Primary Structure of Streptococcal Proteinase

III. ISOLATION OF CYANOCEN BROMIDE PEPTIDES: COMPLETE COVALENT STRUCTURE OF THE POLYPEPTIDE CHAIN*

(Received for publication, August 6, 1975)

JOSEPH Y. TAI,§ ALEXANDER A. KORTT,* T.-Y. LIIU,**§ AND S. D. ELLIOTT†

From the Division of Bacterial Products, Bureau of Biologics, Bethesda, Maryland 20014; the Biology Department, Brookhaven National Laboratory, Upton, New York 11973; and the Rockefeller University, New York, New York 10021

The following sequence has been derived for streptococcal proteinase.

10 Gln-Pro-Val-Val-Lys-Ser-Leu-Leu-Asp-Ser-Lys-Gly-Ile-His-Tyr-Asn-Gln-Gly-
20 Asn-Pro-Tyr-Asn-Leu-Leu-Thr-Pro-Val-Ile-Glu-Lys-Val-Lys-Pro-Gly-Glu-Gln-
30 Ser-Phe-Val-Gly-Gln-Ala-Ala-Tin-Gly-His-Cys-Val-Ala-Thr-Ala-Thr-Ala-Gln-
40 Ile-Met-Lys-Tyr-His-Asn-Tyr-Pro-Asp-Lys-Gly-Leu-Yrs-Lys-Tyr-Thr-Thr-
50 Leu-Ser-Ser-Asn-Pro-Asp-Tyr-Phe-Asp-His-Pro-Lys-Leu-Phe-Ala-Ala-Ile-
60 Ser-Thr-Arg-Gln-Tyr-Asp-Trp-Asn-Asn-Ile-Leu-Pro-Thr-Tyr-Ser-Gly-Arg-Gln-
70 Ser-Gln-Asn-Val-Lys-Met-Ala-Ile-Ser-Glu-Leu-Met-Ala-Asp-Val-Gly-Ile-Ser-
80 Val-Asp-Met-Asp-Tyr-Gly-Pro-Ser-Ser-Gly-Ser-Ala-Gly-Ser-Ser-Arg-Val-Gln-
90 Arg-Ala-Leu-Lys-Glu-Asp-Phe-Gly-Tyr-Asn-Gln-Ser-Val-His-Gln-Ile-Asp-Ang-
100 Gly-Asp-Phe-Ser-Lys-Gln-Asp-Str-Pgl-Glu-Leu-Ser-Gln-
110 Asn-Gln-Pro-Val-Tyr-Glu-Gly-Val-Val-Gly-Lys-Val-Gly-Gly-His-Ala-Phe-Val-
120 Ile-Asp-Gly-Ala-Gly-Arg-Asn-Phe-Tyr-His-Asp-Val-Trp-Gly-Trp-Gly-Gly-
130 Val-Glu-Asp-Ala-Phe-Arg-Leu-Asp-Ala-Leu-Asp-Pro-Arg-Leu-Asn-Asp-
140 Gly-Asp-Gly-Ala-Gly-Asp-Gly-Tyr-Glu-Glu-Asp-Val-Thr-
150

The sequence permits the assignment of the single cysteine residue essential for catalytic action at position 47 from the NH₂ terminus of the protein. The tryptophan residue at the binding site of the enzyme is at position 214. A histidine residue at position 195 has been assigned as the catalytically important entity in the molecule. Streptococcal proteinase and papain, an enzyme with similar properties, are compared with respect to structure and function.

The preceding communications (1, 2) report the sequence studies of cyanogen bromide fragments of streptococcal proteinase.

* This research was carried out at Brookhaven National Laboratory under the auspices of the United States Atomic Energy Commission; † the National Institute of Child Health and Human Development, Section on the Biochemistry of Microbial Structure, National Institutes of Health. Present address, Division of Bacterial Products, Bureau of Biologics, Bethesda, Maryland 20014.

†§ and at the Rockefeller University, New York, N. Y. with the support of United States Public Health Service Grant HL-03929 from National Institutes of Health.

EXPERIMENTAL PROCEDURE

The materials and procedures were those described in the preceding papers (1, 2) with the following additions. Cyanogen bromide was obtained from Aldrich. Iodo[¹³C]iactamamide was purchased from New
**Isolation and Amino Acid Composition of CB Peptides—**

Five peptides were isolated by the procedures summarized in Fig. 1, A and B, and Table I. The amino acid compositions are given in Table II. The results come closest to fitting a molecule of about 300 residues. Some discrepancies merit comment. The amino acid composition of this peptide is 3.

