Acta Chemica Scandinavica B 30 (1976) 979—984

N-Terminal Amino Acid Sequences of Pepsinogens from Dogfish and Seal and of Bovine Pepsinogen B

PER KLEMM, FRITZ POULSEN, MARIANNE K. HARBOE and BENT FOLTMANN

Institute of Biochemical Genetics, University of Copenhagen, Ø. Farimagsgade 2 A, DK-1353 København K, Denmark

The N-terminal amino acid sequences of dogfish pepsinogen (8 residues), seal pepsinogen (26 residues) and bovine pepsinogen B (25 residues) have been determined. The primary structures of all three pepsinogens show considerable homology with those of zymogens of gastric proteases sequenced previously.

At acidic pH the zymogens of the gastric proteases are converted into active enzymes by limited proteolysis during which N-terminal segments of about 45 amino acid residues are removed. It has been suggested that at neutral pH these zymogens are stabilized in inactive conformation through electrostatic interaction between positive charges of basic amino acids in the N-terminal end of the peptide chain and negative charges of aspartyl and glutamyl residues in the enzyme moiety of the zymogens.[1,2]

Previous determinations of the primary structures of porcine pepsinogen, bovine pepsinogen A and bovine prochymosin have shown a high degree of homology in the N-terminal amino acid sequences of these zymogens.[2–4] In order to investigate the general validity of the observations, especially with respect to distribution of basic amino acid residues, we have isolated and analysed pepsinogens from two more distantly related species (dogfish and ring seal) and one more of the bovine pepsinogens.

EXPERIMENTAL

Enzyme assay. To monitor the potential proteolytic activity during the purification of pepsinogen, modifications of the radial diffusion method[5] were used: For assay of seal pepsinogen gels were made of 1.1 % agarose and 0.3 % acid precipitated casein in 0.1 M sodium acetate buffer (pH 4.0). The gels were spread on glass plates to a thickness of 1.5 mm. Wells of 2 mm in diameter were punched in the gels. Before the assay, samples of pepsinogen solutions were converted into pepsin by standing at room temperature for 10 min at pH 2. Five μl of pepsin solutions were placed in the wells, and digestion took place for 20 h at 37 °C. After staining with Coomassie Brilliant Blue R-250 clear zones were observed around the pepsin containing wells. Calibration experiments carried out with porcine pepsin showed that the diameters of the clear zones were proportional to the logarithm of the pepsin concentrations, in these experiments 125 ng of porcine pepsin produced a clear zone of 5.0 mm in diameter.

A similar method was used for estimation of the potential proteolytic activity of dogfish pepsinogen, but in this case the most well defined zones were obtained with a gel layer containing 0.05 % acid denatured hemoglobin in formate buffer (0.1 M, pH 2.8), under such conditions both activation and digestion took place for 20 h at 37 °C. In this system 50 ng of porcine pepsin produced a clear zone of 6.6 mm in diameter.

The potential proteolytic activity is expressed as equivalents of μg of porcine pepsin/ml (ppa).

During the preparation of bovine pepsinogen B the potential enzymatic activity was determined by the milk clotting test.[1]

Purification of pepsinogens. Pepsinogen from ring seal (*Phoca hispida*). The seals were shot at Upernavik (Greenland) and the stomachs shipped in frozen state to the laboratory. Mucosa from 5 stomachs (500 g) was extracted with 2 l of 0.05 M Tris-HCl (pH 7.0) containing 0.1 M NaCl. The crude extract was clarified by formation of aluminium phosphate gel *in situ*.[1] All following operations were carried out at 5 °C. The precipitate was removed by centrifugation, and the supernatant dialysed against

