PRIMARY STRUCTURE OF TRYPSIN INHIBITOR FROM COW COLOSTRUM (COMPONENT B2)

D.Čechová, V.Jonáková and F.Šorm

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

Received September 23rd, 1970

By sequential degradation of peptides obtained from tryptic digest of S-sulfonated or aminoethylated cow colostrum trypsin inhibitor and from chymotryptic digest of S-sulfonated cow colostrum trypsin inhibitor, the following amino acid sequence of the inhibitor was determined: Phe.Gln.Thr.Pro.Pro.Asp.Leu.Cys.Gln.Leu.Pro.Gln.Ala.Arg.Gly.Pro.Cys.Lys.Ala.Ala.Leu.Leu. .Arg.Tyr.Phe.Tyr.Asx.Ser.Thr.Ser.Asn.Ala.Cys.Glu.Pro.Phe.Thr.Tyr.Gly.Gly.Cys.Gln.Gly.Asn. .Asn.Asx.Asn.Phe.Glu.Thr.Thr.Glu.Met.Cys.Leu.Arg.Ile.Cys.Glu.Pro.Pro.Gln.Gln.Thr.Asp. Lys.Ser.[67]

Naturally-occurring polypeptide trypsin inhibitors have been found in many sources. The covalent structures of several inhibitors have been determined[1] and studies on their interaction with trypsin and other proteolytic enzymes are in progress. During the past years attention has been focused on the basic trypsin inhibitor from beef pancreas whose chemical and biological properties have been studied extensively. Inhibitors showing an identical primary structure have been isolated also from other beef organs, such as lungs, submandibulary glands, and liver[2,3]. An inhibitor of acidic character and certain properties identical to those of the basic inhibitor has been isolated from bovine pancreas. This inhibitor, however, shows a considerably different primary structure[4]. Another trypsin inhibitor has been isolated from cow colostrum[5,6]; it resembles in many features both the basic and the acidic trypsin inhibitor from beef pancreas. It is of low molecular weight, has three disulfide bonds and no tryptophan and histidine residues in its molecule which contains a firmly bound nonprotein moiety[6]. As a part of our studies on similarities in protein structures we considered worth while to determine the amino acid sequence of this inhibitor and to investigate the similarity between this inhibitor and other inhibitors of proteases.

EXPERIMENTAL

Material

The cow colostrum inhibitor was obtained by the method described earlier[6]. This study was carried out with inhibitor B2 which represents the dominant component in cow colostrum.

Trypsin was a three-times recrystallized preparation. Chymotrypsin was recrystallized four times. Both enzymes were commercial products of Léčiva, Prague. Pepsin was a two-times crystallized product of Worthington Biochemical Corp. (Freehold, N.J., USA), subtilisin was a lyophilized preparation of Japanese origin. Thermolysin was a product of Calbiochem (Los Angeles, Cal., USA).

Sephadex G-25 was a product of Pharmacia (Uppsala, Sweden). Dansyl chloride was purchased from Calbiochem. Silufol, a thin silica gel layer on aluminum sheet, was a product of Kavalier, Czechoslovakia. All chemicals used in this study were of G.R. purity grade.

Amino Acid Analysis

Samples containing 0·05—0·07 μmol of peptides were hydrolyzed in 1 ml of redistilled constant-boiling hydrochloric acid in evacuated test tubes, 20 h at 110°C. Peptides containing hexosamines, which interfere with tyrosine and phenylalanine on amino acid analysis, were hydrolyzed 48 h. The values of valine and isoleucine were obtained with 70 h hydrolysates. After hydrochloric acid had been dried off in a rotary evaporator, the amino acid analysis was performed by the conventional procedure[7]. Cystine, unless aminoethylated, was determined after the oxidation of samples in performic acid[8] as cysteic acid and methionine as methionine sulfone. The values of serine and threonine were not extrapolated to zero time hydrolysis.

Sulfitolytic Oxidation[9]

The native protein (120 mg) was dissolved in 14 ml of 6M guanidine hydrochloride adjusted to pH 8·6 by ammonia and 1 ml of 0·1M copper sulfate and 0·9 g of crystalline sodium sulfite was added. After the crystals had dissolved the volume of the solution was made up to 20 ml with 6M guanidine hydrochloride at pH 8·6 and the solution sealed off. The reaction mixture was set aside for 3 h at room temperature. The S-sulfonated protein was desalted on a column of Sephadex G-25 medium (4 . 25 cm) equilibrated with 0·21M ammonium carbonate. Before application of the sample the column was washed with 10 ml of 6M urea in the eluting buffer. After the sample had soaked into the column 3 ml of 6M guanidine hydrochloride at pH 8·6 was applied.

Digestion of S-Sulfonated Protein with Trypsin

The solution of the desalted S-sulfo protein (100 mg in 35 ml of 0·021M ammonium carbonate) was directly digested with trypsin (2 mg in 0·1 ml of 0·001M hydrochloric acid). The pH of the sample was adjusted to pH 8·3 by 0·1M ammonium carbonate. The digestion was allowed to proceed 2 h at 37°C and was discontinued by freezing the sample (−20°C). The digest was freeze-dried.