**RESULTS**

**Isolation and Amino Acid Composition of CB Peptides—**

Five peptides were isolated by the procedures summarized in Fig. 1, A and B, and Table I. The amino acid compositions are given in Table II. The results come closest to fitting a molecule of about 300 residues. Some discrepancies merit comment. The initial amino acid composition of the proteinase was calculated on the basis of 1.0 residue of half-cystine. The results come closest to fitting a molecule of about 300 residues. Some discrepancies merit comment. The amino acid composition of this peptide is 3.

**Summary of isolation of cyanogen bromide peptides**

The fractions are those indicated in Fig. 1, A and B. The peptides are numbered in the order of their occurrence in the chains: the amino acid compositions are given in Table II and the sequence in Fig. 2.

---

**Fig. 1. Separation of cyanogen bromide peptides. A, gel filtration of cyanogen bromide cleavage products of streptococcal proteinase (100 mg) on Sephadex G-50 (fine). Column, 2.5 × 300 cm; eluent, 20% acetic acid, flow rate 26 ml/hour. Fractions of 4.2 ml were collected, and 0.1-ml aliquots of each fraction were analyzed with ninhydrin after alkaline hydrolysis. B, chromatography of fraction II (50 mg) from A on sulfoethyl Sephadex C-25. A column (1.6 × 80 cm) was developed with a linear gradient of 0.2 M pyridine acetate, pH 3.1 to 2.0 M pyridine acetate, pH 5.0, 300 ml each, at 25°C. Fractions of 3.0 ml were collected, and 0.1 ml of each fraction was hydrolyzed in alkali and analyzed with ninhydrin. Bars and numbers show pooled fractions.**

**Tables**

**Table I**

| Peptide | Sephadex G-50 | SE-Sephadex | Yield |
|---------|---------------|-------------|-------|
| CB 1    | II            | I           | 88    |
| CB 2    | II            | III         | 93    |
| CB 3    | III           | II          | 68    |
| CB 4    | IV            | III         | 61    |
| CB 5    | I             | I           | 55    |

---

*The abbreviations used are: CB, cyanogen bromide; Cm-methionine, S-carboxamidomethyl methionine.
Primary Structure of Streptococcal Proteinase

1957

SEPHADEX G-66

FIG. 3. Separation of the tryptic peptides from [14C]Cm-methionyl streptococcal proteinase on a Sephadex G-50 column (fine). Column, 2.5 x 190 cm; eluent, 20% acetic acid; flow rate, 12 ml/hour. Fractions of 2.9 ml were collected, of which two 0.1 ml aliquots were used for both ninhydrin analysis after alkaline hydrolysis (solid line) and radioactivity measurement (dashed line).

Fig. 2. Amino acid sequence of streptococcal proteinase. The amino acids are characterized as H (strong former), h (medium former), I (weak former), i (indifferent), b (breaker), and B (strong breaker) of helices in the first row and 5 structures in the second row. Predicted secondary structural regions are enclosed in boxes.

TABLE II
Amino acid compositions of CB peptides and streptococcal proteinase

Results of 22-, 48-, and 72-hour hydrolyses of peptides at 115° in 4 M methanesulfonic acid (4); serine and threonine were corrected by extrapolation. The 72-hour values were used for valine and isoleucine. The integral numbers are the results of the determination of the sequence. The abbreviations used are: Anal., analysis; seq., sequence.

procedure summarized in Figs. 3 and 4, and Table III. The amino acid compositions are given in Table IV. Based on the established sequence in Fig. 2, the overlap tryptic peptides for CB1-CB2 and for CB3-CB4-CB5 were constructed and their amino acid compositions were calculated. As shown in Table IV, isolated peptides have identical compositions with the corresponding constructed peptides.

The amino acid compositions of peptides I-A and I-B indicated that each came from the same segment of the polypeptide chain. Similarly, peptides IIA, B, C, and D could have come only from the same segment of the molecule. The diversities of these peptides are due most likely to the variety of

Downloaded from http://www.jbc.org/ by guest on March 23, 2020
possible alkylation products. Since peptide II contains three methionine residues, different degrees of alkylation would produce a wide variety of charges on the peptide. This is reflected in the complex behavior of these peptides on SP-Sephadex and on paper electrophoresis.

Alignment of Cyanogen Bromide Fragments—The sequence for streptococcal proteinase presented in Fig. 2 is based upon all of the information which we have obtained. The results can be examined by starting with residue 1 and citing the points in the evidence which merit special mention. The earlier studies of streptococcal proteinase (7) established glutamine as the NH2-terminal residue. Only one of the peptides of the CB series, CB1, begins with glutamine.

