THE AMINO ACID SEQUENCE AT THE N TERMINUS OF PEPSSIN*

BY MARTIN B. WILLIAMSON AND JOHN M. PASSMANN†

(From the Department of Biochemistry, Stritch School of Medicine and Graduate School, Loyola University, Chicago, Illinois)

(Received for publication, January 27, 1956)

Attempts to apply the technique of Edman (1), for the stepwise degradation of peptides with the formation of phenylthiohydantoins (PTH), to the study of the N-terminal amino acid sequence of pepsin have resulted in only qualified success. One of the principal drawbacks of this technique arises from the fact that the reactions entailed are usually not complete. Thus, residual amounts of terminal amino acid, which are, theoretically at least, supposed to have been removed during the first degradative step, still remain attached to some of the protein molecules and appear as the products of subsequent degradations of the protein. This also occurs with the succeeding amino acids in the sequence. In the situation in which the succeeding amino acid residues yield different phenylthiohydantoins, an extensive portion of the sequence might be elucidated by this method. However, when succeeding or closely neighboring amino acid residues are the same, the number and arrangement of the identical residues cannot always be determined from qualitative data alone.

By means of the reactions for the formation of PTH and the reaction with dinitrofluorobenzene (DNFB), it was found that the first two amino acids at the N terminus of pepsin were leucine and glycine. These amino acids were followed by what appeared to be more than one aspartyl residue. In order to determine the number of aspartyl residues and the nature of the subsequent amino acids, a quantitative study of the phenylthiohydantoins derived from the amino acids in the N-terminal sequence of pepsin was undertaken.

EXPERIMENTAL

Crystalline pepsin of porcine origin, which contained 15.2 per cent of nitrogen (ash- and moisture-free), was used. By electrophoretic analysis at pH 4.3 (acetate buffer) and an ionic strength of 0.1, 99 per cent of the protein in the pepsin preparation was found to be the enzyme.

* Part of the data in this paper has been presented before the 3rd International Congress of Biochemistry in Brussels, August, 1955.
† Present address, Veterans Administration Research Hospital, Chicago, Illinois.
A weighed portion of pepsin, or degraded pepsin\(^1\) (50 to 100 mg.), was dissolved in 5 ml. of water with the aid of 2 to 3 drops of 1 N NaOH, and 5 ml. of \(\text{N-methylmorpholine}\), 0.5 ml. of phenyl isothiocyanate, and 3 ml. of acetone were added with cooling. The mixture was stirred at room temperature (22–26\(^\circ\)) for various lengths of time. The phenylthiocarbamyl (PTC) derivative of the protein was precipitated by the addition of 25 to 35 ml. of acetone, centrifuged, and washed several times with acetone to remove the last traces of phenyl isothiocyanate. The terminal amino acid derivative was hydrolyzed from the protein and the PTH formed by suspending the precipitate in 7 ml. of 35 per cent acetic acid which was 1.5 N with respect to HCl. The suspension was stirred at 28–30\(^\circ\). The effect of length of hydrolysis will be discussed later. After hydrolysis, the residual protein material was again precipitated and washed with acetone to recover the PTH. The acetone-dried powder was kept in a vacuum desiccator for several days before being weighed for the next degradative cycle.

To the acetone and washings were added 15 to 20 ml. of water and the acetone was removed under reduced pressure at temperatures below 30\(^\circ\). A slight turbidity which usually appeared was filtered off and discarded. The aqueous solution was extracted with ether and the ether was dried and removed under reduced pressure. The residue was suitably diluted with ethyl acetate for chromatography and for quantitative analysis. The concentration of the amino acid derivatives was determined in a Beckman spectrophotometer at 267 \(\text{m} \mu\). The products obtained from each degradation were chromatographed to identify the components extracted. The identity of the unknown compounds was established by comparing their \(R_F\) values with those of authentic phenylthiohydantoins on the same chromatogram. The solvent systems suggested by Sjöquist (2) and by Landmann et al. (3) were used to develop the chromatograms.

The ethyl acetate solutions were occasionally found to become tinted yellow. Fraenkel-Conrat and Harris (4) have indicated that this color arises from polymerization of the degradation products. The yellow color appeared in the solutions of the phenylthiohydantoins when left standing, even in the cold, for several days, when the solutions were made alkaline, or when the solutions were heated for relatively short periods of time. By working rapidly and keeping the solutions cool, the formation of the yellow color could be almost completely avoided.

**RESULTS AND DISCUSSION**

The effect of the length of hydrolysis was studied first. Aliquots of a single PTC-pepsin preparation, which had been formed by allowing phenyl

\(^1\) Degraded pepsin was obtained from the crystalline pepsin which had been treated one or more times according to the technique described.
isothiocyanate to react with pepsin for about 5 hours, were hydrolyzed for 0.5, 1.0, and 1.5 hours. It was found that there is essentially no difference in the yield of PTH after 1.0 and 1.5 hours of hydrolysis (Fig. 1). The subsequent data were obtained from preparations of PTC-pepsin which had been submitted to 1.0 hour of hydrolysis.

