COVALENT STRUCTURE OF BOVINE TRYPSINOGEN. THE POSITION OF THE REMAINING AMIDES

O.Mikeš, V.Holeyšovský, V.Tomášek and F.Šorm

Institute of Organic Chemistry and Biochemistry,
Czechoslovak Academy of Science, Prague, Czechoslovakia

Received June 28, 1966

Our recent paper concerning bovine trypsinogen[1,2] presented a description of the primary structure of this protein, including the disulfide bridges. However, at that time an unequivocal assignment of all amino acid residues had not been made for several amino acid residues (positions 121, 177, 180). Thus it was not known whether these residues occurred in the form of acids or amides. The aspartic acid residue in position 151 has now been correctly determined as asparagine. The aim of the present communication was to remove all these uncertainties and to complete the investigation of the covalent structure of this protein. A preliminary report of this work was presented at the Third Federation Meeting of European Biochemical Societies[3].

Material and Methods

The peptides used were isolated from various digests of the protein and are summarized as follows:

Designation of the Peptide	Hydrolyzate	Reference
C 25, C 39	chymotryptic from S-sulfo-DIP-trypsin*	4
N 28	peptic from DIP-trypsin	5
T 1, T 12, T 30	tryptic from S-sulfo-trypsinogen	6,7

* Abbreviations: DANSYL- = dimethylaminonaphtalenesulfonyl-,
DIP- = diisopropylphosphoryl.

The quantitative analyses of these peptides can be found in the papers quoted. The charges of the peptides were estimated by descending electrophoresis[8] at pH 5.6. Growth tests for asparagine using a strain of Lactobacillus casei were carried out by methods developed in our laboratory[9,10]. Edman degradation of peptides[11] was carried out by the DANSYL-modification according to Gray and Hartley[12], the amino acid derivatives being identified by thin-layer chromatography[13-15]. The aminoethylation of reduced S-sulfocysteine peptides was carried out according to Raftery and Cole[16] (cf.ref.[17]).

Results

The composition and charge of the peptides, the asparagine tests and the position in the resulting structure are shown in Table I.

Peptide C 25 was neutral. Since it contains one basic amino acid (Lys) and one acidic residue (S-sulfo-Cys) the aspartic acid must be present in the amidated form. For the same reason peptide C 39 must contain all the acidic amino acid residues in the amide form. The acidic character of peptide N 28 from the region of the active center must be due not only to the aspartic acid residue, but also to the partly or completely hydrolyzed DIP-serine residue. Therefore, 0.75 μmole of the peptide was subjected to the DANSYL-Edman technique[12] using descending paper electrophoresis after each step. The original peptide gave a grayish-yellow colour with the ninhydrin reagent (0.2 % ninhydrine + 1 % collidine in acetone) typical for N-terminal glycine, and its mobility was 0.83 (relative to the mobility of glutamic acid). After the first degradation step it was blue-purple when detected with the ninhydrin reagent (indicating aspartic acid and not asparagine at the N-terminus) and its relative mobility rose to 0.95 since the acidic peptide was degraded beyond the neutral glycine residue. After the second step, when aspartic acid was split off, the relative mobility of the remaining fragment dropped to 0.53 which corresponds to the presence of a partly hydrolyzed isopropylphosphorylated serine residue in the peptide. After the third step the mobility dropped to zero and the grayish-yellow colour

Table I

Composition of Peptides Defining the Hitherto Unidentified Amides in Trypsinogen

(Cys stands for S-sulfo-cysteine, Ser* stands for isopropylphosphorylated or diisopropylphosphorylated serine)

Peptide Designation	Peptide Charge	Test for Asparagine	Peptide Structure	Position in Trypsinogen	Residue Defined
C 25	0		(Ala,Pro,Ile,Leu,Ser,Asn,Ser,Ser,Cys).Lys	146-155	151
C 39	0		(Leu,Ser,Asn,Ser,Ser,Cys,Lys,Ala,Tyr,Pro,Gly,Gln,Ile, ,Thr,Ser,Asn,Met,Phe)	149-167	151
N 28	-		Gly.Asp.Ser*.Gly(Gly,Pro)Val	181-187	182
T 1	-	+	Asn.Ser.Cys.Gln.Gly.Asp.Ser*.Gly.Gly.Pro.Val.Val.Cys. .Ser.Gly.Lys	177-192	177 180
T 24	-		Val.Ala.Ser.Ile.Ser.Leu.Pro.Thr.Ser.Cys.Ala.Ser.Ala. .Gly.Thr.Gln.Cys.Leu.Ile.Ser.Gly.Trp.Gly.Asn.Thr.Lys	106-131	121
T 30		+	Cys.Leu.Lys.Ala.Pro.Ile.Leu.Ser.Asn.Ser.Ser.Cys.Lys	143-155	151

