fractions were collected until the material giving the purple spots on the paper chromatogram identified as glycine \( (R_F 0.33) \) and alanine \( (R_F 0.47) \) (Fig. 1) were recovered. This elution required 1.5 l. of buffer. The pH of the eluent was then gradually increased by running 0.2 M NaH2PO4 into a 500-ml. mixing chamber, similar to that described by Moore and Stein (19), previously filled with buffer, pH 3.1. When 1.5 to 2 l. of this solution had passed through the column, Peptide B was recovered. Following this zone, several unidentified peptides were recovered and then Peptide A was eluted, in most cases chromatographically pure.

Chromatograms of the first fractions containing Peptide B showed the purple spot at \( R_F 0.08 \) (Fig. 1). By repeated
fractionation on Dowex 50-X4 columns (2.5 × 50 cm.) using the same technique, Peptide B was recovered chromatographically pure.

Desalting Procedure—Although the desalting procedure of Drèze et al. (20) was successfully applied in this work to prevent salt interference in paper chromatography of various amino acid solutions, it was not effective in the separation of salts from Peptides A and B. A procedure based in part on ion exclusion (21, 22) was found applicable for this purpose.

A 1.7 × 7.5 cm. column of Dowex 50-X4 (200 to 400 mesh) in the sodium cycle was washed with 0.2 N NaOH and then with distilled water until free from alkali. 2 ml. of peptide solution adjusted to pH 7 were put on the column. The column was eluted with water and 1- to 2-ml. samples were collected every 5 minutes until 20 samples were obtained. The salts passed through the column before the peptides and were ordinarily separated from them by several salt-free tubes. The column was re-used after washing with water. This procedure has also been successfully applied to the separation of sodium phosphate and sodium citrate-aceate buffers from amino acids. Peptide solutions for identification of amino acids but not those used for structural studies were desalted.

Amino Acid Analysis—Peptide solutions were hydrolyzed in 6 N HCl in an autoclave in capped test tubes at 121° for 6 to 8 hours and repeatedly evaporated to dryness to remove most of the HCl.

Frequent amino acid identifications necessary during this work were accomplished by unidirectional ascending paper chromatography using two solvent systems, ethanol-water-ammonia (50:20:1) and water-saturated phenol at 25° in an atmosphere in equilibrium with 0.05 N ammonium hydroxide. For identification of spots, 0.2 per cent ninhydrin in water-saturated butanol was used, and the ninhydrin and isatin spray of Kolar and Roberts (23) was also found very useful because of its sensitivity to proline and hydroxyproline.

The amino acid contents of the most highly purified fractions of Peptides A and B were determined by ion exchange chromatography of the hydrolyzed peptides. Alanine and glycine were determined by the procedures of Moore and Stein (19, 24) except that sodium phosphate was used as the eluent of the Dowex 50-X4 columns 1 × 100 cm. The colors were read with an Evelyn colorimeter (Rubicon Company) using a No. 505 filter.

From fractions of the same column hydroxyproline was determined by Procedure A and proline by Procedure B of Pies et al. (25). The color formed with proline was found to vary with the buffer and its concentration. Consequently comparisons of color from all column fractions were made against the color produced simultaneoulsy from standard amino acid solutions prepared with the buffer contained in the fraction to be tested. Average recoveries and standard deviations (5 determinations) using samples containing the amino acids in approximately the same quantities as in the samples of hydrolysates analyzed were: alanine, 102 ± 6 per cent; glycine, 104 ± 5 per cent; hydroxyproline, 97 ± 2 per cent; and proline, 96 ± 2 per cent.

Estimation of Peptides—The total amounts of Peptides A and B were determined from the combined chromatographically pure fractions from columns obtained to obtain maximum yields of these peptides. The proline and hydroxyproline contents of hydrolysates of these fractions were determined by the specific methods of Troll and Lindsley (26) and Neuman and Logan (27), respectively. Although the peptide preparations were chromatographically pure, the amounts of Peptides A and B as calculated from the proline content accounted for 93 per cent of the total nitrogen of the preparations. The amount of Peptide B as calculated from the hydroxyproline content accounted for 60 per cent of the total nitrogen of the preparation.

The amount of Peptide A which was recovered from collagen was calculated from the proline value and that of Peptide B from the hydroxyproline value to give the minimum values indicated by the method.

