Volume 251, number 1,2, 94–98 FEB 07372 July 1989

Primary structure of cathepsin D inhibitor from potatoes and its structure relationship to soybean trypsin inhibitor family

M. Mareš, B. Meloun, M. Pavlík, V. Kostka and M. Baudyš

Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Sciences, Flemingovo 2, CS-166 10 Prague, Czechoslovakia

Received 19 May 1989

A novel effective procedure for the purification of cathepsin D inhibitor from potatoes (PDI) was developed. The amino acid sequence of PDI was determined by analysis of the cyanogen bromide digest and of the limited tryptic and chymotryptic digests of the protein. The inhibitor is a single polypeptide chain protein consisting of 188 residues with a simple sugar moiety attached to Asn-19. The tentative disulfide pairings are also suggested. The sequence data clearly indicate that PDI is homologous with the soybean trypsin inhibitor (STI) (Kunitz) family. The active center of PDI for trypsin inhibition was identified as Pro-Val-Arg-Phe in analogy to STI.

Cathepsin D; Proteinase; Enzyme inhibitor; Sequence homology; Amino acid sequence; (Potato)

1. INTRODUCTION

The interest in naturally occurring inhibitors of proteolytic enzymes has been increasing mainly because of their possible use as therapeutic agents. Very little is known so far about inhibitors of aspartic proteinases of protein character. The inhibitor isolated from the roundworm *Ascaris lumbricoides* is a potent inhibitor of pepsin, gastricsin and cathepsin E [1,2]. Intact activation (propart) peptides represent another type of such inhibitors [3]. Inhibitor IA_3 of proteinase A from yeast displays a surprising singularity of inhibition with respect to the other aspartic proteinases [4]. So far no cathepsin D inhibitor of tissue origin has been described in spite of the fact that this major lysosomal aspartic proteinase [5] plays an important role in many physiological or pathophysiological processes such as neurodegenerative changes [6]. The only well characterized cathepsin D inhibitor of protein character is the inhibitor from potato tubers (PDI) [7]. PDI inhibits, however, besides cathepsin D also trypsin. A preliminary chemical characterization of PDI has shown that PDI is a type of trypsin inhibitor not discovered so far in potatoes [8].

The aim of the present study has been to determine the structure characteristics of PDI and thus to provide a basis for the investigation of its relationship with trypsin inhibitors in general and for future studies on the mechanism of cathepsin D inhibition from the molecular viewpoint. Structure studies of the same inhibitor have reached an advanced stage also in Professor Turk's group at the J. Stefan Institute in Ljubljana [9].

Correspondence address: M. Mareš, Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Sciences, Flemingovo 2, CS-166 10 Prague, Czechoslovakia

Abbreviations: CmCys, S-carboxymethylcysteine; SDS-PAGE, polyacrylamide gel electrophoresis in sodium dodecyl sulfate

2. MATERIALS AND METHODS

Potatoes used for isolation of PDI were of commercial origin. Cathepsin D used for the PDI assay [7] was prepared from bovine spleen essentially according to [10] using affinity chromatography on pepstatin-Sepharose as a last step. The FPLC system was from Pharmacia AB, Sweden. Trypsin- and cathepsin D-Sepharose were prepared and trypsin inhibition

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50 © 1989 Federation of European Biochemical Societies

