Complete amino acid sequence of bovine colostrum low-M_r cysteine proteinase inhibitor

Masayuki Hirado, Susumu Tsunasawa, Fumio Sakiyama, Michio Niinobe and Setsuro Fujii

Institute for Protein Research, Osaka University, Suita, Osaka 565, Japan

Received 7 May 1985

The complete amino acid sequence of bovine colostrum cysteine proteinase inhibitor was determined by sequencing native inhibitor and peptides obtained by cyanogen bromide degradation, *Achromobacter* lysylendopeptidase digestion and partial acid hydrolysis of reduced and *S*-carboxymethylated protein. *Achromobacter* peptidase digestion was successfully used to isolate two disulfide-containing peptides. The inhibitor consists of 112 amino acids with an M_r of 12787. Two disulfide bonds were established between Cys 66 and Cys 77 and between Cys 90 and Cys 110. A high degree of homology in the sequence was found between the colostrum inhibitor and human γ-trace, human salivary acidic protein and chicken egg-white cystatin.

Amino acid sequence *Cysteine proteinase inhibitor* *Bovine colostrum*

1. INTRODUCTION

Low-M_r cysteine proteinase inhibitors have been found in various animal tissues [1–3], sera [4–6] and egg whites [7]. Recently, the primary structure of several cysteine proteinase inhibitors has been elucidated [8–15]. These small protein proteinase inhibitors are members of a family which has diverged from a common ancestor [14]. The inhibitors identified so far are classified into two distinct groups, one being cellular associated [8–10] and the other found in extracellular fluids [11–15]. Recently we purified an extracellular basic cysteine proteinase inhibitor from bovine colostrum [16]. Here, we report the complete amino acid sequence of the inhibitor and compare it with other low-M_r cysteine proteinase inhibitors thus far determined.

2. MATERIALS AND METHODS

2.1. *Materials*

Bovine colostrum was kindly supplied by Sanda City Agricultural Cooperative (Sanda, Hyogo, Japan). The silica-bonded reverse-phase columns for high-performance liquid chromatography (HPLC) were purchased from the sources indicated: YMC-Pack A 303 column (0.46 × 25 cm, C18, 300 Å), Yamamura Chemical; LS-410 K column (0.4 × 30 cm, C18, 100 Å), Toyo-Soda; Baker bond wide pore butyl (0.46 × 25 cm, C4, 330 Å), J.T. Baker Chemical; ProRPC HR5/10 column (0.5 × 10 cm, C8, 300 Å), Pharmacia. *Achromobacter* lysylendopeptidase (*A. lyticus* protease I [17]) was a kind gift from Dr T. Masaki of Ibaraki University, Japan. All other chemicals were of the purest grade commercially available.

2.2. *Isolation of inhibitors*

The isolation procedure for low-M_r cysteine proteinase inhibitors from bovine colostrum was essentially the same as reported [16] except that HPLC on a reverse-phase column was used instead of gel filtration on a Sephadex G-50 column as the final step. The pooled fractions containing inhibitory activity from CM-Sephadex chromatography were concentrated to a small volume by ultrafiltration on a YM-5 membrane (Amicon, Lexington, MA) before HPLC on a YMC-Pack A 303 column. The inhibitors were eluted with a linear gradient of 30–50% 2-propanol/acetonitrile

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/85/$3.30 © 1985 Federation of European Biochemical Societies

(7:3, v/v) in 0.1% trifluoroacetic acid for 1 h at a flow rate of 1 ml per min.

2.3. *Peptide fragmentation and separation on reverse-phase HPLC*

Reduced and *S*-carboxymethylated (Rcm) inhibitor [18] was digested with *Achromobacter* lysylendopeptidase (AP) at 37°C for 6 h in 50 mM Tris-HCl, pH 9.5. The molar ratio of protease to substrate was 1:200. Digestion with the intact inhibitor was carried out at pH 7.0. Methionyl bonds were cleaved with a 100-fold molar excess of CNBr over 3 methionine residues in 70% formic acid at room temperature for 18 h. Partial hydrolysis was performed by heating Rcm inhibitor with 0.03 M HCl (pH 2.0) in evacuated sealed tubes at 108°C for 2 h [19].

