AMINO ACID SEQUENCE OF CYANOGEN BROMIDE FRAGMENTS CB4 AND CB6 OF HOG PEPSIN

Ladislav MORÁVEK

Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Sciences, 166 10 Prague 6

Received September 17th, 1979

By the analyses of chymotryptic, tryptic, and thermolytic peptides the amino acid sequence was determined of cyanogen bromide fragment CB4 representing the region of the pepsin chain between the N-terminus and methonine-residue I: Ile-Gly-Asp-Glu-Pro-Leu-Glu-Asn-Tyr-Leu--Asp-Thr-Glu-Tyr-Phe-Gly-Thr-Ile-Gly-Ile-Gly-Thr-Pro-Ala-Gln-Asp-Phe-Thr-Val-Ile-Phe-Asp--Thr-Gly-Ser-Ser-Asn-Leu-Trp-Val-Pro-Ser-Val-Tyr-Cys-Ser-Ser-Leu-Ala-Cys-Ser-Asp-His-Asn-+-Gln-Phe-Asn-Pro-Asp-Asp-Ser-Ser-Thr-Phe-Glu-Ala-Thr-Ser-Gln-Glu-Leu-Ser-Ile-Thr-Tyr-Gly-+-Thr-Gly-Ser-Met. The serine residue (Ser) in position 68 of pepsin is phosphorylated. By sequential analysis of chymotryptic, tryptic, and thermolytic peptides the amino acid sequence was determined of cyanogen bromide fragment CB6 representing the region between methionine residues II and III in the pepsin chain: Asp-Gly-Glu-Thr-Ile-Ala-Cys-Ser-Gly-Gly-Cys-Gln-Ala-+-Ile-Val-Asp-Thr-Gly-Thr-Ser-Leu-Leu-Thr-Gly-Pro-Thr-Ser-Ala-Ile-Ala-Asn-Ile-Gln-Ser-Asp-+-Ile-Gly-Ala-Ser-Glu-Asn-Ser-Asp-Gly-Glu-Met. The aspartic acid residue (Asp) in position 16 of this fragment is identical with the residue reacting with diazo inhibitors which forms a part of the active center of the enzyme. Both half-cystine residues of fragment CB4 and fragment CB6 are linked to one another by a disulfide bond in native pepsin.

In the course of studies on the covalent structure of hog pepsin in this Laboratory the cyanogen bromide digest[1] of the S-sulfo derivative of the protein was prepared. Fractions FCB1 to FCB5 (ref.[2]) containing fragments CB1 to CB6 were obtained by gel filtration of the digest. Sequential analysis of the cyanogen bromide fragments afforded basic information permitting the complete primary structure of the protease, which has been published earlier, to be defined[3]. This paper reports on experimental data of the sequential determination of fragments CB4 and CB6. The investigation of N-terminal fragment CB4 extends the N-terminal 55-residue sequence of pepsin determined earlier[4]. Fragment CB6 represents the second component of fraction FCB5 of the cyanogen bromide digest in addition to fragment CB5 reported earlier[5].

EXPERIMENTAL

Material

Pepsin was a twice crystallized preparation of Worthington Biochemical Corporation, Freehold, N.J., U.S.A. The preparation of S-sulfo-pepsin, its cyanogen bromide cleavage, and the separa-

tion of the cyanogen bromide fragments by gel filtration have been described earlier[2]. The components of fraction FCB5 have been resolved in the preceding study[5]. Fragment CB6 emerged in peak "A" during ion exchange chromatography in urea-containing solutions. In an alternative separation in solutions not containing urea fragment CB6 was found in peak "1Va". This preparation was rechromatographed on DEAE-Sephadex under conditions identical with those employed for the original separation[5]; the resulting preparation was marked "rIVa". Chymotrypsin was prepared by the activation of its zymogen obtained by five crystallizations of a commercial product of Léčiva, Prague. TPCK-trypsin[6] (trypsin treated with 1-chloro-4-phenyl-3-tosylamino-2-butanone) was from Worthington. Thermolysin was a B-grade product of Calbiochem, San Diego, U.S.A. Soybean trypsin inhibitor, ethylene imine, and 2-amino-2-methyl-1,3-propanediol were from Koch-Light Laboratories Ltd. (London, England). Cyanogen bromide, 2-mercaptoethanol, and Dowex 1-X2, 200—300 mesh were from Fluka, Buchs, Switzerland. Sephadex G-100, G-25 fine, and DEAE-Sephadex A-25 were products of Pharmacia, Uppsala, Sweden. Silufol, silica gel layer sheets for thin-layer chromatography, was a product of Kavalier, Sázava, Czechoslovakia. All the remaining chemicals were of analytical purity.

