Volume 235, number 1,2, 271–274 FEB 06159 August 1988

Protein chemical characterization of *Mucor pusillus* aspartic proteinase

Amino acid sequence homology with the other aspartic proteinases, disulfide bond arrangement and site of carbohydrate attachment

Miroslav Baudyš, Stephen Foundling*, Manfred Pavlík, Tom Blundell[+] and Vladimír Kostka

Institute of Organic Chemistry and Bio-chemistry, Czechoslovak Academy of Sciences, Flemingovo nam. 2, CS-166 10 Prague, Czechoslovakia, [+]Laboratory of Molecular Biology, Department of Crystallography, Birkbeck College, University of London, London WC1E 7HX, England and **E.I. du Pont de Nemours & Co., Ltd, Central Research & Development Department, Bldg 328, Rm 350A, Wilmington, DE 19898, USA*

Received 3 June 1988

The amino acid sequence of *Mucor pusillus* aspartic proteinase was determined by analysis of fragments obtained from cleavage of the enzyme by CNBr and limited tryptic digestion. The proteinase is a single polypeptide chain protein containing 361 amino acid residues, cross-linked by two disulfide bonds. A sugar moiety composed of two GlcNAc residues and four neutral sugar residues is asparagine-linked to the chain. The sequence of *M. pusillus* proteinase is highly homologous with the *M. miehei* proteinase (83% identity). The homology with other aspartic proteinases is low (22–24%) and indicates that the *Mucor* proteinases diverged at an early evolutionary phase. The most conservative regions of the molecule are those involved in catalysis and forming the binding cleft and the core region of the molecule.

Aspartic proteinase; Amino acid sequence; *N*-Glycosylation; Disulfide bridge; Evolution; (*Mucor pusillus*)

1. INTRODUCTION

The last decade has witnessed growing interest in studies on the structure-function relationship of aspartic proteinases. Particular attention has been focused on microbial aspartic proteinases (EC 3.4.23.6) [1–4]. Of these enzymes capable to clot milk, two most frequently used in cheese production are the extracellular aspartic proteinases produced by two related species of a zygomycetes fungus, *Mucor pusillus* and *Mucor miehei* (MPP and MMP). Both enzymes have been purified and characterized [5–8]. The primary structure of MMP was elucidated by both protein and nucleic acids sequencing [9–11]. Evidence was obtained showing that the enzyme is synthesized as a zymogen. The original goal of this study was the determination of the amino acid sequence of MPP as a necessary prerequisite of X-ray studies of the enzyme. In the final stage of this study, the structure of the MPP gene was published [12]. This paper describes the covalent structure of MPP and the sites of its different posttranslational modifications. The homology of this structure with aspartic proteinases of other classes and its impact on the three-dimensional structure of the enzyme are discussed.

Correspondence address: M. Baudys, Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Sciences, Flemingovo nam. 2, CS-16610 Prague, Czechoslovakia

2. MATERIALS AND METHODS

MPP was isolated by a modification of the procedures described elsewhere [5,6] using FPLC (Pharmacia) on a Mono Q column for the final purification. Cleavage of the protein in

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50

its carboxymethylated form [13] (CMMPP) was effected by CNBr [14] or TPCK-trypsin (Worthington) after citraconylation of the lysine residues [15]. The resulting fragments were separated by gel filtration on Sephadex G-75 (Pharmacia) in 1% NH_4HCO_3. Large fragments were further purified by FPLC on a Mono Q column essentially according to [16] and by HPLC on an Ultrapore^{TM} RPSC column (Beckman) in 0.1% ammonium acetate buffer, pH 6.5, using a gradient of acetonitrile (Fluka). The homogeneity of the protein sequenced and of the large fragments was examined by SDS-PAGE in 10% gels according to Weber and Osborn [17]. The enzyme, denatured with alkali (pH 10.0, 30 min) and then digested with chymotrypsin (Worthington) in 0.1% NH_4HCO_3 at a substrate to enzyme ratio of 50:1 (w/w) for 2 h at 37°C, served as starting material for the determination of S-S bridges and overlaps. Purified large CNBr fragments were cleaved by staphylococcal V-8 proteinase (Miles) or by TPCK-trypsin. The peptides containing disulfide bridges were detected as described in [18]. Glycopeptides were identified and neutral sugars, amino sugars and fucose were quantitated as described [19]. Small and middle-size peptides were separated by HPLC on an Ultrasphere Octyl (Beckman) or Vydac 218TP54 column (Separations Group) using a Beckman-Altex model 420 liquid chromatograph. The amino acid analyses were carried out in a Durrum D-500 amino acid analyzer. Amino acid sequences were determined either in a Beckman 890C amino acid sequenator [20] or in Applied Biosystems model 470A protein sequencer using the program provided by the instrument manufacturer. The phenylthiohydantoins were identified by HPLC as in [21]. Smaller peptides were sequenced manually [22]. Carboxypeptidase Y (Sigma) digestion of CMMPP in 0.25% SDS and 0.1 M *N*-methylmorpholine, pH 6.0, was performed according to [23].

