The Amino Acid Sequence of the A Protein (α Subunit) of the Tryptophan Synthetase of Escherichia coli

I. Tryptic Peptides*

(Received for publication, February 20, 1967)

JOHN R. GUEST,† BRUCE C. CARLTON,§ AND CHARLES YANOFSKY

From the Department of Biological Sciences, Stanford University, Stanford, California 94305

SUMMARY

Tryptic digestion of the A protein (α subunit) of the tryptophan synthetase of Escherichia coli and separation of tryptic peptides by column and paper chromatography yielded 22 nonoverlapping peptides of the total of 25 anticipated on the basis of total amino acid composition of the A protein. The amino acid composition and sequence of each purified peptide were determined. These peptides account for 213 of the total of 267 amino acid residues in the A protein.

Studies on the primary structure of the A protein (α subunit) of the tryptophan synthetase of Escherichia coli have been directed toward a detailed examination of various aspects of the relationship between the A protein and the corresponding structural gene, the A gene. Previous studies on the structure of the protein have described the amino acid composition (1), the amino- and carboxyl-terminal sequences (2, 3), the peptide patterns obtained by proteolytic digestion of the protein (4), and the colinear relationship between the fine structure map of the A gene and the amino acid sequence of the protein (5). Residues believed to be situated at or near the active center of the protein have also been described (6), and the sequences of amino acids surrounding the three reactive sulfhydryl groups have been determined (7).

In this paper the fractionation, purification, composition, and amino acid sequences of all but two of the peptides in tryptic digests of the A protein are presented. The isolation and sequence of the two remaining peptides are described in the following paper (8).

METHODS AND MATERIALS

Purification and Digestion of A Protein—The purification of wild type A protein was carried out by the procedure of Henning et al. (1). Fractions of A protein with maximum specific enzyme activity (4000 to 5000 units per mg of protein) or crystalline material were used for structure studies. Preparations of this type are homogeneous in the ultracentrifuge but occasionally have a trace contaminant which is detectable by electrophoresis on polyacrylamide gel. Digestion with trypsin was performed subsequent to denaturation in urea by a modification (4) of the procedure described by Rodbell and Fredrickson (9). Batches of 100 to 250 mg of protein were generally used, and digestion was carried out for 4 hours.

Isolation of Tryptic Peptides—The digests were diluted 6-fold with distilled water and applied to a column of Dowex 50-X2 (50 to 100 mesh, H+ form, 3 × 50 cm). Salts and urea were removed by washing with 1 liter of water, and the digestion products were then eluted with 4 N NH₄OH and dried over NaOH and H₂SO₄. Generally 90 to 100% of the starting material was recovered, on a weight basis. The dried material was dissolved in water, and the fine precipitate which appeared on adjustment to pH 4.0 with 1.0 N acetic acid was removed from the soluble digestion products by centrifugation. This precipitate amounted to approximately 30% of the total weight of the protein and was designated the trypsin-resistant core. When the A protein was oxidized with performic acid (10) prior to digestion, the trypsin-resistant core represented a smaller proportion of the starting material and amounted to approximately 10%. In later work, purification of the peptides was often facilitated by the inclusion of an initial separation into two groups containing large or small peptides. This was accomplished by chromatography on Sephadex G-25 (11). According to this procedure, the larger peptides plus trypsin-resistant core were present in Sephadex Fractions A, B, and C, while smaller peptides (plus salt and urea) were recovered from Fractions D and E.

Tryptic peptides were fractionated on Dowex 1-X2 resin by a
modified of the procedure of Rudloff and Braunitzer (12). The soluble peptides were dissolved in 3 to 10 ml of 1% collidine-acetate buffer, pH 8.7 (1%, 2, 4, 6-collidine by volume, adjusted to pH 8.7 with acetic acid), and adjusted to pH 9.0 with 1 N NaOH. The peptide solution was applied to a column of Dowex 1-X2 resin previously equilibrated with deaerated 1% collidine-acetate buffer (0.9 × 60 or 1.8 × 60 cm, depending upon the amount of protein digested). All peptide separations were carried out with the stream-splitting accessory of the Beckman automatic amino acid analyzer. The stream divider pump was adjusted to remove 5 to 12.5% of the total volume of the column effluent for reaction with ninhydrin. Fractions of approximately 20 ml were collected at 10-min intervals, and the flow rate was maintained at 120 ml per hour in all cases. After elution of the initial basic components with the collidine-acetate solution (pH 8.7), a gradient elution system was applied with 500 ml or 1 liter of 1% collidine-acetate buffer in a constant volume mixing flask fed from a reservoir containing successively 0.1, 0.3, 1.0, and 5.0 N deaerated solutions of acetic acid. With protein digests (but not peptide digests), it was found desirable to delay the mixing of ninhydrin with the column effluent to prevent blocking of the reaction coil of the amino acid analyzer when high concentrations of basic components were eluted just after development commenced. Consequently, the ascending portion of the first peak in the elution profile (Fig. 1) represents the entry of ninhydrin and not the peptide front.

Purification and Analysis of Peptides—Individual peak fractions of ninhydrin-reactive material were concentrated on a rotary evaporator at 50° under reduced pressure, and an aliquot from each peak was removed for electrophoretic and chromatographic examination. Samples were applied in bands ⅛ × ⅜ inch on each of two sheets of Whatman No. 3MM filter paper. One of the sheets was subjected to electrophoresis at pH 3.7, and the other to chromatography in sec-butyl alcohol-formic acid-water (7:0.9:2.1), by the standard methods used throughout this work and described previously by Helinski and Yanofsky (4). Those fractions having more than one component were further purified by the most appropriate procedure, either by paper electrophoresis, by paper chromatography, or a combination of the two if necessary. For this purpose the concentrated fractions were applied to Whatman No. 3MM paper sheets 18½ × 22¼ inches as bands 3 to 5 × ⅜ inches along a line 4 inches from the narrow edge of the paper (for electrophoresis) or 3¼ inches from the wide edge (for descending chromatography). Loading was between 0.4 and 2.0 μmole of peptide per inch, and problems arising from the presence of salts (collidine-acetate), especially in the early fractions, were not encountered. The conditions for chromatography and electrophoresis were standardized throughout this work by using marker spots of methyl red (migrating 13 inches in approximately 17 hours) and quinine sulfate (migrating 18 inches in approximately 2 hours at 2400 volts), respectively. The same procedures were followed successively for preparing the two-dimensional peptide patterns (4). The method used for particular fractions can be determined simply by examining the Dowex 1-X2 elution profiles (Fig. 1) and the relative chromatographic and electrophoretic mobilities of the component peptides on the two-dimensional peptide pattern (Fig. 2). The peptides were located by staining guide strips with 0.5% ninhydrin in methanol or with tert-butyl hypochlorite reagent (13), and the peptide bands were eluted with 10 to 25 ml of either 10% acetic acid or 0.5 N NaOH. The eluates were evaporated to dryness over NaOH pellets and concentrated H₂SO₄ in a vacuum desiccator and then dissolved in water. Appropriate aliquots of the peptides were hydrolyzed at 105° in 5.7 N HCl for 40 hours in sealed, evacuated tubes. After removal of the HCl in a vacuum desiccator, the residue was taken up in sodium acetate sample buffer at pH 2.2 and analyzed on the 10- and 50-cm columns of the Beckman amino acid analyzer.

Identification of Peptides on Peptide Pattern—Each tryptic peptide was identified on the tryptic peptide pattern of the A protein (4) in the following manner. The peptide (approximately 0.05 μmole) was applied to a sheet of Whatman No. 3MM paper along with 0.015 μmole of a complete tryptic digest. The mixture was then subjected to the two-dimensional chromatographic and electrophoretic procedure. The paper was stained with ninhydrin, and the peptide in question appeared as a dark spot on a faint background of the entire tryptic peptide pattern of the A protein. By this method, each peptide could be located on the peptide pattern and its analysis compared with the information obtained by the use of the specific staining reactions for arginine, histidine, and tyrosine referred to previously (4).

Treatment of Tryptic Peptides with Other Proteolytic Enzymes—Fragments of tryptic peptides were obtained by digestion with chymotrypsin, papain, and occasionally pepsin and Nagarse, according to procedures described previously (7, 11). These fragments were purified by paper electrophoresis at pH 1.5 (10% formic acid), pH 3.7 (pyridine-acetic acid-water, 1:10:289), or pH 6.5 (pyridine-acetic acid-water, 30:1.2:270), by paper chromatography with the solvent system used above, or by a combination of these procedures (7). In a few cases digests were fractionated by chromatography on Dowex 1-X2 by the method used for the isolation of tryptic peptides. Digests with leucine aminopeptidase and carboxypeptidases A and B were performed as described previously (11).

Edman Degradation of Peptides—The method used originally (2) was later replaced by the improved method (Procedure 2) of Konigsberg and Hill (14). The phenylthiohydantoin derivative liberated after each cycle was sometimes extracted with ether and identified by chromatography with the solvents of Sjöquist (15) and Edman and Sjöquist (16). The identification of amino-terminal residues of peptide fragments isolated directly from ninhydrin-treated peptide patterns was based occasionally on the lower proportion of the amino-terminal residue detected in hydrolysates of the treated peptide.

Characterization of Peptides from Trypsin-resistant Core—The peptides which were believed to be segments of, or associated with, the trypsin-resistant core were defined by several criteria. Peptides which were recovered from digests of performic acid-oxidized protein but were not found in digests of unoxidized protein were regarded as components of the trypsin-resistant core. The same peptides could be purified from samples of isolated core by oxidation and further treatment with trypsin. The acid-precipitable (pH 4), trypsin-resistant residue found after digestion of either the oxidized A protein or the oxidized trypsin-resistant core from the untreated protein contained small amounts of the large peptide TP15, which will be discussed in the following paper (8).

Materials—Crystalline trypsin was obtained as a special five times crystallized preparation (Lot TR 706) from Worthing-
ton. The sources of other enzymes and reagents have been described elsewhere (7, 11).

RESULTS

Isolation of Tryptic Peptides

A representative elution profile of the tryptic peptides of the A protein is shown in Fig. 1. Depending upon the degree of resolution achieved both in the separation into large and small peptide groups on Sephadex and during chromatography on Dowex 1-X2, approximately 20 ninhydrin-reactive peaks were obtained. Although several of the peak fractions yielded single peptides, many of the peaks were found to contain either two or more major peptide components, or they were contaminated with small amounts of other peptides. Purification of mixtures of peptides was generally accomplished by electrophoresis at pH 3.7 or by chromatography in the see-butyl alcohol-formic acid-water system (4). These procedures also served to desalt the peptides.

