PRIMARY STRUCTURE OF PEPTIDES WHICH FORM THE DISULFIDE BONDS OF CHICKEN PEPSIN

Helena KEILOVÁ, Miroslav BAUDYŠ and Vladimír KOSTKA

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

Received November 28th, 1978

The molecule of chicken pepsin is cross-linked by three disulfide bonds. The structures of the half-cystine peptides which form these bonds were determined by the analysis of different enzymic digests of the enzyme. The order of the disulfide bonds in the molecule was elucidated with regard to sequential homologies between chicken pepsin and other acid proteases. The three disulfide bonds of chicken pepsin, numbered from the N-terminus of pepsin, are:

1st bond Ile-Tyr-Cys-Lys-Ser-Ser-Ala-Cys-Ser-Asn-His-Lys;

2nd bond Val-Ala-Cys Phe-Thr-Cys-Gln-Ala;

3rd bond Asp-Leu-Gly-Val-Ser-Ser-Asp-Gly-Glu-Ile-Ser-Cys-Asp-Asp-Ile-Ser-Lys-Leu

-Pro-Asp Val-Leu-Asn-Glu-Asp-Gly-Ser-Cys-Met

All the animal proteases of the carboxyl type which have been studied so far have their molecules cross-linked by three disulfide bonds[1]. Pepsin isolated from chicken forestomachs, the first avian pepsin which has been subjected to structural studies, contains moreover a cysteine residue in its molecule. Studies on peptides from the tryptic digest[2] of chicken pepsinogen have revealed the existence of far-reaching sequential homologies between chicken pepsin and other pepsins, especially as regards the N- and C-terminal part of the molecule. There are also homologies[3] in amino acid sequences of cystine peptides derived from various proteases of the pepsin type.

This study, a part of a complex investigation of chicken pepsin, has been designed to elucidate the amino acid sequences around the half-cystine residues which form the disulfide bonds and to compare the degree of homology between chicken pepsin and other proteases of this type.

EXPERIMENTAL

Material

Chicken pepsinogen was prepared in this Laboratory as described earlier[4]. Chicken pepsin was obtained by the activation of pepsinogen by 1 mM-HCl and the peptides liberated during the

activation were separated by gel filtration on Sephadex G-100. Thermolysin was a 3-times crystallized preparation of Calbiochem, San Diego, Cal., U.S.A. Thiol-Sepharose and Sephadex G-25 fine were from Pharmacia Fine Chemicals, Sweden. Dowex 50X2 (200–400 mesh) was from Fluka, Switzerland. All the chemicals used were of analytical purity.

Methods

High-voltage electrophoresis was carried out on Whatman No 3 MM paper in the vertical arrangement at 1 500 V in the buffer pyridine–acetic acid–water (2 : 0·5 : 495·5, by vol.) at pH 5·6 (ref.[5]) or in the horizontal apparatus at 4 000 V in the buffer formic acid–acetic acid–water (50 : 150 : 800, by vol.) at pH 1·9 (ref.[6]).

The chromatography of peptides on Whatman No 3 MM paper was effected in the system n-butanol–pyridine–acetic acid–water (15 : 10 : 3 : 12, by vol.) (ref.[7]).

The detection of thiol peptides on paper was performed by sodium nitroprusside in the presence of sodium cyanide[8].

The oxidation of peptides was effected by performic acid[9].

The disulfides in peptides were determined colorimetrically using 5,5′-dithio-bis-(2-nitrobenzoic acid) (DTNB, ref.[10]).

Amino acid analysis. The samples were hydrolyzed in 6M-HCl, 20 h at 110°C in evacuated tubes. The amino acid content of the hydrolysate was determined by the method[11] modified according to Benson and Patterson[12] either in Beckman-Spinco Model 120 B Amino Acid Analyzer or in an automatic analyzer of Czechoslovak make (Instrument Development Workshops, Czechoslovak Academy of Sciences, Prague).

The N-terminal amino acids of peptides were determined by the dansyl technique[13].

The manual Edman degradation[14] and the detection of the amino acid phenylthiohydantoins were carried out as described earlier[15].