assayed as described elsewhere [11]. Purification of PDI was performed initially according to [7]; after gel filtration the ion-exchange chromatography step on a Mono Q column equilibrated in 0.1 M ethanolamine-HCl buffer, pH 10.0, was included (fig.1). The homogeneity of PDI and of its large fragments was examined by SDS-PAGE in 12.5% gels as in [12]. The cleavage of carboxymethylated PDI [13] by CNBr, after preceding reduction of methionine sulfoxide present in the protein by mercaptoacetic acid [14], was carried out as in [15]. The CNBr fragments after citraconylation [16] were resolved on Sephadex G-100 (Pharmacia) in 6 M guanidinium chloride. The desalting was carried out on Sephadex G-25 in 0.5% NH_4HCO_3 and both CNBr fragments were further cleaved by trypsin (80:1, w/w; 2 h; 37°C). The resulting subfragments (CB-LT) were separated by gel chromatography on Sephadex G-50 in 0.5% NH_4HCO_3 and then purified by HPLC (Beckman-Altex Model 420 Liquid Chromatograph). An Ultrapore RPSC column (Beckman, USA) in 0.1% ammonium acetate buffer, pH 6.5, and a gradient of acetonitrile were used. Alternatively the carboxymethylated protein was citraconylated and digested with trypsin as in [17]. The resulting fragments (LT) were separated by gel chromatography on Sephadex G-50 in 0.5% NH_4HCO_3, then decitraconylated as described elsewhere [16] and purified by HPLC (see above) whenever necessary. The subsequent cleavage of the LT fragments by trypsin was carried out as described for the CB-LT fragments. The resulting peptides (LT-Tr) were purified by RP HPLC. For the determination of some overlaps, the carboxymethylated protein was also digested with chymotrypsin in 0.5% NH_4HCO_3 at a substrate to enzyme ratio of 50:1 (w/w) for 2 h at 37°C. The peptides were separated by HPLC on a Vydac 218TP54 column (Separations Group, USA). Amino sugars, neutral sugars and fucose were quantified as described elsewhere [17]. The amino acid analyses were carried out using a Durrum D-500 amino acid analyzer. The amino acid sequences were determined on an Applied Biosystems model 470A protein sequencer using the program provided by the instrument manufacturer. The phenylthiohydantoins were identified by HPLC as in [18]. The carboxypeptidase A digestion of carboxymethylated PDI in 0.25% SDS and 0.1 M *N*-methylmorpholine, pH 8.5, was performed according to [19].

3. RESULTS AND DISCUSSION

The first aim of this study was to simplify the process of PDI purification in order to obtain higher yields than by the original procedure [7]. Chromatography on a Mono Q column (fig.1) was found to be sufficient for the isolation of pure PDI by a one step procedure. A large number of other, more basic trypsin inhibitors were separated concomitantly. PDI was eluted from the column as the only symmetrical peak showing both anticathepsin D and antitrypsin activity. To provide evidence that these two activities cannot be separated we repeated the experiments of Keilová [11] with affinity chromatography on cathepsin D- and

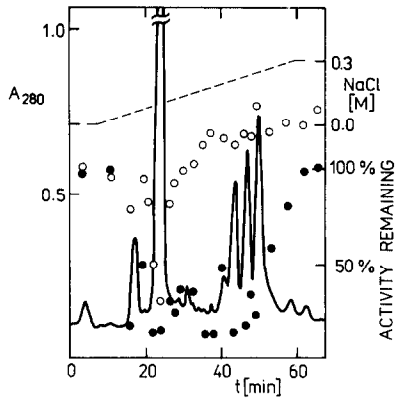

Fig.1. Chromatography of crude PDI on Mono Q column (4.6 × 75 mm). A solution of 0.1 M ethanolamine-HCl, pH 10.0, was used as starting buffer. The lyophilisate (10 mg, after gel filtration) dissolved in 2 ml of starting buffer was applied and eluted at a flow rate of 1 ml/min by a 0.0–0.3 M gradient of NaCl. A_{280} was measured continuously, 1 ml fractions were collected and 100 μl aliquots were assayed for anticatheptic (○) [7] and antitryptic (●) [11] activity.

