POLYMORPHISM OF COW COLOSTRUM PROTEINASE INHIBITOR

Věra Jonáková, Dana Čechová and Otakar Mach

Institute of Molecular Genetics,
Czechoslovak Academy of Sciences, 166 10 Prague 6

Cow colostrum contains three isoinhibitors A, B, and C, which are glycoproteins. In this study isoinhibitor A was isolated and characterized and the structure of its protein moiety compared with the known protein structures of isoinhibitors B and C. It was found that the primary structure of isoinhibitor A is identical with the primary structure of isoinhibitor B except that the C-terminus of its molecule is shorter by five amino acid residues. Four discrete chromatographic forms (I—IV) with different isoelectric points (pI $3 \cdot 8$ — I, $4 \cdot 0$ — II, $4 \cdot 3$ — III, and $4 \cdot 5$ — IV) were isolated by chromatography on SE-Sephadex, Form I is identical with isoinhibitor A. Forms II, III, and IV are represented by mixtures of isoinhibitors A, B, and C with a heterogeneous carbohydrate moiety which affects the total net charge of the individual inhibitor forms.

The low molecular weight inhibitor isolated from cow colostrum exists in several chromatographically distinguishable forms[1,2]. Its molecule consists of one single polypeptide chain[3] cross-linked by three disulfide bridges[4]; the carbohydrate moiety[5] is covalently attached to the polypeptide chain. This moiety represents roughly one third of the molecule, is heterogeneous and different with the individual forms. It was uncertain until recently whether differences in the carbohydrate moiety or in the protein moiety are responsible for the existence of the multiple chromatographic forms of this inhibitor.

In this study the term chromatographic form (or merely form) is used for an inhibitor of uniform net charge irrespectively of the possible heterogeneity of the protein or carbohydrate moiety. The term isoinhibitor is reserved for an inhibitor of uniform structure of its protein moiety, irrespectively of the charge heterogeneity of its molecules.

Recently we reported on the isolation and primary structure of two isoinhibitors from cow colostrum, named isoinhibitor B and isoinhibitor C. These isoinhibitors have an identical amino acid composition except that the third position from the N-terminus of isoinhibitor B is occupied by threonine whereas the same position in isoinhibitor C is occupied by lysine[6]. In this study we described the isolation of still another isoinhibitor from cow colostrum named isoinhibitor A. It differs from isoinhibitor B by the C-terminus of its molecule which is by five amino acid residues shorter. It thus represents an isoinhibitor with one lysine residue in its molecule whose existence has been suggested both by our preliminary data[7] and

in the reports from other laboratories[8,9]. Because of its low molecular weight and relatively easy identification of amino acid replacements in the molecules of its individual forms, the cow colostrum proteinase inhibitor represents a suitable model for studies of the polymorphism of glycoproteins.

EXPERIMENTAL

Material. The inhibitor concentrate was prepared from cow colostrum by a method[1,10] involving the precipitation of contaminating proteins by 2·5% trichloroacetic acid with a minor modification: the mixture was not heated and the precipitation was allowed to proceed at room temperature. The inhibitor was precipitated from the filtrate by ammonium sulfate (80% saturation). SE-Sephadex C—25, Sephadex G—25, and Sepharose 4B were products of Pharmacia Fine Chemicals (Uppsala, Sweden). Bovine trypsin was crystallized twice and TPCK-trypsin was from Worthington Biochemical Corporation (Freehold, U.S.A.). N-α-benzoyl-DL-arginine- -p-nitroanilide hydrochloride (BAPA) was prepared according to Tuppy[11]. Cellogel acetate cellulose sheets and Servalyte AG 2—11 were from Serva (Heidelberg, Germany).

