AMINO ACID SEQUENCE OF THE ACIDIC ACROSIN INHIBITOR (BUSI I B2) FROM BULL SEMINAL PLASMA

Bedřich MELOUN[a], Věra JONÁKOVÁ[b] and Dana ČECHOVÁ[b]

[a] *Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Sciences, 166 10 Prague and*
[b] *Institute of Molecular Genetics, Czechoslovak Academy of Sciences, 166 37 Prague*

Received June 28th, 1982

The molecule of the inhibitor consists of 63 amino acid residues whose sequence is the following: Glu-Ile-Tyr-Phe-Glu-Pro-Asp-Phe-Gly-Phe-Pro-Pro-Asp-Cys-Lys-Val-Tyr-Thr-Glu-Ala-Cys-Thr-Arg-Glu-Tyr-Asn-Pro-Ile-Cys-Asp-Ser-Ala-Ala-Lys-Thr-Tyr-Ser-Asn-Glu-Cys-Thr-Phe-Cys-Asn--Glu-Lys-Met-Asn-Asn-Asp-Ala-Asp-Ile-His-Phe-Gln-His-Phe-Gly-Glu-Cys-Glu-Tyr. The sequential data were obtained by the analysis of peptides isolated from the tryptic and chymotryptic digest of the carboxymethylated inhibitor. The molecular weight of the inhibitor calculated from its amino acid sequence is 7377.

The fertilization process involves at the molecular level[1] the participation of acrosin, a trypsin-like proteinase present in the acrosome of the intact spermatozoa. Efficient acrosin inhibitors of protein character are present in the seminal plasma of numerous animals[2]. Two types of acrosin inhibitors[3] of a relatively low molecular weight have been isolated from bull seminal plasma. These inhibitors are specific of the bull reproductive tract[4]; they differ in their isoelectric point, amino acid composition, and inhibitory properties[3]. The amino acid sequence of the basic inhibitor has been elucidated in a preceding study[5]. The acidic inhibitor was obtained in three forms differing in isoelectric points (pI = 3·9, 3·4, 3·7). The amino acid composition of the acidic acrosin inhibitor designated by the authors[6] BUSI I B2 (pI = 3·7) corresponds to the formula Asx_{10},Thr_4,Ser_2,Glx_9,Pro_4,Gly_2,Ala_4,Cys_6,Val_1,Met_1, Ile_3,Tyr_5,Phe_6,Lys_3,His_2,Arg_1 (in molar ratios). The inhibitor molecule is stabilized by three disulfide bonds.

The amino acid sequence of the inhibitor was derived in this study from overlapping peptides obtained by tryptic and chymotryptic digestion of the S-carboxymethyl derivative of the inhibitor.

EXPERIMENTAL

Material

The acidic acrosin inhibitor (BUSI I B2) was isolated as described before[6]. Bovine TPCK-trypsin treated with chloro-(N-*p*-tosyl-L-phenylalanyl)methane[7] and bovine chymotrypsin were purchased from Worthington Biochem. Corp., Freehold, New Jersey, USA. Thermolysin was a B-grade product of Calbiochem, Los Angeles, CA., USA. 2-Mercaptoethanol was a product of Fluka, Buchs, Switzerland. N-Ethylmorpholine was a product of Koch-Light, Colnbrook, England. Fluram (fluorescamine) was purchased from Serva, Heidelberg, FRG. The reagents used for the thiazolinone degradation were obtained from Pierce Eurochemie, Rotterdam, Holland. Sephadex G-25 fine and G-50 (superfine) were products of Pharmacia, Uppsala, Sweden. Polyamide layer sheets were purchased from BDH Chemical Ltd., Poole, England. TLC-plates (cellulose precoated plates) without the fluorescent indicator were from Merck, Darmstadt and cellulose MN 300 from Macherey-Nagel, Düren, FRG. Whatman No 3MM paper for chromatography was a product of Whatman Biochemicals Ltd., Maidstone, England.