Methods

The preparation and separation of the thermolytic, and tryptic digest of aminoethylated fraction FCB5 (ref.[5]) and of the thermolytic digest of fraction CB2 (ref.[4]) were described before. The final purification of the peptides was carried out by paper chromatography in the system 1-butanol–pyridine–acetic acid–water (15 : 10 : 3 : 12, v/v) and by two electrophoretic procedures on Whatman No 3 or 3MM paper. The separation at pH 1·9 in the system formic acid–acetic acid–water (50 : 150 : 800, v/v) at 70 V/cm was carried out in the horizontal arrangement in the apparatus developed by Prusik and coworkers[7]. The separation at pH 5·6 in the system pyridine–acetic acid–water (5 : 1 : 494, v/v) was carried out at 25 V/cm in the vertical arrangement in the apparatus designed by Mikes[8]. The two-dimensional peptide maps of the chymotryptic digest of samples of high molecular weight material were prepared by a combination of electrophoresis at pH 1·9 (first direction) and paper chromatography in the system described above (second direction). The methods of the detection of peptides, amino acid analysis of the investigated material, and determination of amino acid sequences were described earlier as well as the methods of selective isolation of some half-cystine peptides from fraction FCB5 (ref.[5]).

Preparation of chymotryptic digest of fragment CB4: fraction FCB4 of the cyanogen bromide digest contains one main component only, fragment CB4 (ref.[2]). The lyophilized material (100 mg) was dissolved to a 2% solution in water made alkaline by the addition of 0·1M-NH_4HCO_3 (red color of phenol red used as an indicator) and digested by two portions of chymotrypsin (final molar enzyme to substrate ratio 1 : 50) 4 h at 37°C. The digest was subsequently fractionated by gel filtration on a column of Sephadex G-15 (65 × 5 cm), equilibrated with 0·05M-NH_4HCO_3. The fractions (20 ml/15 min) were evaluated by paper chromatography of aliquots (0·2 ml). The final purification of the peptides was effected by paper chromatography and by the two electrophoretic procedures described above.

Preparation of chymotryptic digest of fraction CB6: The material in S-sulfo form was digested with chymotrypsin at a molar enzyme to substrate ratio of 1 : 50, 4 h at 37°C, in a solution made alkaline by 0·1M-NH_4HCO_3 (red. color of phenol red used as an indicator). Portions (2 mg) of the chymotryptic digest were subjected to two-dimensional separation by electrophoresis (pH 1·9) and paper chromatography. The spots of peptides stained with ninhydrin were cut out, eluted by water, and their amino acid composition determined. The map of the chymotryptic digest of fragment CB6 is shown in Fig. 1. The amino acid composition of the peptides derived

from the region of fragment CB4 is shown in Table I and of peptides derived from fragment CB6 in Table II.

Purification of fragment CB4 *and of other fragments of pepsin by ion-exchange chromatography*: Selected fractions of the cyanogen bromide digest, obtained by gel filtration, were purified by ion-exchange chromatography on DEAE-Sephadex A-25. The course of the separation is shown in Fig. 2. Fragment CB4 was obtained in this case from pooled fractions FCB3 and FCB4 after separation by gel filtration without recycling. Fragment CB1 was a preparation purified by recycling[2]. Fragment CB2-RAE-T11b was obtained in a study on tryptic fragments from the N-terminal region of pepsin[9].

TABLE I

Amino Acid Composition of Peptides Derived from the Region of Fragment CB4

The peptides are marked by the symbol of the enzyme used for the cleavage: "Th" and "C" denote the thermolytic and chymotryptic hydrolysate, respectively. (The unlisted peptides Th1 through 13 and C1 through 10 belong to the N-terminal region of the fragment[4] determined before.) The peptides were analyzed after 20 h hydrolysis; the values are not corrected. None of the peptides contained lysine, arginine, or tryptophan. Peptides Th15, Th16, C12 and C13 contain the phosphoserine residue reacting with the Hanes–Isherwood reagent[11].

Designation of peptide	Number of amino acid residues													
	His	Cys	Asp	Thr	Ser	Glu	Pro	Gly	Ala	Val	Ile	Leu	Tyr	Phe	Hse[a]
Th14	—	—	3·1	1·0	1·9	—	0·8	—	—	—	—	—	—	0·9	—
Th15	—	—	—	1·0	0·9	3·0	—	—	1·0	—	—	—	—	1·0	—
Th16	—	—	—	1·0	1·8	2·7	—	—	1·0	—	—	0·9	—	0·9	—
Th17	—	—	—	—	—	1·1	—	—	—	—	—	—	—	1·0	—
Th18	—	—	—	—	1·0	—	—	—	—	—	—	1·0	—	—	—
Th19	—	—	—	2·8	1·2	—	—	2·8	—	—	0·9	—	0·5	—	1·0
Th20	—	—	—	2·0	1·1	—	—	2·0	—	—	1·0	—	0·9	—	—
Th21	—	—	—	0·9	—	—	—	1·0	—	—	—	—	—	—	0·7[b]
Th22	—	—	—	1·0	—	—	—	—	—	—	1·0	—	0·6	—	—
Th23	—	—	—	2·0	1·2	—	—	3·0	—	—	—	—	—	—	0·8
Th24	—	—	—	1·0	—	—	—	1·1	—	—	1·0	—	0·9	—	—
Th25	—	—	—	1·0	1·0	—	—	1·9	—	—	—	—	—	—	—
C11	1·0	1·7	5·1	1·1	4·6	1·1	1·0	—	1·1	—	—	0·9	—	2·2	—
C12	—	—	—	1·9	2·0	2·7	—	—	1·0	—	0·8	1·0	0·6	—	—
C13	—	—	—	1·1	1·1	2·8	—	—	1·0	—	—	1·0	—	—	—
C14	—	—	—	1·1	1·0	—	—	1·9	—	—	—	—	—	—	0·8
C15	—	—	—	2·0	1·0	—	—	2·9	—	—	1·0	1·0	—	—	1·1
C16	—	—	—	1·9	1·0	—	—	3·6	—	—	1·0	1·0	0·5	—	1·2