Column procedures. The purification of the inhibitor on a column of trypsin immobilized on Sepharose 4B was identical with the procedure described before[2]. The inhibitor thus obtained was desalted on a column of Sephadex G—25 equilibrated with 0·2% formic acid, lyophilized, and used for chromatography. The latter was performed on a 4 × 21 cm column of SE-Sephadex C—25 equilibrated with 0·01M sodium formate whose pH had been adjusted to 3·5 by formic acid. The sample (2 g) was dissolved in the equilibration buffer (150 ml) and placed onto the column. After the application of the sample and elution of unadsorbed proteins by the equilibration buffer (1 000 ml) the individual chromatographic forms of the inhibitor were eluted by a linear gradient of ionic strength (Fig. 1), desalted on a column of Sephadex G—25 in 0·2% formic acid, and subsequently lyophilized.

Measurement of inhibitory activities. The inhibition of trypsin was measured in terms of decrease of enzymatic activity after incubation of the enzyme with the inhibitor solution for 20 min in Tris-HCl buffer at pH 7·8 before the addition of the substrate (BAPA). The remaining tryptic activity was measured spectrophotometrically[12,13] as the change of absorbance at 405 nm in Opton (Zeiss, Germany) spectrophotometer. The measurements were made at 25°C in 0·1M Tris-HCl buffer at pH 7·8 containing 0·02M-CaCl$_2$. One enzymatic unit (U) represents[13] the hydrolysis of 1 μmol of substrate in 1 min. One inhibitory unit (IU) reduces the activity of two enzymatic units by 50%. The specific activity indicates the number of inhibitory units (BAPA) per 1 mg of protein (IU/mg).

Determination of isoelectric points by electrofocusing. Isoelectric focusing was carried out[14,15] with the individual chromatographic forms of the inhibitor (5 mg samples) in a 110 ml glass column (LKB, Uppsala) for 24 h. The anodic electrolyte (1% phosphoric acid) was placed in the upper part of the column and the cathodic electrolyte (0·8% sodium hydroxide) at the bottom of the column. Sucrose (5 to 50%) was used as a stabilizing density gradient. The initial electric output of the column did not exceed 5 W. Servalyte 2—11 (2%) was used to develop the pH-gradient. The contents of each tube were subjected to the measurement of pH by a combined electrode, optical density at 280 nm, and antitryptic activity.

Electrophoresis on cellulose acetate. The electrophoresis of the individual forms was carried out in 0·05M sodium acetate adjusted to pH 4·3 by acetic acid or in 0·2M-Tris-HCl/EDTA buffer at pH 9·2. The protein (0·05 mg) was applied along a 1 cm line (0·05 mg of each form when

a mixture of forms was applied); time 30 min, gradient 40 V/cm. The proteins were stained with amido black.

Resolution of tryptic digests of oxidized forms (I–IV). The individual forms (2 mg of each) were oxidized by performic acid[16]. Tryptic digestion was effected by TPCK-trypsin (1 : 50, w/w) in 0·1M ammonium carbonate at pH 8·3. The digestion was terminated after 90 min by acidification to pH 3·0 by formic acid and the samples were taken to dryness in a dessicator. The separation of the tryptic digests was carried out on paper Whatman No 3MM. Electrophoresis in acetic acid–pyridine–water (1 : 4 : 250) at pH 5·6 (100 V/cm, 4°C) was carried out in the first direction; after the paper had been dried, chromatography in n-butyl alcohol–pyridine–acetic acid–water (15 : 10 : 3 : 12) was carried out in the second direction at right angles to the first one. The peptides were stained by 0·2% solution of ninhydrin in acetone.

Isolation and characterization of peptides from tryptic digest of oxidized form I. The oxidation and tryptic digestion were carried out as described above except that 10 mg of the protein was used. The digest was resolved by high voltage electrophoresis and bands of paper containing peptides were cut out, stitched to paper Whatman No 3MM and subjected to descending chromatography. The peptides separated were cut out and eluted by distilled water. The eluates were taken to dryness in a dessicator. The amino acid analysis of the individual forms was carried out after hydrolysis (6M-HCl, 20 h, 110°C) by the conventional procedure[17] in an amino acid analyzer manufactured by the Instrument Development Workshops, Czechoslovak Academy of Sciences, Prague. Cysteine was determined as cysteic acid after oxidation by performic acid[16]. The N-terminal amino acid of samples of the individual forms and the amino acid sequence of tryptic peptide 5, isolated from form I, were determined by the phenyl isothiocyanate method[18]. The amino acid of phenylthiohydantoins were identified[19] by thin-layer chromatography on Silufol sheets.