3. RESULTS AND DISCUSSION

The strategy employed for the sequence determination is schematically summarized in fig.1. We started with CNBr cleavage which afforded 5 specific fragments. The C-terminal fragment aggregated heavily during gel filtration even in 6 M guanidine. This phenomenon was effectively utilized for its purification. Small and middle-size fragments were sequenced directly. The structure of the two largest CNBr fragments was determined by direct sequencing and by analysis of their tryptic and V-8 digests. Additional sequence information was obtained from tryptic digestion of citraconylated CMMPP, yielding 6 specific fragments (fig.1). All these fragments were characterized by extensive N-terminal sequence analyses. For the isolation of overlaps of the CNBr fragments and corroboration of their sequences the chymotryptic digest of the intact MPP was used. The digest was concomitantly utilized for the determination of the halfcystine pairing and identification of the site of attachment of the saccharide moiety. The results obtained enabled us to order all fragments unambiguously. The resulting amino acid sequence is presented in fig.2. The molecule contains 361 residues including 2 disulfide bonds (Cys-51–Cys-57, Cys-272–Cys-316). The polypeptide chain of MPP contains one glycosylation site, Asn-188, with the adjacent signal glycosylation sequence for asparagine [24]. There are two other potential glycosylation sites, Asn-79 and Asn-113, which are not glycosylated as shown by sequence analysis. This finding is in agreement with the carbohydrate composition of MPP showing the presence of 2 glucosamines, 3 mannoses and probably 1 fucose residue. The carbohydrate moiety of MPP is thus most likely represented by a simple type of Asn-linked oligosaccharide containing the mannose core only and is fucosylated on the Asn-linked GlcNAc [24]. In contrast, MMP has 2 glycosylation sites, Asn-79 and -188 [10].

Our complete amino acid sequence of MPP is in

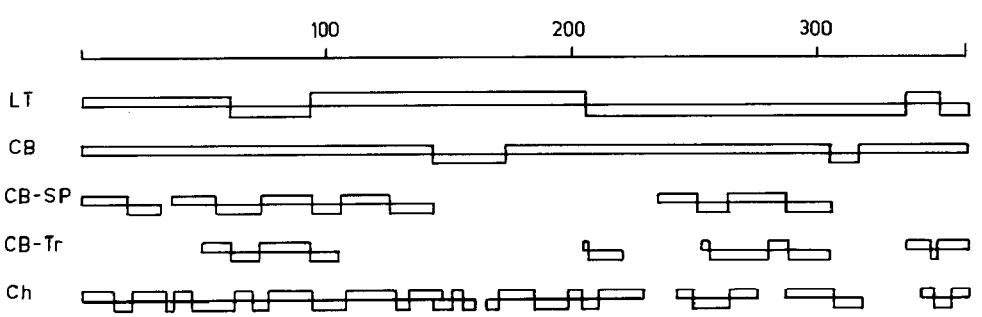

Fig.1. Summary of sequence evidence. Bars represent individual fragments, their length is proportional to their size. LT, limited tryptic fragments; CB, cyanogen bromide fragments, and their (CB-Tr) tryptic and (CB-SP) *Staphylococcus aureus* V8 proteinase subfragments; Ch, chymotryptic peptides.