trypsin-Sepharose columns. In both cases the protein peak corresponding to PDI purified by affinity chromatography coincided with the peak of both anticathepsin D and antitrypsin activity. The preparation obtained by chromatography on the Mono Q column was therefore considered homogeneous. Its M_r value according to SDS-PAGE is 22000. The following amino acid analysis of PDI, i.e. Asp 25.5; Thr 8.1; Ser 12.7; Glu 11.5; Pro 15.0; Gly 17.9; Ala 7.0; Cys 5.8; Val 15.0; Met 0.9; Ile 11.5; Leu 20.5; Tyr 8.3; Phe 11.2; Lys 9.2; Arg 8.1; Trp 1.0 is in good agreement both with the M_r value determined by SDS-PAGE and with the sequence data (the calculated M_r is 20582 without sugars). PDI is a glycoprotein containing 2 $GlcNH_2$ residues, 4–5 neutral sugar residues and 0.8 fucose residues. The carbohydrate moiety of PDI attached to Asn-19 is thus a simple type, Asn-linked oligosaccharide [20]. The other site of post-translational modification is the only methionine residue of PDI. Alkaline hydrolysis of PDI [14] provided unambiguous evidence of the presence of methionine sulfoxide to methionine of 1:1; this finding explains incomplete cleavage at the methionine residue by CNBr. The fact that the conversion of MeSO to Met is incomplete even after exhaustive reduction suggests that the sulfonium salt of Met rather than MeSO alone is present.

We were able to sequence the intact carboxymethylated protein up to the 58th residue. To obtain additional information, limited tryptic digestion and CNBr cleavage of the carboxymethylated PDI were carried out in parallel experiments. All the bonds were cleaved as expected with the exception of that involving Arg-154. The failure to obtain cleavage at this site (fig.2) can be accounted for by the fact that Arg-154 is surrounded by three acid residues (CmCys-153, Asp-155 and -156). Both CNBr fragments were further subjected to limited tryptic digestion to obtain information on the order of the LT fragments (fig.2). The LT fragments after decitraconylation were cleaved by trypsin. All these data, together with the knowledge of the peptides resulting from chymotryptic digestion of PDI, enabled us to derive the primary structure of PDI shown in fig.3. This structure is homologous to the structure of soybean trypsin inhibitor (STI) [21] and to that of the related winged bean trypsin inhibitor (WTI) [22]. From sequence homologies it was recently shown that the doubleheaded α-amylase/subtilisin inhibitor from barley [23] as well as the inhibitors of cysteine proteinases from potatoes [24] also belong to the same structural superfamily. This structural arrangement thus seems to be very widespread among different kinds of inhibitors. The homology of PDI and STI is more pronounced in N- and C-terminal parts of the molecule but it is considerably lower in other parts. We used the first two half-cystines (Cys-48 and -93) forming a disulfide in STI as an 'alignment marker'. The active site for trypsin inhibition should be represented by Arg-67 in analogy with STI. The invariant Pro-65 residue in position P_3 [25] also seems to be important and our alignment was made to meet this criterion. The secondary structure prediction by several methods and the resulting prediction [26] indicate the presence of only one helical region (res.76–82), which is inserted in PDI in contrast to STI. From the experimentally determined tertiary structure of STI [27] there is no helical region in this inhibitor; this indicates a common core structure of PDI and STI. The high predicted content of β-structures in PDI is also in accordance with the tertiary structure of STI. Many insertions and deletions on the surface of the molecules compared had to be introduced, however, to obtain an appropriate alignment (fig.3).

STI has two disulfides of which the second forms a short loop [27] on the surface (Cys-142–Cys-153). PDI should contain three disulfides, the second and third lying in the region of the second STI disulfide; hence, the alignment of half-cystines in this part of the molecule should be considered as only tentative (the pairing Cys-142–Cys-153 being a counterpart of the second disulfide of STI).

All these facts indicate that the structures of PDI and STI are not simply superimposable. It is of interest that the sequence Cys-Pro-Phe-Cys present in PDI (res.150–153) is homologous to the active center of glutaredoxin which plays a role in the reduction of ribonucleotides [28]. Recent literature has provided data on the multivalency of STI, at least as far as chymotrypsin binding is concerned [29]. There is no doubt that the homologous α-amylase/subtilisin inhibitor from barley is doubleheaded. PDI is both a trypsin and cathepsin D inhibitor. It is too early to speculate whether the