RESULTS

Using chromatography on SE-Sephadex and a linear gradient of ionic strength, four well separated forms showing antitryptic activity (Fig. 1) were obtained from the material prepared by affinity chromatography (Table I). The yield after separation and lyophilization was 76·0 mg (7%) of form I, 462·5 mg (41%) of form II, 441 mg (39·4%) of form III, and 139 mg (12·4%) of form IV. The total recovery was 56% of the protein applied to the column. The recoveries and specific inhibitory activities of the colostrum inhibitor after the individual isolation steps are shown in Table I.

The amino acid composition of forms I to IV is given in Table II. None of the forms contains tryptophan, histidine, and valine; the low content of basic amino acids of all forms is the cause of the low isoelectric points determined by isoelectric focusing (Table III). The homogeneity of the forms obtained was demonstrated by electrophoresis on cellulose acetate where all zones moved as a single zone both in acidic (pH 4·3) and basic (pH 9·2) medium. The individual forms differ, however, in mobility, as shown in Fig. 2A and B. All forms are N-terminated with phenylalanine. Their homogeneity is evidenced also by the symmetry of the peaks of the protein and of inhibitory activity during electrofocusing of the individual forms.

TABLE I

Isolation of Proteinase Inhibitors from Cow Colostrum

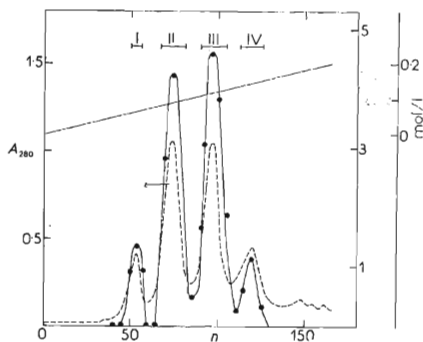| Fraction	Recovery of activity %	Specific activity IU/mg
Colostrum	100	—
Filtrate after treatment of colostrum by 2·5% trichloroacetic acid	68	0·02
Material after affinity chromatography	61	1·20
Material after SE-Sephadex[a]	53	1·85

[a] All forms.

Fig. 1

Chromatography of Cow Colostrum Trypsin Inhibitor on SE-Sephadex at pH 3·5

The inhibitory activity is expressed in 1U per ml of solution. I to IV chromatographic forms. The individual inhibitor forms were eluted by a linear gradient of ionic strength (0 to 0·2M-NaCl) in the equilibration buffer (1 500 ml). Fractions (n) of 18·2 ml were collected at 15 min intervals.

The positions of the peptides on the peptide maps (Fig. 3A—D) show that peptides 2, 3, and 4 occupy identical positions in all forms. It was confirmed by amino acid analysis that the composition of the peptides isolated is also identical. The N-termi-

TABLE II

Amino Acid Composition of Forms I—IV of Cow Colostrum Trypsin Inhibitor

The values were obtained with 20 h hydrolysates and are not corrected. Analyses were calculated assuming the presence of 4 alanine residues.

| Amino acid	Number of residues per mol of protein			
	I	II	III	IV
Lysine	0·95	1·25	1·63	2·38
Arginine	3·00	3·00	2·97	3·00
Aspartic acid	6·90	7·82	8·10	8·00
Threonine	5·10	5·46	5·60	4·90
Serine	1·90	2·36	2·42	2·48
Glutamic acid	8·90	9·80	10·00	9·62
Proline	7·00	6·65	7·00	6·44
Glycine	4·00	4·00	4·05	4·08
Alanine	4·00	4·00	4·00	4·00
Cysteic acid	5·98	5·89	6·05	5·95
Methionine	1·00	1·00	1·05	0·95
Isoleucine	0·95	0·96	0·98	1·05
Leucine	4·67	5·14	5·40	5·19
Tyrosine	2·84	2·89	2·96	2·90
Phenylalanine	4·03	3·99	4·05	4·00
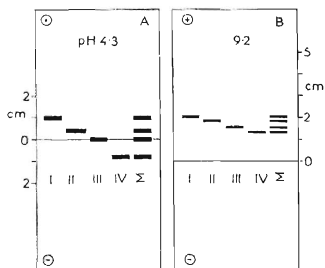
FIG. 2