Enzymic Digests of Chicken Pepsin

Thermolysin digest. Pepsin (700 mg) was dissolved in 30 ml of 0·1M-Tris-HCl buffer, pH 8·0. Thermolysin was added in two portions (at 0 and 6 h); the final enzyme to substrate ratio was 1 : 30 (w/w). The digestion was allowed to proceed 18 h at 38°C. A small quantity of sediment was centrifuged off and the digest was placed on a column of Sephadex G-25 equilibrated with 0·1% NH_4HCO_3 which was also used for the elution. The absorbance of the fractions was measured at 230 and 280 nm; an aliquot was taken from each fraction for the determination of SH-groups by DTNB. According to the results of these analyses the effluent was divided into 13 fractions. These fractions were lyophilized and their composition examined by the technique of peptide maps (electrophoresis at pH 5·6 in the first direction, chromatography in the second direction). Peptides containing disulfides were detected on the map by the nitroprusside reagent. Certain fractions were subjected to ion-exchange chromatography on Dowex 50X2 (column 1·8 × 35 cm), equilibrated with 50 mM pyridine–formic acid buffer (pH 3). The peptides were eluted by a linear gradient of pyridine buffers: 50 mM pyridine–formic acid (pH 3)–0·2M pyridine–acetic acid (pH 4) (500 ml + 500 ml) and 0·2M pyridine–acetic acid (pH 4)–2M pyridine (500 ml + 500 ml). The flow rate was 30 ml/h.

Chymotryptic digest. Pepsin (200 mg) was dissolved in 18 ml of water and the pH of the solution was adjusted to 10·7 by 0·1M-NaOH. After 30-min inactivation the pH was decreased to 8·5 by 0·3M-HCl. Chymotrypsin (4 mg) and soy-bean trypsin inhibitor (4 mg) were added and the mixture was incubated 18 h at 38°C. The digest was rotary evaporated to dryness and then

subjected to electrophoretic separation at pH 5·6. Cystine peptides were again detected by the nitroprusside reagent.

Thermolysin digest of chicken pepsin immobilized on Thiol-Sepharose. Thiol-Sepharose (12 g), washed and activated[16] was equilibrated with 0·1M phosphate buffer, pH 7·8, containing 0·3M-NaCl and 1 mM-EDTA. Pepsin (140 mg) was dissolved in the same buffer and bonded to Thiol-Sepharose according to Jegorov and coworkers[16]. The quantity of pepsin attached, calculated from the quantity of thiopyridone determined spectrophotometrically at 343 nm, was 85 mg. The pepsin attached was digested with thermolysin in 0·1M-Tris-HCl buffer, pH 8 (enzyme to substrate ratio 1 : 10, w/w) 18 h at 38°C, with rocking-wise stirring and in an atmosphere of nitrogen. Subsequently the column was exhaustively washed and the material attached to the support through the SH-group displaced by 0·1% NH_4HCO_3 containing 20 mM mercaptoethanol. This fraction was lyophilized and analyzed by paper electrophoresis and chromatography.

RESULTS

The main source of cystine peptides for their isolation was the thermolysin digest. As can be seen in Fig. 1, this digest represented a rich mixture of peptides. Peptides

TABLE I

Amino Acid Composition of Peptides Containing Cystine or Cysteine and Isolated from Enzymatic Digests of Chicken Pepsin

The analyses were performed on samples hydrolyzed 20 h. None of the peptides contained arginine, proline or tryptophan.

Amino acid	mol/mol peptide							
	Th-1	Th-2	Th-3	Th-4	Th-5	Th-6	Th-7	Ch-1
Lysine							1·3	1·0
Histidine								0·9
Cysteic acid[a]	1·8	1·8	1·9	0·8	1·9	0·8	0·8	1·9
Aspartic acid		5·1	3·9	2·0				1·1
Methionine sulfone[a]		0·9	0·8					
Threonine	0·9			1·1	0·9	0·9		
Serine		3·8	2·0	1·9			1·6	2·8
Glutamic acid	1·0	2·0	1·0	1·1	1·0	1·0		
Glycine		2·1	1·1	2·0				
Alanine	0·9				1·0	1·0		1·0
Valine	0·9	1·8			1·0			
Isoleucine		0·8	0·8				1·2	
inecLeu		1·0	1·0					
Tyrosine							1·0	
Phenylala nine	1·0				1·0	1·0		

[a] Determined after oxidation of peptide in performic acid.

which gave a positive reaction with DTNB and were treated further, were contained in fractions I–VIII.

Fraction I contains unhydrolyzed or partly hydrolyzed pepsin and a mixture of peptides. From this fraction, homogeneous peptide Th-1, Val, Ala, Cys_2, Phe, Thr, Gln (Table I), was isolated by electrophoresis and paper chromatography. This peptide was oxidized and the two peptides formed, Val(Ala, $CysO_3H$) and Phe(Thr, $CysSO_3H$, Gln) were separated by electrophoresis at pH 5·6. The N-terminal groups of the peptides, valine and phenylalanine, respectively, were determined by dansylation. The main bulk of the peptide was found in fractions V–VIII and will therefore be discussed below.