Table II

Covalent Structure of Bovine Trypsinogen

```
                    *            I
Val.Asp.Asp.Asp.Asp.Lys.Ile.Val.Gly.Gly.Tyr.Thr.Cys.Gly.Ala.Asn.Thr.Val.Pro.Tyr.Gln.Val.Ser.Leu.Asn.Ser.Gly.Tyr.His.Phe.
                5                   10                  15                  20                  25                  30
II                                          * III
.Cys.Gly.Gly.Ser.Leu.Ile.Asn.Ser.Gln.Trp.Val.Val.Ser.Ala.Ala.His.Cys.Tyr.Lys.Ser.Gly.Ile.Gln.Val.Arg.Leu.Gly.Gln.Asp.Asn.
                35                  40                  45                  50                  55                  60

.Ile.Asn.Val.Val.Glu.Gly.Asn.Gln.Gln.Phe.Ile.Ser.Ala.Ser.Lys.Ser.Ile.Val.His.Pro.Ser.Tyr.Asn.Ser.Asn.Thr.Leu.Asn.Asn.Asp.
                65                  70                  75                  80                  85                  90
                                                                                            IV
.Ile.Met.Leu.Ile.Lys.Leu.Lys.Ser.Ala.Ala.Ser.Leu.Asn.Ser.Arg.Val.Ala.Ser.Ile.Ser.Leu.Pro.Thr.Ser.Cys.Ala.Ser.Ala.Gly.Thr.
                95                  100                 105                 110                 115                 120
    V                                                                               VI
.Gln.Cys.Leu.Ile.Ser.Gly.Trp.Gly.Asn.Thr.Lys.Ser.Ser.Gly.Thr.Ser.Tyr.Pro.Asp.Val.Leu.Lys.Cys.Leu.Lys.Ala.Pro.Ile.Leu.Ser.
                125                 130                 135                 140                 145                 150
            VII                                                     VIII                                    IX
.Asn.Ser.Ser.Cys.Lys.Ser.Ala.Tyr.Pro.Gly.Gln.Ile.Thr.Ser.Asn.Met.Phe.Cys.Ala.Gly.Tyr.Leu.Glu.Gly.Gly.Lys.Asn.Ser.Cys.Gln.
                155                 160                 165                 170                 175                 180
        *                   X                                                       XI
.Gly.Asp.Ser.Gly.Gly.Pro.Val.Val.Cys.Ser.Gly.Lys.Leu.Gln.Gly.Ile.Val.Ser.Trp.Gly.Ser.Gly.Cys.Ala.Gln.Lys.Asn.Lys.Pro.Gly.
                185                 190                 195                 200                 205                 210
                    XII
.Val.Tyr.Thr.Lys.Val.Cys.Asn.Tyr.Val.Ser.Trp.Ile.Lys.Gln.Thr.Ile.Ala.Ser.Asn.
                215                 220                 225             229
```

Disulfide bridges: I(13)–VI(143); II(31)–III(47); IV(115)–XII(216); V(122)–X(189); VII(154)–VIII(168); IX(179)–XI(203)

---

* Residues participating in the active site of trypsin[24–27]

with ninhydrin reagent corresponds to N-terminal glycine. The results of chromatography of DANSYL-derivatives of amino acids are included in Table I.

Peptide T 1 contains 2 residues of aspartic acid and gives a positive test for asparagine. Since residue 182 in peptide N 28 and also in other peptides[4,18] was determined as aspartic acid, the positive asparagine test in the peptide T 1 should be ascribed to residue No. 177. In order to determine the character of the residue in position No. 180, 0.375 μmole of fragment T 1 ox S 7 (prepared[19] by cleavage of peptide T 1 with subtilisin) was gradually degraded using the modified[12] Edman method. After the third step 20 % of the material was treated with leucine aminopeptidase and the glutamine split off was determined as its DANSYL-derivative by thin-layer chromatography. The determination of the other part of the sequence of peptide T 1 has been published elsewhere[18]. To determine residue 121, 20 μmole of peptide T 24 were hydrolyzed with 8 mg of pepsin (37°C, 6 h). Using chromatography on Dowex 50 and on paper the heptapeptide T 24-pA3 of composition (Ser,Ala,Gly,Thr,Gln,S-sulfo-Cys,Leu) was isolated. One μmole was reduced with 2-mercaptoethanol, aminoethylated with ethylenimine and directly desalted on a column of Sephadex G 10 (2.5 x 39 cm). The resulting product was of basic character on descending paper electrophoresis[8]. Since its hydrolysate contained only one basic residue (aminoethylcysteine) and one acidic amino acid (Glu) (cf.[18]), the latter must be present in the form of an amide. The determination of the amino acid sequence of the whole peptide, T 24, has been published in[18]. Peptide T 30 gave a positive test for asparagine and hence the only aspartic acid in position 151 must be in the form of an amide. This conclusion is in agreement with the results obtained with peptides C 25 and C 39.