[a] Hse stands for homoserine; [b] determined as methionine (ref.[10]).

RESULTS AND DISCUSSION

Amino Acid Sequence of Fragment CB4

It was shown[2] in an earlier study of the cyanogen bromide digest of S-sulfo-pepsin that fragment CB4 represents the N-terminal region of pepsin. Because of incomplete cleavage at the site of the first methionine residue, a large fragment CB2 was isolated in which the N-terminal fragment CB4 is linked with another fragment, namely with fragment CB3. Fragment CB4 contains two half-cysteine residues and the only histidine residue of pepsin. In an earlier study[9] of the N-terminal region of pepsin

TABLE II

Amino Acid Composition of Peptides Derived from the Region of Fragment CB6

The peptides are marked by the symbol of the enzyme used for the cleavage: "T", "Th", and "C" denote the tryptic, thermolytic and chymotryptic digest, respectively. The peptides were analyzed after 20 h hydrolysis. None of the peptides contained lysine, histidine, arginine, or tryptophan.

Designation of peptide	Number of amino acid residues													
	Cys	Asp	Thr	Ser	Glu	Pro	Gly	Ala	Val	Ile	Leu	Tyr	Phe	Hse[a]
Th1	—	1·0	0·9	—	1·0	—	1·0	—	—	—	—	—	—	—
Th2	1·8	—	—	1·1	1·1	—	2·0	1·9	—	1·0	—	—	—	—
Th3	—	0·9	1·9	1·2	—	—	1·2	—	0·8[c]	0·7[c]	—	—	—	—
Th4	—	—	—	—	—	—	—	—	—	—	1·0	—	—	—
Th5	—	—	1·9	1·1	—	0·9	1·0	1·0	—	—	1·0	—	—	—
Th6	—	1·1	—	—	—	—	—	1·0	—	1·0	—	—	—	—
Th7	—	1·1	—	1·0	1·0	—	—	—	—	0·9	—	—	—	—
Th8	—	2·2	—	1·8	2·1	—	1·9	1·0	—	0·9	—	—	—	1·0
Th9	—	1·0	1·9	0·8	—	—	1·0	—	1·0[c]	0·8[c]	1·0	—	—	—
Th10	—	2·3	—	2·0	2·0	—	1·1	0·8	—	—	—	—	—	0·8[d]
C1	1·9	2·0	2·9	2·0	2·1	—	4·3	2·2	0·8	1·7	2·0	—	—	—
C2	—	1·0	1·8	1·0	—	1·1	1·0	2·0	—	1·1	—	—	—	—
C3	—	3·2	—	2·7	2·8	—	1·9	1·1	—	2·0	—	—	—	1·0
C4	2·0	1·1	1·0	1·1	2·0	—	3·0	1·1	—	1·0	—	—	—	—
C5	—	1·2	1·8	0·9	—	—	1·0	0·8	0·8	0·9	2·0	—	—	—
T1	1·2[b]	0·9	0·9	1·1	1·1	—	2·9	0·8	—	1·0	—	—	—	—
T2	0·8[b]	1·0	0·9	—	1·0	—	1·0	1·0	—	1·0	—	—	—	—
T3	1·0[b]	—	—	0·8	—	—	2·1	—	—	—	—	—	—	—

[a] Hse stands for homoserine; [b] determined as aminoethylcysteine; [c] determined after 70 h hydrolysis; [d] determined as methionine[10].

the large fragment CB2 was reduced and aminoethylated, cleaved with trypsin at the two aminoethylcysteine residues, and the resulting three fragments were isolated. The material obtained permitted the N-terminal 55-residue sequence

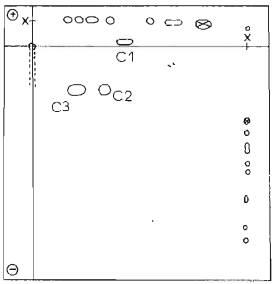

FIG. 1

Positions of Peptides on Chymotryptic Peptide Map of Fragment CB6 (in S-sulfo form)

The chymotryptic digest of fragment CB6 was resolved by electrophoresis at pH 1·9 in the first direction and by paper chromatography in the second direction. Neutral red and ε-DNP-lysine, the markers for paper chromatography and electrophoresis, are denoted by ⊗. A reference amino acid mixture was applied for each corresponding separation run at sites marked x. The peptides are marked by symbols used in Table I.