Methods

The amino acid analyses were carried out according to Spackman and coworkers[8]. The samples of peptide material were hydrolyzed in 100 μl of 6M-HCl containing 0·02% of 2-mercaptoethanol[9] *in vacuo* for 20 h at 110°C. The analyses of the hydrolysates were performed in Durrum D-500 amino acid analyzer (Palo Alto, CA., U.S.A.) at a sensitivity of 1·0 A. The peptides (0·1% solutions in 0·2M-NH_4HCO_3) were digested with trypsin or chymotrypsin at an enzyme to substrate ratio of 1 : 40 for 2 h at 37°C; subsequently the same quantity of the enzyme was added and the digestion was continued for 2 additional hours. The thermolysin digestion of the peptides (0·1% solutions in 0·2M N-ethylmorpholine acetate, pH 8·1) was allowed to proceed for 4 h at 50°C. All digests were lyophilized before further treatment. The stock 0·1% solutions of the enzymes (trypsin, chymotrypsin, thermolysin) in water were storred at −20°C and used for the period of one month. To perform limited cleavage of the peptides in an acid medium[10], the peptides (80 nmol) were dissolved in 1·5 ml of dilute hydrochloric acid (pH 2·0) and the solutions heated 150 min at 108°C. Hydrochloric acid was dried off in a desicactor over KOH. The analytical and micropreparative separation of peptide mixtures was affected on Merck TLC glass plates (20 × 20 cm) coated with a cellulose layer. Plates of equal dimensions were used for the preparation of TLC layers from MN 300 cellulose. The electrophoretic separation of peptide mixtures was carried out on these plates in the apparatus of Pastuska[11] at a voltage of 500 V, 1 h in buffers S1—S3: acetic acid–pyridine–water (20 : 2 : 178, v/v), pH 3·5, buffer[12] S1; acetic acid–pyridine–water (1 : 20 : 179, v/v), pH 6·5, buffer[13] S2; N-ethylmorpholine–acetic acid–water (1·6 : 0·3 : 198·1, v/v), pH 8·0, buffer[14] S3. The chromatographic separation was carried out in the mixture[15] 1-butanol–pyridine–acetic acid–water (15 : 10 : 3 : 12, v/v), system S4. N(2,4-dinitrophenyl)ethylenediamine, whose R_F in this system is slightly higher than that of leucine, was employed as a visual marker. The analytical peptide maps were prepared by separation in one of the above systems (S1—S3) in the first direction and by chromatography in system S4 in the second direction. For preparative separation of peptide mixtures on Whatman No 3MM paper descending electrophoresis[16] at a potential gradient of 30 V/cm and high voltage electrophoresis[17] at a potential gradient of 85 V/cm were used. Preparative electrophoresis was carried out in the buffer acetic acid–formic acid (98%)–water (150 : 50 : 800, v/v, pH 1·9, buffer[17] S5), in buffer[12] S1 at pH 3·5, in the buffer acetic acid–pyridine–water (1 : 5 : 994, v/v, pH 5·6, buffer[16] S6), and in buffer[14] S3 at pH 8·0. For preparative peptide maps 100 nmol of the enzymatic digest

of the peptide was applied to one sheet of Whatman No 3MM paper. The separation in the first direction was carried out in systems S1, S3, S5, or S6 and chromatography in system S4 was used for the second direction. The peptides on TLC maps were stained with the reagent[18] prepared by dissolving 300 mg of ninhydrin in a stock solution containing 100 mg of cadmium acetate, 10 ml of water, 4 ml of acetic acid, and 96 ml of acetone. The peptides were stained on preparative maps with 0·001% solution of fluorescamine[19] in acetone; the paper was dipped in 5% pyridine solution in order to adjust the pH before the staining. The fluorescing spots were traced by pencil at 366 nm, the corresponding parts of paper were cut out, and the chromatograms were stained with 0·2% solution of ninhydrin in acetone. The peptides were eluted from paper by 0·1M-NH_4OH. The peptides (10–15 nmol) were sequenced by the DABITC (4-N,N--dimethylaminoazobenzene 4'-isothiocyanate double couplig) method[20]. Certain peptides containing S-carboxymethylcysteine were available in oxidized form for the sequencing. The oxidized form of the DABTH-S-carboxymethylcystein (DABTH = 4-N,N-dimethylaminoazobenzene--4'-thiohydantoin), coincided with the spot of DABTH.$CySO_3H$ in the system introduced by Chang and coworkers[20]. DABTH-Gln or DABTH-Asn present in the individual sequential steps were converted for the most part into DABTH-Glu and DABTH-Asp respectively, by the addition of 20 μl of 6M-HCl and heating 1 h at 54°C. These derivatives could then be easily distinguished from DABTH-Ser and DABTH-Thr which are separated on the polyamide layers only with difficulties. Additional sequential procedures were used with peptides containing serine or threonine. They were cleaved stepwise by the phenylisothiocyanate method according to Sauer[21] and the newly formed N-terminal residues were determined by dansylation[22]. The DNS derivatives (DNS = 1-dimethylaminonaphthalene-5-sulfonyl) of amino acids were identified in 1·5% formic acid[23] in the first direction and in the system benzene–acetic acid[34] (9 : 2, v/v) in the second direction. DNS-carboxymethylcysteine was located by additional chromatography (in the second direction) in the system[25] 1M-NH_4OH–ethanol (1 : 1, v/v).