Structural Studies—The N-terminal residues and the size of the peptides were determined by the DNP method using the preparation and hydrolysis of DNP peptides given by Cowgill and Pardee (28). The DNP amino acids recovered from hydrolysates of the DNP peptides were identified by comparative ascending paper chromatography using both ethanol-water-ammonia (50:20:1) and isopropanol-ethanol-water-ammonia (100:80:20:1) as solvents. The free amino acids were identified by the paper chromatographic methods described above.

The sequence of the peptide was established by the PTC method for stepwise degradation from the amino end (29). The quantities of the PTH derivatives were estimated using the molar extinction coefficients of the various derivatives. The identities of the PTH derivatives were established both from the identity of the free amino acids recovered upon acid hydrolysis of the PTH derivatives using the paper chromatographic methods described above and by direct paper chromatographic identification of the PTH derivatives using Solvent A and Solvent B (29) with known PTH derivatives prepared from the amino acids as controls. The structures of the peptides were also studied by applying the DNP method to that portion of each peptide which remained after removal of the N-terminal residue by the PTC method and by applying the C-terminal method of Turner and Schmerzler (30).

1 The abbreviations used are: DNP, dinitrophenyl; PTC, phenylisothiocyanate; and PTH, phenylthiohydantoin.
Two Tripeptides from Collagen

Table I

Results of amino acid analysis of Peptide A

Each sample of the hydrolysate analyzed contained $3.58 \times 10^6$ gram atoms of nitrogen. The results are expressed in micromoles and in per cent of the sampled nitrogen recovered.

<table>
<thead>
<tr>
<th>Alanine</th>
<th>Glycine</th>
<th>Proline</th>
<th>Total recovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>pmole</td>
<td>%</td>
<td>pmole</td>
<td>%</td>
</tr>
<tr>
<td>1.17</td>
<td>32.6</td>
<td>1.10</td>
<td>30.8</td>
</tr>
<tr>
<td>1.17</td>
<td>32.7</td>
<td>1.14</td>
<td>31.8</td>
</tr>
<tr>
<td>1.22</td>
<td>34.1</td>
<td>1.22</td>
<td>34.1</td>
</tr>
<tr>
<td>1.10</td>
<td>30.7</td>
<td>1.06</td>
<td>29.6</td>
</tr>
</tbody>
</table>

Table II

Results of amino acid analysis of Peptide B

The samples of hydrolysates analyzed in determinations (1) and (3) contained $3.58 \times 10^6$ gram atoms of nitrogen and that analyzed in determination (2) contained $3.54 \times 10^6$ gram atoms of nitrogen. The results are expressed in micromoles and in per cent of the sampled nitrogen recovered.

<table>
<thead>
<tr>
<th>Determination No.</th>
<th>Glycine</th>
<th>Hydroxyproline</th>
<th>Proline</th>
<th>Total recovery</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pmole</td>
<td>%</td>
<td>pmole</td>
<td>%</td>
</tr>
<tr>
<td>1</td>
<td>1.16</td>
<td>32.4</td>
<td>0.93</td>
<td>26.0</td>
</tr>
<tr>
<td>2</td>
<td>1.21</td>
<td>34.2</td>
<td>0.95</td>
<td>26.8</td>
</tr>
<tr>
<td>3</td>
<td>1.13</td>
<td>31.6</td>
<td>0.93</td>
<td>26.0</td>
</tr>
</tbody>
</table>

Table III

Results of stepwise degradation from amino end of peptides by phenylisothiocyanate method

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Position</th>
<th>Identity</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Theoretical</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>pmole</td>
</tr>
<tr>
<td>A</td>
<td>N-terminal</td>
<td>Glycine</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>Middle</td>
<td>Proline</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>C-terminal</td>
<td>Alanine</td>
<td>0.50</td>
</tr>
<tr>
<td>B</td>
<td>N-terminal</td>
<td>Glycine</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>Middle</td>
<td>Proline</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>C-terminal</td>
<td>Hydroxyproline</td>
<td>0.50</td>
</tr>
</tbody>
</table>

RESULTS AND DISCUSSION

Paper chromatographic analysis of the acid hydrolysates showed Peptide A to be composed of alanine, glycine, and proline and Peptide B to be composed of glycine, hydroxyproline, and proline.

The results of the amino acid analysis of Peptide A are given in Table I and those for Peptide B are given in Table II. These results showed each peptide to be composed of equimolar parts of its constituent amino acids. Control experiments performed to account for the somewhat low values for hydroxyproline in Peptide B failed to indicate a loss on hydrolysis. The hydroxyproline content of the hydrolysate was not found to decrease upon increasing the time of hydrolysis from 8 to 16 hours. Also, hydroxyproline was completely recovered from a mixture of glycine, hydroxyproline, proline, and buffer approximating the composition of the hydrolysates when this mixture was subjected to the same conditions of hydrolysis as the peptide.