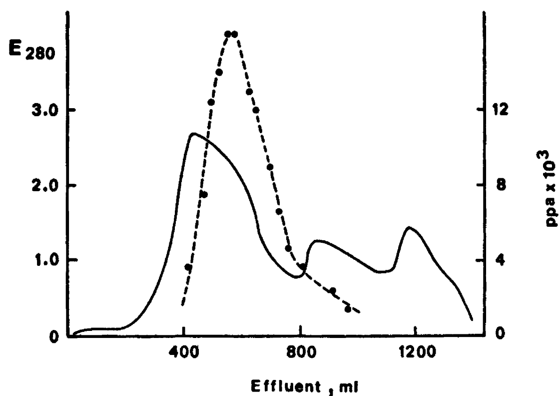

Fig. 1. Ion exchange chromatography of seal pepsinogen. Column: DEAE cellulose (2.5 × 43 cm). Elution: Linear gradient of 500 ml of 0.05 M Tris-HCl (pH 7.0) and 500 ml of 0.5 M NaCl in the same buffer. Flow rate 39 ml/h. Full line: E_{280}. Dashed line: Potential proteolytic activity (ppa).

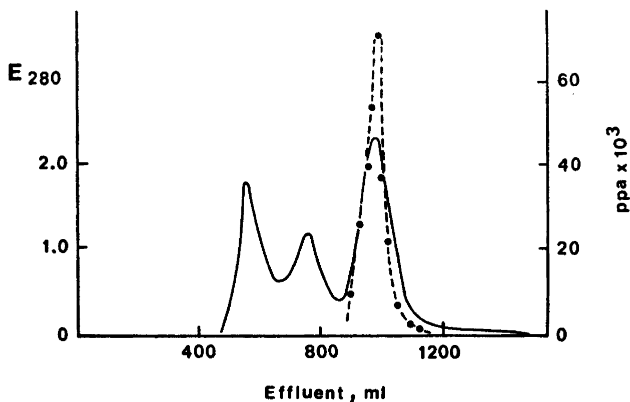

Fig. 2. Gel filtration of seal pepsinogen. Column: Sephadex G-100 (5 × 90 cm). Eluent: 0.05 M Tris-HCl (pH 7.0). Flow rate 70 ml/h. Full line: E_{280}. Dashed line: Potential proteolytic activity.

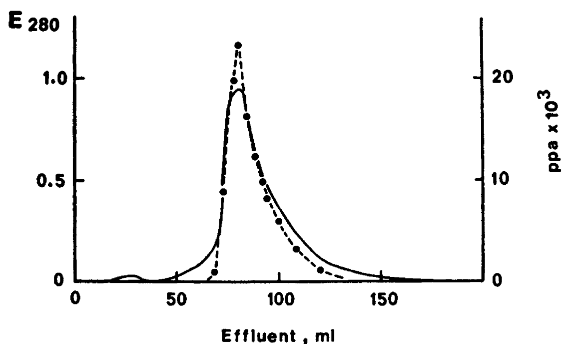

Fig. 3. Ion exchange chromatography of seal pepsinogen. Column: DEAE cellulose (0.9 × 15 cm). Elution: Linear gradient of 75 ml of 0.05 M piperazine-HCl (pH 6.0) and 75 ml 0.2 M NaCl in the same buffer. Flow rate 4 ml/h. Full line: E_{280}. Dashed line: Potential proteolytic activity.

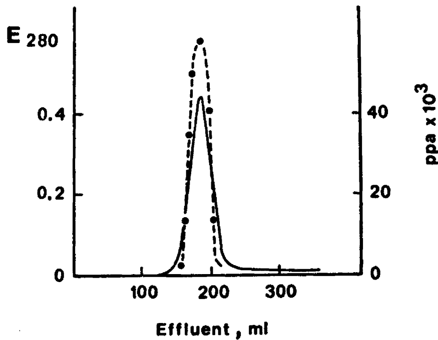

Fig. 4. Gel filtration of seal pepsinogen. Column: Sephadex G-75 (2.5 × 100 cm). Eluent: 0.02 M piperazine-HCl (pH 6.0). Flow rate 18 ml/h. Full line: E_{280}. Dashed line: Potential proteolytic activity.