The amino acid composition of each of the peptides obtained from tryptic digests of performic acid-oxidized A protein is presented in Table I. The location of these peptides is indicated on the elution profile of Fig. 1 and on the peptide pattern shown in Fig. 2. Tryptic Peptide TP15 was somewhat insoluble at pH 4.0 and was often found as a component of the trypsin-resistant core. When native protein was digested with trypsin, several peptides were not recovered from the soluble fraction; these were TP22, TP24, TP25, TP26, and TP29, and, together with TP15, they were presumed to make up the larger trypsin-resistant residue obtained from unoxidized protein. In addition, the peptides containing methionine (TP7, TP9, and TP17) were eluted slightly earlier than the corresponding oxidized peptides. A total of 29 peptides were obtained, including a free serine (TP1), a free lysine (TP21), four peptides (TP3, TP10, TP27, and TP28) later found to contain trypsin-susceptible bonds which had not been cleaved, and one chymotryptic fragment (TP24) of a tryptic peptide (TP8).

The final recoveries of individual peptides varied from an estimated 10% to 90%, based on the amount of protein digested. Individual yields will not be presented because we feel that this information would be somewhat misleading. With most peptides several purification steps were required, and fractions with the least contamination with other peptides, regardless of yield, were generally selected for further purification at each step. One of the major factors that was responsible for poor final recoveries in certain cases was the tendency for some peptides to adhere to paper during purification. This was particularly noticeable for the larger peptides.

Sequences of Amino Acids in Tryptic Peptides

The peptides will generally be discussed in an arbitrarily assigned numerical order. The sequential order of the peptides in the protein is indicated in Table II. The derived sequence of amino acids is followed by the total number of amino acids, given in parentheses. The composition of each peptide is presented as molar ratios of the component amino acids, calculated by the method of Hirs, Moore, and Stein (17). In certain instances, when the content of 1 or more residues was obviously low, these residues were not included in the calculation of the average molar equivalent for 1 residue, e.g., in analyses in residual peptides in the Edman degradations. This treatment was used with the methionine and tyrosine residues of a few peptides when excessive destruction during acid hydrolysis was suspected (it was not used for serine or threonine, for which recoveries of 85 to 95% were normally obtained). This treatment was also used for isoleucine in peptides containing the sequence Ile-Ile, and sometimes when a residue was suspected as being amino-terminal in a peptide which had been treated with ninhydrin. The presence of residues in amounts less than 0.15 mole per mole of peptide was ignored and not reported. The composition is followed immediately by the stepwise Edman degradation and, when recorded, the figure in parentheses represents the recovery at each stage as a percentage of the amount of peptide used in that step. The amino acid thought to be lost at each stage is indicated by boldface type, and, when investigated, the solvent system and phenylthiohydantoin derivative identified by paper chromatography are included. The results of digestion with leucine aminopeptidase and with carboxypeptidase A plus B are presented as moles of amino acid released per mole of peptide digested. A similar nomenclature is used for peptide fragments obtained by further proteolytic digestion or by partial acid hydrolysis. The specificities of trypsin and chymotrypsin were used extensively for identifying the carboxyl-terminal residues, and very often the penultimate residue was identified by difference.

Peptide TP1 Details are given following TP28.

Peptide TP2 (Residues 199 and 200): Leu-Lys (2 Residues)

Like all the peptides in the initial peak (Fig. 1B), this peptide was purified by paper electrophoresis at pH 3.7 followed by paper chromatography by the standard procedure.

Composition: Leu, 0.91; Lys, 1.00.

Edman degradation: Stage 1 (83%): Leu, 0.21; Lys, 1.00.

The specificity of trypsin requires that lysine be the carboxyl-
Amino acid analyses are given in the text.

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>TP-1</th>
<th>TP-2</th>
<th>(TP-3)</th>
<th>TP-4</th>
<th>TP-5</th>
<th>TP-6</th>
<th>TP-7</th>
<th>TP-8</th>
<th>TP-9</th>
<th>(TP-10)</th>
<th>TP-11</th>
<th>TP-12</th>
<th>TP-13</th>
<th>TP-14</th>
<th>TP-15</th>
<th>TP-16</th>
<th>TP-17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysine</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Histidine</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arginine</td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cysteic acid</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threonine</td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serine</td>
<td>1</td>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proline</td>
<td>3</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycine</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alanine</td>
<td>3</td>
<td>2</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valine</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methionine</td>
<td>1</td>
<td></td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isoleucine</td>
<td></td>
<td>1</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leucine</td>
<td></td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyrosine</td>
<td>1</td>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>2</td>
<td>2 (22)</td>
<td>9</td>
<td>11</td>
<td>18</td>
<td>7</td>
<td>8</td>
<td>3 (21)</td>
<td>7</td>
<td>20</td>
<td>20</td>
<td>9</td>
<td>34</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>TP-18</th>
<th>TP-19</th>
<th>TP-20</th>
<th>TP-21</th>
<th>TP-22</th>
<th>TP-23</th>
<th>TP-24</th>
<th>TP-25</th>
<th>TP-26</th>
<th>(TP-27)</th>
<th>(TP-28)</th>
<th>(TP-29)</th>
<th>Total</th>
<th>Protein composition</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysine</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>13</td>
<td>14</td>
<td>(14.2)</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Histidine</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>4 (4.3)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arginine</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>12</td>
<td>(11.7)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cysteic acid</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>22</td>
<td>24</td>
<td>(23.6)</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>10</td>
<td>(9.7)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threonine</td>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>12</td>
<td>(11.6)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serine</td>
<td></td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>29</td>
<td>31</td>
<td>(31.4)</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>1</td>
<td></td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>17</td>
<td>18</td>
<td>(18.3)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Proline</td>
<td></td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>19</td>
<td>20</td>
<td>(20.2)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycine</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>14</td>
<td>16</td>
<td>(16.0)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alanine</td>
<td></td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>40</td>
<td>41</td>
<td>(40.9)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valine</td>
<td></td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>17</td>
<td>18</td>
<td>(18.3)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methionine</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>(4.9)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isoleucine</td>
<td></td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>27</td>
<td>28</td>
<td>(27.9)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leucine</td>
<td></td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>27</td>
<td>28</td>
<td>(28.9)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyrosine</td>
<td></td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>12</td>
<td>(12.0)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>14</td>
<td>14</td>
<td>(14.0)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Parentheses denote overlapping tryptic peptides, the compositions of which are not included in the total number of amino acid residues recovered.

* See the accompanying paper (8) for a description of the isolation and characterization of these peptides.

terminal residue, and a single step of the Edman degradation confirmed the sequence indicated.

**Peptide TP3 (Residues 199 through 220):** Leu-Lys-Glu-Tyr-Asn-Asp-Ala-Ala-Pro-Leu-Gln-Gly-Phe-Gly-Ile-Ser-Ala-Pro-(Asx, Glx)-Val-Lys (22 Residues)

Composition: Asp, 1.97; Ser, 0.93; Glu, 3.15; Pro, 2.87; Gly, 2.13; Ala, 2.95; Val, 1.05; Ile, 0.96; Leu, 2.08; Tyr, 0.83; Phe, 1.02, Lys, 2.05.

Edman degradation: Stage 1: Asp, 2.00; Ser, 0.94; Glu, 3.06; Pro, 2.94; Gly, 2.15; Ala, 2.96; Val, 0.99; Ile, 0.97; Leu, 1.24; Tyr, 0.50; Phe, 1.04; Lys, 0.69. (D; PTH-Leu.)

Carboxypeptidase: Val, 0.96; Lys, 1.00.

The abbreviations used are: Asx and Glx, aspartic acid or asparagine and glutamic acid or glutamine, respectively, when it is not known whether the residue occurs as the free acid or its amide; PTH-, phenylthiohydantoin-; PTC-, phenylthiocarbamyl-.
Leucine aminopeptidase: Traces of leucine and lysine.

Chymotryptic Digestion—Digestion with chymotrypsin released three fragments which were readily separated by paper electrophoresis at pH 3.7.

TP3-C1: Asp-Ala-Ala-Pro-Leu-Gln-Gly-Phe—Based on its composition, electrophoretic and chromatographic behavior, a characteristic yellow-gray staining reaction with ninhydrin, and its origin (from TP3), this peptide fragment was identified as one which had been sequenced previously (18).

TP3-C2: Gly-Ile-Ser-Ala-Pro (Asx, Glx, Val) Lys—Composition: Asp, 1.04; Ser, 0.96; Gln, 1.12; Pro, 1.01; Gly, 1.00; Ala, 1.09; Val, 0.82; Ile, 0.97; Lys, 1.02.

Edman degradation: Stage 1 (71%): Asp, 1.06; Ser, 0.86; Gln, 1.09; Pro, 1.10; Gly, 0.21; Ala, 1.06; Val, 0.91; Ile, 0.88; Lys, not determined. (D; F; PTH-Gly.)

Stage 2 (75%): Asp, 1.03; Ser, 0.76; Gln, 1.08; Pro, 0.98; Gly, 0.19; Ala, 0.97; Val, 0.92; Ile, 0.17; Lys, not determined. (F; PTH-Ile or Val.)

Stage 3 (97%): Asp, 0.98; Ser, 0.29; Gln, 1.08; Pro, 1.00; Ala, 0.04; Val, 0.08; Lys, not determined. (F; PTH-Ser.)

Stage 4 (75%): Asp, 1.00; Ser, 0.18; Gln, 1.10; Pro, 0.93; Ala, 0.49; Val, 0.97; Lys, not determined. (F; PTH-Ala.)

Stage 5 (74%): Asp, 1.03; Gly, 1.12; Pro, 0.47; Ala, 0.35; Val, 0.58; Lys, not determined.

TPS-C3: Leu-Lys-Glu-Tyr—Composition: Glu, 1.03; Leu, 1.00; Tyr, 0.92; Lys, 1.05.

Edman degradation: Stage 1 (80%): Glu, 1.00; Leu, 0.04; Tyr, 1.00; Lys, 0.22. (A, D, F; e-PTC-PTH-Lys, PTH-Leu.)

Stage 2 (82%): Glu, 1.20; Tyr, 0.80; Lys, 0.14. (D; no phenylthiohydantoin observed.)

Stage 3 (86%): Glu, 0.57; Tyr, 1.00. (F; PTH-Glu.)

Carboxypeptidase: Tyr, 0.50; Leu, 0.76; Lys, 0.72; Glu, 0.55.

Leucine aminopeptidase: Glu, 0.65; Leu, 0.58; Tyr, 0.58; Lys, 0.43.