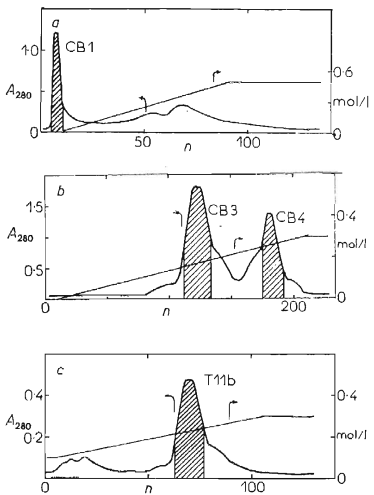

FIG. 2

Ion-Exchange Chromatography of Pepsin Fragments

a Chromatography of fraction FCB1. The sample (100 mg) was applied to a column of DEAE-Sephadex A-25 (30 × 1·5 cm) equilibrated with 0·05M Tris-HCl (pH 6·0), containing 8M urea. Elution: linear concentration gradient of NaCl in the same buffer (total volume 500 ml). Fractions 6 ml/15 min, *n* fraction number. *b* Chromatography of pooled fractions FCB3 and FCB4. The material (1·8 g) was separated on a column of DEAE-Sephadex A-25 (100 × 2·5 cm), equilibrated with 0·1M Tris-HCl (pH 7·5) containing 8M urea. Elution: linear concentration gradient of NaCl in the same buffer (total volume 3000 ml). Fractions 15 ml/20 min, *n* fraction number. *c* Chromatography of fraction CB2-RAE-T11b (T11b for short, ref.[9]). The sample (100 mg) was separated on a column of DEAE-Sephadex A-25 (30 × 1·5 cm) equilibrated with 0·05M Tris-HCl (pH 7·0) containing 8M urea. Elution: linear concentration gradient of NaCl in the same buffer (total volume 600 ml). Fractions 6 ml/15 min, *n* fraction number.

TABLE III

Amino Acid Composition of Fragment CB4

The sample was analyzed after 20 h hydrolysis; the values for threonine and serine were extrapolated to zero time of hydrolysis. Arginine and lysine were not detected in the sample. The results are expressed as the number of amino acid residues.

Amino acids	His	Cys	Asp	Thr	Ser	Glu	Pro	Gly	Ala	Val	Ile	Leu	Tyr	Phe	Hse	Trp
CB4	0·7	1·7	11·0	8·5	10·8	7·7	3·8	7·3	3·1	3·2[a]	5·0[a]	5·2	3·8	4·6	0·9	+[b]
In sequence of CB4	1	2	11	9	11	8	4	7	3	3	5	5	4	5	1	1

[a] Values after 70 h hydrolysis; [b] the presence of tryptophan was determined by the qualitative test only.

TABLE IV

Sequence of Peptides from the Region of Fragment CB4

Designation of peptide	Amino acid sequence of peptide
Th14	Phe-Asn-Pro-Asp-Asp-Ser-Ser-Thr
Th15	Phe-Glu-Ala-Thr-Ser-Gln-Glx
Th16	(Phe,Glx,Ala,Thr,Ser,Glx,Glx,Leu,Ser)
Th17	Phe-Glu
Th18	Leu-Ser
Th19	Ile-Thr-Tyr-Gly(Thr,Gly,Ser,Hse,Thr,Gly)
Th20	Ile-Thr-Tyr-Gly-Thr-Gly-Ser
Th21	Met-Thr-Gly
Th22	Ile-Thr-Tyr
Th23	Gly-Thr-Gly-Ser-Hse-Thr-Gly
Th24	Ile-Thr-Tyr-Gly
Th25	Gly-Thr-Gly-Ser
C11	(His,Cys_2,Asx_5,Thr,Ser_5,Glx,Pro,Ala,Leu,Phe)Phe
C12	Glu-Ala-Thr-Ser-Gln-Glu-Leu-Ser(Ile,Thr)Tyr
C13	(Glx,Ala,Thr,Ser,Glx,Glx)Leu
C14	(Gly,Thr,Gly,Ser)Hse
C15	Gly(Thr,Gly,Ser,Hse,Thr,Gly,Ile)Leu
C16	Gly-Thr-Gly(Ser,Hse,Thr,Gly)(Ile,Leu,Gly)Tyr