Schematic Representation of Individual Inhibitor Forms on Cellulose Acetate Sheets

I, II, III, IV — individual forms, Σ-mixture of individual forms.

nal peptide exists in variants 1 and 1' and the C-terminal peptide in variants 5 and 5' (Table III). Form I only has one N-terminal peptide 1 and one C-terminal peptide 5. This shows that form I contains one inhibitor only; we named it isoinhibitor A. Forms II, III, and IV are mixtures of isoinhibitors. The survey of all the forms isolated of all the known cow colostrum trypsin isoinhibitors is given in Table III and IV.

The treatment of the tryptic digest of oxidized form I afforded peptide 1 (0·224 μmol), 2 (0·240 μmol), 3 (0·314 μmol), 4 (0·320 μmol), and 5 (0·290 μmol). The quantitative amino acid analyses of isolated peptides 1 to 4 confirmed that these peptides are identical with peptides derived from the same positions of isoinhibitor B (ref.[3]). The amino acid composition of peptide 5, which is Glu (2·02), Pro (2·00), CySO$_3$H (0·95), and Ile (0·74), shows that this peptide is by five amino acid residues shorter than the corresponding peptide derived from isoinhibitors B and C; this was also evidenced by the determination of the amino acid sequence of this peptide, Ile-Cys--Glu-Pro-Pro-Gln (residues 57 to 62). A comparison of the structures of all three isoinhibitors A, B, and C is given in Table V.

DISCUSSION

The naturally occurring, low molecular weight proteinase inhibitor from cow colostrum is a glycoprotein. Approximately one third of its molecule is represented by the carbohydrate moiety containing fucose, mannose, galactose, galactosamine, glucosamine, and neuraminic acid[5]. The sugar moiety is attached to the protein molecule at one site only (ref.[20]), namely at Asx 27. The inhibitor is as a rule eluted from ion exchange columns in multiple peaks showing inhibitory activity, in forms which often differ in amino acid composition[1,2,6,21].

When using the original method of isolation of cow colostrum proteinase inhibitor we obtained first isoinhibitor B containing two lysines in its molecule[1] and later isoinhibitor C with three lysines[6]. We had not been able to isolate in pure state the isoinhibitor with one lysine residue even though we obtained evidence of its existence during our experiments. In pure state this inhibitor is named form I. The survey of all isoinhibitors is given in Table IV. The structure of isoinhibitor A is identical with the structure of isoinhibitor B except that isoinhibitor A is shorter by five amino acid residues at its C-terminus (Table IV). Since we found isoinhibitor A also in cow colostrum treated after the milking with 2·5% trichloracetic acid to inactivate all the proteinases, we postulate that the decrease of the length of the polypeptide chain is coded for genetically.

The four forms of the inhibitor are homogeneous with respect to their total net charge as evidenced by electrophoresis on cellulose acetate strips (Fig. 2) and by single symmetrical peaks of both protein and inhibitory activity, demonstrated during the determination of the isoelectric points by isoelectrofocusing. The existence of four

electrophoretically homogeneous fractions has been reported by Pineiro and co-workers[22] who showed the presence of four zones not only in the intact inhibitor but also in the material from which neuraminic acid had been removed by neuraminidase. We demonstrated by the analysis of our four forms obtained (Fig. 3A — D, Table III) that form *I* only is pure isoinhibitor A; the remaining forms represent mixtures of isoinhibitors.