Reduction and carboxymethylation of inhibitor: The inhibitor (30 mg) was reduced by 2--mercaptoethanol in 9M urea[26] dissolved in 0·5M Tris-HCl buffer, pH 8·7 and the carboxymethylation was carried out according to Crestfield and coworkers[27]. The S-carboxymethyl derivative of the inhibitor was separated from the reaction products by gel filtration on a column of Sephadex G-25 (2·7 × 47 cm) in 0·05M-NH_4OH. The protein-containing fraction was lyophilized.

N-*Terminal amino acid analysis of inhibitor*. The degradation of 50 nmol of the carboxymethylated inhibitor was carried out by the DABITC method using a five-fold higher quantity of all reagents than that described by the authors[20] for the usual procedure.

C-*Terminal end group analysis of the inhibitor*: The inhibitor (50 nmol) was treated with hydrazine[28] as described elsewhere[29].

Tryptic digest of carboxymethylated inhibitor: The inhibitor (15 mg) was dissolved in 1·5 ml of 0·15M-NH_4HCO_3 and digested with trypsin (0·5 mg) dissolved in 0·005M-HCl (0·3 ml). The digestion was allowed to proceed 4 h at 37°C. The digest without any concentration was applied onto a column of Sephadex G-50 superfine (1·6 × 57 ml) which was subsequently eluted by 0·15M--NH_4HCO_3. Fractions of 4 ml were collected at 7 min intervals. The course of the fractionation was monitored by measuring the effluent absorbance at 280, 254, and 230 nm. All peptides absorbed at 254 nm, were not resolved, however. The fraction containing a small quantity of a high molecular weight product was discarded. The main portion of the peptides was lyophilized and then separated chromatographically by repeated development in system S4. The individual peptide fractions obtained were separated electrophoretically in system S6.

Chymotryptic digest of carboxymethylated inhibitor: The inhibitor (5 mg) was dissolved in 5 ml of 0·2M N-ethylmorpholine acetate, pH 8·1, and 0·1 mg of chymotrypsin dissolved in 100 µl of water was added to the solution. The hydrolysis of the inhibitor was allowed to proceed 5 h at 37°C. The hydrolysate was lyophilized and separated chromatographically in system S4 using repeated development. The fraction showing an R_F-value similar to alanine was then re-resolved electrophoretically in system S3.

RESULTS AND DISCUSSION

The S-carboxymethyl derivative of the inhibitor, whose amino acid composition was identical with that of the active inhibitor[6], was prepared as starting material for sequential work. The analysis of the N-terminal region of the inhibitor molecule by the DABITC method[20] permitted the sequence of residues No 1 – 12 to be determined as Glu-Ile-Tyr-Phe-Glu-Pro-Asp-Phe-Gly-Phe-Pro-Pro. Hydrazinolysis of the intact inhibitor revealed tyrosine as the only C-terminal residue.

The amino acid sequence of the inhibitor was derived from the analysis of peptides which had been obtained from the tryptic and chymotryptic digest of the carboxymethylated inhibitor. The results of the sequential analysis of all the peptides isolated are given in Table I and the amino acid composition of the peptides in Table II. The relations among these peptides which permitted the complete amino acid sequence of the acidic acrosin inhibitor (BUSI I B2) to be determined are shown in Fig. 1. The optimal conditions of treatment of the enzymatic digests of the inhibitor or of its main fragments were studied by the method of peptide maps prepared on thin layers of cellulose.