of pepsin[4] covering a considerable part of fragment CB4, to be determined. For additional sequential studies fragment CB4 was purified by ion-exchange chromatography; its amino acid composition is given in Table III. The required sequential information on the C-terminal region of fragment CB4 afforded peptides from its chymotryptic digest and peptides from the thermolytic digest of fragment CB2. The amino acid sequences of these peptides are given in Table IV, the manner in which the amino acid sequence of fragment CB4 was deduced is shown in Fig. 2. The entire region studied is covered by three main chymotryptic peptides, C11, C12 and C14, containing characteristic amino acid residues. With the latter, thermolytic peptides sequentially analyzed and obtained earlier[10] (some of them in duplicate) could be aligned. Peptide C11 provides an overlap of the N-terminal sequence determined earlier at the site of the two half-cystine residues and the histidine residue, thus linking together thermolytic peptides Th14 and Th15–Th17. Peptide C12 is linked to peptide C11, identical with peptide CB2–RAE-T11b-C21, in which the phosphoserine residue (Ser 68) was located[9,12]. When this peptide was sequenced it was shown that the neighboring residue No 69, originally reported[9] as a glutamic acid, is a glutamine. To peptide C12 is linked chymotryptic peptide C14 comprising homoserine residue No 80 which represents the carboxyl-terminus of fragment CB4. The neighborhood of this residue, originally a methionine, was studied in more detail since its bond ..Met-Thr.. was cleaved only partly by cyanogen bromide even though the methionine residue was converted into the homoserine residue. This is evidenced by peptides Th19, Th20, C15, and C16, isolated from enzymatic digests of fragment CB2, in which the bond ..Hse-Thr.. was found uncleaved. The results of studies of the same region obtained by Soviet authors[13] are identical with ours, except for the order of residues 60–61 which these authors determined as ..Ser-Asp.., in accordance with Tang and coworkers[14] who, however, revised this sequence later[15] to ..Asp-Ser.. identical with our findings[3]. Earlier data[4] together with knowledge of the C-terminal region determined in this study, permit the complete amino acid sequence of fragment CB4 to be formulated as follows: Ile-Gly-Asp-Glu-Pro-Leu-Glu--Asn-Tyr-Leu-Asp-Thr-Glu-Tyr-Phe-Gly-Thr-Ile-Gly-Ile-Gly-Thr-Pro-Ala-Gln-Asp--Phe-Thr-Val-Ile-Phe-Asp-Thr-Gly-Ser-Ser-Asn-Leu-Trp-Val-Pro-Ser-Val-Tyr-Cys--Ser-Ser-Leu-Ala-Cys-Ser-Asp-His-Asn-Gln-Phe-Asn-Pro-Asp-Asp-Ser-Ser-Thr-Phe--Glu-Ala-Thr-Ser-Gln-Glu-Leu-Ser-Ile-Thr-Tyr-Gly-Thr-Gly-Ser-Met.

Amino Acid Sequence of Fragment CB6

When the cyanogen bromide digest of S-sulfo-pepsin was fractionated it was observed[2] that fraction FCB5 contains two components: Fragment CB5 and fragment CB6. Both fragments were resolved during additional studies[5] on fraction FCB5. Tryptic, chymotryptic, and thermolytic peptides were isolated from the digests of aminoethylated fraction FCB5. Peptides belonging to fragment CB5 were selected

according to a comparison with peptides prepared from pure fragment CB5 and its amino acid sequence was thus determined. The remaining peptides from the digests of fraction FCB5 belong to fragment CB6 whose amino acid sequence is determined in this study. When fraction FCB5 was subjected to additional fractionation[5], fragment CB6 was found in peak „A", in another experiment in peak „IVa", designated „rIVa" after rechromatography. The amino acid composition of both these preparations is shown in Table V. The absence of tryptophan in fragment CB6 was proved by the qualitative test only. All five[16] tryptophan residues of pepsin were detected in other regions of its chain[17]. A characteristic feature of this fragment is the absence of aromatic amino acid residues. As can be seen in the chymotryptic peptide map of the fragment (Fig. 1) three chymotryptic peptides only are formed in a higher yield. The amino acid sequence of fragment CB6 was derived from the structures of peptides isolated from the tryptic, chymotryptic, and thermolytic hydrolysate of aminoethylated fraction FCB5. Some of the peptides were isolated in duplicate in earlier work[5,10], are listed, however, once only. Peptides corresponding to the spots on the peptide map (Fig. 1) were isolated from the chymotryptic digest of pure fragment CB6 (Fig. 1). The amino acid composition of all peptides which belong to fragment CB6 is shown in Table II. The sequential data on these peptides are summarized in Table VI. Disulfide peptide „SSI" was obtained in earlier work[18].

The data from which the sequence of fragment CB6 was derived are shown schematically in Fig. 3. The fragment is characterized by the N-terminal sequence Asp-Gly--Glu-Thr-Ile-Ala.. and by the C-terminal homoserine residue. The N-terminal sequence is contained in peptides Th1 and Th2 determined in full; the continuation is pro-

TABLE V

Amino Acid Composition of Fragment CB6

The samples were analyzed after 20 h hydrolysis, the values are not corrected. None of the samples contained lysine, histidine, arginine, or tryptophan. The results are expressed as the number of amino acid residues.

Amino acids	Cys	Asp	Thr	Ser	Glu	Pro	Gly	Ala	Val	Ile	Leu	Tyr	Phe	Hse
Fraction A (ref.[5])	1·6	6·0	4·7	6·0	4·8	1·2	6·6	4·6	1·0[a]	4·8[a]	2·3	0·4	0·3	1·0
rIVa (ref.[5])	1·6	6·0	5·0	5·8	5·1	1·1	6·7	4·9	0·9[a]	4·8[a]	2·2	0·3	0·2	1·0
In sequence of CB6	2	6	5	6	5	1	7	5	1	5	2	0	0	1

[a] Determined after 70 h hydrolysis.