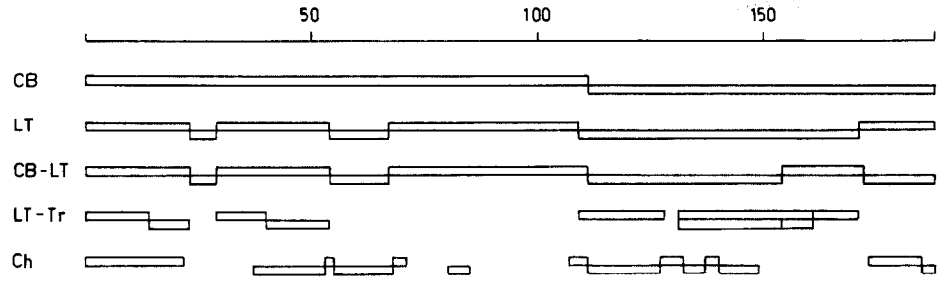

Fig.2. Summary of sequence data. Bars represent individual fragments, their length is proportional to their size. CB, cyanogen bromide fragments; LT, limited tryptic fragments; CB-LT, limited tryptic fragments prepared from CB fragments; LT-Tr, tryptic peptides prepared from LT fragments after decitraconylation; Ch, chymotryptic peptides.

```
                          10                  + 20                  30
PDI   E S P L P K P V L D T N G K E L N P N S S Y R I I S I G R G
STI             D F V L D N E G N P L E N G G T Y Y I L S - D I T
WTI             E P L L D S E G E L V R N G G T Y Y L L P - D R W

                          40                  50
PDI   A L G G D V Y L G K S P N S D A P C P D G V F R Y N - - - -
STI   A F G G - - - I R A A P T G N E R C P L T V V Q S R N E L D
WTI   A L G G G - - I E A A A T G T E T C P L T V V R S - - - - -

              60              *   70                  80
PDI   - S D V G P S G T P V R F I P L S G G I F E D Q L L N I Q F
STI   K G I G T I I S S P Y R I R F I A E G H P - - - L S L K F D
WTI   - P N E V S V G E P L R I S S Q L R S G F I P D Y S L V R I

              90                  100                 110
PDI   N I A T V K L C V S Y T - I W K V G N L N A Y F R T M L L E
STI   S F A V I M L C V G I P T E W S V V E D L P E G P A V K I G
WTI   G F A N P P K C A P S P - W W T V V E D Q P Q Q P S V K L S

                  120                 130                 140
PDI   T G G T I G Q A D S S Y F K I V K L S - - - N F G Y N L L Y
STI   E N K D A M - - - D G W F R L E R V S D D E F N N Y K L V F
WTI   E L K S T K F - - D Y L F K F E K V T S - K F S S Y K L K Y

                      150                 160
PDI   C P I T P P F L C P F C R D D N F C A K V G V V I Q N - - G
STI   C P Q Q A E D D K - - C G D - - - - - - I G I S I D H D D G
WTI   C A K R D T - - - - - C K D - - - - - - I G I Y R D Q K - G

      170                 180
PDI   K R R L A L V N E N P L D V L F Q E V
STI   T R R L V V S K N K P L V V Q F Q K L D K E S L
WTI   Y A R L V V T D E N P L V V I F K K V E S S
```

Fig.3. Alignment of amino acid sequence of potato cathepsin D inhibitor (PDI) with sequence of soybean trypsin inhibitor (STI) and winged bean trypsin inhibitor (WTI). The residues are numbered according to PDI. Identical residues are set in boldface type. Glycosylated Asn-19 is marked by a cross, the putative active site for trypsin inhibition (Arg-67) by an asterisk.

active site for trypsin inhibition is also responsible for the cathepsin D inhibition or whether there is another binding site responsible for the inhibition of cathepsin D. A possible candidate could be the bond Lys-91–Leu-92 in the vicinity of the second half-cystine; in STI there is a Met-Leu bond at the corresponding site which is cleaved by subtilisin [30].

Acknowledgements: We are grateful to Mr J. Zbrožek for the amino acid analyses, and to Ms J. Zaoralová and Ms A. Kulhánková for skilful technical assistance.