Five peptides designated T1, T2, T3, T4, and T5 were observed on the peptide map (Fig. 2) of the tryptic digest of the CM-inhibitor. Peptides T4 and T5 are very little separated from one another; a considerably better resolution was achieved by paper chromatography in system S4. All peptides expected with respect to the known amino acid composition of the inhibitor and to the specificity of trypsin were isolated from the tryptic digest. The partial amino acid sequence only of peptide T1 was determined; this sequence was verified by the analysis of its thermolytic fragments. Peptide T1 is more easily soluble in 50% pyridine than in water. Peptides T2 and T3 were sequenced completely. When peptide T2 was treated by the DABITC technique a partial oxidation of the carboxymethylcysteine residue was observed. The sequence of the C-terminal part of peptide T3 was verified by analysis of dipeptide T3A1 isolated from the limited acid hydrolysate of the parent peptide. Sequential analysis of peptide T4 yielded its partial structure only, including the cysteine residue; the products obtained in the subsequent steps could not be identified. Peptide T4 was therefore cleaved with chymotrypsin and fragments T4C1 and T4C2 were obtained. The latter was sequenced completely without difficulties. The sequence of the last tryptic peptide, T5, consisting of 17 amino acid residues emerged from the analysis of the parent peptide and its chymotryptic fragments; the sequence of the last four residues remained undetermined.

TABLE I

Amino acid sequence of peptides isolated from tryptic (T) and chymotryptic (C) digest of carboxymethylated inhibitor, including data on fragments obtained by thermolytic (Th) or acid (A) clevage of peptides

Designation of peptides	Position	Amino acid sequence
T1	1—15	Glu-Ile-Tyr-Phe(Glx,Pro)Asp-Phe(Gly,Phe,Pro,Pro,Asx,Cys,Lys)
T1Th1	1—3	Glu-Ile-Tyr
T1Th2	4—7	(Phe, Glx, Pro, Asx)
T1Th3	8—15	(Phe, Gly, Phe, Pro, Pro, Asx, Cys, Lys)
T2	16—23	Val-Tyr-Thr-Glu-Ala-Cys Thr-Arg
T3	24—34	Glu-Tyr-Asn-Pro-Ile-Cys-Asp-Ser-Ala-Ala-Lys
T3A1	33—34	Ala-Lys
T4	35—36	Thr-Tyr-Ser-Asn-Glu-Cys-Thr-Phe-Cys(Asx,Glx,Lys)
T4C1	35—42	(Thr,Tyr,Ser,Asx,Glx,Cys,Thr,Phe)
T4C2	43—46	Cys-Asn-Glu-Lys
T5	47—63	Met-Asn-Asn-Asp-Ala-Asp-Ile-His-Phe-Gln(His,Phe,Gly,Glx,Cys, Glx,Tyr)
T5C1	47—55	(Met,Asx,Asx,Asx,Ala,Asx,Ile,His,Phe)
T5C2	56—63	Gln-His-Phe-Gly(Glx,Cys,Glx,Tyr)
C1	1—17	Glu-Ile-Tyr(Phe,Glx,Pro,Asx,Phe,Pro,Pro,Asx,Cys,Lys,Val,Tyr)
C1Th1	1—3	Glu-Ile-Tyr
C1Th2	4—15	(Phe,Glx,Pro,Asx,Phe,Gly,Phe,Pro,Pro,Asx,Cys,Lys)
C1Th3	16—17	Val-Tyr
C1Th4	4—9	Phe-Glu-Pro-Asp-Phe-Gly
C1Th5	9—15	Gly-Phe-Pro-Pro-Asp-Cys-Lys
C1Th6	8—15	(Phe,Gly,Phe,Pro,Pro,Asx,Cys,Lys)
C1A1	14—17	Cys-Lys-Val-Tyr
C2	18—25	Thr-Glu-Ala-Cys(Thr,Arg,Glx,Tyr)
C2T1	24—25	Glu-Tyr
C3	26—36	Asn-Pro-Ile-Cys(Asx,Ser,Ala,Ala,Lys,Thr,Tyr)
C3T1	35—36	Thr-Tyr
C4	37—42	Ser-Asn-Glu-Cys(Thr,Phe)
C5	43—55	Cys-Asn-Glu-Lys-Met-Asn-Asn-Asp-Ala-Asp-Ile(His,Phe)
C6	56—63	Gln-His-Phe-Gly(Glx,Cys,Glx,Tyr)
C5+6	43—63	Cys-Asn(Glx,Lys,Met,Asx,Asx,Asx,Ala,Asx,Ile,His,Phe,Glx,His, Phe,Glx,Cys,Glx,Tyr)
C5+6Th1	43—46	Cys-Asn-Glu-Lys
C5+6Th2	47—52	Met-Asn-Asn-Asp-Ala-Asp
C5+6Th3	53—54	Ile-His
C5+6Th4	55—57	Phe-Gln-His
C5+6Th5	58—63	Phe-Gly-Glu-Cys-Glu-Tyr