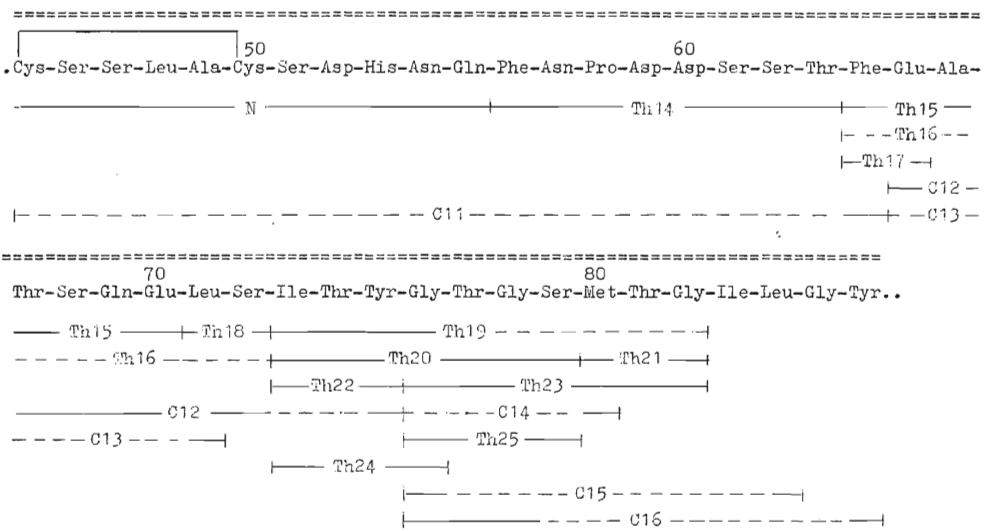

FIG. 3

Amino Acid Sequence of C-terminal Region of Fragment CB4

The peptides are marked by lines, determined sequences by full lines, undetermined sequences by dashed lines. The N-terminal 55-residue sequence of pepsin determined before[4] is marked by symbol "N".

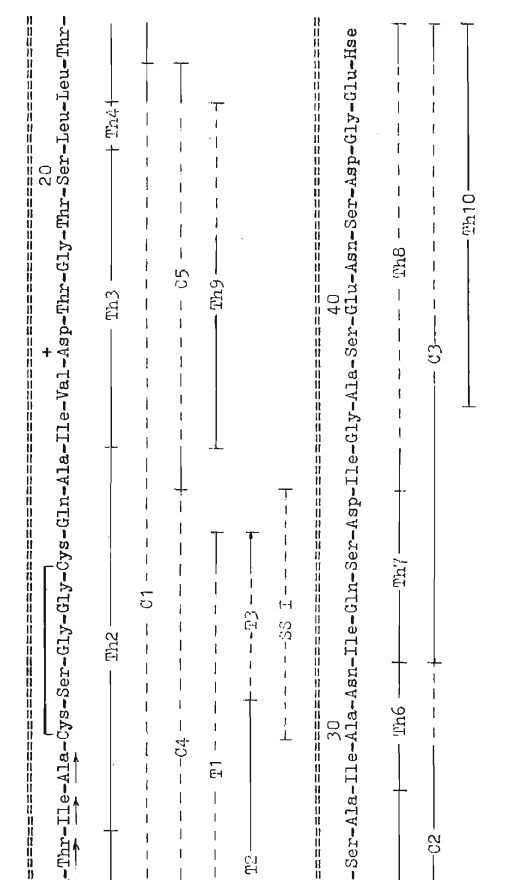

FIG. 4
Amino Acid Sequence of Fragment CB6

vided by peptide Th2. This N-terminal region contains two half-cysteine residues in the sequence ..Cys-Ser-Gly-Gly-Cys-Gln... Disulfide peptide SSI identical with this part of the sequence shows that both half-cystines of fragment CB6 form a disulfide loop in native pepsin. The arrangement of the sequence of fragment CB6 facilitate the three main chymotryptic peptides Cl through C3 involving the region of the whole fragment. The sequential overlaps permitted us to arrange all peptides into two continuous segments (residues 1–31 and 32–46). The bond ..Asn-Ile.. (res. 31–32) was cleaved both by thermolysin and by chymotrypsin. The large tryptic peptide (res. 12–46), expected to arise from the cleavage of aminoethylated material, was not obtained. The link between residues 31 and 32 follows from the agreement between the total sum of residues in both continuous sequential segments and the amino acid composition of fragment CB6; the latter is also in agreement with the sum of peptides of its peptide map, C1 to C3. This conclusion is also evidenced

TABLE VI

Sequence of Peptides from the Region of Fragment CB6
Peptides Th10 and SS1 were obtained in earlier work[10,18].