Peptides obtained by the above procedures were fractionated by reverse-phase HPLC using appropriate reverse-phase columns with 2-propanol/acetonitrile (7:3, v/v) in 0.1% trifluoroacetic acid as a solvent. Columns used for peptide purification were an LS-410K column for AP digests, a ProRPC HR5/10 column for CNBr fragments and a Baker-bond wide pore butyl column for partial acid hydrolysate.

2.4. *Amino acid and sequence analyses*

Amino acid analyses were carried out with a Hitachi 835S amino acid analyzer after hydrolysis of samples at 110°C in vacuo for 24 h in 5.7 M HCl or in 4 M methanesulfonic acid [20]. Automated Edman degradation was performed with a 470A protein sequencer (Applied Biosystems) using a standard program for sequencing [21]. PTH-amino acids were determined by reverse-phase HPLC with isocratic elution as in [22].

3. RESULTS AND DISCUSSION

3.1. *Isolation of cysteine proteinase inhibitors*

The result of purification using reverse-phase HPLC is presented in table 1. Two species of cysteine proteinase inhibitor were finally purified to homogeneity (fig.1). Each protein in peaks I and II displayed the same inhibitory activity to cathepsin H and gave a single band corresponding to M_r 11000 on SDS-polyacrylamide gel electrophoresis. The protein in peak I is always a major component and that in peak II minor. Only in an experiment shown here was a considerable amount of the inhibitor isolated from peak II. Amino acid compositions of both inhibitors were the same except that one additional residue each of glycine and valine was found in the protein of peak II (table 2). Edman degradation revealed that the inhibitor of the peak II had a glycine residue as amino-terminus which was followed by the same 23 amino acids found in the N-terminal portion of the major component of inhibitor. However, the amino acid sequence of the residual portion remains to be established.

3.2. *Determination of amino acid sequence of the inhibitor*

The complete amino acid sequence of the major component of bovine colostrum cysteine proteinase inhibitor is presented in fig.2. The strategy of sequencing analysis is as follows. Degradation

Table 1

Purification of bovine colostrum low-M_r cysteine proteinase inhibitor

Step	Total protein (mg)	Total inhibitory activity (units)	Specific activity (units/mg)	Yield (%)	Purification (-fold)
CM-Sephadex	0.81	1874.3	2314.0	9.2	2386
Reverse phase HPLC					
peak I	0.15	798.6	5324.0	3.9	5489
peak II	0.11	571.8	5343.9	2.8	5509

Data for purification before CM-Sephadex chromatography are those presented in table 1 of [16]

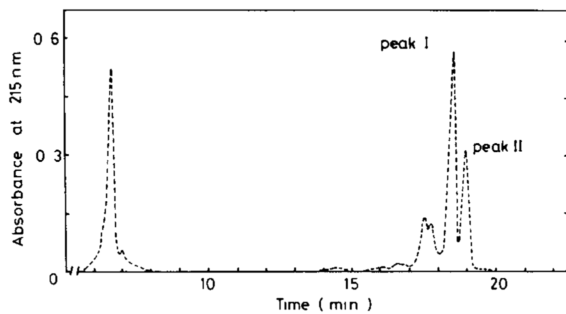

Fig.1. Reverse-phase HPLC of crude inhibitor fractionated on a CM-Sephadex column. Samples ($120\,\mu g/100\,\mu l$) were applied to a YMC-Pack A 303 ODS column. For experimental details, see text.