Studies of mutationally altered A proteins have revealed amino acid substitutions involving the 2 glycine residues of TP3 (5, 19). The order of the chymotryptic fragments of TP3 has been deduced as (TP3-C3)-(TP3-C1)-(TP3-C2), based on the partial sequence for the remaining 4 residues (Asx, Glx) Val-Lys, come from release of valine and lysine from TP3 by carboxypeptidases A plus B, and the specificity of trypsin, which requires the lysyl residue to be carboxyl-terminal. The discovery of TP13, which has the composition of TP3 minus 1 residue each of valine and lysine, indicates that TP3 is an overlapping tryptic peptide linking TP2 and TP13, and that it occurs because of the poor susceptibility of the Lys-Glu bond to trypsin digestion. The complete sequence of TP13 and TP3-C2 was determined by analogy with the corresponding Peptides TP13 and TP13-C2 (see following peptide).

Peptide TP4—Details are given following TP13.

Peptide TP13 (Residues 201 through 220): Glu-Tyr-(Asx, Glx, Pro, Gly, Ala, Leu, Phe)-(Ser, Gly, Ile)-Ala-Pro-Asp-Gln-Val-Lys (20 Residues)

Peptide TP13 and Peptide TP5 were not completely separated by chromatography on Dowex 1-X2 (Fig. 1A), but each peptide was isolated in pure form from the combined peak fractions by paper electrophoresis at pH 3.7.
**Table II**

**Amino acid sequences of tryptic peptides of the A protein of tryptophan synthetase**

TP24 (Residues 175 through 178) is omitted from this list because it is probably a chymotryptic fragment of TP8.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Residue positions in protein</th>
<th>Amino acid sequence (and number of residues)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP9</td>
<td>1-3</td>
<td>Met-Glu-Arg (3)</td>
</tr>
<tr>
<td>TP14</td>
<td>4-12</td>
<td>Tyr-Glu-Ser-Leu-Phe-Ala-Glu-Leu-Lys (9)</td>
</tr>
<tr>
<td>TP19</td>
<td>13-14</td>
<td>Glu-Arg (2)</td>
</tr>
<tr>
<td>TP27</td>
<td>15</td>
<td>Glu Arg Lys (3)</td>
</tr>
<tr>
<td>TP21</td>
<td>15</td>
<td>Lys (1)</td>
</tr>
<tr>
<td>TP10</td>
<td>15-35</td>
<td>Lys-Glu-Gly-Ala-Phe-Val-Pro-Phe-Val-Thr-Leu-Gly-Asp-Pro-Gly-Ile-Glu-Gln-Ser-Leu-Lys (21)</td>
</tr>
<tr>
<td>TP12</td>
<td>16-35</td>
<td>Glu-Gly-Ala-Phe-Val-Pro-Phe-Val-Thr-Leu-Gly-Asp-Pro-Gly-Ile-Glu-Gln-Ser-Leu-Lys (20)</td>
</tr>
<tr>
<td>TP15</td>
<td>36-69</td>
<td>Ile-Asp-Thr-Leu-Ile-Glu-Ala-Gly-Ala-Leu-Glu-Leu-Gly-Ile-Pro-Ser-Asp-Pro-Leu-Ala-Ala-Gly-Pro-Thr-Ile-Gln-Asn-Ala-Thr-Leu-Arg (34)</td>
</tr>
<tr>
<td>TP25</td>
<td>70-88</td>
<td>Ala-Phe-Ala-Ala-Glu-Val-Thr-Pro-Ala-Ala-Gln-Cys-Phe-Glu-Met-Leu-ALA-Leu-Ile-Arg (19)</td>
</tr>
<tr>
<td>TP26</td>
<td>89-108</td>
<td>Gln-Lys-His-Pro-Thr-Ile-Pro-Ile-Gly-Leu-Leu-Met-Tyr-Ala-Asn-Leu-Val-Phe-Asn-Lys (20)</td>
</tr>
<tr>
<td>TP23</td>
<td>109-119</td>
<td>Gly-Ile-Asp-Glu-Phe-Tyr-Ala-Gln-Cys-Glu-Lys (11)</td>
</tr>
<tr>
<td>TP22</td>
<td>120-139</td>
<td>Val-Gly-Val-Asp-Ser-Val-Leu-Ala-Asp-Val-Pro-Val-Gln-Glu-Ser-Ala-Pro-Phe-Arg (20)</td>
</tr>
<tr>
<td>TP16</td>
<td>140-144</td>
<td>Ala-Glu-Leu-Leu (5)</td>
</tr>
<tr>
<td>TP29</td>
<td>145-163</td>
<td>His-Asn-Val-Ala-Pro-Ile-Phe-Cys-Pro-Asn-Ala-Ala-Asp-Asp-Leu-Leu-Ang (19)</td>
</tr>
<tr>
<td>TP11</td>
<td>164-170</td>
<td>Gln-Ile-Ala-Asp-Tyr-Gly-Ary (7)</td>
</tr>
<tr>
<td>TP18</td>
<td>171-178</td>
<td>G1y-Tyr-Thr-Tyr-Leu-Leu-Ser-Leu-Ang (8)</td>
</tr>
<tr>
<td>TP17</td>
<td>178</td>
<td>Ala-Glu-Val-Thr-Glu-Ala-Glu-Arg (9)</td>
</tr>
<tr>
<td>TP18</td>
<td>188-198</td>
<td>Ala-Ala-Pro-Leu-Asn-His-Leu-Leu-Val-Lys (11)</td>
</tr>
<tr>
<td>TP2</td>
<td>199-200</td>
<td>Leu-Lys (2)</td>
</tr>
<tr>
<td>TP3</td>
<td>199-220</td>
<td>Leu-Lys-Glu-Tyr-Asn-Ala-Ala-Pro-Pro-Leu-Gly-Gln-Glu-Phe-Ile-Ala-Pro-Asp-Gln-Val-Lys (12)</td>
</tr>
<tr>
<td>TP13</td>
<td>201-220</td>
<td>Glu-Tyr-Asn-Ala-Ala-Pro-Pro-Leu-Gly-Gln-Glu-Phe-Ile-Ser-Ala-Pro-Asp-Gln-Val-Lys (20)</td>
</tr>
<tr>
<td>TP6</td>
<td>221-238</td>
<td>Ala-Ala-Ile-Asp-Ala-Gly-Ala-Ala-Gly-Ile-Asp-Ser-Val-Ile-Asp-Ala-Pro-Asp-Leu-Leu-Ang (18)</td>
</tr>
<tr>
<td>TP17</td>
<td>239-249</td>
<td>Ile-Ile-Glu-Gln-His-Asn-Ile-Glu-Pro-Glu-Lys (11)</td>
</tr>
<tr>
<td>TP7</td>
<td>256-262</td>
<td>Val-Phe-Val-Pro-Met-Lys (7)</td>
</tr>
<tr>
<td>TP20</td>
<td>263-266</td>
<td>Ala-Ala-Thr-Arg (4)</td>
</tr>
<tr>
<td>TP28</td>
<td>263-267</td>
<td>Ala-Ala-Thr-Arg-Ser (5)</td>
</tr>
<tr>
<td>TP1</td>
<td>267</td>
<td>Ser (1)</td>
</tr>
</tbody>
</table>

* Overlapping tryptic peptide.

Composition: Asp, 2.18; Ser, 0.85; G1u, 2.98; Pro, 2.84; Gly, 2.08; Ala, 3.12; Val, 0.90; Ile, 1.04; Leu, 1.13; Tyr, 0.76; Phe, 1.04; Lys, 1.13.

Edman degradation: Stage 1 (106%): Asp, 2.12; Ser, 0.90; G1u, 2.35; Pro, 2.84; Gly, 2.02; Ala, 3.00; Val, 1.01; Ile, 0.99; Leu, 1.06; Tyr, 0.80; Phe, 0.96; Lys, not determined.

Leucine aminopeptidase: Traces of glutamic acid and tyrosine were released.

**Chymotryptic Digestion**—Three chymotryptic fragments were recovered from digests of TP13 by paper chromatography. Two of the fragments occupied the same positions on two-dimensional peptide patterns as chymotryptic fragments of TP3 (TP13-C1 (TP3-C1) and TP13-C2 (TP3-C2)), and TP13-C1 gave a yellow-gray reaction with ninhydrin.

**TP13-C1** (Asp, G1u, Pro, Gly, Ala, Leu, Phe)—Composition: Asp, 0.96; G1u, 1.10; Pro, 1.90; Gly, 1.06; Ala, 1.95; Leu, 1.07; Phe, 0.94.

This peptide has precisely the same composition as TP3-C1.

**TP13-C2** (Ser, Gly, Ile)—A1a-Pro-Asp-Gln-Val-Lys—Composition: Asp, 1.08; Ser, 0.93; G1u, 1.00; Pro, 1.11; Gly, 0.98; Ala, 1.06; Val, 1.00; Ile, 0.93; Lys, 1.00.

This peptide also has the same composition as the corresponding chymotryptic fragment of TP3. In order to complete the sequence of TP3 and TP13, this peptide fragment was examined further (see below).

**TP13-C3**: G1u-Tyr—Composition: G1u, 1.10; Tyr, 0.90.

Carboxypeptidase: Tyr, 0.74; G1u, 0.70.

Leucine aminopeptidase: G1u, 0.59; Tyr, 0.55.

This peptide was acidic at pH 6.5, confirming the presence of glutamic acid.

**Papain Digestion of TP13-C2**

The following peptides were recovered from a papain digest of TP13-C2, together with a small proportion of undigested peptide.

**TP13-C2-Pa1**: (Ser, Gly, Ile)—Composition: Ser, 0.96; G1u, 0.95; Ile, 1.09.

**TP13-C2-Pa2**: Ala-Pro-Asp-Gln-Val-Lys—Composition: Asp, 1.12; G1u, 1.08; Pro, 0.95; Ala, 0.99; Val, 0.91; Lys, 0.94.

Edman degradation: Stage 1 (102%): Asp, 0.99; G1u, 1.10; Pro, 1.03; Ala, 0.08; Val, 0.88; Lys, not determined.

Stage 2 (100%): Asp, 1.03; G1u, 1.04; Pro, 0.17; Val, 0.92; Lys, not determined.

Stage 3 (94%): Asp, 0.20; G1u, 1.04; Val, 0.94; Lys, 0.40.

(F; PTH-Asp.)

At this stage a portion of the degraded peptide was treated with carboxypeptidases A plus B and leucine aminopeptidase, releasing Val, 0.64; G1u, 0.49. Lysine was not determined.