REFERENCES

[1] Abu-Erreish, G.M. and Peanasky, R.J. (1974) J. Biol. Chem. 249, 1558–1565.
[2] Keilová, H. and Tomášek, V. (1972) Biochim. Biophys. Acta 284, 461–464.
[3] Baudyš, M., Pichová, I., Pohl, J. and Kostka, V. (1985) in: Aspartic Proteinases and their Inhibitors (Kostka, V. ed.) pp.309–318, Walter De Gruyter, Berlin.
[4] Dreyer, T., Valler, M.J., Kay, J., Charlton, P. and Dunn, B.M. (1985) Biochem. J. 231, 777–779.
[5] Barrett, A.J. (1977) in: Proteinases in Mammalian Cells and Tissues (Barrett, A.J. ed.) pp.209–248, Elsevier, Amsterdam.
[6] Matus, A. and Green, G.D.J. (1987) Biochemistry 26, 8083–8086.
[7] Keilová, H. and Tomášek, V. (1976) Collect. Czech. Chem. Commun. 41, 489–497.
[8] Rupová, L., Keilová, H. and Tomášek, V. (1977) Collect. Czech. Chem. Commun. 42, 2279–2286.
[9] Meško, P., Kopitar, M., Drobnič-Košorok, M. and Turk, V. (1988) The 18th Linderstrom-Lang Conference on Aspartic Proteinases, Elsinore, Denmark, Book of Abstracts, pp.72.
[10] Keilová, H. (1971) in: Tissue Proteinases (Barrett, A.J. and Dingle, T.J. eds) pp.45–57, Elsevier, Amsterdam.
[11] Keilová, H. and Tomášek, V. (1976) Collect. Czech. Chem. Commun. 41, 2440–2447.
[12] Laemmli, U.K. (1970) Nature 227, 680–685.
[13] Hirs, C.H.W. (1967) Methods Enzymol. 11, 199–203.
[14] Houghten, R.A. and Li, C.H. (1983) Methods Enzymol. 91, 549–559.
[15] Gross, E. (1967) Methods Enzymol. 11, 238–255.
[16] Dixon, H.B.F. and Perham, R.N. (1968) Biochem. J. 109, 312–314.
[17] Baudyš, M. and Kostka, V. (1983) Eur. J. Biochem. 136, 89–99.
[18] Pavlík, M. and Kostka, V. (1985) Anal. Biochem. 151, 520–525.
[19] Ambler, R.P. (1972) Methods Enzymol. 25, 143–154.
[20] Kornfeld, R. and Kornfeld, S. (1985) Annu. Rev. Biochem. 54, 631–664.
[21] Kim, S.H., Hara, S., Hase, S., Ikenaka, T., Toda, H., Kitamura, K. and Kaizuma, N. (1985) J. Biochem. 98, 435–448.
[22] Yamamoto, M., Hara, S. and Ikenaka, T. (1983) J. Biochem. 94, 849–863.
[23] Svendsen, I., Hejgaard, J. and Mundy, J. (1986) Carlsberg Res. Commun. 51, 43–50.
[24] Popovič, T., Brzin, J., Drobnič-Košorok, M. and Turk, V. (1988) 14th International Congress of Biochemistry, Prague, Czechoslovakia, Book of Abstracts Mo, p.59.
[25] Schechter, L. and Berger, I. (1967) Biochem. Biophys. Res. Commun. 27, 157–162.
[26] Dzionara, M., Robinson, S.M.L. and Wittman-Liebold, B. (1977) Hoppe-Seyler's Z. Physiol. Chem. 358, 1003–1024.
[27] Sweet, R.M., Wright, H.T., Janin, J., Chothia, C.H. and Blow, D.M. (1974) Biochemistry 13, 4212–4228.
[28] Klintrot, I.M., Höög, J.O., Jörnvall, H., Holmgren, A. and Luthman, M. (1984) Eur. J. Biochem. 144, 417–423.
[29] Bosterling, B. and Quast, U. (1981) Biochim. Biophys. Acta 657, 58–72.
[30] Laskowski, M., jr, Kato, I., Leary, T.R., Schrode, J. and Sealock, R.W. (1974) in: Proteinase Inhibitors – Bayer Symposium V (Fritz, H. et al. eds) pp.597–611, Springer, Heidelberg.