Fraction II was chromatographed on paper and pure cystine peptide Th-2, Cys_2, Asp_5, Ser_4, Glu_2, Gly_2, Val_2, Ile, Leu, Met (Table I) was obtained. Oxidation cleaved this peptide to two peptides which were separated by paper chromatography. Both these peptides were N-terminated by valine and the total number of amino

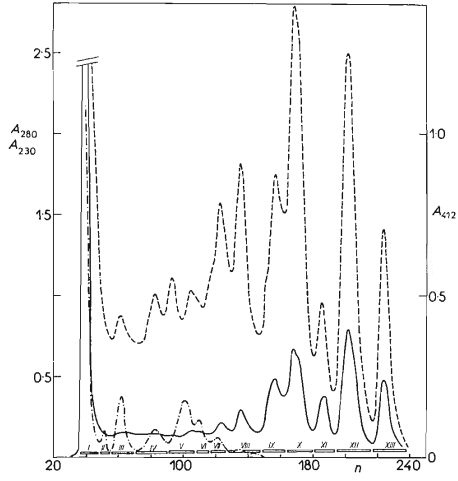

FIG. 1

Gel Filtration of Thermolysin Digest of Chicken Pepsin on Sephadex G-25 Fine

Column 4·4 × 200 cm, eluted by 0·1% NH_4HCO_3. Flow rate 48 ml/h. The fractions were collected every 15 min. ——— A_{280nm}, – – – – A_{230nm}, –·–·– A_{412nm} (detection with DTNB), *n* number of fractions.

acids in both peptides corresponded to the amino acid composition of the unoxidized peptide. The amino acid sequences of these peptides, determined by Edman degradation, are Val-Ser-Ser-Asp-Gly-Glu-Ile-Ser-Cys-Asp-Asp and Val-Leu-Asn-Glu-Asp--Gly-Ser-Cys-Met.

A peptide of high anodic mobility, Th-3, was isolated from fraction III by electrophoresis at pH 5·6 and purified further by paper chromatography. The amino acid composition of peptide Th-3 was $Asp_4,Ser_2,Glu,Gly,Cys_2,Met,Ile,Leu$ (Table I). After oxidation this peptide afforded two peptides which were separated by electrophoresis at pH 5·6 and purified by paper chromatography. The N-terminal groups of the peptides determined by dansylation were isoleucine and leucine. The sequences of these peptides elucidated by stepwise degradation are Ile-Ser-Cys-Asp-Asp and Leu-Asn-Glu-Asp-Gly-Ser-Cys-Met. This shows that the peptides are derived from the same disulfide bridge as the peptides isolated from fraction II. The data on this bridge can be complemented by the alignment of two overlapping peptides: Asp-Leu--Gly-Val-Ser-Ser-Asp-Gly-Glu-Ile-Ser-Cys-Asp-Asp-Ile-Ser-Lys, isolated from the tryptic digest of S-sulfo-pepsinogen[2], and Ile-Ser-Lys-Leu-Pro-Asp- from the thermolysin digest[2]. Summarizing all the data this disulfide bridge can be formulated as follows:

Thermolysin digest: Val-Ser-Ser-Asp-Gly-Glu-Ile-Ser-Cys-Asp-Asp

Thermolysin digest: Ile-Ser-Cys-Asp-Asp

Tryptic digest: Asp-Leu-Gly-Val-Ser-Ser-Asp-Gly-Glu-Ile-Ser-Cys-Asp-Asp-Ile-Ser-Lys

Thermolysin digest: Ile-Ser-Lys--Leu-Pro-Asp

Thermolysin digest: Leu-Asn-Glu-Asp-Gly-Ser-Cys-Met

Thermolysin digest: Val-Leu-Asn-Glu-Asp-Gly-Ser-Cys-Met

Disulfide bridge:

Asp-Leu-Gly-Val-Ser-Ser-Asp-Gly-Glu-Ile-Ser-Cys-Asp-Asp-Ile-Ser-Lys-Leu-Pro-Asp
Val-Leu-Asn-Glu-Asp-Gly-Ser-Cys-Met

Fraction IV was rechromatographed on a column of Sephadex G-25 in 0·1% NH_4HCO_3 and the DTNB positive fraction was separated from a mixture of small peptides. This fraction was then resolved into 4 zones by paper chromatography. The zone of the lowest R_F-value was found to contain cysteic acid by amino acid analysis. The peptide present in this zone, Th-4, was homogeneous when subjected to electrophoresis at pH 1·9 and its amino acid composition was $CySO_3H,Asp_2,Thr,Ser_2,Glu,Gly_2$ (Table I). The failure to reveal the N-terminal end group either by dansylation or by the phenylthiohydantoin technique indicated the presence of a nonreactive amino acid residue in this position. According to its amino acid

composition the peptide does not correspond to any of the three bridges; this leads us to assume that it is derived from the neighborhood of the odd, cysteine residue.