The peptide map of the chymotryptic digest of the CM-inhibitor (Fig. 3) showed the presence of seven peptides well resolved in system S4. The separation of peptides C2, C3, and C4 by electrophoresis was achieved only when the buffer at pH 8·0 (S3) was used. The separation at other pH-values (1·9, 3·5, 5·6) was insufficient. The amino acid sequence of the first three residues only of peptide C1 was determined; the main

TABLE II

Amino acid composition of peptides isolated from tryptic (T) and chymotryptic (C) digest of carboxylated inhibitor including fragments obtained by thermolytic (Th) or acid (A) cleavage of peptides

Designation of peptides	Number of amino acid residues
T1	$Cys_{0.7}$,$Glu_{1.9}$,$Pro_{2.9}$,$Gly_{.11}$,$Ile_{C.6}$,$Tyr_{C.9}$,$Asp_{1.9}$,$Phe_{2.8}$,$Lys_{C.9}$
T1Th1	$Glu_{1.1}$,$Ile_{0.9}$,$Tyr_{1.0}$
T1Th2	$Asp_{1.0}$,$Glu_{0.8}$,$Pro_{1.1}$,$Phe_{1.2}$
T1Th3	$Cys_{C.9}$,$Asp_{1.0}$,$Pro_{2.0}$,$Gly_{1.0}$,$Phe_{1.7}$,$Lys_{1.0}$
T2	$Cys_{0.6}$,$Thr_{2.3}$,$Glu_{1.3}$,$Ala_{1.2}$,$Val_{0.8}$,$Tyr_{1.0}$,$Arg_{1.0}$
T3	$Cys_{0.8}$,$Asp_{2.0}$,$Ser_{1.0}$,$Glu_{1.0}$,$Pro_{1.0}$,$Ala_{1.8}$,$Ile_{0.8}$,$Tyr_{1.0}$,$Lys_{1.0}$
T3A1	$Ala_{1.0}$,$Lys_{1.1}$
T4	$Cys_{1.6}$,$Asp_{2.3}$,$Thr_{1.8}$,$Ser_{1.1}$,$Glu_{2.0}$,$Tyr_{0.7}$,$Phe_{0.9}$,$Lys_{1.0}$
T4C1	$Cys_{0.7}$,$Asp_{1.0}$,$Thr_{1.6}$,$Ser_{0.9}$,$Glu_{1.2}$,$Tyr_{0.9}$,$Phe_{1.0}$
T4C2	$Cys_{0.5}$,$Asp_{1.1}$,$Glu_{1.1}$,$Lys_{1.0}$
T5	$Cys_{1.1}$,$Asp_{4.0}$,$Glu_{3.2}$,$Gly_{1.0}$,$Ala_{0.9}$,$Met_{0.6}$,$Ile_{0.8}$,$Tyr_{0.9}$,$Phe_{2.0}$,$His_{1.7}$
T5C1	$Asp_{3.7}$,$Ala_{1.0}$,$Met_{0.8}$,$Ile_{0.8}$,$Phe_{1.0}$,$His_{0.9}$
T5C2	$Cys_{0.9}$,$Glu_{2.8}$,$Gly_{1.1}$,$Tyr_{0.7}$,$Phe_{1.4}$,$His_{0.9}$
C1	$Cys_{1.0}$,$Asp_{2.1}$,$Glu_{1.9}$,$Pro_{3.0}$,$Gly_{1.1}$,$Val_{1.0}$,$Ile_{C.7}$,$Tyr_{1.7}$,$Phe_{3.1}$,$Lys_{1.0}$
C1Th1	$Glu_{0.9}$,$Ile_{1.0}$,$Tyr_{1.0}$
C1Th2	$Cys_{1.0}$,$Asp_{2.0}$,$Glu_{1.2}$,$Pro_{2.7}$,$Gly_{1.3}$,$Phe_{2.6}$,$Lys_{1.1}$
C1Th3	$Val_{1.0}$,$Tyr_{1.0}$
C1Th4	$Asp_{1.0}$,$Glu_{1.0}$,$Pro_{0.8}$,$Phe_{1.5}$
C1Th5	$Cys_{1.0}$,$Asp_{1.2}$,$Pro_{2.0}$,$Gly_{1.0}$,$Phe_{1.1}$,$Lys_{1.0}$
C1Th6	$Cys_{1.0}$,$Asp_{1.0}$,$Pro_{2.0}$,$Gly_{1.0}$,$Phe_{1.8}$,$Lys_{0.9}$
C1A1	$Cys_{0.7}$,$Val_{1.0}$,$Tyr_{1.0}$,$Lys_{1.3}$
C2	$Cys_{1.0}$,$Thr_{1.8}$,$Glu_{1.9}$,$Ala_{1.2}$,$Tyr_{0.4}$,$Arg_{0.8}$
C2T1	$Glu_{1.0}$,$Tyr_{0.9}$
C3	$Cys_{1.0}$,$Asp_{1.9}$,$Thr_{1.0}$,$Ser_{0.9}$,$Pro_{1.1}$,$Ala_{1.9}$,$Ile_{C.9}$,$Tyr_{1.0}$,$Lys_{1.0}$
C3T1	$Thr_{0.9}$,$Tyr_{1.0}$
C4	$Cys_{1.6}$,$Asp_{1.0}$,$Thr_{0.9}$,$Ser_{0.8}$,$Glu_{1.2}$,$Phe_{1.0}$
C5	$Cys_{1.1}$,$Asp_{5.1}$,$Glu_{0.8}$,$Ala_{1.0}$,$Met_{C.6}$,$Ile_{C.7}$,$Phe_{1.2}$,$His_{0.9}$,$Lys_{C.9}$
C6	$Cys_{0.7}$,$Glu_{2.8}$,$Gly_{1.1}$,$Tyr_{0.4}$,$Phe_{1.0}$,$His_{C.7}$
C5+6	$Cys_{2.0}$,$Asp_{4.7}$,$Glu_{3.7}$,$Gly_{C.8}$,$Ala_{1.C}$,$Met_{C.8}$,$Ile_{C.9}$,$Tyr_{C.8}$,$Phe_{2.0}$,$His_{1.7}$,$Lys_{1.1}$