Stage 4 (85%): Asp, 0.18; G1u, 0.55; Val, 1.00; Lys, 0.40. (F; PTH-Gln.)
**Fig. 3. Summary of the amino acid sequence represented by tryptic Peptides TP2, TP13, and the overlapping Peptide TP3.** The resultant sequence of amino acids in this segment of the A protein is given in the second line.

**TP13-C2-Pa4: Val-Lys—Composition: Val, 1.00; Lys, 1.00.**
**TP13-C2-Pa4: (Ala, Pro, Glx, Asx)—Composition: Asp, 0.90; Glu, 1.08; Pro, 1.18; Ala, 0.84.**

The similarity between TP13 and TP3 was apparent from a comparison of their compositions. Compared with TP3, Peptide TP13 lacks 1 residue each of leucine and lysine, and it has an amino-terminal glutamyl residue, not the leucyl residue of TP3. Two of the three chymotryptic fragments of TP13 and TP3 have identical compositions, and the sequences of the third, TP13-C3 (Glu-Tyr) and TP3-C3 (Leu-Lys-Glu-Tyr), indicate that TP13 arises by tryptic cleavage of the Lys-Glu bond, which remains intact in TP3.

Further sequence studies with TP13 were restricted to elucidating the identity and order of the aspartyl or asparaginyl and glutamyl or glutaminyl residues in the carboxyl-terminal region, which had not been accomplished in previous studies with the analogous fragment of TP3. For this purpose the chymotryptic Fragment TP13-C2 was digested with papain, and the results of Edman degradation and exopeptidase digestion with the important papain fragment, TP13-C2-Pa2, gave the sequences Ala-Asp-Gln-Val-Lys. The other papain fragments provide confirmatory evidence for the established sequence of the TP3-C2 region of the A protein.

The partial sequence for TP13 is presented in Fig. 3, together with the partial sequence derived earlier for TP3. It is clear that TP3 overlaps TP2 and TP13, and by combining the two series of results the complete sequence for this segment of the A protein can be deduced.

**Peptide TP4 (Residues 179 through 187): Ala-Gly-Val-Thr-Gly-Ala-Glu-Asn-Arg (9 Residues)**

Peptides TP4 and TP6 were eluted from Dowex 1-X2 in the same fraction (Fig. 1A); both peptides were readily separated and purified by paper electrophoresis at pH 3.7.

**Composition:** Asp, 1.06; Thr, 0.96; Glu, 1.14; Gly, 2.00; Ala, 1.87; Val, 1.00; Arg, 1.00.

Edman degradation: Stage 1: Asp, 1.08; Thr, 0.89; Glu, 1.12; Gly, 1.94; Ala, 1.16; Val, 1.00; Arg, 0.95. (A, D, F; PTH-Ala.)

Stage 2: Asp, 1.14; Thr, 0.78; Glu, 1.12; Gly, 1.29; Ala, 1.18; Val, 0.86; Arg, 0.94. (A, D, F; PTH-Gly.)

Stage 3: Asp, 1.00; Thr, 0.76; Glu, 1.02; Gly, 1.16; Ala, 1.05; Val, 0.29; Arg, 0.89. (A, F; PTH-Val.)

Stage 4: Asp, 1.00; Thr, 0.29; Glu, 1.13; Gly, 1.11; Ala, 0.93; Val, 0.24; Arg, 0.83. (A, F; no phenylthiohydantoin derivative observed.)

Carboxypeptidase: Arg, 1.05; Asn, 0.59; Ala, 0.55; Glu, 0.43; Gly, 0.30; Thr, 0.27; Val, 0.23.

Leucine aminopeptidase: Asn, 1.19; Thr, 1.07; Glu, 1.04; Gly, 2.03; Ala, 2.08; Val, 1.17; Arg, 1.02.

The presence of asparagine and glutamic acid was deduced from the amino acid analyses of the exopeptidase digests because the peptide contains no serine.

**Chymotryptic Digestion**—The following fragments were obtained by electrophoresis at pH 3.7.

**TP4-C1: (Asx, Thr, Glx, Gly, Ala, Val)—Composition: Asp, 1.00; Thr, 0.86; Glu, 1.13; Gly, 2.13; Ala, 1.87; Val, 1.00.**

This fragment was occasionally isolated from trypsin digests of the A protein, presumably because of chymotryptic activity of the proteolytic enzyme. The chymotrypsin-susceptible bond probably involves the asparaginyl residue, the only residue in TP4 which could provide a bond of known chymotryptic susceptibility (22).

**Partial Acid Hydrolysis**—Peptide TP4 was dissolved in concentrated HCl and heated in an evacuated, sealed tube at 37° for 24 hours. The acid was removed by evaporation over NaOH pellets in a desiccator, and the products were subjected to two-dimensional chromatography and electrophoresis. Peptides were located with detection strength ninhydrin reagent (0.05 in methanol) and analyzed after elution. The following fragments were obtained: Ala; Val; (Thr, Gly); (Glx, Asx); (Ala, Gly, Val); and (Ala, Gly).

**Papain and Nagarse Fragments**—Papain digestion of the peptide gave one important peptide. The peptide was eluted from a peptide pattern that had been stained with ninhydrin reagent (0.05 in methanol) and analyzed after elution. The following fragments were obtained: Ala; Val; (Thr, Gly); (Glx, Asx); (Ala, Gly, Val); and (Ala, Gly).

**Nagarse digestion of the peptide gave one relevant fragment which was eluted from a ninhydrin-stained peptide pattern.** The composition was Ala, 0.40; Glu, 1.08; Asp, 1.12; Arg, 0.80. The alanine is assigned the amino-terminal position because of the loss resulting from ninhydrin treatment. The peptide was neutral at pH 6.5; this finding plus the results of the leucine aminopeptidase treatment of TP4 suggest that glutamic acid and asparagine are present.

**Nagarse digestion of the peptide gave one relevant fragment which was eluted from a ninhydrin-stained peptide pattern.** The composition was Gly, 0.60; Asx, 1.10; Arg, 0.90. The peptide was neutral at pH 6.5. These findings place glutamic acid at the amino-terminal position of this fragment and suggest that either glutamic acid or aspartic acid is present as the amide.
The first four steps of the Edman degradation of the original Peptide TP4 showed the amino-terminal segment to be Ala-Gly-Val-Thr. Digestion with the carboxypeptidases and leucine aminopeptidase clearly indicated the presence of an asparaginyl residue, and this is the only residue in TP4 which could provide a susceptible bond for the release of arginine by chymotrypsin.

This gives the carboxyl-terminal sequence Asn-Arg, which can be extended to Ala-Glu-Asn-Arg when the important arginine-containing papain and Nagarse fragments are considered. The outstanding glycyl residue is placed between these amino- and carboxyl-terminal segments by difference and by the existence of a dipeptide (Thr, Gly) among other confirmatory peptides in partial acid hydrolysates of TP4.

Stage 5 (84%): Asp, 0.98; Glu, 2.42; Pro, 1.02; Ile, 1.02; Lys, 0.34; His, 0.34.
Stage 6 (87%): Asp, 0.92; Glu, 2.30; Pro, 1.07; Ile, 0.47; Lys, 0.33; His, 0.30.
Stage 7 (93%): Asp, 0.68; Glu, 2.50; Pro, 1.00; Ile, 0.46; Lys, 0.39; His, 0.30.

Carboxypeptidase: Lys, 1.03.
Leucine aminopeptidase: Ile, 2.55; Asn + Gin, 1.48; Glu, 1.11; Lys and His, not determined.

The intact peptide had a negative charge at pH 6.5.

Papain Digestion—Papain digests of TP5 gave the peptide pattern illustrated in Fig. 4A, and the corresponding fragments were isolated in the normal manner by chromatography on Dowex 1-X2 (Fig. 4B).

Peptide TP5 (Residues 239 through 249): Ile-Ile-Glu-Gln-Asn-Ile-Glu-Pro-Glu-Lys (11 Residues)

Chromatography on Dowex 1-X2 gave samples of Peptide TP5 which were contaminated by TP13. However, pure TP5 was readily obtained by paper electrophoresis at pH 3.7. Composition: Asp, 1.06; Glu, 4.00; Pro, 1.10; Ile, 2.55; Lys, 0.98; His, 0.88.

This analysis was obtained after acid hydrolysis for 72 hours, and the isoleucine content was not included in the calculation of the average molar equivalent.

Edman degradation: Stage 1 (100%): Asp, 1.04; Glu, 4.00; Pro, 1.04; Ile, 1.78; Lys, 0.88; His, 0.93. (A, D, F; PTH-Ile.)
Stage 2 (86%): Asp, 1.00; Glu, 3.98; Pro, 1.12; Ile, 0.92; Lys, 0.60; His, 0.89. (F, PTH-Ile.)
Stage 3 (83%): Asp, 0.99; Glu, 3.38; Pro, 1.01; Ile, 1.10; Lys, 0.58; His, 0.95. (A, D, F; PTH-Glu.)
Stage 4 (110%): Asp, 0.98; Glu, 2.45; Pro, 1.02; Ile, 0.86; Lys, 0.54; His, 0.78. (A, D, F; PTH-Gln.)
Assuming that the single lysine residue is carboxyl-terminal as required by tryptic specificity, papain Fragments TP5-Pa1, -Pa4, and -Pa6, which contain lysine, must represent the carboxyl-terminal segment. Studies of these three papain fragments give the sequence Glx-His-(Asx, Ile)-Glu-Pro-Glu-Lys, and from a consideration of the total composition of TP5 the first 2 residues must be identical with the Gln-His residues of the aminoterminal segment. Conflicting evidence for the (Asx, Ile) sequence region was obtained. The sixth and seventh cycles of the Edman degradation, performed with the original peptide, suggest the sequence Ile-Asx, whereas sequence studies with TP5-Pa2 and TP5-Pa3 (particularly the carboxypeptidase digestion evidence) indicate that the sequence is Asn-Ile. The latter evidence was judged more reliable.

Peptide TP6 (Residues 221 through 226): Ala-Ala-Ile-Asp-Ala-Gly-Ala-Ala-Glu-Ala-Ile-Ser-Glu-Ser-Ala-Ile-Val-Lys-
(13 Residues)

Sequence studies with this peptide are published elsewhere (11).

Peptide TP7 (Residues 225 through 226): Val-Phe-Val-Gln-Pro-Met-Lys-
(7 Residues)

This peptide was isolated by paper electrophoresis (pH 3.7) and paper chromatography from the initial fractions of the Dowex 1-X2 fractionation of the smaller peptides (Fig. 1B) and purified, when necessary, by electrophoresis at pH 3.7.