Designation of peptide	Amino acid sequence of peptide
Th1	Asp-Gly-Glu-Thr
Th2	Ile-Ala-Cys-Ser-Gly-Gly-Cys-Gln-Ala
Th3	Ile-Val-Asp-Thr-Gly-Thr-Ser
Th4	Leu
Th5	Leu-Thr-Gly-Pro-Thr-Ser-Ala
Th6	Ile-Ala-Asn
Th7	Ile-Gln-Ser-Asp
Th8	Ile(Gly,Ala,Ser,Glx,Asx,Ser,Asx,Gly,Glx)Hse
Th9	Ile-Val-Asp-Thr-Gly(Thr,Ser,Leu)
Th10	Ala-Ser-Glu-Asn-Ser-Asp-Gly-Glu-Met
C1	Asp(Cys_2,Asx,Thr_3,Ser_2,Glx_2,Gly_4,Ala_2,Val,Ile_2,Leu)Leu
C2	Thr-Gly-Pro-Thr-Ser-Ala-Ile(Ala,Asx)
C3	Ile-Gln-Ser-Asp-Ile-Gly-Ala-Ser(Glx,Asx,Ser,Asx,Gly,Glx)Hse
C4	Asp(Gly_3,Glx_2,Thr,Ile,Ala,Cys_2,Ser)
C5	Ala(Ile,Val,Asx,Thr,Gly,Thr,Ser,Leu)Leu
T1	Asp(Gly,Glx,Thr,Ile,Ala,Aec[a],Ser,Gly,Gly)Aec
T2	Asp-Gly-Glu-Thr-Ile-Ala-Aec
T3	(Ser,Gly,Gly)Aec
SS I	(Cys,Ser,Gly,Gly,Cys,Gln)

[a] Aec denotes aminoethylcysteine.

by the amino acid composition of peptide RAEP-tA 422 from the tryptic digest of aminoethylated pepsin[19] which involves the region of residues 12–46 of cyanogen bromide fragment CB6 in uncleaved form.

Soviet authors reported the complete amino acid sequence of the corresponding cyanogen bromide fragment[20] which is in full agreement with our data[3]. The University of Oklahoma group reported on sequential heterogeneity in this segment of the pepsin chain. These authors insert after Ala(30) of our fragment CB6 (of fragment CB4 according to their nomenclature) another isoleucine residue present in some molecules; they did not determine, however, which one of the two variants is prevailing[21]. When analyzing their fragment they found 4·9(5) isoleucine residues[22] even though the variant with an extra isoleucine in position 31 would require the

TABLE VII

Amino Acid Composition of Other Fragments of Pepsin

The samples were analyzed after 20 h hydrolysis. The results are expressed as the number of amino acid residues.

Amino acid	Fragment			
	CB1	CB2	CB3	T11b
Lysine	0·8	—	—	—
Histidine	—	1·1	—	0·7
Arginine	1·8	—	—	—
Cysteic acid[a]	—	1·9	—	—
Aspartic acid	5·0	25·0	14·0	18·7
Threonine	2·0	17·0[b]	7·8[b]	12·1[b]
Serine	2·4	28·1[b]	16·9[b]	22·4[b]
Glutamic acid	2·2	16·0	8·2	12·1
Proline	1·8	8·4	4·6	5·7
Glycine	3·1	21·7	14·6	16·5
Alanine	2·6	7·3	4·1	5·3
Valine	4·6	12·9[c]	9·8[c]	9·0[c]
Isoleucine	2·0	12·3[c]	7·9[c]	8·6[c]
Leucine	3·0	18·0	12·7	14·3
Tyrosine	2·1	11·4	7·4	8·5
Phenylalanine	2·0	9·7	5·0	7·2
Homoserine	—	1·8	0·8	1·9
Tryptophan[d]	+	+	+	—

[a] Oxidized sample; [b] values extrapolated to zero time of hydrolysis; [c] values after 70 h hydrolysis; [d] the presence of tryptophan was determined by the qualitative test only.

presence of 6 isoleucines. The presence of this extra isoleucine is not indicated either by the composition of tryptic fragment RAEP-tA 422 (ref.[19]) or by the sequence published by the Soviet authors[20], supported in this region by automated sequential degradation of 38 amino acid residies[23]. In this study we also employed for the preparation of the enzymatic digests fraction FCB5 purified by gel filtration which cannot be expected to separate the two variants; the peptides obtained were homogeneous, however. The corresponding site of the prochymosin sequence, published by Foltmann and coworkers[24] is highly homologous with our sequence:

		30
CB6	..Ser-Ala-Ile-Ala-Asn-Ile-Gln-Ser-Asp-Ile..	
		275
Prochymosin	..Ser-Asp-Ile-Leu-Asn-Ile-Gln-Gln-Ala-Ile..	

Residue Asp(16) of fragment CB6 is a part of the sequence ..Ile-Val-Asp-Thr-Gly--Thr-Ser-Leu.. (residues 14–24), characterizing the aspartic acid residue of pepsin[25–27] which reacts with diazo inactivators and forms a part of the active center of this enzyme.

Amino Acid Composition of Other Fragments of Pepsin

Table VII shows the amino acid composition of other fragments of pepsin, rechromatographed mostly in this study (Fig. 2).