Table 2

Amino acid compositions of bovine colostrum low-M_r cysteine proteinase inhibitors

Amino acid	Protein in peak I (residues/mol[a])	Protein in peak II
Asp	12.8 (13)[b]	13.0
Thr	3.6 (4)	3.7
Ser	8.0 (9)	8.0
Glu	13.5 (14)	13.8
Pro	3.3 (3)	3.0
Gly	5.4 (5)	6.1
Ala	6.0 (6)	6.0
Half Cys	3.6 (4)	3.8
Val	10.7 (14)[c]	11.7
Met	2.3 (3)	2.3
Ile	1.0 (1)	1.1
Leu	10.1 (10)	10.3
Tyr	2.9 (3)	3.0
Phe	6.3 (6)	6.2
Lys	6.1 (6)	6.3
His	2.7 (2)	2.3
Trp	0.9 (1)	1.0
Arg	8.0 (8)	8.1
Total	(112)	

[a] Values after 24 h hydrolysis with 4 M methanesulfonic acid

[b] Numbers in parentheses are theoretical values based on amino acid sequence

[c] Discrepancy between experimental and calculated values may be due to the incomplete hydrolysis of 4 Val–Val bonds in the sequence

of the inhibitor with CNBr gave 3 peptide fragments (CB-I–III), which were separated by reverse-phase HPLC and sequenced on a gas-phase sequencer. Sequence information of the first 51 residues obtained by direct sequencing of the native inhibitor established the arrangement of these 3 fragments as (CB-I)–(CB-II)–(CB-III).

When Rcm inhibitor was digested with AP, 7 peptides (CM-AP-I–VII) were produced and purified completely by reverse-phase HPLC. These peptides were aligned based on their complete or partial amino acid sequences, amino acid compositions and structural information obtained from direct sequencing of the native inhibitor and CB-I–III. To overlap peptide CM-AP-V, Rcm inhibitor was partially hydrolyzed with dilute acid and peptide CM-DA was isolated. The determination of its N-terminal 18 residues located the CM-AP-V in the N-terminal half of the peptide CM-DA.

Two cystine-containing peptides, AP-III and AP-V, were detected instead of CM-AP-III and CM-AP-IV, and CM-AP-VI and CM-AP-VII,

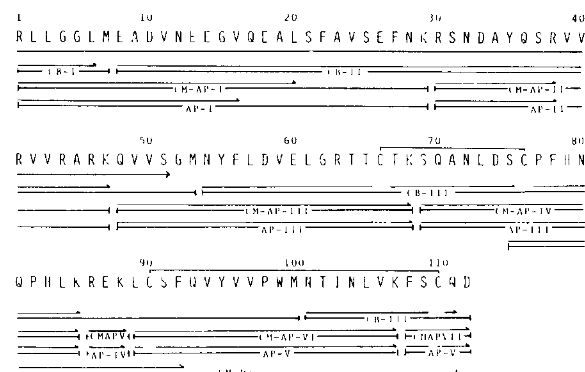

Fig.2. Amino acid sequence of bovine colostrum low-M_r cysteine proteinase inhibitor (peak I). CB, CNBr fragments; AP and CM-AP, *Achromobacter* lysylendopeptidase-digested fragments of the native protein (AP) and Rcm protein (CM-AP); CM-DA, dilute acid hydrolysate of Rcm protein. Amino acids identified by automated Edman degradation are shown by arrows. (⟶) Direct sequencing of the native protein, (—⇀) sequencing of peptide fragment. Dashed lines denote identification of PTH-cysteic acid after performic acid-oxidation peptide. Amino acids not identified as PTH-amino acids are shown by thin dotted lines. Amino acid compositions were determined for all the peptide fragments isolated.