Reduction of Disulfide Groups and Aminoethylation of Cysteine Residues

The method developed by Raftery and Cole[10] was used. The protein (30 mg) was dissolved in 1·5 ml of guanidine–Tris buffer (5·5M guanidine hydrochloride, 0·3M Tris, 0·2% of disodium salt of ethylenediaminetetraacetic acid, pH adjusted to pH 8·5 by hydrochloric acid). The solu-

tion was flushed with nitrogen and dithiothreitol was added (final concentration 0·1M). The reaction mixture was allowed to stand 2 h at room temperature in the atmosphere of nitrogen then cooled down to 4°C, and ethylene imine was added (final concentration 0·3M). After 30 min at 4°C the mixture was diluted to 3 ml and applied to a column of Sephadex G-25 (2 × 10 cm), equilibrated with 0·021M ammonium carbonate, and desalted.

Tryptic Digestion of Aminoethylated Inhibitor

The solution of the desalted aminoethylated protein in 0·021M (total volume 15 ml) was directly digested with trypsin (0·6 mg in 100 µl of 0·001M hydrochloric acid). The pH of the solution was adjusted to pH 8·3 by 0·1M ammonium carbonate and the digestion was allowed to proceed, 4 h at 37°C. The digestion was discontinued by rapid freezing and the digest was freeze-dried.

Chymotryptic Digestion of S-Sulfonated Inhibitor

The protein (10 mg) was dissolved in 0·2 l of 0·021M ammonium carbonate and chymotrypsin was added (0·2 mg in 20 µl of distilled water). The digestion was allowed to proceed 3 h at 37°C. The digestion was discontinued by freezing and the digest was taken to dryness in a desiccator (oil pump).

Isolation of Peptides from Tryptic Digest of S-Sulfonated Inhibitor

The tryptic digest of the S-sulfonated trypsin inhibitor was divided into two halves. Each of them was applied as a thin band on Whatman No 3 paper and subjected to electrophoretic separation in a buffer at pH 1·9 (50 ml of 85% formic acid — 150 ml of 99% acetic acid, made up to 1 l with water) at a potential of 3000 V and 4°C (ref.[11]). After the end of the separation, the middle part of the electropherogram was cut out and the margins were stained with 0·2% ninhydrin.

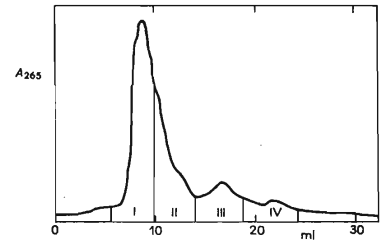


FIG. 1

Gel Filtration nf Tryptic Digest of Aminoethylated Inhibitor on Sephadex G-25 Medium at pH 7·5

A_{265}, absorbance at 265 nm, volume of effluent in ml. Eluant, 0·021M ammonium carbonate; I, II, III, IV, designation of pooled fractions.

Bands corresponding to individual peptides were cut out. Bands C and D were eluted directly. Band C contained pure peptide T2, band D pure peptide T3. Bands A and B were stitched to Whatman No 3 paper and subjected to descending chromatography (1-butanol–pyridine–acetic acid–water, 15 : 10 : 3 : 12, by vol.). Peptide T4 from band A remained at the origin while the impurities showed all a higher R_F. Band B yielded two zones containing peptides T 1 and T 5. The peptides were eluted. All peptides were obtained in corresponding quantities, *i.e.* peptide T 2 and T 4 in approximately 50% yield, peptides T 1 and T 5 in approximately 45% yield, and peptide T 3 in approximately 40% yield.

Isolation of Peptides from Tryptic Digest of Aminoethylated Inhibitor

The digest (27 mg) in 0·021M ammonium carbonate (1 ml) was subjected first to gel filtration over a Sephadex G-25 (medium) column (0·8 × 87 cm) equilibrated with 0·021M ammonium carbonate which also served as eluting buffer. The course of the elution was monitored at 265 nm by an automatic recorder of ultraviolet absorption (Fig. 1). Fractions corresponding to individual peaks were pooled and freeze-dried. The pooled fractions (designated I through IV) were then fractionated either by high-voltage paper electrophoresis at 3000 V in the buffer at pH 1·9 (see above) or at pH 5·6 (16 ml of pyridine and 4 ml of acetic acid made up to 1 l with water). The other procedures used for the isolation of peptides were the same as those used for the isolation of tryptic peptides from the digest of the S-sulfonated protein.

Isolation of Peptides from Chymotryptic Digest of S-Sulfonated Protein

The sample was applied to paper as a 2 cm wide band and subjected to electrophoresis at pH 1·9. Descending chromatography was used for the second direction. The peptides were detected by staining with 0·01% ninhydrin. The spots of peptides were cut out and eluted with water. In subsequent work a 30—50% loss of the N-terminal amino acid had to be taken into account.