Since the described isolation procedure also involves the precipitation of proteins by 2·5% trichloroacetic acid, we cannot exclude the possibility that the existence of the forms is caused by destruction of the carbohydrate moiety in acid media. Since the isolation procedures which do not use trichloroacetic acid also yielded multiple forms[2,5,22] and since there is no difference in the elution diagrams obtained by chromatography on SE-Sephadex of inhibitor concentrates prepared with and without trichloroacetic acid[7], we assume that the heterogenicity of the carbohydrate moiety of the colostrum inhibitor originates in the synthesis of this glycoprotein.

TABLE III

Characterization of Forms I—IV of Cow Colostrum Trypsin Inhibitor

Form	I	II	III	IV
Isoelectric point	3·8	4·0	4·3	4·5
Content of lysine residues per mol of protein	1·0	1·3	1·6	2·4
Peptides in tryptic digest	1,2,3,4,5	1,2,3,4,5,5′	1′,1,2,3,4,5,5′	1′,1,2,3,4,5,5′
Isoinhibitors in given form	A	A + B	A + B + C	A + B + C

TABLE IV

Characterization of Isoinhibitors A, B, and C isolated from Cow Colostrum

Isoinhibitor	A	B	C
Number of amino acid residues	62	67	67
Number of lysine residues	1	2	3
Peptides in tryptic digest	1,2,3,4,5	1,2,3,4,5′	1′,2,3,4,5′

The occurrence of polymorphism of glycoproteins can obviously be accounted for also by differences in both the protein and the carbohydrate moieties. The colostrum inhibitor represents because of its low molecular weight and relatively easy identificability of mutants a suitable model for studies on the polymorphism of glycoproteins. The results of the analyses of electrophoretically homogeneous forms point to complications involved in the interpretation of the results of electrophoretic
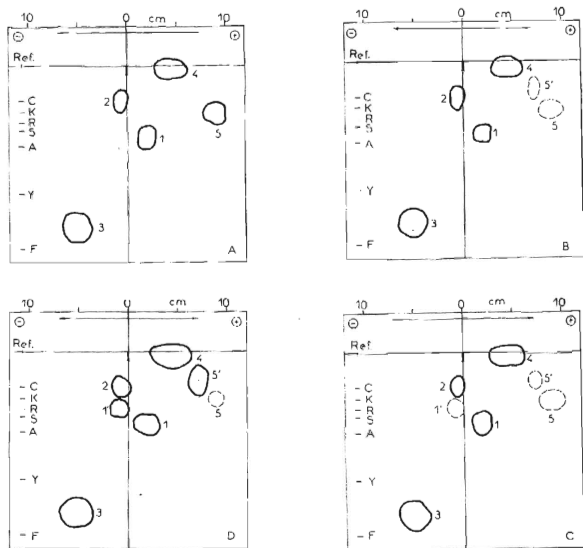

FIG. 3

Peptide Maps of Tryptic Digests of Oxidized Forms

A) Form I, B) II, C) III, and D) IV. The high voltage electrophoresis was carried out in the horizontal arrangement, chromatography in the vertical arrangement. Ref. — reference mixture; C cysteic acid, K lysine, R arginine, S serine, A alanine, Y tyrosine, F phenylalanine. 1 to 5 peptides from tryptic digest in order from the N-terminus of the molecule of isoinhibitor A; 1′ N-terminal peptide of isoinhibitor C; 5′ C-terminal peptide of isoinhibitor B and C. The peptides present in smaller quantities are marked by a dashed line.

Isoinhibitor A

```
1    *                          10                            20              3
Phe-Thr-Pro-Pro-Asp-Leu-Cys-Cys-Gln-Leu-Pro-Gln-Ala-Arg-Gly-Pro-Cys-Lys-Ala-Ala-Leu-Leu-Arg-
|————————————————————————1————————————————————————|           |——3——|

          +                30                            40
-Tyr-Phe-Tyr-Asx-Ser-Thr-Ser-Asn-Ala-Cys-Glu-Pro-Phe-Thr-Tyr-Gly-Gly-Cys-Gln-Gly-Asn-Asn-Asx-
|——————————————————4——————————————————|

       50                              60        62
-Asn-Phe-Glu-Thr-Thr-Glu-Met-Cys-Leu-Arg-Ile-Cys-Glu-Glu-Pro-Pro-Gln
|————————4————————|  |—————5—————|
```