The isolation and sequence studies of the active site peptide containing the sole half-cystine residue in the protein (10) provides a point of reference for aligning CB1 and CB2. The [14C]Cm-methionine peptides I-A and I-B are from the same segment of the proteinase. These peptides have a composition that is identical with the sum of the COOH-terminal tryptic peptide of CB1 and the NH2-terminal residue (lysine) of CB2. Moreover, the NH2-terminal sequence of 1-A and 1-B determined by the dansyl procedure has the sequence Val-Lys, identical with the active site peptide (10).

From the composition of the peptides [14C]Cm-Met II (Table IV) and the sequence information obtained for the regenerated peptide II, Met-Ala-Ile-Ser as was shown under "Experimental Procedure" this peptide must extend from the COOH-terminal portion of CB2 through CB3 and into CB4. CB5, which contains no homoserine, is the COOH-terminal peptide. Liu and Elliott (8) found proline to be the COOH-terminal residue. Based on the above alignment data and the knowledge that CB5 represents the COOH-terminal fragment of streptococcal proteinase, the alignment of the five cyanogen bromide fragments is as follows: 1-2-3-4-5.

DISCUSSION

The complete amino acid sequence of streptococcal proteinase is shown in Fig. 2. Under each amino acid are the helical (a) potentials in the first row and b structure potentials (b) in the second row, based on protein conformation prediction rules derived by Chou and Fasman (11, 12). The conformational parameters Pa and Pb for amino acids in helices and b structures, respectively, were utilized to predict helical regions and b structures in the proteinase. Inspection of the models (Fig. 2) indicated that eight a helical regions are detected and accounted for 25% of the amino acid residues in the protein. There are 10 b sheet regions in the molecule accounting for 22% of the amino acid residues in the protein. The result in Fig. 2

| Peptide  | Sephadex G-50  | SP-Sephadex C 25 |
|----------|----------------|------------------|
|          | Figure Fraction | Figure Fraction  | Electrophoresis at pH  |
| CM-Met-I-A | 3 4 IV | 1.9 | 3% |
| CM-Met-I-B | 3 1 | 3.5 | 2% |
| CM-Met-II-A | 3 4 | 3.5 | 3% |
| CM-Met-II-B | 3 1 | 3.5 | 6% |
| CM-Met-II-C | 3 4 | 3.5 | 6% |
| CM-Met-II-D | 3 4 | 3.5 | 1% |

**TABLE IV**

Amino acid composition of Cm-Met peptides from streptococcal proteinase

| Amino Acid | CM-Met-I-A | CM-Met-I-B | Sequence (31-57) | CM-Met-II-A | CM-Met-II-B | CM-Met-II-C | CM-Met-II-D | Sequence (114-142) | II-Ra Regeneration Peptide |
|------------|------------|------------|------------------|------------|------------|------------|------------|------------------|-------------------------|
| Lysine     | 1.89(2)    | 1.70(2)    | 2                | 1.20(1)    | 0.80(1)    | 0.92(1)    | 0.79(1)    | 1                | 0.95(1)                |
| Histidine  | 1.08(1)    | 0.83(1)    | 1                | 3.84(3)    | 3.41(3)    | 2.95(3)    | 2.76(3)    | 3                | 3.11(3)                |
| Arginine   | 1.24(1)    | 1.00(1)    | 1                | 6.85(7)    | 6.76(7)    | 7.01(7)    | 6.54(7)    | 7                | 6.20(7)                |
| Aspartic Acid | 4.36(4) | 4.94(4) | 4                | 2.15(1)    | 2.64(1)    | 3.54(1)    | 3.33(1)    | 1                | 1.40(1)                |
| Serine     | 0.81(1)    | 1.10(1)    | 1                | 1.10(1)    | 1.02(1)    | 0.86(1)    | 1.15(1)    | 1                | 1.14(1)                |
| Glutamic Acid | 2.30(3) | 4.02(3) | 3                | 4.39(4)    | 4.18(4)    | 4.14(4)    | 4.30(4)    | 4                | 3.20(3)                |
| Proline    | 4.66(5)    | 5.16(5)    | 5                | 3.15(3)    | 3.02(3)    | 3.40(3)    | 3.33(3)    | 3                | 3.20(3)                |
| Glycine    | 0.81(1)    | 1.10(1)    | 1                | 1.24(1)    | 1.34(1)    | 1.08(1)    | 1.15(1)    | 2                | 2.00(2)                |
| Alanine    | 2.71(2)    | 2.65(3)    | 3                | 1.54(2)    | 1.80(2)    | 2.06(2)    | 1.91(2)    | 2                | 2.00(2)                |
| Valine     | N.D.       | N.D.       | 1                | N.D.       | N.D.       | N.D.       | N.D.       | 3                | 2.15(3)                |
| Methionine | N.D.       | N.D.       | 1                | N.D.       | N.D.       | N.D.       | N.D.       | 3                | 2.15(3)                |
| Isoleucine | 0.87(1)    | 0.88(1)    | 1                | 1.84(2)    | 1.94(2)    | 1.94(2)    | 1.76(2)    | 2                | 2.05(2)                |
| Leucine    | 0.91(1)    | 0.96(1)    | 1                | 0.78(1)    | 0.69(1)    | 0.78(1)    | 0.69(1)    | 1                | 1.10(1)                |
| Tyrosine   | 1.00(1)    | 1.00(1)    | 1                | 0.88(1)    | 0.68(1)    | 0.88(1)    | 0.68(1)    | 1                | 1.08(1)                |
| Phenylalanine | 0.80(1) | 0.89(1) | 1                | 0.80(1)    | 0.89(1)    | 1          | 0.80(1)    | 1                | 1.08(1)                |