Determination of Terminal Amino Acids

The Dansyl technique[12] as modified by Gross and Labouess[13] was used for the determination of N-terminal amino acids. Dansyl-(1-dimethylaminonaphthalene-S-sulfonyl)amino acids were identified by thin layer chromatography on silica gel. One-dimensional repeated development was carried out in the following systems: chloroform–1-butanol–toluene–acetic acid (7 : 1 : 1 : 1, by vol.), 1-butanol–toluene–25% ammonia (8 : 1 : 1, by vol.), ethanol–toluene–25% ammonia (7 : 2 : 1, by vol.), tert-amyl alcohol–pyridine–toluene–1-butanol (6 : 2 : 1 : 1, by vol.)[14].

C-terminal amino acids were determined by hydrazinolysis[15].

Sequential Degradation

The modified technique of Niall and Edman[16] was used. The sample (1 μmol) was dissolved in 1 ml of buffer at pH 9 (pyridine–dimethylamine–water, 15 : 1, 18 : 10, by vol., adjusted to pH 9 by trifluoroacetic acid). After the addition of 50 μl of phenyl isothiocyanate the solution was flushed with nitrogen and allowed to stand 1 h at 40°C in a stoppered tube. The reaction mixture was then extracted repeatedly with benzene and lyophilized. The lyophilisate was dissolved in 50 μl of trifluoroacetic acid and set aside for 15 min. The liberated thiazoline was extracted by 1,2-dichloroethane. The latter was removed by evaporation in the stream of nitrogen. To the residue 30 μl of 1M hydrochloric acid was added and the mixture was heated 15 min at 80°C. The formed hydantoin was extracted by ethyl acetate. After drying of the sample in the stream of nitrogen, the dry residue of the hydantoin was dissolved in 1,2-dichloroethane.

Phenylthiohydantoins obtained in each degradation step were systematically analyzed by thin layer chromatography[17] using ascending chromatography in the systems described above. Leucine was distinguished from isoleucine in the system xylene–acetone (5 : 1, by vol.)[18] with glycerin as stationary phase. The spots of phenylthiohydantoins on the thin layers were inspected in the light of a mercury lamp and detected by iodine vapors[19] and then by ammonia vapors[19].

When the sequential degradation was used in combination with the Dansyl technique[20], the following procedure was employed. The solution of the peptide (containing according to the assumed number of steps 0·025 μl of peptide per step) in a conical test tube (3 ml, with ground-glass stopper) was taken to dryness *in vacuo*. The dry residue was dissolved in 200 μl of 50% aqueous pyridine and 100 μl of the solution of phenyl isothiocyanate was added. After air in the test tube had been replaced by nitrogen, the test tube was stoppered and the reaction allowed to proceed 3 h at 37°C. Excess reagent was removed by three-times repeated extraction with benzene. After each addition of benzene the solution was centrifuged and the organic phase removed. After the third extraction the aqueous layer in the test tube was frozen and taken to dryness in a desiccator *in vacuo* (oil pump). The dry peptide was dissolved in 200 μl of anhydrous trifluoroacetic acid, the test tube was filled with nitrogen and stoppered. The liberation of the hydantoin was allowed to proceed 1 h at 37°C. After the hydrolysis had been completed, the acid was evaporated with the aid of a water aspirator. The sample was dissolved in 50% pyridine, the corresponding portion was withdrawn for dansylation, and the rest was subjected to the subsequent degradation step. The N-terminal amino acid was determined by the Dansyl technique as described above.

Enzymatic Digestion of Peptides

Chymotryptic digestion of peptide T4. The peptide (0·6 μmol) was digested with chymotrypsin (1 : 50, w/w) in 0·021M ammonium carbonate at pH 8·3, 2 h at 37°C. The reaction was discontinued by freezing and rapid drying of the sample.

Subtilisin digestion. Peptides T4, T4C4, T5, AE 10 (0·1–0·3 μmol) were digested with subtilisin (1 : 50–1 : 100, w/w) at pH 7·0 and 37°C for 4 to 8 h.

Peptic digestion of peptide T4C4. The peptide (0·3 μmol) was digested with pepsin (0·3 mg) at pH 2·1 and 37°C for 24 h. The reaction was discontinued by the addition of 1 drop of pyridine and the digest was taken to dryness in a desiccator.

Aminoethylation and tryptic digestion of peptides T4 and S1. The method of Raftery and Cole[10] was used. To S-sulfonated peptide T4 (4 μmol) dissolved in 2 ml of 0·021M ammonium carbonate, 75 μl of mercaptoethanol was added and the reduction was allowed to proceed 30 min at room temperature in a test tube with ground-glass stopper. After the addition of ethylene imine, the substitution was completed in 1 h. Aminoethylated peptide T4 designated AET4 (2·2 ml) was desalted on a column of Sephadex G-10 (1·3 . 22 cm) equilibrated with 0·021M ammonium carbonate. To the desalted solution (4 ml) of peptide AET4 in 0·021M ammonium carbonate, trypsin was added (1 : 50, w/w) and the digestion was allowed to proceed under the conditions described above.