1 10 20
Glu-Ile-Tyr-Phe-Glu-Pro-Asp-Phe-Gly-Phe-Pro-Pro-Asp-Cys-Lys-Val-Tyr-Thr-Glu-Ala-
T1
T2
T1Th1
T1Th2
T1Th3
C1
C2
C1Th1
C1Th2
C1Th3
C1Th4
C1Th5
C1Th6
N
C1A1
30 40
Cys-Thr-Arg-Glu-Tyr-Asn-Pro-Ile-Cys-Asp-Ser-Ala-Ala-Lys-Thr-Tyr-Ser-Asn-Glu-Cys-
(T2)
T3
T4
T3A1
T4C1
(C2)
C3
C4
C2T1
C3T1

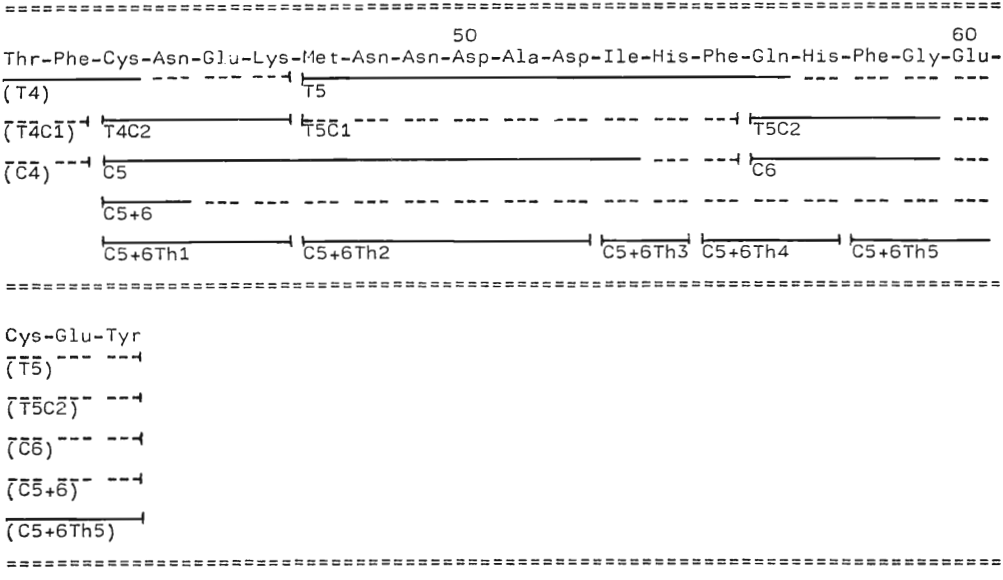

FIG. 1

Amino acid sequence of acid acrosin inhibitor (BUSI I B2). The peptides are designated by horizontal bars; sequences determined completely are marked by full lines, undetermined sequences by dashed lines. The symbols designating the individual peptides are explained in the legend to Table I. N designates the sequence determined with the uncleaved CM—inhibitor by the DABITS method