![Graph showing recovery of PTH derivative of leucine from PTC-pepsin in moles per mole of pepsin, after different times of hydrolysis.](image)

**Fig. 1.** Recovery of the PTH derivative of leucine from PTC-pepsin in moles per mole of pepsin, after different times of hydrolysis. The pepsin had been treated for 5 hours with phenyl isothiocyanate and hydrolyzed in 35 per cent acetic acid-1.5 N HCl.

### Table I

**Recovery of Phenylthiohydantoins from Pepsin Serially Treated with Phenyl Isothiocyanate and Hydrolysed**

<table>
<thead>
<tr>
<th>Degradation</th>
<th>No. of preparations</th>
<th>Treatment with phenyl isothiocyanate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0.5 hr.</td>
</tr>
<tr>
<td>1st</td>
<td>5</td>
<td>0.70</td>
</tr>
<tr>
<td>2nd</td>
<td>3</td>
<td>0.74</td>
</tr>
<tr>
<td>3rd</td>
<td>3</td>
<td>0.59</td>
</tr>
<tr>
<td>4th</td>
<td>2</td>
<td>0.50</td>
</tr>
<tr>
<td>5th</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

All the values are given in terms of moles of phenylthiohydantoin measured per mole of pepsin or degraded pepsin.

After treating pepsin, or degraded pepsin, with phenyl isothiocyanate for different lengths of time, it was found that the maximal yields of PTH were recovered after 4 hours of treatment. Table I shows the yields obtained from several consecutive degradations after varying the time of treatment. It did not seem profitable to try more extended treatment of the protein to increase the yield.

Only the product of the first degradation of intact pepsin can be expected to be uncontaminated with the derivatives of other amino acids. Subsequent degradations should yield mixtures containing the principal PTH,
154 AMINO ACID AT PEPSIN \( N \) TERMINUS

which would be the derivative of the amino acid whose position in the sequence corresponds to the number of degradations to which the protein has been subjected, and the residual phenylthiohydantoins, which would be the derivatives of the amino acids which had not gone through the complete degradative cycle during prior degradations. Chromatographic analysis of the products of the several degradations showed this to be true.

Assuming that the yield of the PTH derived from any amino acid remains constant under the same conditions of preparation, the data in Table I can be used to estimate the amount of the principal PTH and the residual phenylthiohydantoins obtained from each degradation. It then becomes possible to determine whether adjacent amino acids are the same.

The fraction of the total possible terminal amino acid \( (F_1) \) recovered as the PTH from the first degradation can be obtained directly from the data in Table I. The fraction of the total possible yield of the PTH derivative of the amino acid occupying the \( n \)th position in the sequence \( (F_n) \), obtained from the \( n \)th degradation, can be calculated from the following series:

\[
F_n = \frac{m_n}{M_{n-1}} - \frac{(1 - A_1 - \cdots - A_{n-1})F_1}{M_{n-1}} - \frac{(A_1 + \cdots A_{n-1} - B_2 - \cdots - B_{n-1})F_2}{M_{n-1}} - \cdots \frac{(L_{n-2} + L_{n-1} - M_{n-1})F_{n-1}}{M_{n-1}}
\]

where \( m_n \) is the actual amount of PTH measured as recovered from degradation \( n \), in moles of PTH per mole of protein; \( F_1, F_2, \ldots, F_{n-1}, \) etc., are the fractions of the total possible PTH derivative to be recovered from the principal amino acid during degradations 1, 2, \ldots, \( n - 1 \); \( A_1, A_2, \ldots, A_{n-1} \) are the amounts of the terminal amino acid calculated to be recovered as the PTH derivative, in moles per mole of protein, from degradations 1, 2, \ldots, \( n - 1 \); and \( A, B, \ldots, L, \) and \( M \) are the designations for the amino acids in the order of their appearance in the sequence.

The amount of PTH derivative from an amino acid which may be expected to be recovered from degradations subsequent to the one in which it was the principal amino acid may be calculated from the following equation:

\[
M_n = F_n(\Sigma L_{n-1} - \Sigma M_{n-1})
\]

where \( F_n \) is obtained from the previous equation, \( \Sigma L_{n-1} \) is the sum of the recoveries of the derivative of the immediately preceding amino acid in the sequence in all prior degradations, in moles per mole of protein, and \( \Sigma M_{n-1} \) is the sum of the recoveries of the PTH derivative of the amino acid in question in all prior degradations, in moles per mole of protein. Table II shows the recoveries expected to be obtained for the first four amino acids in the sequence, resulting from six consecutive degradations,
based on the data from the preparations which had been treated for 4 hours with phenyl isothiocyanate.