The complete covalent structure of bovine trypsinogen including all the amide positions is presented in Table II.

Discussion

The results presented here resolve the uncertainties with regard to the determination of amides mentioned in our previous papers[1,2], and correct the residue

in position 151 to asparagine. The presence of asparagine in peptide T 30 had already been tentatively identified in a previous paper[7]. The structure of trypsinogen presented by Walsh and Neurath[20] includes 26 amides. The analysis for total amides reported by Viswanatha[21] showed 29 amides in trypsin in agreement with the analysis of DIP-trypsin by Zmrhal[22]. Our covalent structure of trypsinogen (Table II) includes 30 amides and differs from the structure proposed by Walsh and Neurath[20] in the replacement of serine (position 84) with proline (within the range from 84 to 87) as well as in the presence of amides in positions 58, 67, 68 and 151. Our total structure differs from the structure quoted by Hartley and co-workers[23] in the Ser/Pro replacement, in the reversal of the dipeptide sequence Thr.Gln (residues No. 120 and 121), and in the presence of amides in positions 58, 68 and 151.

Details of our final investigation of the amino acid sequence are now in press[4,18]. The knowledge of the complete covalent structure of bovine trypsinogen represents a solid base for an exact study of the tertiary structure of this protein and of the relationship between structure and function.

References

1. Holeyšovský V., Tomášek V., Mikeš O., Danilova A.S., Šorm F., Collection Czechoslov.Chem.Commun. 30, 3936 (1965)
2. Mikeš O., Tomášek V., Holeyšovský V., Šorm F., Biochim.Biophys.Acta 117, 281 (1966)
3. Mikeš O., Tomášek V., Holeyšovský V., Šorm F., IIIrd Federat.Meeting of European Biochemical Societies, Warsaw, 5 April, 1965
4. Mikeš O., Müller G.H., Holeyšovský V., Tomášek V., Šorm F., Collection Czechoslov.Chem.Commun., in the press
5. Holeyšovský V., Mesrob B., Tomášek V., Šorm F., Collection Czechoslov. Chem.Commun., to be published
6. Holeyšovský V., Alexijev B., Tomášek V., Mikeš O., Šorm F., Collection Czechoslov.Chem.Commun. 27, 2662 (1962)
7. Tomášek V., Mikeš O., Holeyšovský V., Šorm F., Collection Czechoslov. Chem.Commun. 29, 3122 (1964)
8. Mikeš O., Collection Czechoslov.Chem.Commun. 22, 831 (1957)
9. Mikeš O., Schuh V., Šorm F., Collection Czechoslov.Chem.Commun. 24, 583 (1959)
10. Mikeš O., Šorm F., Collection Czechoslov.Chem.Commun. 24, 1525, 1897 (1959)
11. Edman P., Acta Chem.Scand. 4, 283 (1950); 7, 700 (1953).
12. Gray W.R., Hartley B.S., Biochem.J. 89, 59P, 379 (1963)
13. Seiler N., Wiechman J., Experientia 20, 559 (1964)
14. Deyl Z., Rosmus J., J.Chromat. 20, 514 (1965)
15. Mesrob B., Holeyšovský V., J.Chromat. 21, 135 (1965)
16. Raftery M.A., Cole R.D., this Journal 10, 467 (1963)

17. Zmrhal Z., Vaněček J., Mészáros M., Sorm F., Collection Czechoslov.Chem. Commun. 30, 1339 (1965)
18. Mikeš O., Tomášek V., Holeyšovský V., Šorm F., Collection Czechoslov.Chem. Commun., in the press
19. Alexijev B., Holeyšovský V., Šorm F., Compt.Rend.Acad.Bulgare Sci. 15, 755 (1962); Collection Czechoslov.Chem.Commun. 29, 1296 (1964)
20. Walsh K.A., Neurath H., Proc.Natl.Acad.Sci. 52, 884 (1964)
21. Vishwanatha T., Lawson W.B., Witkop B., Biochim.Biophys.Acta 40, 216 (1960)
22. Zmrhal Z., Collection Czechoslov.Chem.Commun. 28, 2356 (1963)
23. Hartley B.S., Brown J.R., Kauffman D.L., Smillie L.B., Nature 207, 1157 (1965)
24. Schaffer N., Lang R.P., Simat L., Drisko R.W., J.Biol.Chem. 230, 185 (1958)
25. Dixon G.H., Kauffman D.L., Neurath H., J.Biol.Chem. 233, 1373 (1958); J.Am. Chem.Soc. 80, 1260 (1958)
26. Tomášek V., Severin E.S., Šorm F., Biochem.Biophys.Research Commun. 20, 545 (1965)
27. Hofmann T., Scrimger S.T., Federat.Proc., in the press (private communication from T.Hofmann).