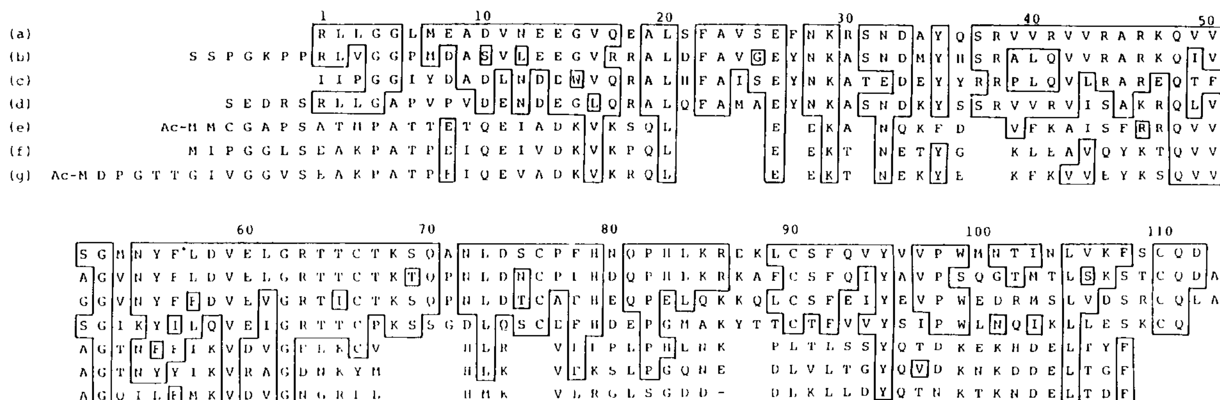

Fig.3. An alignment of amino acid sequences of 7 low-M_r cysteine proteinase inhibitors. (a) Bovine colostrum inhibitor (peak I); (b) human γ-trace [13]; (c) human SAP-1 [15]; (d) egg-white cystatin [11,12]; (e) rat liver TPI [8]; (f) human leucocytes Stefin [9] and (g) rat epidermal TPI [10]. Boxed residues are those that are identical between other inhibitors and the bovine colostrum inhibitor. The alignment of Glu 26 in 3 intracellular inhibitors is tentative. Ac-, acetyl-.

respectively, on reverse-phase HPLC of the AP digest of native inhibitor. Direct sequencing revealed that AP-III corresponds to CM-AP-III and CM-AP-IV, and AP-V to CM-AP-VI and CM-AP-VII. After performic acid oxidation, AP-III and AP-V each gave 2 mol cysteic acid per mol peptide by amino acid analysis. PTH-cysteic acid was also identified by Edman degradation at the positions corresponding to *S*-carboxymethyl cysteine previously determined for Rcm derivatives of AP-III and AP-V. We conclude therefore that two disulfide bonds exist between Cys 66 and Cys 76 and Cys 90 and Cys 110.

3.3. *Comparison of amino acid sequences of inhibitors from various sources*

Seven sequences of low-M_r cysteine proteinase inhibitors are compared by aligning each amino acid with the highest homology (fig.3). Among the extracellular inhibitors, the colostrum inhibitor is highly homologous to human γ-trace, a basic protein, with 59% homology. A significant degree of sequence homology was also found between extracellular and intracellular inhibitors. The colostrum protein has about 20% homology to rat liver protein.

Nine amino acid residues, Leu 20, Glu 26, Lys 29, Gln 48, Gly 52, Val 59, Gly 62, Tyr 95 and Leu 105, are conserved in the seven low-M_r inhibitors hitherto sequenced. These residues and particularly those which are in a high homology region comprising residues 42–62 are likely to be involved in the inhibitory function of these proteins. It should be noted that two of the proposed reactive-site residues of human plasma α_2-thiol protease inhibitor [22] are found in the amino acids mentioned above.

Bovine colostrum inhibitor (p*I* 10.0) and human γ-trace (p*I* 9.0) are basic proteins in contrast to 5 other inhibitors. However, it is difficult to assume that the basicity itself plays a particular functional role in the inhibition of cysteine proteinase because of a relatively high sequence homology (45–56%) between acidic (p*I* 4.6–6.5) and basic extracellular inhibitors.