The same method of aminoethylation and tryptic digestion was used also with S-sulfonated peptide T1 designated AET1.

Fractionation of Enzymatic Digests of Peptides

The combination of high-voltage electrophoresis at pH 1·9 or 5·6 with descending chromatography was used. The peptides were mostly eluted after being stained with 0·01% ninhydrin

save for a few exceptions (chymotryptic digest of peptide T4, tryptic digest of peptides AET4 and AET1, and subtilisin digest of peptide AE 10) when the digests were resolved, the corresponding zones cut out and, if necessary, stitched to a sheet of paper and subjected to chromatography.

Determination of Glutamine and Asparagine

The presence of acidic amino acids or of their amides in small peptides was judged by their mobility at pH 5·6. When the peptides were subjected to Edman degradation, glutamine and asparagine residues were determined either directly as the corresponding phenylthiohydantoins or indirectly, *i.e.* the presence of acidic residues was judged by the change of electrophoretic mobility of the peptide after the removal of the N-terminal amino acid. After the removal of the latter, the corresponding portion of the sample, which contained the peptide shorter by the N-terminal amino acid, was set aside. After the degradation of the whole peptide had been completed, individual samples were applied on paper in order of the degradation steps and their electrophoretic mobility on paper at pH 5·6 was invsetigated. In case that the liberated amino acid bore a positive or a negative charge, the mobility of the subsequent peptide was changed. The peptides were

TABLE I

Quantitative Amino Acid Analyses of Peptides T1—T5 from Tryptic Digest of S-Sulfonated Trypsin Inhibitor

Amino acid	Residues per mol					Total
	T-1	T-2	T-3	T-4	T-5	
Alanine	0·96	—	1·70[a]	1·00	—	4
Aspartic acid	1·00	—	—	5·72	1·00	8
Arginine	0·95	—	1·00	1·00	—	3
Half-cystine	1·07	1·00	—	3·17	1·00	6
Glutamic acid	2·98	—	—	4·02	3·00	10
Glycine	—	0·50[a]	—	2·87	—	4
Isoleucine	—	—	—	—	0·69[a]	1
Leucine	2·07	—	2·30	1·01	—	5
Lysine	—	0·93	—	—	0·65	2
Methionine	—	—	—	1·00	—	1
Phenylalanine	0·55[a]	—	—	2·79	—	4
Proline	3·07	1·24	—	0·96	1·83	7
Serine	—	—	—	2·12	0·92	3
Threonine	0·96	—	—	3·76	0·94	6
Tyrosine	—	—	—	2·45[a]	—	3
Total	*14*	*4*	*5*	*33*	*11*	*67*

[a] The value of the N-terminal amino acid is decreased due to its partial reaction with ninhydrin.

detected by staining with 0·2% ninhydrin. When only a small amount was available, the peptides were first dansylated and the electrophoretic mobility of the dansyl peptides examined in the buffer at pH 6·5 (16 ml of pyridine and 4 ml of acetic acid made up to 1 l with water). The peptides were detected in ultraviolet light after drying of the paper.

RESULTS

Preparation of Tryptic Peptides from S-Sulfonated Protein

Tryptic digestion of the S-sulfonated protein afforded 5 peptides (Fig. 2). After their isolation their homogeneity was proved by N-terminal end-group analysis. The peptides were obtained in corresponding quantities. Quantitative analysis of the peptides has shown that these peptides account for the entire molecule of the inhibitor (Table I).

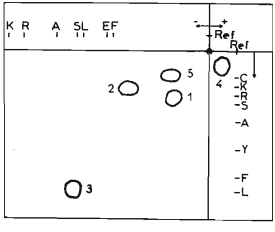


FIG. 2

Peptide Map of Tryptic Digest of S-Sulfonated Inhibitor

Horizontally high-voltage electrophoresis at 3000 V and pH 1·9; vertically descending chromatography in the system 1-butanol–pyridine–acetic acid–water (15 : 10 : 3 : 12); REF, reference amino acids; 1—5, designation of tryptic peptides T 1—T 5.

Amino Acid Sequence of Short Peptides T2 and T3

The amino acid sequence was determined by Edman degradation in combination with the Dansyl technique as follows