0.05 M Tris-HCl (pH 7.0) and applied to a column of DEAE cellulose. Inert material was washed from the column with 500 ml of 0.05 M Tris-HCl (pH 7.0) before elution with a linear gradient from the starting buffer to 0.5 M NaCl in the starting buffer. The elution was completed with 350 ml of 0.5 M NaCl in the starting buffer. The results are illustrated in Fig. 1. The effluent from 364 ml to 780 ml was concentrated to 35 ml by ultrafiltration.

After gel filtration on Sephadex G-100 as shown in Fig. 2, effluent from 896 to 1064 ml was concentrated to 10 ml by ultrafiltration and applied to a column of DEAE cellulose (0.9 × 15 cm) equilibrated with 0.05 M piperazin-HCl (pH 6.0). Elution took place with a linear gradient from the equilibration buffer to 0.2 M NaCl in the same buffer (Fig. 3). Effluent from 76 to 110 ml was concentrated to 10 ml by ultrafiltration and gel filtration was carried out on a column of Sephadex G-75 as illustrated in Fig. 4. Two aliquots of the peak fraction, each containing *ca.* 15 nmol of pepsinogen, were freeze-dried and sequenced in parallel.

Pepsinogen from dogfish (*Squalus acanthias*). The dogfish was caught in the North Sea, stomachs were frozen immediately after landing and shipped frozen to the laboratory. Mucosa from 10 stomachs (180 g) was extracted with 600 ml of 0.05 M Tris-HCl (pH 7.5). Clarification with aluminium phosphate followed by centrifugation and dialysis was performed as described above. The following purification was carried out by a procedure similar to that published.[6] The zymogen was adsorbed to a column of DEAE cellulose (2.5 × 29 cm). Inert material was washed from the column with 150 ml 0.03 M of Tris-HCl (pH 7.5), and elution took place with a linear gradient (500 ml of 0.03 M Tris-HCl (pH 7.5) + 500 ml of 0.5 M NaCl in the same buffer). The major pepsinogen component was subjected to gel filtration on Sephadex G-100 (0.3 M Tris-HCl, pH 7.5) and finally purified by chromatography on a column of DEAE cellulose (0.9 × 13 cm), eluted with a linear gradient of 75 ml of 0.01 M sodium phosphate (pH 5.0) and 75 ml of 0.02 M NaCl in the same buffer. Of the peak fraction 0.9 ml (containing 16 nmol of pepsinogen) was freeze-dried and used for sequencing.

Bovine pepsinogen B is one of the minor components in bovine gastric mucosa. The starting material was extract of cow's stomachs obtained from Chr. Hansen's Laboratory Ltd., Copenhagen. The zymogen was prepared as described by Antonini and Ribadeau Dumas[7] except for the use of an additional step of chromatography on polylysine Sepharose 4B.[3] Bovine pepsinogen B was converted into pepsin B by activation at pH 2 for 30 min in an ice-bath. The activation peptides were purified by chromatography on DEAE cellulose and paper electrophoresis as previously described.[3] Activation peptides are denoted Act and numbered from the *N*-terminus of pepsinogen.

Sequencing. For Edman degradation of proteins the method of Weiner *et al.*[8] was modified to allow identification of the liberated thiazolinone derivatives by conversion into free amino acids.[9] The reaction was carried out in test tubes of 12 × 70 mm by the following procedure:

Coupling: dissolve 15 – 25 nmol of protein in 150 μl of 0.5 M $NaHCO_3$ (pH 9.8), add 20 μl of 10 % sodium dodecyl sulfate. Flush with nitrogen for 1 min, add 10 μl of phenylisothiocyanate, flush again with nitrogen for 1 min, and seal with parafilm. Place the test tube in a water bath at 50 °C and shake thoroughly at intervals of 5 min. After 60 min place the test tube in a freezer (– 20 °C) for 4 to 5 min and add 1 ml of acetone (stored in the freezer). Shake thoroughly and collect precipitate by centrifugation, withdraw and discard the yellowish supernatant. Wash precipitate once more with 1 ml of cold acetone, after centrifugation and withdrawal of supernatant, evaporate the last amount of acetone by flushing with nitrogen while vortexing the tube to distribute the precipitate over the lower part of the tube. Place the tube horizontally in a desiccator, and dry under vacuum for 20 min at 60 °C.