ACKNOWLEDGEMENTS

We are grateful to Miss Yoshiko Yagi for amino acid analyses and Dr Hiromu Sugino, a former member of Takeda Pharmaceutical Industries, for giving us an opportunity to use a gas-phase sequencer at the beginning of this investigation.

REFERENCES

[1] Lenney, J.F., Tolan, J.R., Sugai, W.J. and Lee, A.G. (1979) Eur. J. Biochem. 101, 153–161.
[2] Hirado, M., Iwata, D., Niinobe, M. and Fujii, S. (1981) Biochim. Biophys. Acta 669, 21–27.
[3] Takeda, A., Kobayashi, S. and Samejima, T. (1983) J. Biochem. 94, 811–820.

[4] Iwata, D., Hirado, M., Niinobe, M. and Fujii, S. (1982) Biochem. Biophys. Res. Commun. 104, 1525–1535.

[5] Lenney, J.F., Liao, J.R., Sugg, S.L., Gopala-Krishnan, V., Wong, H.C.H., Ouye, K.H. and Chan, P.W.H. (1982) Biochem. Biophys. Res. Commun. 108, 1581–1587.

[6] Hirado, M., Niinobe, M. and Fujii, S. (1983) Biochim. Biophys. Acta 757, 196–201.

[7] Barrett, A.J. (1981) Methods Enzymol. 80, 771–778.

[8] Takio, K., Kominami, E., Wakamatsu, N., Katunuma, N. and Titani, K. (1983) Biochem. Biophys. Res. Commun. 115, 902–908.

[9] Machleidt, W., Borchart, U., Fritz, H., Brzin, J., Ritonja, A. and Turk, V. (1983) Hoppe-Seyler's Z. Physiol. Chem. 364, 1481–1486.

[10] Takio, K., Kominami, E., Bando, Y., Katunuma, N. and Titani, K. (1984) Biochem. Biophys. Res. Commun. 121, 149–154.

[11] Turk, V., Brzin, J., Longer, M., Ritonja, A. and Eropkin, M. (1983) Hoppe-Seyler's Z. Physiol. Chem. 364, 1487–1496.

[12] Schwabe, C., Anastasi, A., Crow, H., McDonald, J.K. and Barrett, A.J. (1984) Biochem. J. 217, 813–817.

[13] Grubb, A. and Lofberg (1982) Proc. Natl. Acad. Sci. USA 79, 3024–3027.

[14] Barrett, A.J., Davies, M.E. and Grubb, A. (1984) Biochem. Biophys. Res. Commun. 120, 631–636.

[15] Isemura, S., Saitoh, E. and Sanada, K. (1984) J. Biochem. 96, 489–498.

[16] Hirado, M., Niinobe, M. and Fujii, S. (1984) J. Biochem. 96, 51–58.

[17] Masaki, T., Fujihashi, T., Nakamura, K. and Soejima, M. (1981) Biochim. Biophys. Acta 660, 51–55.

[18] Crestfield, A.M., Moore, S. and Stein, W.H. (1963) J. Biol. Chem. 238, 622–627.

[19] Inglis, A., Mckern, N., Roxburgh, C. and Strike, P. (1980) Methods in Peptide and Protein Sequence Analysis (Birr, C. ed.) pp.329–343, Elsevier, Amsterdam, New York.

[20] Simpson, R.J., Neuberger, M.R. and Liu, T.Y. (1976) J. Biol. Chem. 251, 1936–1940.

[21] Hankapiller, M.W., Hewick, R.M., Dryer, W.J. and Hood, L.E. (1983) Methods Enzymol. 91, 399–413.

[22] Tsunasawa, S., Kondo, J. and Sakiyama, F. (1985) J. Biochem. 97, 701–704.

[23] Ohkubo, I., Kurachi, K., Takasawa, T., Shiokawa, H. and Sasaki, M. (1984) Biochemistry 23, 5691–5697.