Fractions V–VIII were analyzed electrophoretically at pH 1·9 and found all to contain material giving a positive nitroprusside test; these fractions were pooled and chromatographed on a column of Dowex 50-X2. The individual fractions were examined by paper chromatography. Two peptides only of the very rich mixture gave a positive nitroprusside test, one peptide located in the approximately neutral region and one very little basic peptide which was not treated further. The neutral cystine peptide was purified by paper chromatography and electrophoresis at pH 1·9 and homogeneous peptide Th-5 (Table I) was obtained. Its N-terminal amino acids determined by dansylation were valine and phenylalanine. After oxidation two peptides, Val-Ala-Cys and Phe-Thr-Cys-Gln, were obtained which were separated by paper electrophoresis at pH 5·6. The same peptides were also isolated from fraction I. Another half-cystine peptide N-terminated by phenylalanine, Phe(Thr,Cys, Glu,Ala) (Th-6), was also isolated by paper chromatography from the thermolysin digest of pepsin attached to Thiol-Sepharose. The amino acid sequence around the disulfide bridge is therefore the folowing:

Thermolysin digest:	Val-Ala-Cys
Thermolysin digest:	Phe-Thr-Cys-Gln
Thermolysin digest of attached pepsin:	Phe(Thr,Cys,Gln,Ala)
Disulfide bridge:	Val-Ala-Cys Phe-Thr-Cys-Gln-Ala

The last bridge was obtained by the analysis of the chymotryptic digest. The latter was resolved first by electrophoresis at pH 5.6 and the presence of cystine detected in the most basic zone by the nitroprusside test. The material contained in this zone was eluted and resolved by paper chromatography. The cystine peptide was purified electrophoretically at pH 1·9. The amino acid composition of the homogeneous peptide obtained, CH-1, was $Lys,Cys_2,His,Asp,Ser_3,Ala$ (Table I). The peptide was not cleaved into two peptides by oxidation. It was therefore subjected to stepwise degradation and the sequence Cys-Lys-Ser-Ser-Ala(Cys,Ser,Asn,His) was established. The sequence of the sixth to ninth residue was deduced by comparison with the sequence of peptide TP-10 from the tryptic digest of S-sulfo-pepsinogen[2], Ser-Ser-Ala-Cys-Ser-Asn-His-Lys, which is derived from the same site of the polypeptide chain. Peptide Th-7 isolated from the thermolysin digest of pepsin attached to Thiol-Sepharose, Ile(Tyr,Cys,Lys,Ser,Ser) permits the sequence to be extended as follows:

Chymotryptic digest: Cys-Lys-Ser-Ser-Ala(Cys,Ser,Asn,His)

Thermolysin digest: Ile(Tyr, Cys, Lys, Ser, Ser)

Tryptic digest: Ser-Ser-Ala-Cys-Ser-Asn-His-Lys

Disulfide bridge: Ile-Tyr-Cys-Lys-Ser-Ser-Ala-Cys-Ser-Asn-His-Lys

DISCUSSION

The molecule of chicken pepsin, like the molecules of other acid proteases[3], contains three disulfide bonds of which the first and second one (from the N-terminus) form relatively short loops whereas the third one links more distant parts of the chain forming a long loop.

A comparison of the disulfide peptides from chicken pepsin with analogous peptides from other pepsins and chymosin is shown in Fig. 2. The structures of the proteases are aligned according to Foltman and Pedersen[3].

The peptides derived from the neighborhood of the first loop show a high degree of homology. Thus, *e.g.* chicken pepsin and chymosin share the same pentapeptide, Ile-Tyr-Cys-Lys-Ser; the positions occupied in this sequence by Tyr, Cys, and Ser are constant in all the proteases compared. Likewise constant are the position of dipeptide Ala-Cys and the position of histidine or Asn-His, respectively. It is known that this cystine peptide plays an important role in the activity of hog pepsin[17].