Peptide TP8 (Residues 221 through 223): Ala-Ala-Ile-Asp-Ala-Gly-Ala-Ala-Glu-Ala-Ile-Ser-Glu-Ser-Ala-Ile-Val-Lys-
(18 Residues)

This peptide was also isolated from other samples of this peptide and of chymotryptic fragments and from the liberation, by leucine aminopeptidase, of 2 valyl and 1 phenylalanyl residues per molecule of TP7, and the presence of glutamine and not the free acid is indicated by the electrophoretic mobility of the peptide at pH 6.5.

Peptide TP8 (Residues 221 through 223): Ala-Ala-Ile-Asp-Ala-Gly-Ala-Ala-Glu-Ala-Ile-Ser-Glu-Ser-Ala-Ile-Val-Lys-
(18 Residues)

TP5-C1: Gly-Tyr-Thr-Pro-Leu-Leu-Leu-Arg—Composition: Thr, 1.11; Ser, 0.81; Gly, 1.08; Leu, 1.91; Tyr, 1.90; Arg, 1.10.

The tyrosine content was not included in the calculation of the average molar equivalent.

Edman degradation: Stage 1: Gly, 0.24; Leu, 1.90; Tyr, 0.63; Arg, 1.12; Thr, Ser, analyses lost. (A, F; PTH-Gly.)

Chymotryptic Digestion—Four chymotryptic fragments were separated by electrophoresis at pH 3.7.

TP5-C2: Gly-Thr-Pro-Leu-Leu-Leu-Arg—Composition: Thr, 1.06; Gly, 1.06; Tyr, 1.88.

Edman degradation: Stage 1: Thr, 1.00; Gly, 0.42; Tyr, 1.54. (F; PTH-Gly.)

The Tyr-Thr bond of this chymotryptic fragment was cleaved only after an extended period of digestion with chymotrypsin.

TP5-C3: Leu-Leu-Leu-Leu-Arg—Composition: Ser, 1.02; Leu, 1.88; Arg, 1.12.

Edman degradation: Stage 1: Ser, 1.08; Leu, 1.00; Arg, 0.90. (A; PTH-Leu.)

This peptide was also isolated as TP24.

TP5-C4: Ser-Arg—Composition: Ser, 1.33; Leu, 0.89; Arg, 0.78.

Digestion with chymotrypsin was the basis for defining the sequence of this peptide. The amino-terminal residues of TP8-C1, Gly Tyr, were sequenced by the Edman degradation, and the sequence of the remaining residues, Thr-Tyr, was deduced from the required specificity for chymotryptic hydrolysis. Edman degradation of the original peptide indicates that TP8-C1 represents the amino-terminal segment of TP8, and the partial sequences of the chymotryptic fragments TP8-C2, TP8-C3, and TP8-C4 provide convincing evidence for

Peptide TP7 (Residues 225 through 226): Val-Phe-Val-Gln-Pro-Met-Lys-
(7 Residues)

Edman degradation: Stage 1: Val, 1.05; Pro, 0.95; Val, 1.91; Met, 0.62; Phe, 1.38; Lys, 1.06.

Edman degradation: Stage 2: Glu, 0.66; Pro, 1.10; Val, 0.14; Met, 0.90; Lys, not determined. (A; F; PTH-Gln.)

Chymotryptic Digestion—The products of digestion of TP7 with chymotrypsin were separated by paper electrophoresis at pH 3.7.

TP7-C1: Val-Phe—Composition: Val, 1.05; Phe, 0.96.

This peptide was isolated by paper electrophoresis (pH 3.7) and paper chromatography from the initial fractions of the Dowex 1-X2 fractionation of the smaller peptides (Fig. 1B) and purified, when necessary, by electrophoresis at pH 3.7.

TP7-C2: Val-Gln—Composition: Glu, 1.03; Pro, 1.14; Val, 0.89; Met, 0.87; Lys, 1.07.

Edman degradation: Stage 1: Glu, 1.07; Pro, 1.18; Val, 0.14; Met, 0.78; Lys, not determined. (A; PTH-Val.)

Stage 2: Glu, 0.66; Pro, 1.10; Val, 0.14; Met, 0.90; Lys, not determined. (A; F; PTH-Gln.)

TP7-C3: Val-Gly—Composition: Glu, 1.10; Pro, 1.00; Val, 0.90; Met, 1.00.

The Edman degradation showed that the amino-terminal segment is Val-Phe, and this is identical with the chymotryptic Fragment TP7-C1. The carboxy-terminal segment is represented by the lysine-containing chymotryptic Fragment TP7-C2, Val-Gln (Pro, Met) Lys, the first 2 residues being sequenced by the Edman method. The sequence of the (Pro, Met) region can be deduced from the discovery of the chymotryptic Fragment TP7-C3, in which methionine must be carboxy-terminal to provide a chymotrypsin-sensitive bond. This conclusion is further supported by the isolation of a peptide, (Glu, Pro, Val, Phe) homoserine lactone, as a tryptic fragment of a polypeptide liberated from the A protein by specific cleavage of methionyl bonds with CNBr. The amino-terminal sequence of TP7 is

Fig. 5. Summary of the sequence studies on tryptic Peptide TP5.

* J. R. Guest, unpublished observations.
the sequence of the carboxyl-terminal segment, Leu-Leu-Ser-Arg. Tryptic digests of the A protein were often found to contain substantial quantities of TP8-C1 and TP8-C2 (TP24) (see following peptide), presumably because of the extraordinary susceptibility of the Tyr-Leu bond.

**Peptide TP24 (Residues 175 through 178): Leu-Leu-Ser-Arg (4 Residues)**

This peptide was found to have the same composition, partial sequence, and position on the peptide pattern as TP8-C2, and it is presumed to arise by hydrolysis of the Tyr-Leu bond of TP8, owing to the weak chymotryptic activity of trypsin.

**Peptide TP9 (Residues 1 through 3): Met-Gln-Arg (3 Residues)**

The composition and sequence of this peptide have been reported previously; it is the amino-terminal tryptic peptide of the A protein (3).

**Peptide TP10 (Residues 15 through 35): Lys-Glu-(Gly, Ala)-Phe-Val-Pro-Phe-Val-Thr-Leu-Gly-Asp-Pro-(Glx, Gly, Ile)-(Ser, Leu)-Lys (21 Residues)**

This is a large peptide which was isolated by Dowex 1-X2 chromatography of the Sephadex fractions containing the larger peptides of the tryptic digest of the A protein (Fig. 1B). The peptide, which gave a weak reaction with ninhydrin, was separated from small amounts of contaminating material by paper electrophoresis at pH 3.7. The composition of TPIO shows the presence of 2 lysyl residues, and the first step of the Edman degradation suggested that 1 of these is the amino-terminal residue; the second step gave a glutamine residue. Three of the five chymotryptic fragments contain lysine, and the amino-terminal segment is clearly represented by two of these, TP10-C2 and TP10-C5, because they contain lysyl and glutamyl residues. It follows that the third lysine-containing fragment, TP10-C1 (TP10-C5 and TP10-C4), was located with the tert-butyl hypochlorite reagent.

**Peptide TP9-C1: (Ser, Leu)-Lys—Composition: Ser, 0.84; Leu, 1.01; Lys, 1.00.**

Since the sequence of TP10 commences with Lys-Glu, it is clear that this fragment (TP10-C1) must represent the carboxyl-terminal region of the peptide.

**TP10-C2: (Glx, Pro, Gly, Ala, Val, Phe, Lys)—Phe—Composition: Glu, 0.99; Pro, 1.05; Gly, 1.12; Ala, 1.02; Val, 0.95; Phe, 1.92; Lys, 0.94.**

**TP10-C3: Val-Thr-Gly-Asp-Pro-(Glx, Gly, Ile)—Composition: Asp, 1.03; Thr, 0.91; Glu, 1.98; Pro, 2.05; Val, 0.92; Ile, 1.05; Leu, 1.03.**

Edman degradation: Stage 1 (91%): Asp, 1.01; Thr, 1.02; Glu, 2.03; Pro, 1.11; Gly, 1.97; Val, 0.03; Ile, 0.90; Leu, 0.96. Stage 2 (84%): Asp, 1.04; Thr, 0.18; Glu, 2.02; Pro, 0.94; Gly, 2.03; Ile, 0.98; Leu, 0.98. Stage 3 (90%): Asp, 1.06; Glu, 2.12; Pro, 0.86; Gly, 1.94; Ile, 1.01; Leu, 0.24. Stage 4 (120%): Asp, 1.03; Glu, 2.11; Pro, 0.92; Gly, 1.21; Ile, 0.94; Leu, 0.16. Stage 5 (79%): Asp, 0.17; Glu, 2.03; Pro, 0.88; Gly, 1.22; Ile, 0.88. (F; PTH-Asp.) Stage 6 (88%): Asp, 0.28; Glu, 2.30; Pro, 0.23; Gly, 0.95; Ile, 0.74. Carboxypeptidase: No significant release of amino acids was observed. Leucine aminopeptidase: Val, 0.83; Thr, 0.82; Leu, 0.77; Gly, 0.44.

**TP10-C4: Val-Pro-Phe—Composition: Pro, 0.96; Val, 1.12; Phe, 0.91.**

Edman degradation: Stage 1 (24%): Asp, 1.03; Val, 0.59; Phe, 0.97. TP10-C5: (Glx, Gly, Ala, Lys)—Phe—Composition: Glu, 1.00; Gly, 1.21; Ala, 1.00; Phe, 0.91; Lys, 0.91. The composition of TP10 shows the presence of 2 lysyl residues, and the first step of the Edman degradation suggested that 1 of these is the amino-terminal residue; the second step gave a glutamate residue. Three of the five chymotryptic fragments contain lysine, and the amino-terminal segment is clearly represented by two of these, TP10-C2 and TP10-C5, because they contain lysyl and glutamate residues. It follows that the third lysine-containing fragment, TP10-C1 (TP10-C5 and TP10-C4), is contained by the composition of TP10-C2, must correspond to the carboxyl-terminal segment of TP10 to satisfy the tryptic specificity. By difference, the remaining fragment, TP10-C3, corresponds to the middle segment of the original peptide.
defining an amino-terminal valyl residue for TP10-C4, the partial sequence for the amino-terminal segment of TP10 becomes Lys-Glu-(Gly, Ala)-Phe-Val-Pro-Phe. A partial sequence for the midportion of TP10 comes from the Edman degradation of TP10-C3, in which 6 out of the 10 residues were placed. No decision about the carboxyl-terminal residue of this fragment was made, but, because the leucyl residue occupies the third position, it must be assumed that 1 of the remaining glutamyl residues is amidated and thus provides a chymotrypsin-susceptible bond. The structure of TP10 derived from the sequence information is illustrated in Fig. 7. The presence of 2 lysine residues indicates that this is an overlapping tryptic peptide which links TP12 (see below) and TP21 (free lysine). Just as with TP3, the poor susceptibility of the Lys-Glu bond to hydrolysis by trypsin appears to be responsible for its occurrence.