T2 Gly - Pro - Cys - Lys

T3 Ala - Ala - Leu - Leu - Arg

Amino Acid Sequence of Peptide T1

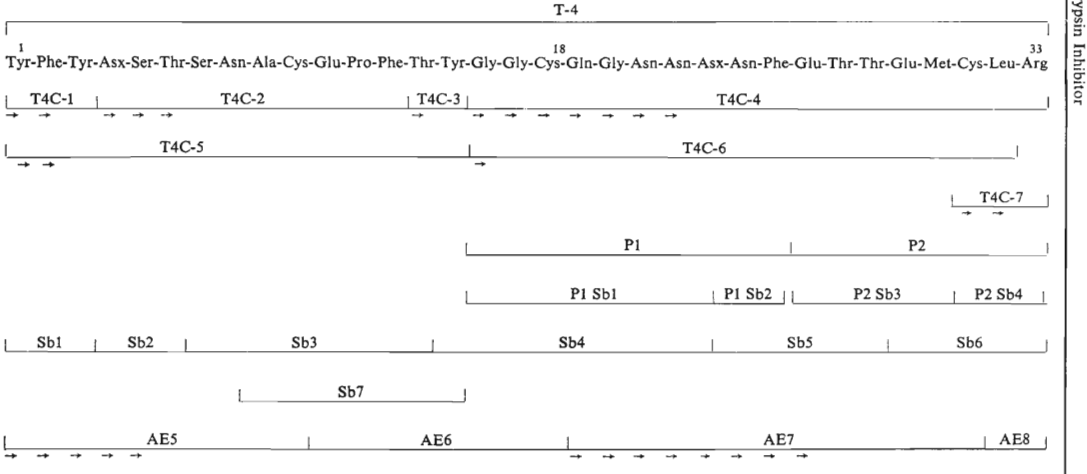
This peptide is N-terminal and as the only one had N-terminal phenylalanine which is the N-terminal amino acid in the whole molecule. The larger part of the amino acid sequence of the peptide was determined by Edman degradation as follows

T1 Phe - Gln - Thr - Pro - Pro - Asp - Asp - Leu - Cys - Gln - Pro - Pro - Gln - (Ala,Arg)
(Leu Leu)

TABLE II

Cleavage of Tryptic Peptide T-4

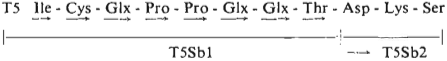
T4C-1 through T4C-7, peptides from chymotryptic digest; P1, P2, peptides arising from peptic digest of peptide T4C-4; Sb1 through Sb7, peptides from subtilisin digest of peptide T4; AE5 through AE8, peptides from tryptic digest of aminoethylated peptide T4; ⟶ sequence of amino acids confirmed by Edman degradation; P1Sb1, P1Sb2, and P1Sb3, P2Sb4, peptides from subtilisin digests of peptic peptides.



Proline together with leucine was detected in the tenth and eleventh step. The quantitative amino acid analysis shows that the peptide contains three proline residues and two leucine residues. The final structure of the peptide was determined from the tryptic digest of the aminoethylated protein and confirmed by the analysis of the tryptic digest of aminoethylated peptide T1.

Amino Acid Sequence of Peptide T5

The C-terminus of this peptide is occupied by serine which is the C-terminal amino acid of the inhibitor. Peptide T5 thus represents the C-terminus of the molecule. The amino acid sequence of the peptide was determined by Edman degradation in combination with the Dansyl technique and by the analysis of subtilisin fragments. Peptide T5Sb1 was strongly acidic while peptide T5Sb2 was neutral. This indicates the presence of aspartic acid in peptide T5Sb2, which was found as N-terminal.



Amino Acid Sequence of Peptide T4

The peptide contains 33 residues, *i.e.* one half of the molecule of the inhibitor. Its hydrolysate has a dark color and the presence of hexosamines was shown by quantitative amino acid analysis. The sequence of 18 amino acid residues in the peptide was determined by Edman degradation. The peptide was then digested with chymotrypsin. The cleavage of all sensitive bonds was not quantitative and therefore 7 peptides (T4C1 – T4C7) (Table II) were isolated. The sequence of several amino acids in the isolated peptides was determined by Edman degradation in combination with the Dansyl technique. Our attention was focused on the unknown region 19 – 33, *i.e.* on peptide T4C4. The peptide was digested with pepsin and the two peptic fragments (P1, P2) isolated were cleaved further by subtilisin. Peptide P1 was cleaved to neutral peptide Asn-Phe and acidic peptides whose N-terminal amino acids were determined by the Dansyl technique. From the subtilisin digest of peptide P2, neutral peptide (Cys,Leu,Arg) and strongly acidic peptide (Gly,Thr,Thr,Glu,Met) were obtained.

When whole peptide T4 was digested with subtilisin, the peptides were obtained, *i.e.* acidic peptide (Glu,Met,Cys,Leu,Arg), Sb6, which indicates the absence of glutamine, and another acidic peptide (Asx,Phe,Glx,Thr,Thr), Sb 5. The presence of sequence Asn-Phe has been determined before, the peptide thus contains glutamic acid and not its amide. Peptide Asx-Ser-Thr shows a zero R_F-value on chromatography. When subjected to electrophoresis at pH 5·6, it gives several components of the same amino acid composition, obviously due to the presence of amino sugars.