A high degree of homology is also observed with the second loop (from the N-terminus of the molecule) which immediately precedes the catalytic site[3]. Tripeptides Val-Ala-Cys and Cys-Gln-Ala, *e.g.* are identical in chicken pepsin and chymosin. As can be seen in Fig. 2 there are three amino acid residues between the half-cystines in all the pepsin-like enzymes shown, whereas only two residues were found in chicken pepsin. As of now it is impossible to determine whether there is a deletion or whether the missing amino acid residue was split off during hydrolysis.

The third bridge from the N-terminus forms a long loop and minor homologies only can be found in its neighborhood. Chicken pepsin is no exception in this respect – even though certain homologies can be observed.

The problem of the cysteine-containing peptide remains to be solved. We tried to isolate this peptide from the thermolysin digest of pepsin attached to Thiol-Sepharose. However, we observed that the disulfide bridges were also partly interrupted during this treatment; hence, the peptides derived from these bridges were found also when the material attached was displaced from the column by mercaptoethanol. Moreover, active pepsin present in the mixture also hydrolyzed glutathione. The sequence of the thiol peptide will therefore be determined in the course of systematic elucidation of the complete linear structure of the enzyme.

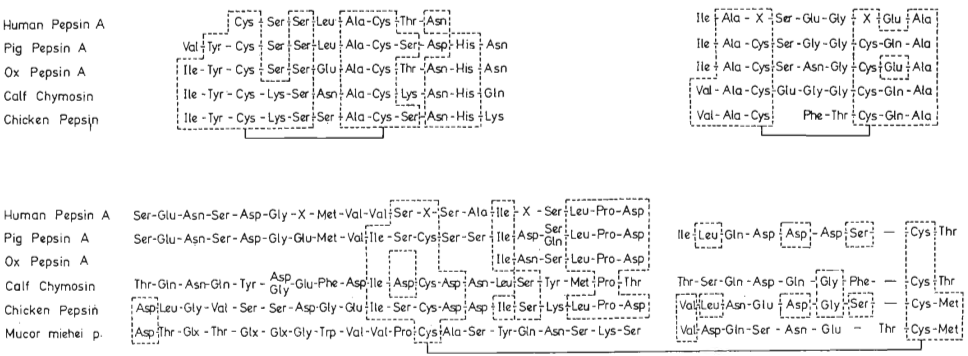

FIG. 2

Amino Acid Sequences Around Disulfide Bonds of Chicken, Human and Porcine Pepsin, of Chymosin and *Mucor miehei* Protease

The amino acid sequences of the proteins compared with chicken pepsin are taken from[3]. ------ encircles residues identical in two or more enzymes.

We acknowledge greatly the careful technical assistance of Mrs E. Bulantová. We thank Mr K. Grüner for the Edman degradation experiments and Mr J. Zbrožek for the amino acid analyses.

REFERENCES

1. Bohak Z.: J. Biol. Chem. *244*, 4638 (1969).
2. Baudyš M., Keilová H., Kostka V.: VIII. Celostátní biochemický sjezd, Špindlerův Mlýn 1978, Abstr. II-3.
3. Foltmann B., Pedersen V. B.: Advan. Exp. Med. Biol. *95*, 3 (1977).
4. Kostka V., Keilová H., Grüner K., Zbrožek J.: This Journal *42*, 3691 (1977).
5. Mikeš O.: This Journal *22*, 831 (1957).
6. Prusík Z., Keil B.: This Journal *25*, 2049 (1960).
7. Waley S. G., Watson J.: Biochem. J. *55*, 328 (1963).
8. Toennies G., Kolb J.: J. Anal. Chem. *23*, 823 (1951).
9. Hirs C. H. W.: J. Biol. Chem. *219*, 611 (1959).
10. Anderson W. L., Wetlaufer D. B.: Anal. Biochem. *6*, 493 (1975).
11. Spackman D. H., Stein W. H., Moore S.: Anal. Chem. *30*, 1190 (1958).
12. Benson jr J. V., Patterson J. A.: Anal. Chem. *37*, 1108 (1965).
13. Gray W. R., Hartley B. S.: Biochem. J. *89*, 379 (1963).
14. Edman P.: Acta Chem. Scand. *4*, 283 (1950).
15. Kostka V., Morávek L., Šorm F.: Eur. J. Biochem. *13*, 447 (1970).
16. Egorov T. A., Svenson A., Rydén L., Carlsson J.: Biochemistry *72*, 3029 (1975).
17. Nakagawa Y., Perlmann G. E.: Arch. Biochem. Biophys. *144*, 59 (1971).

Translated by V. Kostka.