Peptide TP11 Details are given following TP12.

Peptide TP12 (Residues 16 through 36): Glu-Gly-Ala-Phe-Val-Pro-Phe-(Val, Thr)-(Leu, Gly)-Asp-Pro-Gly-Ile-Glu-Gln-Ser-Leu-Lys (20 Residues)

The peptide was isolated in relatively pure form by Dowex 1-X2 chromatography of the Sephadex fractions containing the larger peptides of a tryptic digest of the A protein. Further purification was usually effected by paper electrophoresis at pH 3.7.

Composition: Asp, 1.12; Thr, 0.96; Ser, 0.92; Glu, 3.03; Pro, 2.16; Gly, 2.92; Ala, 1.24; Val, 1.33; Ile, 0.95; Leu, 1.86; Phe, 1.86; Lys, 1.00.

Edman degradation: Stage 1 (75%): Asp, 1.20; Thr, 0.93; Ser, 1.00; Glu, 2.26; Pro, 1.86; Gly, 2.98; Ala, 1.17; Val, 1.98; Ile, 1.00; Leu, 1.98; Phe, 1.80; Lys, not determined.

Stage 2 (35%): Asp, 1.27; Thr, 0.89; Ser, 1.06; Glu, 2.43; Pro, 1.84; Gly, 2.33; Ala, 1.22; Val, 1.90; Ile, 0.98; Leu, 1.92; Phe, 1.62; Lys, not determined.

Stage 3 (100%): Asp, 1.20; Thr, 0.88; Ser, 1.27; Glu, 2.36; Pro, 1.48; Gly, 2.22; Ala, 0.60; Val, 1.70; Ile, 0.05; Leu, 1.83; Phe, 1.93; Lys, not determined.

Stage 4 (120%): Asp, 1.13; Thr, 0.90; Ser, 0.96; Glu, 2.28; Pro, 1.96; Gly, 2.34; Ala, 0.45; Val, 1.96; Ile, 0.96; Leu, 1.80; Phe, 1.23; Lys, not determined.

Carboxypeptidase: Leu, 0.80; Ser (+Gln?), 0.56; Lys, not determined.

Leucine aminopeptidase: Traces of Glu, Gly, and Ala.

Chymotryptic Digestion—A chymotryptic digest of TP12 gave the peptide pattern outlined in Fig. 6B and the following fragments were eluted from similar peptide patterns.

TP12-C1: Ser-Leu-Lys—Composition: Ser, 0.88; Leu, 1.00; Lys, 1.20.

Edman degradation: Stage 1 (71%): Ser, 0.27; Leu, 1.00; Lys, not determined.

TP12-C2: Glu-(Pro, Gly, Ala, Val, Phe)-Phe—Composition: Glu, 0.70; Pro, 1.10; Gly, 0.96; Ala, 1.00; Val, 1.20; Phe, 1.50.

Edman degradation: Stage 1 (100%): Glu, 0.41; Pro, 0.96; Gly, 1.18; Ala, 1.04; Val, 1.02; Phe, 1.84.

The presence of glutamic acid was inferred from the electrophoretic mobility at pH 0.5.

TP12-C3: (Asp, Thr, Glx, Pro, Glx, Val, Ile, Leu)—Composition: Asp, 1.05; Thr, 0.92; Glu, 2.12; Pro, 1.05; Gly, 2.09; Val, 0.84; Ile, 0.99; Leu, 0.97.

It can be seen that these fragments are analogous to those isolated from similar digests of TP10.

Papain Digestion—To elucidate the sequence of TP12, papain fragments (Fig. 6C) were isolated from the corresponding digest by electrophoresis at pH 3.7 and chromatography where necessary.

TP12-P1: Asp-Pro-Gly-Ile-Glu—Composition: Asp, 0.92; Glu, 1.02; Pro, 0.06; Gly, 1.07; Ile, 1.01.
Edman degradation: Stage 1 (106%): Asp, 0.16; Glu, 0.00; Pro, 1.04; Gly, 1.02; Ile, 0.94; (F; PTH-Asp).

Stage 2 (81%): Asp, 0.13; Glu, 0.99; Pro, 0.21; Gly, 1.11; Ile, 0.89.

Stage 3 (70%): Gly, 1.03; Gly, 0.47; Ile, 0.96.

Stage 4 (53%): Glu, 1.00; Gly, 0.31; Ile, 0.31. Direct analysis of the unhydrolyzed residual peptide at this stage of the degradation yielded a small amount of glutamic acid.

TP12-Pa2: (Gly, Leu)—Composition: Gly, 1.11; Leu, 0.90.

TP12-Pa3: Leu—Gly—Composition: Leu, 0.93; Lys, 1.07.

TP12-Pa4: (Glx, Gly)—Composition: Glu, 0.90; Gly, 1.10.

TP12-Pa5: (Thr, Pro, Ala, Val, Phe)—Composition: Thr, 1.06; Pro, 0.95; Ala, 1.05; Val, 1.75; Phe, 2.13.

TP12-Pa6: (Ser, Leu)—Lys—Composition: Ser, 0.93; Leu, 1.07; Lys, 1.00.

TP12-Pa7: Gln—Ser—Leu—Lys—Composition: Ser, 0.90; Glu, 1.00; Leu, 1.04; Lys, 1.07.

Edman degradation: Stage 1 (74%): Ser, 0.97; Glu, 0.48; Leu, 1.00; Lys, not determined. (F; PTH-Gln.)

Carboxypeptidase plus leucine aminopeptidase: Glx + Ser, 1.12; Leu, 0.66; Lys, not determined.

TP12-Pa8: Val—Pro—Thr, Val, Phe)—Composition: Thr, 0.87; Pro, 1.22; Val, 1.90; Phe, 0.99.

Edman degradation: Stage 1 (101%): Thr, 0.92; Pro, 1.07; Val, 1.09; Phe, 0.99.

Stage 2 (93%): Thr, 1.00; Pro, 0.34; Val, 1.10; Phe, 0.90.

The first four steps of the Edman degradation with TP12 showed that the sequence of the amino-terminal segment is Glu—Gly—Ala—Phe. The aggregate composition of the three chymotryptic fragments is equivalent to the composition of the original peptide, and the fragments can be placed in the order (TP12-C2)—(TP12-C3)—(TP12-C1), because TP12-C2 contains the amino-terminal residues and TP12-C1 can be designated as the carboxyl-terminal segment by virtue of its lysyl residue.

The details of the partial sequence of TP12 were obtained with the papain fragments. Fragments TP12-Pa3 and TP12-Pa6 confirm the sequence of TP12-C1, and this sequence is extended in turn by TP12-Pa7 to give this sequence for the carboxyl-terminal segment, Glx—Ser—Leu—Lys. At the amino terminus, papain Fragment TP12-Pa4 must correspond to the amino-terminal Glu—Gly sequence, because the only other glycyl residue is associated with a leucyl residue, which is known to be located in the central region (TP12-C3). Papain Fragments TP12-Pa5 and TP12-Pa8 link the chymotryptic Fragment TP12-C2 with TP12-C3, and the Edman degradation with TP12-Pa8 completes the sequence of amino acids in the TP12-C2 region: Glu—Gly—Ala—Phe—Val—Pro—Phe. The remaining fragments, TP12-Pa1 and TP12-Pa2, belong to the central region (the TP12-C3 segment) of TP12. The sequence of TP12-Pa1 is established by the Edman degradation as Asp—Pro—Gly—Ile—Glu, but the relative order of TP12-Pa1 and TP12-Pa2 only becomes apparent when considered in conjunction with the similar sequences in TP10.

The sequence information for TP12 is summarized in Fig. 7, together with the partial sequence derived for TP10. As shown, this information is sufficient to establish the complete sequence for the corresponding segment of the A protein. Compared with TP12, TP10 has a greater electrophoretic mobility at pH 3.7 and it has a weaker affinity for Dowex 1-X2, as would be expected for a peptide having an additional basic residue.

Peptide TP11 (Residues 164 through 170): Gln—Ile Ala Sor Tyr Gly—Arg (7 Residues)

This peptide was purified, like all the other small basic peptides, by paper electrophoresis at pH 3.7 followed by paper chromatography.

Composition: Ser, 0.83; Glu, 1.01; Gly, 1.13; Ala, 1.25; Ile, 0.81; Tyr, 0.94; Arg, 1.02.

Edman degradation: Stage 1: Ser, 0.95; Glu, 0.44; Gly, 1.16; Ala, 1.20; Ile, 0.87; Tyr, 0.76; Arg, 1.04. (F; PTH-Gln.)

Stage 2: Ser, 0.78; Glu, 0.44; Gly, 1.21; Ala, 1.23; Ile, 0.33; Tyr, 0.64; Arg, 1.15. (D, F; PTH-Ile.)

Stage 3: Ser, 0.55; Glu, 0.33; Gly, 0.08; Ala, 0.36; Ile, 0.29; Tyr, 0.46; Arg, 1.02. (A, D; PTH-Ala, PTH-Ile.)

Leucine aminopeptidase: Ser + Gln, 1.72; Tyr, 0.91; Ile, 0.74; Ala, 0.73; Gly, 0.61; Arg, not determined.

Chymotryptic Digestion—The following fragments were obtained.

TP11-C1: (Glx, Ile, Ala, Ser)—Tyr—Composition: Ser, 0.89; Glu, 1.00; Ala, 1.09; Ile, 1.03; Tyr, 1.00.

TP11-C2: Gly—Arg—Composition: Gly, 0.94; Arg, 1.06.

The first three steps of the Edman degradation with TP11 showed that the amino-terminal segment is Glx—Ile—Ala. The presence of a glutaminyl residue was confirmed by the results of leucine aminopeptidase digestion of the original peptide. The composition of the chymotryptic fragment, TP11-C1, indicates that tyrosine occupies the fifth position. By difference, therefore, serine is placed before tyrosine. The remaining segment corresponds to the Gly—Arg dipeptide, TP11-C2, which is sequenced according to the specificities required by trypsin.