TABLE III

Quantitative Amino Acid Analyses of Peptides C1—C13 from Chymotryptic Digest of S-Sulfonated Trypsin Inhibitor

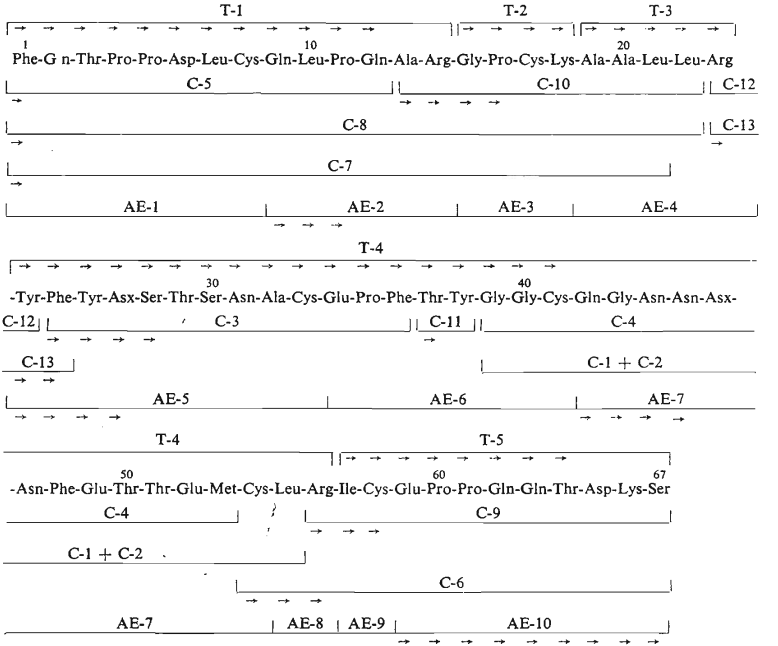
Amino acid	Residues per mol												
	C-1	C-2	C-3	C-4	C-5	C-6	C-7	C-8	C-9	C-10	C-11	C-12	C-13
Alanine			1·00				2·50	2·80		2·50[a]			
Aspartic acid	3·86	3·57	1·65	3·64	1·11	1·20	1·17	1·10	1·04				
Arginine						0·93	0·82	0·90	0·74[a]	0·86		0·80[a]	0·65[a]
Half-cystine	1·87	1·86	0·98	0·90	1·00	1·70[a]	1·62	1·75	0·91	0·78			
Glutamic acid	2·96	2·67	1·14	2·63	3·21	3·10	3·00	3·00	3·11				
Glycine	2·54[a]	2·50[a]		2·30[a]			0·80	1·10		1·12			
Isoleucine						0·97			0·1				
Leucine	1·04	1·05			2·00	1·00	3·17	4·00	2·00				
Lysine						0·84	0·67	0·69	1·00	0·65			
Methionine	1·04	1·02		0·74									
Phenylalanine	0·96	1·00	1·52[a]	1·00	0·53[a]		0·30[a]	0·45[a]					1·00
Proline			1·05		3·17	1·87	3·20	3·50	2·10	0·88			
Serine			2·20			0·84			0·93				
Threonine	1·91	2·10	1·13	1·60	1·09	1·27	1·00	1·00	1·00		0·60[a]		
Tyrosine			0·68								0·70	0·91	0·75

[a] The value of the N-terminal amino acid is decreased due to its partial reaction with ninhydrin.

TABLE IV

Primary Structure of Trypsin Inhibitor from Cow Colostrum

T1 to T5, peptides from tryptic digest, C1 to C13, peptides from chymotryptic digest of S-sulfonated inhibitor; AE- to AE10, peptides from tryptic digest of aminoethylated inhibitor. ⟶ sequence of amino acids confirmed by Edman degradation.



Tryptic Digest of Aminoethylated Protein

According to the results of amino acid analysis the protein was aminoethylated to 95%. Tryptic digestion yielded ten main peptide zones designated AE1-AE10 (Table IV).

Peptides AE2, AE5, AE7, and AE10 were submitted to amino acid analysis. The sequence of several amino acid residues in peptides AE2, AE5, and AE7 was determined by Edman degradation in combination with the Dansyl technique.

AE2 Glx → - Leu → - Pro → - (Glx, Ala, Arg)

AE5 Tyr → - Phe → - Tyr → - Asx → (Ser, Thr, Ser, Asx, Ala, AECys)

AE7 Glx → - Gly → - Asx → - Asx → (Asx, Asx, Phe, Glx, Thr, Thr, Glx, Met, AECys)

Peptide AE10 was cleaved by subtilisin to acidic peptide Sb1 (Glx,Pro,Pro,Glx,Glx, Thr) and neutral peptide Sb2 (Asp,Lys,Ser). After its isolation peptide Sb1 was subjected to Edman degradation. After the liberation of the N-terminal amino acid in each step, an aliquot of the peptide was set aside. At the end of the degradation analysis the mobilities of all peptides with N-terminal amino acid both intact and dansylated were determined. Both procedures showed that only the N-terminal amino acid of the original intact peptide was negatively charged. Peptide Sb 2 was likewise subjected to Edman degradation. The complete amino acid sequence of peptide AE10 is the following:

AE10 Glu - Pro - Pro - Gln - Gln - Thr - Asp - Lys - Ser

Tryptic Digest of Aminoethylated Peptides T4 and T1

Peptides T4 and T1 were subjected also to aminoethylation and tryptic digestion. We obtained also peptides which were identical to those isolated from the tryptic