Fragment CB1 represetns the C-terminal region of pepsin showing a high associating tendency in solution[2]. When chromatographed under the conditions described in Fig. 2 the fragment was not adsorbed to the resin in a weakly acidic medium and was eluted in the first peak. The retarded peak contained residual contaminants which could not be removed by the procedures employed before and (because of the presence of arginine and lysine) also additional material derived from the region of fragment CB1. The amino acid composition of rechromatographed fragment CB1 is in good agreement with the complete sequence of this fragment[28]. Fragment CB2 was freed of contaminants by recyclization on Sephadex G-100 (ref.[2]). When the N-terminal region of pepsin was studied aminoethylated fragment CB2 was digested with trypsin[9]. The cleavage at two aminoethylcysteine residues gives rise to three peptides of which the first two (from the N-terminus) afforded the main bulk of information for the determination of the 55-residue N-terminal sequence of pepsin[4]. The third of these peptides is fragment CB2-RAE-T11b whose amino acid composition was in the original paper[9] calculated on the assumed presence of one homoserine residue only. In this study the fragment was rechromatographed and again analyzed. It comprises the C-terminal region of fragment CB4 and the region of fragment CB3 extending to the second homoserine residue (residues 51–199) of pepsin[3].

I thank Mrs J. Viková for careful technical assistance, Mr K. Grüner and Mrs A. Čermáková for the Edman degradation experiments. I thank Mr J. Zbrožek, Miss V. Himrová and Mrs E. Držková for the amino acid analyses of the peptides.

REFERENCES

1. Gross E., Witkop B.: J. Biol. Chem. *237*, 1856 (1962).
2. Morávek L., Kostka V.: This Journal *38*, 304 (1973).
3. Morávek L., Kostka V.: FEBS (Fed. Eur. Biochem. Soc.) Lett. *43*, 207 (1974).
4. Morávek L.: This Journal *39*, 1933 (1974).
5. Morávek L.: This Journal *43*, 344 (1978).
6. Kostka V., Carpenter F. H.: J. Biol. Chem. *239*, 1799 (1964).
7. Prusík Z., Štěpánek J.: J. Chromatogr. *87*, 73 (1973).
8. Mikeš O.: This Journal *25*, 2044 (1960).
9. Morávek L.: FEBS (Fed. Eur. Biochem. Soc.) Lett. *23*, 237 (1972).
10. Morávek L., Kysilka Č.: This Journal *39*, 2301 (1974).
11. Hanes C. S., Isherwood F. A.: Nature (London) *164*, 1107 (1949).
12. Vakhitova E. A., Amirkhanyan M. M., Stepanov V. M.: Biokhimiya *35*, 1164 (1970).
13. Ostoslavskaya V. I., Vakhitova E. A., Baratova L. A., Belyanova L. P., Katrukha S. P., Kovalishin Ya. F., Lapuk Ya. I., Revina L. P., Surova I. A., Stepanov V. M.: 9th FEBS Meeting, Budapest, August 25—30, 1974, Abstr. f1a1.
14. Tang J., Sepulveda P., Marciniszyn J. jr, Chen K. C. S., Huang W.-Y., Tao N., Liu D., Lanier J. P.: Proc. Nat. Acad. Sci. U.S.A. *70*, 3437 (1973).
15. Sepulveda P., Marciniszyn J. jr, Liu D., Tang J.: J. Biol. Chem. *250*, 5082 (1975).
16. Kostka V., Neradová V.: This Journal *39*, 2287 (1974).
17. Morávek L., Kostka V.: This Journal *39*, 2709 (1974).
18. Keil B., Morávek L., Šorm F.: This Journal *32*, 1968 (1967).
19. Kostka V.: This Journal *43*, 2942 (1978).
20. Surova I. A., Baratova L. A., Belianova L. A., Stepanov V. M.: Biokhimiya *41*, 104 (1976).
21. Marciniszyn J. jr, Sepulveda P., Huang W.-Y., Lanier J. P., Tang J.: J. Biol. Chem. *250*, 5076 (1975).
22. Chen K. C. S., Tao N., Tang J.: J. Biol. Chem. *250*, 5068 (1975).
23. Bejanova L. P., Baratova L. A., Pugatcheva I. B., Revina L. P., Surova O. A., Stepanov V. M.: Biokhimiya *40*, 762 (1975).
24. Foltmann B., Pedersen V. B., Jacobsen H., Kauffman D., Wybrandt G.: Proc. Nat. Acad. Sci. U.S.A. *74*, 2321 (1977).
25. Bayliss R. S., Knowles J. R., Wybrandt G. B.: Biochem J. *113*, 377 (1969).
26. Fry K. T., Kim O. K., Spona J., Hamilton G. A.: Biochemistry *9*, 4624 (1970).
27. Stepanov V. M., Vaganova T. I.: Biochem. Biophys. Res. Commun. *31*, 825 (1968).
28. Kostka V., Morávek L., Šorm F.: Eur. J. Biochem. *13*, 447 (1970).

Translated by V. Kostka.