The third amino acid in the terminal sequence of pepsin must be aspartic acid, since the PTH derivative of this amino acid first appears in the products after the third degradation. The aspartic acid derivative is the only product of the fourth degradation, and appreciably more PTH is found after this degradation than is calculated to arise from the residual hydantoins from previous degradations. It follows that the fourth position in the sequence is also occupied by an aspartic acid residue. The fifth degradation again yielded only the aspartic acid derivative. However, the

<table>
<thead>
<tr>
<th>Degradation</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>5th</th>
<th>6th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total PTH measured</td>
<td>0.84</td>
<td>0.88</td>
<td>0.75</td>
<td>0.72</td>
<td>0.34</td>
<td>PTH derivatives of amino acids identified by chromatography</td>
</tr>
<tr>
<td>Terminal amino acid</td>
<td>0.84</td>
<td>0.14</td>
<td>0.20</td>
<td>0.05</td>
<td>0.11</td>
<td>Leu</td>
</tr>
<tr>
<td>2nd amino acid</td>
<td>0.74</td>
<td>0.20</td>
<td>0.55</td>
<td>0.29</td>
<td>0.17</td>
<td>Leu</td>
</tr>
<tr>
<td>3rd amino acid</td>
<td>0.29</td>
<td>0.38</td>
<td>0.32</td>
<td>0.32</td>
<td>0.32</td>
<td>Gly</td>
</tr>
<tr>
<td>4th amino acid</td>
<td>0.38</td>
<td>0.32</td>
<td>0.32</td>
<td>0.32</td>
<td>0.32</td>
<td>Asp</td>
</tr>
<tr>
<td>5th amino acid</td>
<td>0.38</td>
<td>0.32</td>
<td>0.32</td>
<td>0.32</td>
<td>0.32</td>
<td>Gly</td>
</tr>
<tr>
<td>6th amino acid</td>
<td>0.38</td>
<td>0.32</td>
<td>0.32</td>
<td>0.32</td>
<td>0.32</td>
<td>Asp</td>
</tr>
</tbody>
</table>

All the values are given in terms of moles of PTH per mole of pepsin or degraded pepsin.

total yield of PTH is less than that calculated for the residual phenylthiohydantoins from earlier degradations. It must then be concluded that the fifth position in the sequence cannot be occupied by an aspartyl residue.

The method described for preparing the phenylthiohydantoins does not yield a PTH derivative of the amino acid which occupies the fifth position in the terminal sequence of pepsin. It has been reported that PTH derivatives of the basic amino acids cannot be extracted from an acid solution with organic solvents, but may be extracted from an alkaline solution (4). When extraction from an alkaline solution was attempted, extensive polymerization occurred so that neither could quantitative measurements of the PTH be made, nor definite identification of the products be achieved by chromatography.

The identity of the fifth amino acid in the sequence was established by
treat pepsin, which had been degraded four times by the technique described above, with DNFB in the presence of excess NaHCO₃, and by hydrolyzing the resulting addition product of pepsin in 5 N HCl for 12 hours. The hydrolysate was extracted with ether and then with n-butanol. The dinitrophenyl derivatives in the extracts were identified by chromatography in several solvent systems (5, 6). A composite chromatogram of the extracts is shown in Fig. 2. These data support the view that the fifth amino acid in the sequence is histidine.

![Composite chromatogram of the dinitrophenyl derivatives of pepsin which had been degraded four times prior to the formation of the dinitrophenylamino acids.](http://www.jbc.org/)

**Fig. 2.** Composite chromatogram of the dinitrophenyl derivatives of pepsin which had been degraded four times prior to the formation of the dinitrophenylamino acids. Solvent system: pyridine, isoamyl alcohol, NH₄OH. Column A, authentic amino acid, 1, DNP-Asp; 2, DNP-Arg; 3, dinitrophenol. Column B, DNP derivatives of histidine. Column C, ether extract of the hydrolysate of DNP-degraded pepsin, 1, DNP-Asp; 2, dinitrophenol. Column D, n-butanol extract of the hydrolysate of DNP-degraded pepsin, 1 an unidentified DNP derivative, probably ε-mono-DNP-Lys; 2, 3, DNP-His.

Aliquots of the ether extract were chromatographed and the dinitrophenylaspartic acid was extracted from the chromatogram. Measurement of the absorption at 350 mµ indicated that 0.56 mole of aspartic acid per mole of pepsin had remained attached to the protein after the fourth degradation. Some of this residual amino acid probably arose from the third, as well as the fourth, amino acid in the sequence.

Qualitative identification of the products formed during the sixth and seventh degradations indicated the presence of both aspartic and glutamic acid derivatives. Since the residual PTH derivative of aspartic acid would be expected to appear in the products of even the sixth degradation (see Table II), it seems probable that the sixth amino acid in the sequence is glutamic acid.
SUMMARY

The N-terminal sequence of pepsin was studied with the use of the technique for the stepwise degradation of peptides with the formation of phenylthiohydantoins and the technique for the labeling of the terminal amino acid of a protein with dinitrofluorobenzene. The terminal sequence was found to be Leu-Gly-Asp-Asp-His-Glu.

BIBLIOGRAPHY

THE AMINO ACID SEQUENCE AT THE N TERMINUS OF PEPSIN
Martin B. Williamson and John M. Passmann


Access the most updated version of this article at http://www.jbc.org/content/222/1/151.citation

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/222/1/151.citation.full.html#ref-list-1