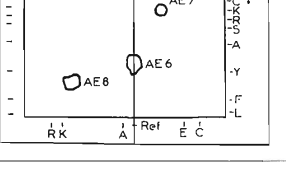


FIG. 3

Peptide Map of Tryptic Digest of Aminoethylated Peptide T 4

Horizontally high-voltage electrophoresis at 3000 V and pH 5·6; vertically descending chromatography in the system 1-butanol–pyridine–acetic acid–water (15 : 10 : 3 : 12); REF, reference amino acids; AE 5—AE 8, designation of peptides obtained from tryptic digest of aminoethylated peptide T 4.

digest of the aminoethylated inhibitor. Thus, *e.g.* from aminoethylated peptide T4 we obtained peptides AE5 through AE8 (Fig. 3, Table II). The isolation of peptides AE5 and AE7 enabled us to derive their complete amino acid sequence. The latter was determined by Edman degradation.

AE5 $\overset{1}{\underrightarrow{\text{Tyr}}}$ - $\underrightarrow{\text{Phe}}$ - $\underrightarrow{\text{Tyr}}$ - $\overset{4}{\underrightarrow{\text{Asx}}}$ - $\underrightarrow{\text{Ser}}$ - (Thr, Ser, Asx, Ala, AECys)

AE7 $\overset{1}{\underrightarrow{\text{Gln}}}$ - $\underrightarrow{\text{Gly}}$ - $\underrightarrow{\text{Asn}}$ - $\underrightarrow{\text{Asn}}$ - $\underrightarrow{\text{Asx}}$ - $\underrightarrow{\text{Asn}}$ - $\underrightarrow{\text{Phe}}$ - $\underrightarrow{\text{Glu}}$ - (Thr, Thr, Glx, Met, AECys)

We have not been able to determine by sequential degradation whether position 4 in peptide AE5 and position 5 in peptide AE7 is occupied by the acids or by their amides, even though the analysis was repeated several times. Quantitative analysis of peptide AE5 showed the presence of a considerable quantity of hexosamines. By quantitative analysis we also found that to peptide AE7 correspond in amino acid composition two peptides which showed different R_F-values on descending chromatography (Fig. 3).

Tryptic digestion of aminoethylated peptide T1 yielded peptide AE1 and peptide AE2. The latter was subjected to Edman degradation in combination with the Dansyl technique and the position of leucine and proline at position 10 and 11 with respect to the N-terminus of the molecule of the inhibitor was thus verified (Table IV).

AE2 $\underrightarrow{\text{Glx}}$ - $\underrightarrow{\text{Leu}}$ - $\underrightarrow{\text{Pro}}$ - (Glx, Ala, Arg)

Chymotryptic Digest

The chymotryptic digestion of the S-sulfonated inhibitor yielded 13 peptides (Fig. 4). All these peptides were subjected to quantitative amino acid analysis (Table III). Peptides C1 and C2, which have an identical amino acid composition, show different R_F-values on descending chromatography. The remaining quantity of individual peptides was used for the determination of their partial amino acid sequence by the combination of Edman degradation with the Dansyl technique. The obtained peptides account for the entire molecule of the inhibitor and provide information on the order of tryptic peptides (Table IV).

DISCUSSION

The obtained results permit us to derive the complete primary structure of the trypsin inhibitor from cow colostrum (Table IV). It has been known that cow colostrum contains several inhibitors which differ very little from each other[6]. This study was carried out only with inhibitor B2 which had been found to be entirely homogeneous. When analyzing the peptides which arise from tryptic or chymotryptic digestion

of this inhibitor we found that some of them form on the chromatogram a main zone and an additional, weaker zone. In cases where enough material was available we subjected also these additional zones to quantitative analysis. We found, *e.g.* that the weaker zone of peptide T1 is identical with the main zone in amino acid composition, save for the content of alanine which is by one residue lower in the main zone. Peptide T4 appeared also in several zones. All zones, however, showed the same amino acid composition. The differences in electrophoretic mobility at pH 5·6 and/or in R_F-value on paper chromatography may be accounted for most likely by the differences in the nonprotein moiety, unless, of course, these differences are due to the presence of artifacts formed during the treatment of the material.

The presence of the nonprotein moiety of different character was observed also with chymotryptic peptides C1 and C2 which show the same amino acid composition (Table III) yet different R_F-values on chromatography (Fig. 4). The same phenomenon was observed with peptide AE7 arising from tryptic digestion of aminoethylated peptide T4 (Fig. 3, Table II). Peptide AE7 is derived from the same region of the molecule of the inhibitor as peptides C1 and C2. So far we have found one nonprotein component of the cow colostrum trypsin inhibitor, *i.e.* hexosamines. In the course of this study we were able to isolate peptide Asx-Ser-Thr to which this component is bound. In our opinion the bond involves residue at position No 27 (Asx) which we have not been able to ascribe either to aspartic acid or asparagine. An analogous case is position No 46 where again, due to the bond to the nonprotein moiety, we cannot distinguish between aspartic acid and asparagine. There were also difficulties with the determination of amino acids at position No 10 and 11 of the inhibitor molecule. When peptide T1 was analyzed by Edman degradation,