Cyclization and cleavage: cool to room temperature and add 200 μl of trifluoroacetic acid. Flush with nitrogen for 30 s and seal with double layer of parafilm. Leave the tube for 5 min at 50 °C. Evaporate trifluoroacetic acid by flushing with nitrogen in the hood while vortexing the tube to distribute the precipitate (NB, use gloves). Dry the precipitate under vacuum in a desiccator with solid NaOH for 10 min at 60 °C.

Extraction and identification: cool the tube to 0 °C in an ice-bath. Add 100 μl of ice-cold water and extract three times with 500 μl

of ice-cold butyl acetate saturated with water. Dry the water-phase and the protein is now ready for the next cycle of Edman degradation. To remove any contaminating peptides or protein from the combined extracts of butyl acetate, extract this with 100 µl of water. Evaporate the butyl acetate in a desiccator at 60 °C. Add 200 µl of hydriodic acid and seal the tube under vacuum. The thiazolinone is now converted into free amino acid by reaction for 24 h at 127 °C.

Edman degradation of peptides was carried out by coupling and cyclization as described in Ref. 10. After evaporation of trifluoroacetic acid, the dried residue was dissolved in 50 µl of ice-cold water; extraction was performed with 200 µl of butyl acetate and the further treatment was as described above.

The amino acids were analysed quantitatively on a Durrum D-500 amino acid analyzer. To recover methionine as metsulfone, Act-3 was oxidized with performic acid, otherwise the known artifacts of the conversion [9] were taken into account.

The *N*-termini were also analysed by dansylation.[10]

Cyanogen bromide cleavage of bovine pepsinogen B was carried out in 70 % formic acid. Fragments were separated by gel filtration (Sephadex G-50, eluent: 0.05 M acetic acid). The purified fragments are denoted CB and numbered from the *N*-terminus of pepsinogen.

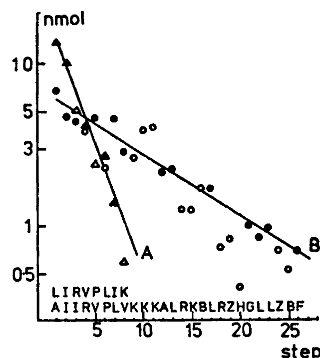

Fig. 5. Recoveries in nmol of amino acids obtained after conversion with hydriodic acid, semilogarithmic plots against steps of Edman degradation. A: Dogfish pepsinogen. B: Ring seal pepsinogen. Solid marks illustrate amino acids of high recovery [9] (Ala, Gly, Ile, Leu, Phe or Val), straight lines are drawn on basis of these. Letters above steps are amino acids in single letter code.[11]

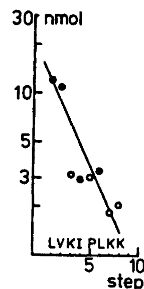

Fig. 6. Recoveries of amino acids after Edman degradation of bovine pepsinogen B. Notation as in Fig. 5.

RESULTS AND DISCUSSION

The *N*-terminal amino acid sequences of dogfish and ring seal pepsinogen were obtained by direct Edman degradation of the proteins, while bovine pepsinogen B was analysed by Edman degradation of the protein up to residue No. 9.

Table 1. Analyses of activation peptides from bovine pepsinogen B. Amino acid compositions are expressed in stoichiometric ratios. Recoveries after Edman degradation are given directly in nmol.