part of the material was subjected to thermolytic cleavage. Peptides C1Th1, C1Th4, and C1Th5 isolated from the latter digest were sequenced. The sequential data are in agreement with the structure of tryptic peptide T1. The sequence of the C-terminal region of peptide C1 was verified by sequential analysis of thermolytic fragment C1Th3 and of fragment C1A1, isolated from the partial acid hydrolysate of peptide C1. Peptide C2 was sequentially degraded up to the fourth amino acid residue, *i.e.* CM-cysteine; information on its C-terminal sequence was obtained by analyzing peptide C2T1 from the tryptic digest of peptide C2. By a similar approach peptide C3 was sequentially elucidated after the isolation of tryptic fragment C3T1. The CM-cysteine residue of peptide C4 had partly been converted into its oxidized form. The structure of the methionine-containing peptide C5 was determined up to the penultimate residue. Peptide C6 is identical to fragment T5C4 as regards its quantitative amino acid composition; it was also unambiguously sequenced up to the fourth amino acid residue, *i.e.* glycine. In addition to the main peptides a small quantity of peptide C5+6, incompletely digested by chymotrypsin, had also been isolated from the chymotryptic digest; this peptide was subsequently hydrolyzed by thermo-

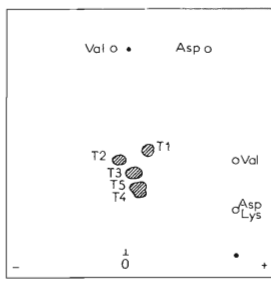

FIG. 2

Peptide map of tryptic digest of carboxymethylated inhibitor First direction — electrophoresis at pH 6·5, second direction ascending chromatography in the system n-butanol–pyridine–acetic acid–water (15 : 10 : 3 : : 12, v/v) on sheets coated with MN 300 cellulose. o origin, • reference mixture of amino acids (Lys, Asp, Val). The quantity of digest used for the preparation of the map was 10 nmol

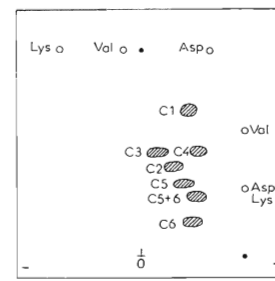

FIG. 3

Peptide map of chymotryptic digest of carboxymethylated inhibitor. First direction — electrophoresis at pH 8·0. The other data are identical to those given in the legend to Fig. 2

lysin. A series of peptides C5 + 6 Th1 – 5 had been isolated and characterized by qualitative amino acid analysis only; their quantity, however, was sufficient for sequential analysis. The thermolytic peptides C5 + 6 Th2 – 5, together with the partial structures of peptides T5 and C6 define the structure of the C-terminal region of the inhibitor, *i.e.* of residues No 47 – 63. Under the conditions of chymotryptic digestion used not all the bonds were cleaved which were expected to undergo hydrolysis with respect to the known specificity of the enzyme.

Peptides T1 – T5, and C1 – C6, isolated from the tryptic and chymotryptic digest of the inhibitor permit its complete amino acid sequence to be determined. The order of the tryptic peptides is the result of the following reasoning: The structure of peptide T1 is identical to the N-terminal region of the inhibitor determined by an independent procedure. Peptide C1 containing the only valine residue of the inhibitor molecule, determines the order of tryptic peptides T1 and T2. Chymotryptic peptide C2, characteristic because of its arginine content, defines the order of tryptic peptides T2 and T3. Likewise, chymotryptic peptide C3 containing two alanine residues affords overlaps permitting peptides T3 and T4 to be linked one to another. Another peptide, 35, contains the only methionine of the inhibitor and joins together the regions of the inhibitor molecule represented by tryptic peptides T4 and T5.