The N-terminal studies by the DNP method showed each peptide to be a tripeptide with glycine in the N-terminal position. The only materials other than DNP glycine visible in the paper chromatograms of the DNP derivative recovered from the hydrolysate of each DNP peptide were small amounts of materials which appeared to be degradation products formed during hydrolysis. No spots corresponding to the DNP derivatives of the other constituent amino acids could be found in either case. No glycine was present among the free amino acids recovered from the hydrolysate of the DNP derivative of either peptide. The only free amino acids recovered from the hydrolysate of the DNP derivative of Peptide A were alanine and proline. Those isolated from the DNP derivative of Peptide B were hydroxyproline and proline.

The results of the structural studies by the PTC method for stepwise degradation of peptides from the amino end are given in Table III. Peptide A was proven to be glycylprolylalanine and Peptide B was proven to be glycylprolylhydroxyproline. The N-terminal data for each peptide by the PTC method confirm the results obtained from the N-terminal studies by the DNP method. The yields obtained were those expected from the method (29). The identity of the N-terminal and middle position of each peptide was established both by identifying the free amino acid liberated from the hydrolysis of the PTH derivative and by direct identification of the PTH derivative. The C-terminal alanine of Peptide A was not estimated. The C-terminal hydroxyproline of Peptide B was estimated by the method of Neuman and Logan (27). The identity of the C-terminal residue of each peptide was established both before and after hydrolysis of the portion of the peptide remaining after the first two residues were removed.

The middle and C-terminal positions of each peptide were confirmed by the results obtained from the application of the DNP method to the dipeptide which remained after the N-terminal glycine was removed from each peptide by the PTC method. Paper chromatograms of the DNP amino acid derivatives isolated from each peptide showed no spots corresponding to any of the possible amino acids except proline. Small amounts of breakdown products, probably resulting from destruction of the DNP proline during hydrolysis, were also present. A comparatively small amount of free proline was found in the hydrolysate of the DNP derivatives of both dipeptides, along with free alanine in that derived from Peptide A and free hydroxyproline in that derived from Peptide B. This proline was formed during the hydrolysis from DNP proline (7). The C-terminal position of Peptide A was also shown to be alanine by the method of Turner and Schmerzer (30). The method could not be applied to Peptide B since the reaction does not proceed when proline or hydroxyproline is in the C-terminal position (9).

The amounts of Peptides A and B isolated from two collagen digests, expressed in terms of per cent of collagen nitrogen recovered as the particular peptide, are for A (Gly. Pro. Ala) 5.1, 5.5 and for B (Gly. Pro. Hypro) 4.7, 5.8. The per cent of the total amount in collagen of each amino acid respectively represented by the peptides are A, glycine 7, proline 20, and alanine 23, and B, glycine 7, proline 20, and hydroxyproline 23 based...
on the nitrogen and amino acid content of the collagen used (17, 27).

The sequence Gly. Pro. Hypro has been postulated as an important order of amino acids in collagen (7) and it as well as Gly. Pro. Ala are possible sequences as indicated by x-ray analysis (11). Both of these peptides, Gly. Pro. Hypro and Gly. Pro. Ala, have been isolated from partial hydrolysates of gelatin (7) or steer hide collagen (9) in amounts which account for 0.4 to 1.3 per cent of the proline content of the protein. The dipeptide Gly. Pro has been isolated in significant amounts from partial hydrolysates of gelatin (7).

The results of this study indicate that the sequence Gly. Pro represents a minimum of 14 per cent of the glycine and 40 per cent of the proline content of the collagen. Of the amount of Gly. Pro accounted for by this study about half occurs as Gly. Pro. Hypro and half as Gly. Pro. Ala.

**SUMMARY**

1. Two tripeptides identified as glycylprolylalanine and glycylprolylhydroxyproline were isolated from an enzymatic digest of collagen. Together these peptides accounted for about 10 per cent of the nitrogen, 23 per cent of the alanine, 14 per cent of the glycine, 23 per cent of the hydroxyproline, and 40 per cent of the proline content of the collagen.  

2. A simple column procedure based in part on ion exclusion for the desalting of peptide and amino acid solutions is described.

**REFERENCES**

Two Tripeptides from an Enzymatic Digest of Collagen
Ralph E. Schrohenloher, James D. Ogle and Milan A. Logan