Step of Edman degradation	Act-1 Amino acid composition		Act-1 Recovery nmol	Act-2 Amino acid composition		Act-2 Recovery nmol	Act-3 Amino acid composition		Act-3 Recovery nmol
1	Leu	2.0	11.4	Phe	0.9	9.2	Ile	1.0	14.8
2	Val	0.9	9.4	Lys	1.0	9.0	Met	0.7	14.8
3	Lys	3.1	7.6	Ser	0.9	2.1	Lys	1.9	13.6
4	Ile	1.0	6.3	Ile	0.9	5.2	Glu	1.1	10.6
5	Pro	0.8	5.8	Arg	1.0	2.9	Lys		11.3
6	Leu		5.0	Glu	1.0	7.2	Gly	1.0	14.2
7	Lys		3.9				Leu	1.1	13.8
8	Lys		2.5						

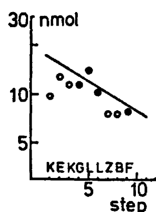

Fig. 7. Recoveries of amino acids after Edman degradation of CB-2 from bovine pepsinogen B. Notation as in Fig. 5. This fragment continues in another *ca.* 25 residues.

The recoveries from the sequencing are illustrated as semilogarithmic plots:[9] Figs. 5 and 6 show the results from entire proteins, and Fig. 7 shows the N-terminal sequence of CB-2 from bovine pepsinogen B, the sequences being given in single letter code.[11] The analyses of activation peptides from bovine pepsinogen B are shown in Table 1. CB-1 was identified only from amino acid composition and by N-terminal Dns-Leu. The final sequences are summarized in Table 2.

Quantitative results from sequencing experiments are presented without correction for "background", due to "out of step" reactions and other side reactions. It is a general experience that such reactions will lead to decreasing yield during the sequencing, it is further well known that such difficulties may vary much from protein to protein, and in addition to this a large and uncontrolled loss often occurs in the first step of Edman degradation.[9,12] Our results reflect the same difficulties. The repetitive yields were calculated to be 70 and 92 % by sequencing dogfish and seal pepsinogen; this allowed an unambiguous interpretation of 8 and 26 steps, respectively. The repetitive yield of bovine pepsinogen B was 74 %, in this case samples after step 9 were lost due to failure of the amino acid analyzer.

By conversion of thiazolinone derivatives into free amino acids serine represents a special problem, since it is converted into low yields of alanine. After conversion, high yield of alanine was found in the first step of Edman degradation of seal pepsinogen. The N-terminus of another sample was therefore investigated by the dansylation method; alanine was found as predominant N-terminal amino acid, but minor amounts of serine were also observed.

Table 2. N-Terminal amino acid sequences of zymogens of gastric proteases.

	5 10 15 20 25
Dogfish pepsinogen	Leu-Ile-Arg-Val-Pro-Leu-Ile-Lys/
Ring seal pepsinogen	Ala-Ile-Ile-Arg-Val-Pro-Leu-Val-Lys-Lys-Lys-Ala-Leu-Arg-Lys-Asx-Leu-Arg-Glx-His-Gly-Leu-Leu-Glx-Asx-Phe/
Bovine pepsinogen B	Leu-Val-Lys-Ile-Pro-Leu-Lys-Lys-Phe-Lys-Ser-Ile-Arg-Glu-Ile-Met-Lys-Glu-Lys-Gly-Leu-Leu-Glx-Asx-Phe/
	Act-1 (Leu¹–Lys⁸); Act-2 (Phe⁹–Glu¹⁴); Act-3 (Ile¹⁵–Leu²¹)
	CB-1 (Leu¹–Met¹⁶); CB-2 (Lys¹⁷–Phe²⁵)
Porcine pepsinogen A[2]	Leu-Val-Lys-Val-Pro-Leu-Val-Arg-Lys-Lys-Ser-Leu-Arg-Gln-Asn-Leu-Ile-Lys-Asp-Gly-Lys-Leu-Lys-Asp-Phe/
Bovine pepsinogen A[3]	Ser-Val-Val-Lys-Ile-Pro-Leu-Val-Lys-Lys-Lys-Ser-Leu-Arg-Gln-Asn-Leu-Ile-Glu-Asn-Gly-Lys-Leu-Lys-Glu-Phe/
Bovine prochymosin[4]	Ala-Glu-Ile-Thr-Arg-Ile-Pro-Leu-Tyr-Lys-Gly-Lys-Ser-Leu-Arg-Lys-Ala-Leu-Lys-Glu-His-Gly-Leu-Leu-Glu-Asp-Phe/
Common	Pro Leu Lys Arg Gly Leu Phe