Peptides T1 – T5 and C1 – C6 isolated account for all the amino acid residues of the inhibitor. The majority of sequential work was carried out with tryptic peptides using the DABITC method. The C-terminal parts of some peptides, however, could not be determined, most likely because they had been extracted during the DABITC degradation. These peptides were therefore sequenced by the dansyl-Edman procedure.

The acid acrosin inhibitor BUSI I B2 consists of a single peptide chain containing 63 amino acid residues. The molecular weight of the inhibitor calculated from sequential data is 7377 and its amino acid composition is in accordance with the reported analytical data[6].

We wish to acknowledge the excellent technical assistance of Mrs A. Kulhánková. We are indebted to Mrs E. Dršková, Mrs V. Himrová and Mr J. Zbrožek for the amino acid analyses. We wish to thank Dr K. Jošt, Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Sciences, for the synthesis of 4-N,N-*dimethylaminoazobenzene* 4'-*isothiocyanate.*

REFERENCES

1. Hartree E. F.: Biochem. Soc. Trans. *5*, 375 (1977).
2. Fritz H., Schiessler H., Schill W. B., Tschesche H., Heimburger N., Wallner O. in the book: *Proteases and Biochemical Control* (E. Reich, D. B. Rifkin, E. Shaw, Eds), p. 737. Cold Spring Harbor Laboratory 1975.
3. Čechová D., Jonáková V.: Methods Enzymol. *80*, 792 (1981).
4. Veselský L., Čechová D.: Hoppe-Seyler's Z. Physiol. Chem. *361*, 715 (1980).
5. Meloun B., Čechová D.: This Journal *44*, 2710 (1979).

6. Čechová D., Jonáková V., Havránková M., Sedláková E., Mach O.: Hoppe-Seyler's Z. Physiol. Chem. *360*, 1759 (1979).
7. Kostka V., Carpenter F. H.: J. Biol. Chem. *357*, 401 (1964).
8. Spackman D. H., Stein W. H., Moore S.: Anal. Chem. *30*, 1190 (1958).
9. Keutmann H. T., Potts J. T. jr: Anal. Biochem. *29*, 175 (1969).
10. Inglis A., McKern N., Strike P. in the book: *Methods in Peptide and Protein Sequence Analysis* (Chr. Birr, Ed.), p. 329. Elsevier/North Holland Biochemical Press, Amsterdam, New York, Oxford 1980.
11. Pastuska G., Trinks H.: Chem.-Ztg. *85*, 535 (1961).
12. Ryle A. P., Sanger F., Smith L. F., Kitai R.: Biochem. J. *60*, 541 (1955).
13. Offord R. E.: Nature (London), *211*, 591 (1966).
14. Guidotti G., Hill R. J., Konigsberg W.: J. Biol. Chem. *337*, 2184 (1962).
15. Walley S. G., Watson J.: Biochem. J. *55*, 328 (1953).
16. Mikeš O.: This Journal *22*, 831 (1957).
17. Prusík Z., Štěpánek J.: J. Chromatogr. *87*, 73 (1973).
18. Heathcote J. G., Haworth C.: J. Chromatogr. *43*, 84 (1969).
19. Vandekerckhove J., VanMontagu M.: Eur. J. Biochem. *44*, 279 (1974).
20. Chang J. Y., Brauer D., Wittmann-Liebold B.: FEBS (Fed. Eur. Biochem. Soc.) Lett. *93*, 206 (1978).
21. Sauer R. T., Niall H. D., Hogan M. L., Keutmann H. T., O'Riordan J. L. H., Potts J. T. jr: Biochemistry *13*, 1994 (1974).
22. Gray W. R., Hartley B. S.: Biochem. J. *89*, 379 (1963).
23. Woods K. R., Wang K. T.: Biochim. Biophys. Acta *133*, 369 (1967).
24. Lee M. L., Safille A.: J. Chromatogr. *116*, 462 (1976).
25. Hartley B. S.: Biochem. J. *119*, 805 (1970).
26. Pospíšilová D., Meloun B., Frič I., Šorm F.: This Journal *32*, 4108 (1967).
27. Crestfield A. M., Moore S., Stein C. H. W.: J. Biol. Chem. *238*, 622 (1963).
28. Akabori S., Ohno K., Narita K.: Bull. Soc. Chem. Jap. *25*, 214 (1952).
29. Meloun B., Vaněček J., Šorm F.: Chem. Listy *52*, 523 (1958).

Translated by V. Kostka.