*a* From combined peptides Cm-Met-II-A and Cm-Met-II-B to regenerate methionine residues.

*Threonine and serine were corrected for 5 and 10% approximate destruction during acid hydrolysis.*

*High values of glutamic acid were due to the mixture of glutamic acid and homoserine which were not separated by the amino acid analyzer used in this experiment.

*N.D., not determined.*
showed no significant repeating sequences in the molecule. The uneven distribution of residues over the chain is marked; tryptophan, phenylalanine, glycine, and arginine are predominantly in the COOH-terminal half of the protein, whereas 9 out of 10 threonine residues are in the NH₂-terminal half.

The sequence in Fig. 2 permits the location of residues in protease that have special reactivities. The catalytically essential sulphydryl group (10) is located at position 47, which is at the beginning of β sheet region 47-56. Robinson (13) showed that treatment of streptococcal protease with 2-hydroxy-5-nitrobenzyl bromide resulted in the modification of a single tryptophan residue and a concomitant loss of enzymic activity. From a tryptic hydrolysate, he isolated a peptide with a chymotryptic-like cleavage at an aromatic residue at position 208 (2). Thus, the tryptophan residue believed to be at the binding site of the enzyme is Trp₄₀⁸.

Liu showed (14) that 1 histidine residue is specially reactive toward α-N-bromoacetyl-L-arginine methyl ester and that the alkylated protein is inactive. Tai and Liu* have assigned His₁₄₇ at the boundary of the helix-coil region 195-202 to be the catalytically essential histidine residue in the protease. The propensity for polar active site residues at the helix-coil boundaries of enzymes has been attributed (11) to the fact that these regions are more flexible than the rigid inner helix core, thus facilitating substrate binding as well as enzyme catalysis.

Streptococcal protease is a sulphydryl protease similar to papain. The two enzymes have similarities as well as differences (8, 15). Similarities include an active site with a highly reactive —SH group (14, 16, 17, 18), similar specificity toward the phenylalanine-chain of insulin (19, 20), and activity in 8 M urea (7, 21). The differences reside in size (253 amino acid residues for protease and 219 residues for papain), in the absence of S-S cross-links in the streptococcal protease, and in the absence of any special similarities in the amino acid compositions of the two proteins. The protease exhibits a high ratio of esterase to peptidase activity (22), whereas papain hydrolyzes both types of substrates at about the same rate (23).

The primary structural studies of the two enzymes revealed the following similarities. The highly reactive —SH group in both enzymes is in the NH₂-terminal portion of the protein (papain, Cys₁₁₀; protease Cys₁₁₀), while the tryptophan residue believed to be at the binding site of the enzyme is positioned at the COOH-terminal portion of the molecule (papain, Trp₁₅₄; protease, Trp₁₅₄). The catalytically crucial histidine residue in papain (His₁₄₇) and in the protease (His₁₄₇) is at about the same distance (18 or 19 amino acid residues) from the crucial tryptophan residue in each protein. Moreover, the sequence around the reactive cysteine of both enzymes is Gly X Cys (Fig. 6) as is the case for all known thiol proteases (24). In both enzymes the reactive tryptophan residue is next to a glycine residue (Fig. 7), and the catalytically important histidine residue is next to an alanine residue.

* Unpublished observation.
Primary structure of streptococcal proteinase. III. Isolation of cyanogen bromide peptides: complete covalent structure of the polypeptide chain.

J Y Tai, A A Kortt, T Y Liu and S D Elliott

J. Biol. Chem. 1976, 251:1955-1959.

Access the most updated version of this article at http://www.jbc.org/content/251/7/1955

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/251/7/1955.full.html#ref-list-1