Peptide TP13 Details are given following TP3.

Peptide TP14 (Residues 4 through 12): Tyr—Glu—Ser—Leu—Phe—Ala—Gln—Leu—Lys (9 Residues)

Peptide TP14 was isolated from the Sephadex fractions containing the larger peptides by chromatography on Dowex 1-X2 (Fig. 1A), followed by purification by paper electrophoresis at pH 3.7.

Composition: Ser, 0.80; Glu, 2.28; Ala, 1.16; Leu, 1.86; Tyr, 0.86; Phe, 0.98; Lys, 0.98.

Edman degradation: Stage 1: Ser, 0.69; Gly, 0.54; Leu, 0.95; Phe, 0.98; Lys, not determined. (A, D, and F; PTH-Tyr.)

Stage 2 (80%): Ser, 0.90; Glu, 1.37; Ala, 1.11; Leu, 1.98; Tyr, 0.02; Phe, 1.01; Lys, not determined. (A, F; PTH-Glu.)

Stage 3 (95%): Ser, 0.40; Glu, 1.16; Ala, 1.06; Leu, 1.88; Phe, 0.89; Lys, not determined.

Stage 4 (91%): Ser, 0.32; Glu, 1.14; Ala, 0.95; Leu, 1.35; Phe, 0.91; Lys, not determined.

Carboxypeptidase: Ser + Gln, 1.06; Ala, 0.98; Leu, 1.84; Phe, 1.03; Lys, 0.60.

Leucine aminopeptidase: Tyr, 0.44; Ser + Gln, 0.40; Leu, 0.40; Phe, 0.32; Glu, 0.29; Ala, 0.22; Lys, 0.09.

Chymotryptic Digestion—The following fragments were obtained.

TP14-C1: (Ser, Glx, Leu, Tyr, Phe)—Composition: Ser, 0.98; Glu, 1.15; Leu, 1.10; Tyr, 0.79; Phe, 1.00.

TP14-C2: Ala—Gln—Leu—Lys—Composition: Glu, 1.08; Ala, 0.85; Leu, 1.02; Lys, 1.04.

Edman degradation: Stage 1 (94%): Glu, 1.04; Ala, 0.11; Leu, 0.96; Lys, not determined.
Stage 2 (89%): Glu, 0.40; Ala, 0.10; Leu, 1.00; Lys, 0.55.
(F; PTH-Gln.)
Carboxypeptidase: Lys, 1.00; Leu, 0.96; Gin, 0.60; Ala, 0.47.
Leucine aminopeptidase: Ala, 1.03; Leu, 1.01; Gin, 0.90; Lys, 0.92.
The sequence of 4 of the amino-terminal residues was established by the Edman degradation, and the discovery of the chymotryptic fragment TP14-C2 extended the sequence of the amino-terminal segment to Tyr-Glu-Ser-Leu-Phe. The carboxy-terminal segment of TP14 corresponds to the other chymotryptic fragment, TP14-C2. The first 2 residues of this fragment are established by the Edman degradation, and the remainder of the sequence can be deduced by assuming that the leucyl residue occupies the carboxy-terminal position, to satisfy the tryptic specificity, and placing the leucyl residue by difference. The carboxy-terminal segment is therefore Ala-Gln-Leu-Lys, and the presence of a single glutaminyl residue is confirmed by exopeptidase digestion of TP14-C2. The final sequence supports the observed electrophoretic neutrality of TP14 at pH 6.5.

**Peptide TP16 (Residues 36 through 69):**

See the accompanying paper (8).

**Peptide TP17 (Residues 140 through 144):** Gln-Ala-Ala-Leu-Arg (5 Residues)

This basic peptide was isolated by paper electrophoresis (pH 3.7) and paper chromatography from the early fractions eluted from Dowex 1-X2 (Fig. 1B). Composition: Glu, 0.78; Ala, 2.14; Leu, 1.06; Arg, 1.00. Edman degradation: Stage 1: Glu, 0.41; Ala, 1.91; Leu, 1.05; Arg, 1.05. (A; PTH-Glx, -Ala.)

Stage 2: Glu, 0.36; Ala, 1.32; Leu, 0.96; Arg, 1.04. (A; PTH-Glx, -Ala; F; PTH-Ala.)

Stage 3: Glu, 0.48; Ala, 0.93; Leu, 0.86; Arg, 1.00. Leucine aminopeptidase: Gin, 0.96; Leu, 0.93; Arg, not determined, Ala, lost.

This sequence is based on three steps of Edman degradation, the remaining residues being placed by assuming that arginine occupies the carboxy-terminal position to satisfy the tryptic specificity. The presence of glutamine is confirmed by digesting TP16 with leucine aminopeptidase.

**Peptide TP18 (Residues 250 through 255):** Met-Leu-Ala-Ala-Leu-Lys (6 Residues)

When the Sephadex fractions containing the larger peptides of tryptic digests of oxidized A protein were eluted from Dowex 1-X2, the main components of the initial basic fractions were TP17 and TP18 (Fig. 1A). They were readily separated by paper electrophoresis at pH 3.7.

Composition: Ala, 2.02; Met, 0.75; Leu, 2.00; Lys, 0.98. Edman degradation: Stage 1 (70%): Ala, 1.98; Met, 0.25; Leu, 2.04; Lys, not determined.

Stage 2 (83%): Ala, 2.00; Met, 0.23; Leu, 1.31; Lys, 0.40.

Stage 3 (79%): Ala, 1.26; Met, 0.26; Leu, 1.00; Lys, not determined.

Chymotryptic Digestion—The following fragments were recovered from the peptide pattern of a chymotrypsin digest of TP17, together with a small quantity of the unaltered tryptic peptide. Because of reaction with ninhydrin, relatively low values were obtained for residues presumed to occupy amino-terminal positions.

**Fig. 8. Peptide patterns of a chymotrypsin digest of TP17 (A) and a papain digest of TP17 (B).** Minor peptides are represented by open circles (see also the legend to Fig. 2).

*TP17-C1: (Ala, Leu)-Lys—Composition: Ala, 1.55; Leu, 1.00; Lys, 1.00.*

*TP17-C2: Met-Leu—Composition: Met, 0.40; Leu, 1.00.*

*TP17-C3: Lys—Composition: Ala, 1.79; Leu, 0.98; Lys, 1.02.*

*TP17-Pa1: (Ala, Leu)-Lys—Composition: Ala, 0.79; Leu, 0.64; Lys, 0.94.*

*TP17-Pa2: (Ala, Met, Leu)—Composition: Ala, 1.05; Met, 0.64; Leu, 0.94.*

*TP17-Pa3: (Ala, Met, Leu)—Composition: Ala, 2.00; Met, 0.50; Lys, 1.00.*

*TP17-Pa4: (Ala, Leu)-Lys—Composition: Ala, 2.00; Leu, 0.77; Lys, 1.24.*

Traces of Ala, Leu, Lys, and (Ala, Leu) were also obtained.

Three steps of the Edman degradation with TP17 gave Met-Leu Ala as the amino-terminal segment. This sequence was confirmed and extended by the isolation of numerous chymotryptic and papain fragments. The respective peptide patterns are illustrated in Fig. 8, and the structures of the corresponding fragments are summarized in Fig. 9A. They provide more than sufficient evidence to support the proposed sequence for TP17.

**Peptide TP18 (Residues 188 through 198):** Ala-Ala-Leu-Pro-Leu-Am-His-Leu-Val-Ala-Lys (11 Residues)

The isolation of this peptide is described above (see TP17). Composition: Asp, 1.00; Pro, 1.03; Ala, 2.99; Val, 0.96; Leu, 2.95; Lys, 1.16; His, 0.89.

Edman degradation: Stage 1 (91%): Asp, 1.00; Pro, 1.03; Ala, 2.99; Val, 0.96; Leu, 2.95; Lys, 1.16; His, 0.89.

Stage 2 (100%): Asp, 1.00; Pro, 0.98; Ala, 1.43; Val, 1.02; Leu, 2.85; Lys, 0.56; His, 0.78. (A; F, D; PTH-Ala.)

Stage 2 (100%): Asp, 1.00; Pro, 0.98; Ala, 1.43; Val, 1.02; Leu, 2.85; Lys, 0.56; His, 0.78. (A; PTH-Ala.)

Stage 3 (88%): Asp, 1.00; Pro, 0.98; Ala, 1.25; Val, 1.00; Leu, 2.09; Lys, His, not determined.

Stage 4 (83%): Asp, 0.97; Pro, 0.61; Ala, 1.17; Val, 0.92; Leu, 1.97; Lys, 0.35; His, 0.65.

Stage 5 (91%): Asp, 0.95; Pro, 0.32; Ala, 1.13; Val, 0.91; Leu, 1.47; Lys, 0.30; His, 0.94.
A. Tryptic Peptide TP17

\[
\begin{align*}
\text{C1} & \quad \text{C2} \quad \text{C3} \\
\text{Met-Leu} & \quad \text{Ala} \quad \text{Ala} \quad \text{Leu} \quad \text{Lys} \\
\text{Pa4} & \quad \text{Pa3} \quad \text{Pa2} \quad \text{Pa1} \\
\end{align*}
\]

B. Tryptic Peptide TP18

\[
\begin{align*}
\text{C1} & \quad \text{C2} \quad \text{C3} \\
\text{Ala-Ala} & \quad \text{Leu-Pro-Leu} \quad \text{Asn} \quad \text{His} \quad \text{Leu} \quad \text{Val-Ala} \quad \text{Lys} \\
\text{Pa4} & \quad \text{Pa3} \quad \text{Pa2} \quad \text{Pa1} \\
\end{align*}
\]

**Fig. 9.** Summary of the amino acid sequences of tryptic peptides TP17 (A), and TP18 (B).

**Stage 6 (87%):** Asp, 0.56; Ala, 1.16; Val, 0.84; Leu, 1.37; Lys, 0.36; His, 0.64.

Carboxypeptidase: Ala, 1.47; Leu, 1.17; His, 1.05; Val, 0.84; Asn, 0.66; Lys, not determined.

Leucine aminopeptidase: Ala, 1.33.

Chymotryptic Digestion—Peptides were isolated from peptide patterns of the chymotryptic digest of TP18 (see also the legend to Fig. 2).

**TP18-C1:** Asn-(Ala, Val, Leu, His)-Lys—Composition: Asp, 0.47; Ala, 1.16; Val, 1.05; Leu, 1.05; Lys, 0.92; His, 0.83.

This fragment exhibited a characteristic yellow ninhydrin reaction, which, together with the low value for aspartic acid obtained after acid hydrolysis, indicate that asparagine is the amino-terminal residue (see also carboxypeptidase data with intact TP18).