In seal pepsinogen residue No. 13 (step No. 12) was only identified after conversion to free amino acid (good yield of alanine, *cf.* Fig. 5B). The corresponding residue in bovine pepsinogen B was unambiguously identified since only serine was found by amino acid analysis of Act-2. The electrophoretic mobilities of Act-2 and Act-3 show the presence of glutamyl residues.[13] In seal pepsinogen information about amides at positions 17, 20, 25, and 26 is lost in the present way of sequencing. For position 25 and 26 this is also the case in bovine pepsinogen B, but these minor ambiguities do not interfere with the pattern of common structure that emerges from the sequences.

To facilitate comparison among the *N*-terminal sequences of the gastric zymogens sequenced up till now, Table 2 also includes the previously published sequences.[2–4] In five of the zymogens we find basic amino acids at positions 5, 10, 12, and 15. Dogfish pepsinogen has only been sequenced up to residue No. 10, but again with basic amino acids at positions Nos. 5 and 10. In this part of the sequence only Pro No. 7 and Leu No. 8 are identical in all the zymogens; however, the majority of the remaining residues show substitutions of very conservative character like Ile/Val. This means that we may predict that these sequences will fit very similar tertiary structures in all the zymogens.

In this context it is noteworthy that among the peptides liberated during activation of bovine pepsinogen A, a peptide consisting of 17 residues (No. 2 to No. 18 in Table 2) was identified as an inhibitor of the milk clotting action of pepsin.[3]

The information is still scattered, but it comprises sequences of three zymogens from one species (cow) and three more from other vertebrate species of which two (ring seal and dogfish) are distantly related to cow. Thus we may tentatively conclude that in the *N*-terminal sequence of the gastric zymogens a pattern of basic amino acid residues with spacings mainly of apolar amino acid residues are necessary for stabilizing the zymogen molecules in inactive conformation at neutral pH. This pattern appears to be common from the cartilaginous fishes (dogfish) to mammalia whether these are carnivorous (seal) or herbivorous (cow).

Acknowledgement. The work was supported by the Danish Natural Science Research Council and by the Carlsberg Foundation.

REFERENCES

1. Foltmann, B. *C. R. Trav. Lab. Carlsberg 35* (1966) 143.
2. Ong, E. B. and Perlmann, G. E. *J. Biol. Chem. 243* (1968) 6104.
3. Harboe, M., Andersen, P. M., Foltmann, B., Kay, J. and Kassell, B. *J. Biol. Chem. 249* (1974) 4487.
4. Pedersen, V. B. and Foltmann, B. *Eur. J. Biochem. 55* (1975) 95.
5. Løwenstein, H. and Ingild, A. *Anal. Biochem. 71* (1976) 204.
6. Merrett, T. G., Bar-Eli, E. and Van Vunakis, H. *Biochemistry 8* (1969) 3696.
7. Antonini, J. and Ribadeau Dumas, B. *Biochimie 53* (1971) 321.
8. Weiner, A. M., Platt, T. and Weber, K. *J. Biol. Chem. 247* (1972) 3242.
9. Smithies, O., Gibson, D., Fanning, E. M., Goodfliesh, R. M., Gilman, J. G. and Ballantyne, D. L. *Biochemistry 10* (1971) 4912.
10. Gray, W. R. *Methods Enzymol. 25* (1972) Chapter 26a.
11. IUPAC-IUB Commission on Biochemical Nomenclature *Biochem. J. 113* (1969) 1.
12. Niall, H. D. *Methods Enzymol. 27* (1973) Chapter 36.
13. Offord, R. E. *Nature (London) 211* (1966) 591.

Received June 1, 1976.