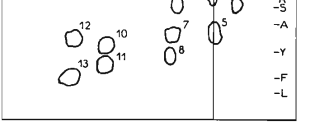


FIG. 4

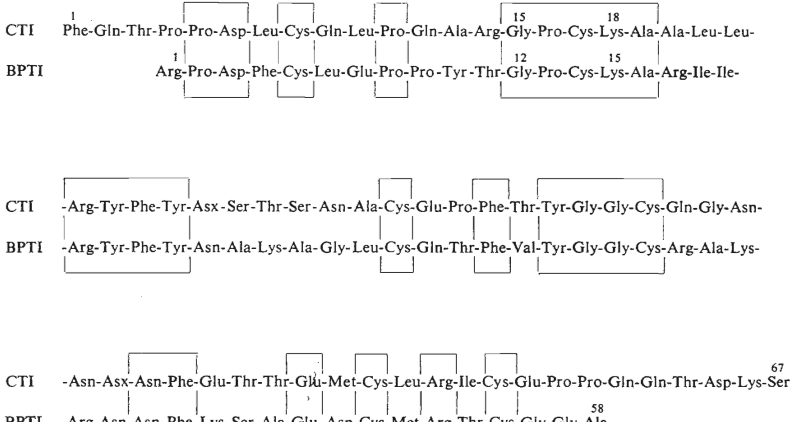
Peptide Map of Chymotryptic Digest of S-Sulfonated Inhibitor

Horizontally high-voltage electrophoresis at 3000 V and pH 1·9; vertically descending chromatography in the system 1-butanol–pyridine–acetic acid–water (15 : 10 : 3 : 12); REF, reference amino acids; 1—13 designation of chymotryptic peptides C1—C13.

TABLE V

Similarities in Known Structures of Trypsin Inhibitors

The occurrence of identical amino acids is marked by a square. CTI trypsin inhibitor from cow colostrum, BPTI basic pancreatic trypsin inhibitor.



leucine together with proline were always found at position No 10 and No 11. Originally we postulated the existence of variability of proline and leucine at this site of the molecule[21]. Tryptic digestion of the aminoethylated protein and of aminoethylated peptide T1 afforded short peptides which helped us to solve this ambiguity and to assign leucine to position No 10 and proline to position No 11.

The primary structure of two trypsin inhibitors from beef pancreas has already been determined[1,4]. A comparison of these structures with that of the trypsin inhibitor from cow colostrum (Table V) shows that the latter is more similar to the pancreatic basic inhibitor. Originally there existed one gene responsible for the synthesis of both these inhibitors, from which two different genes had formed by duplication and independent mutations in the process of phylogenetic development.

REFERENCES

1. Dayhoff M. O., Eck R. V.: *Atlas of Protein Sequence and Structure 1969*. National Biomedical Research Foundation, Silver Spring 1969.
2. Vogel R., Trautschold I., Werle E.: *Natürliche Proteinasen-Inhibitoren*. G. Thieme, Stuttgart 1966.
3. Chauvet J., Acher R.: Int. J. Protein Research *II*, 165 (1970).
4. Greene L. J., Bartelt D. C.: J. Biol. Chem. *244*, 2646 (1969).
5. Laskowski M. jr, Mars P. H., Laskowski M.: J. Biol. Chem. *198*, 745 (1952).
6. Čechová D., Jonáková-Švestková V., Šorm F.: This Journal *35*, 3085 (1970).
7. Spackman D. H., Stein W. H., Moore S.: Anal. Chem. *30*, 1190 (1958).
8. Hirs C. H. W.: J. Biol. Chem. *219*, 611 (1956).
9. Franěk F., Zorina O. M.: This Journal *32*, 3229 (1967).
10. Raftery M. A., Cole R. D.: J. Biol. Chem. *241*, 3457 (1966).
11. Prusík Z., Keil B.: This Journal *25*, 2049 (1960).
12. Gray W. R., Hartley B. S.: Biochem. J. *89*, 59P (1963).
13. Gross C., Labonesse B.: J. Biochem. (Tokyo) *7*, 463 (1969).
14. Novotný J., Franěk F.: Chem. listy *62*, 995 (1968).
15. Akabori S., Ohno K., Narita K.: Bull. Chem. Soc. Japan *25*, 214 (1952).
16. Niall H., Edman P.: J. Gen. Physiol. *45*, 185 (1962).
17. Grüner K.: Chem. listy, in press.
18. Edman P., Sjöquist J.: Acta Chem. Scand. *10*, 1507 (1956).
19. Sjöquist J.: Acta Chem. Scand. *7*, 447 (1953).
20. Gray W. R., Hartley B. S.: Biochem. J. *89*, 379 (1963).
21. Dayhoff M. O., Eck R. V.: *Atlas of Protein Sequence and Structure 1970*. National Biomedical Research Foundation, Silver Spring 1970, in press.

Translated by V. Kostka.