**TP18-C2:** Val-Ala-Lys—Composition: Ala, 1.06; Val, 0.96; Lys, 0.98.

Edman degradation: Stage 1 (96%): Ala, 1.00; Val, 0.10; Lys, 0.76.

**TP18-C3:** Asn-His-Leu—Composition: Asp, 0.54; Leu, 1.14; His, 0.86.

This fragment also gave a characteristic yellow color with ninhydrin, which suggested the presence of an amino-terminal asparagine residue.

**TP18-C4:** (Pro, Ala, Leu)-Leu—Composition: Pro, 1.03; Ala, 1.92; Leu, 2.05.

**TP18-C5**—This fragment had the same composition as TP18-C1 and was therefore presumed to be the deamidated derivative formed during digestion.

**Papain Digestion**—Papain fragments were also eluted from a peptide pattern of papain digests of TP18 similar to that illustrated in Fig. 10B.

**TP18-Pa1:** Lys

**TP18-Pa2:** Ala-Ala—The position of this peptide on the peptide pattern indicated that it was not free alanine but probably a dipeptide derived from the first 2 amino-terminal residues.

**TP18-Pa3:** (Ala, Val, Leu)—Composition: Ala, 1.10; Val, 1.09; Leu, 1.06.

**TP18-Pa4:** His-(Ala, Val, Leu)—Composition: Ala, 1.05; Val, 1.00; Leu, 0.96; His, 0.53.

Edman degradation: Stage 1 (100%): Ala, 1.05; Val, 0.08; Leu, 0.96; His, 0.14.

Fragments containing proline or leucine were not recovered, possibly because they stained poorly with ninhydrin and the tert-butyl hypochlorite reagent was not used here.

The first five steps of Edman degradation of the original Peptide TP18 showed the amino-terminal segment to be Ala-Ala-Leu-Pro-Leu. Assuming the single lysine in TP18 to be carboxyl-terminal, as required by tryptic specificity, chymotryptic Fragment TP18-C2 containing this lysyl residue must represent the carboxyl-terminal segment of TP18, Val-Ala-Lys. Since the valyl residue in this segment is the only valine in TP18, the compositions of TP18-Pa3 and TP18-Pa4 extend this sequence 2 residues further, yielding His-Leu-Val-Ala-Lys. These amino- and carboxyl-terminal segments account for the total composition of TP18, except for 1 asparaginyl residue, which must therefore be placed between them. This residue is identified as an asparagine from the carboxypeptidase data with the intact peptide. The position of the asparaginyl residue is confirmed by the low values for aspartic acid obtained in acid hydrolysates of the two peptides, TP18-C1 and TP18-C3, following ninhydrin reaction and elution from paper. The structure of TP18 and the corresponding chymotryptic and papain fragments are summarized in Fig. 9B. Tryptic Peptide TP18 is located at two positions on the peptide pattern (Fig. 2), the minor component, TP18b, is probably formed by deamidation of the asparagine residue.

**Peptide TP19** Details are given following TP27.

**Peptide TP20** Details are given following TP28.

**Peptide TP21** Details are given following TP27.
Peptide TP22 (Residues 180 through 189): Val-Gly-Val-Asp-Ser-Val-Leu-Val-Ala-Asp-Val-Pro-Val-Glu-Glu-Ser-Ala-Pro-Phe-Arg (20 Residues)

The sequential analysis of this peptide was published previously in a paper concerned with the amino acid sequences surrounding the sulphydryl groups of the A protein (7).

Peptide TP23 (Residues 109 through 119): Gly-Ile-Asp-Glu-Pheryl-Ala-Gln-Cys-Glu-Lys (11 Residues)

The sequence studies with this peptide have been reported previously in a discussion of the amino acid sequences surrounding the sulphydryl groups of the A protein (7).

Peptide TP24 Details are given following TPS.

Peptide TP25 (Residues 70 through 88): Ala-Phe-Ala-Ala-Gly-Val-Thr-Pro-Ala-Gln-Cys-Phe-Glu-Met-Leu-Ala-Leu-Ile-Arg (19 Residues)

The sequence of this peptide was reported previously in a discussion of the sequence surrounding the sulphydryl groups of the A protein (7).

Peptide TP26 (Residues 99 through 108)

See the accompanying paper (8).

Peptide TP27 (Residues 18 through 15): Glu-Arg-Lys (3 Residues)

This peptide was isolated by the standard procedures (Figs. 1B and 2).

Composition: Glu, 0.93; Lys, 1.09; Arg, 0.97.

Edman degradation: Stage 1: Glu, 0.65; Lys, 0.97; Arg, 1.03.

Stage 2: Glu, 0.38; Lys, 1.00; Arg, 0.74.

The occurrence of 2 basic residues indicates that this peptide is an overlapping trypsic peptide, and its origin can readily be appreciated by referring to Fig. 7. In the amino-terminal region of the sequence illustrated, there are two adjacent trypsin-sensitive bonds, and it would appear that cleavage of one of these lowers the susceptibility of the other, trypsin being only a poor exopeptidase. As a consequence of this, and also because of the anticipated resistance of the Lys-Glu bond, two overlapping tryptic peptides (TP27 and TP10) are found, as well as the discrete fragments TP19 (see below), TP21 (see below), and TP12. The assignment of glutamic acid rather than glutamine is based on the presence of glutamic acid in TP10 and the electrophoretic mobility of TP27 at pH 6.8, which suggests 1 net positive charge.

Peptide TP19 (Residues 13 and 14): Glu-Arg (2 Residues)

This peptide was isolated by Dowex 1-X2 chromatography and purified by paper electrophoresis at pH 3.7.

Composition: Glu, 0.98; Arg, 1.02.

The sequence of this peptide can be inferred from the specificity of trypsin, which requires that arginine occupy the carboxyl-terminal position. The peptide was neutral at pH 6.5 and presumably contains the free form of glutamic acid.

Peptide TP21 (Residue 15): Lys (1 Residue)

This digestion product was separated from the small basic peptides eluted from Dowex 1-X2 by paper electrophoresis and paper chromatography by the standard procedure. It was identified as free lysine by amino acid analysis (the lysine content was unchanged by acid hydrolysis), and by its chromatographic and electrophoretic behavior. Free lysine probably arises by tryptic cleavage of the lysyl residue, which is amino-terminal in TP10 and carboxyl-terminal in TP27 (see above). The discovery (26) of an overlapping chymotryptic peptide (PP) linking TP10 and TP27 indicates that in both cases the same lysine residue is involved, although it does not exclude the added possibility of the liberation of a free lysine residue from another segment of the A protein. The relationship among TP10, TP27, and TP21 is illustrated in Fig. 7.

Peptide TP28 (Residues 263 through 267): Ala-Ala-Thr-Arg-Ser (4 Residues)

The sequence studies with this peptide were reported previously (3). It represents the carboxyl-terminal segment of the A protein. It appears that the occurrence of such a peptide in trypsic digests is due to the poor exopeptidase activity of trypsin, resulting in incomplete cleavage of the Arg-Ser bond. The tryptic Peptide TP20 and free serine (TP1) (see following peptides) are recovered as major fragments in trypsin digests of the A protein, and they comprise the carboxyl-terminal segment of the protein (3).

Peptide TP29 (Residues 265 through 267): His-Asn-Val-Ala-Pro-Ile-Phe-Ile-Cys-Pro-Pro-Asn-Ala-Asp-Asp-Leu-Arg (19 Residues)

The sequence studies with this peptide were reported earlier in a discussion of the amino acid sequences surrounding the sulphydryl groups of the A protein (7).

Discussion

By using the procedures described, a total of 29 peptides were isolated from trypsin digests of the A protein of tryptophan synthetase (Table I). Of these, sequence studies revealed that four were overlapping tryptic peptides arising from incomplete hydrolysis of specific bonds involving lysine or arginine residues, and one was a chymotryptic fragment of a tryptic peptide. The amino acid sequences of the 29 peptides are presented in Table II. Also summarized are the sequence studies with TP15 and TP26 described in the accompanying paper (8). Except for
these two peptides, all of the peptides described have been isolated reproducibly from many batches of A protein. It can be seen in Table II that among the nonoverlapping peptides there are 12 lysine-containing peptides, 11 arginine-containing peptides, and a carboxyl-terminal fragment (TP1, serine). They comprise a total of 267 amino acid residues (Table I). There is reasonable agreement between the total amino acid composition of these tryptic peptides and the total number of residues of each amino acid in the A protein, estimated after acid hydrolysis (Table I). Moreover, as is discussed in the accompanying paper (8), the agreement would be better had a slightly lower molecular weight been taken for estimation of total amino acid composition.

The standard procedures for determining amino acid sequences proved to be very satisfactory in the present work. The improved method of Edman degradation described by Konigsberg and Hill (14) was used for the greater part of this work, and five or six cycles could generally be completed without ambiguity, in contrast to the earlier methods used, for example, with TP8, TP11, and TP16. Most of the peptides yielded to the combined approach of the subtractive Edman technique and fragmentation with chymotrypsin or papain. Confirmatory evidence for the established sequences was often obtained by digestion with carboxypeptidase or leucine aminopeptidase, especially in the case of proline-containing peptides. These exopeptidases proved to be most useful in differentiating between aspartic and glutamic acids and their amides. Assignation of amide groups also rested upon chromatographic identification of the PTH-amino acid liberated by Edman degradation and the electrophoretic mobility of peptides or peptide fragments at a neutral pH. Reservations about the reliability of appointing amide groups have been discussed previously (7); with regard to their occurrence in the native protein, positive evidence for their presence in a particular peptide is more reliable than evidence for their absence, since this could simply be due to deamidation either during digestion or during isolation.

In the following papers, information on the remaining peptides and the sequence of the peptides in the protein will be presented.

Acknowledgments—The authors are indebted to George Pegelow, Patricia Greene, Deanna Thorpe, and Donald Vinicor for skilled technical assistance.

Peptide TP26 contains 2 lysine residues.

REFERENCES

The Amino Acid Sequence of the A Protein (α Subunit) of the Tryptophan Synthetase of *Escherichia coli*: 1. Tryptic Peptides

John R. Guest, Bruce C. Carlton and Charles Yanofsky


Access the most updated version of this article at [http://www.jbc.org/content/242/22/5397](http://www.jbc.org/content/242/22/5397)

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/242/22/5397.full.html#ref-list-1](http://www.jbc.org/content/242/22/5397.full.html#ref-list-1)