Isoinhibitor B

```
1    *                    10              14
Phe-Gln-Thr-Pro-Pro-Asp-Leu-Cys-Gln-Leu-Pro-Gln-Ala-Arg············
|——————————————1'——————————————|

              60              57
·······Ile-Cys-Glu-Glu-Pro-Pro-Gln-Gln-Thr-Asp-Lys-Ser
                    |————————5'————————|
```

Isoinhibitor C

```
1    *                    10              14
Phe-Gln-Lys-Pro-Pro-Asp-Leu-Cys-Gln-Leu-Pro-Gln-Ala-Arg············
|——————————————1——————————————|

              60              57
·······Ile-Cys-Glu-Glu-Pro-Pro-Gln-Gln-Thr-Asp-Lys-Ser
                    |————————5————————|
```

TABLE V

Alignment of Structure of Isoinhibitor A with Known Structures of Isoinhibitors B and C

1 to 5, order of tryptic peptides in isoinhibitor A; 1', N-terminal peptide of isoinhibitor C; 5', C-terminal peptide of isoinhibitors B and C; • site of amino acid replacement; + site of attachment of carbohydrate moiety.

methods which represent so far the most common method of screening of enzymatic mutants. It is obvious that the electrophoretic methods are not a reliable tool of identification of isozymes of glycoprotein character.

REFERENCES

1. Čechová D., Jonáková-Švestková V., Šorm F.: This Journal *35*, 3085 (1970).
2. Čechová D.: This Journal *39*, 647 (1974).
3. Čechová D., Jonáková V., Šorm F.: This Journal *36*, 3342 (1971).
4. Čechová D., Ber E.: This Journal *39*, 680 (1974).
5. Tschesche H., Klauser R., Čechová D., Jonáková V.: Hoppe-Seylers' Z. Physiol. Chem. *356*, 1759 (1975).
6. Jonáková V., Čechová D.: This Journal *42*, 759 (1977).
7. Čechová D., Jonáková V.: Unpublished results.
8. Kucich U.: *Thesis*. State University of New York at Buffalo 1972.
9. Laskowski M.: Private communication.
10. Laskowski M. jr, Laskowski M.: J. Biol. Chem. *190*, 563 (1951).
11. Tuppy H., Wiesbauer U., Winterberger E.: Hoppe-Seylers' Z. Physiol. Chem. *329*, 278 (1962).
12. Nagel W., Willig F., Peschke W., Schmidt F. H.: Hoppe-Seylers' Z., Physiol. Chem. *340*, 1 (1966).
13. Fritz H., Trautschold I., Werle E.: Methods Enzym. Anal. *2*, 1064 (1974).
14. Svenson H.: Arch. Biochem. Biophys. *1*, 132 (1962).
15. Vestertag O., Svenson H.: Acta Chem. Scand. *20*, 820 (1966).
16. Hirs C. H.: J. Biol. Chem. *219*, 611 (1956).
17. Spackman D. H., Stein W. H., Moore S.: Anal. Chem. *30*, 1190 (1958).
18. Edman R., Hanschen A. in the book: *Molecular Biology and Biophysics* (A. Kleinzeller, F. G. Springer, H. G., Wittmann, Eds), Vol. 8, p. 251. Springer, Berlin—Heidelberg—New York 1975.
19. Grüner K.: Chem. Listy *64*, 1160 (1970).
20. Klauser R., Čechová D., Tschesche H.: Hoppe-Seylers' Z. Physiol. Chem. *359*, 173 (1978).
21. Kress L. F., Martin S., Laskowski M.: Biochim. Biophys. Acta *229*, 836 (1971).
22. Pineiro A., Ortega F., Uriel J.: Biochim. Biophys. Acta *379*, 201 (1975).