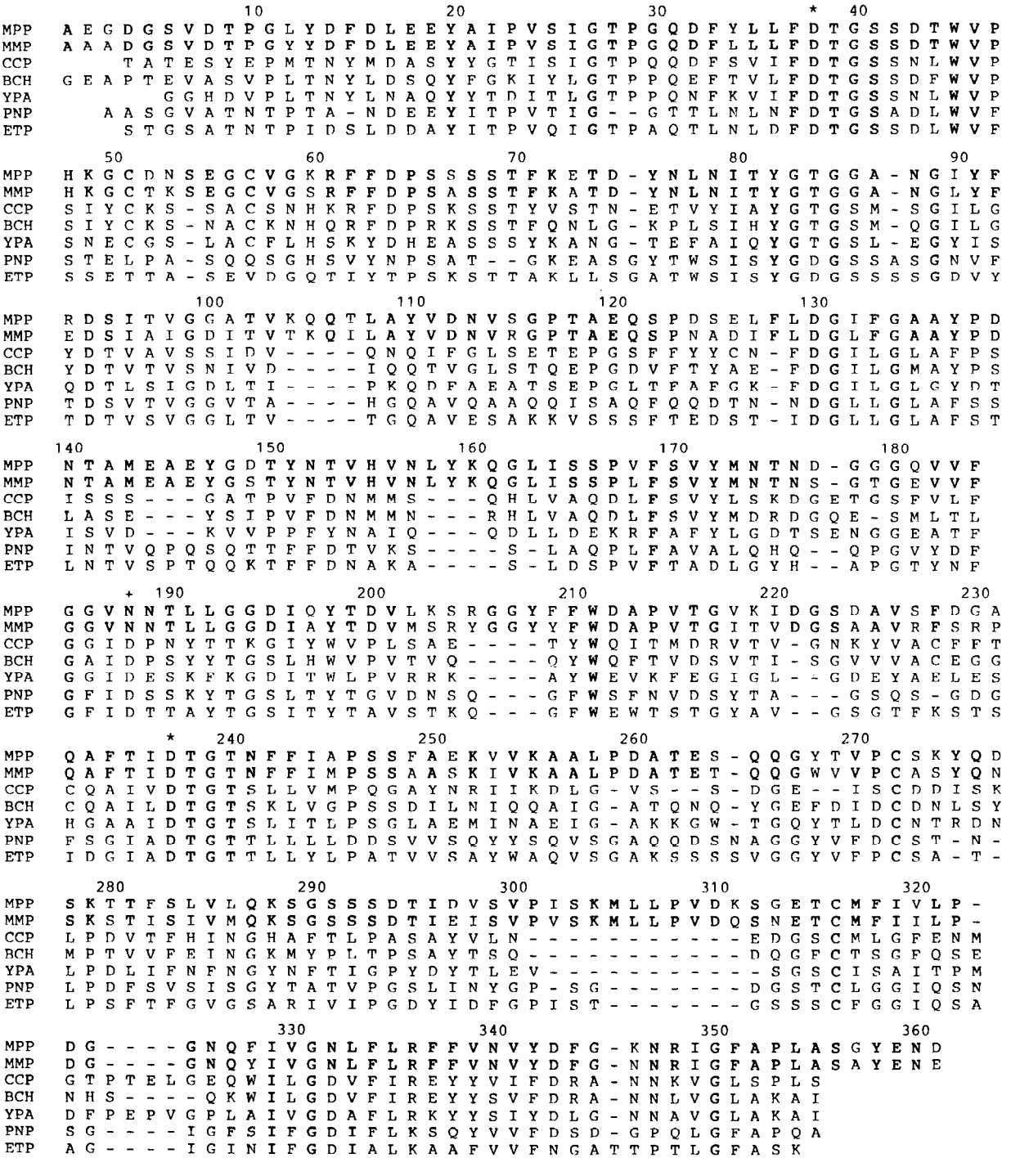

Fig.2. Comparison of the sequence of *Mucor pusillus* aspartic proteinase (MPP) with those of related proteinases from *Mucor miehei* (MMP) [9–11], chicken pepsin (CPP) [16], bovine chymosin (BCH) [26], penicillopepsin (PNP) [3], endothiapepsin (ETP) [2] and yeast proteinase A (YPA) [25]. The residues are numbered according to MPP. Identical residues in MPP and MMP and also in the other proteinases are set in boldface type. The catalytically active aspartic acid residues are marked by an asterisk, glycosylated Asn 188 by a cross.

perfect agreement with the gene structure of MPP published recently [12]. The only exception is the C-terminal amino acid, Asn-361, according to the gene sequence. We have shown the amino acid at 361 to be Asp, its recovery after carboxypeptidase digestion, however, was not 100% with respect to the recovery of the other amino acids of the C-terminal sequence. One possible explanation for this is that Asn is the C-terminal amino acid during proteosynthesis and that deamidation occurs later as a posttranslational event.

It follows from the alignment (fig.2) that MPP and MMP are very closely related (83% homology). The degree of homology between MPP and MMP and other aspartic proteinases is essentially the same and low (22–24%). Moreover, other representatives of lower eukaryotes, endothiapepsin [2], penicillopepsin [3] and proteinase A from yeast [25] are even less related to MPP (22% identity) than to the aspartic proteinases of vertebrates (25–33%). This indicates that MPP and MMP diverged at an early phase in the evolution of aspartic proteinases. This finding is in agreement with general biologic taxonomy since organisms producing MPP and MMP are the only representatives of lower fungi in the set considered. As expected, the strongly conserved regions of aspartic proteinases around the two catalytically essential aspartic residues (fig.2) of the active site are conserved in MPP, too. Strongly conservative are also flap region [3] which may play a significant role in the catalysis by stabilizing the transition state [1], and the regions around residues 130, 170 and 330 (MPP numbering) which were shown to form an extensive binding cleft in endothiapepsin and penicillopepsin [1,3]. The localization of the two S-S bonds of MPP is topologically equivalent to their counterparts in other aspartic proteinases [26]. All these facts clearly indicate that the fundamental tertiary structure of aspartic proteinases is also retained in the case of MPP.

Acknowledgements: We are indebted to Mrs J. Zaoralová for expert technical assistance and to Mr J. Zbrožek for the amino acid analyses.

REFERENCES

[1] Blundell, T.L., Cooper, J., Foundling, S.I., Jones, D.M., Atrash, B. and Szelke, M. (1987) Biochemistry 26, 5585–5590.
[2] Blundell, T.L., Jenkins, J., Pearl, L.H., Sewell, T. and Pederson, V. (1985) in: Aspartic Proteinases and their Inhibitors (Kostka, V. ed.) pp.151–161, De Gruyter, Berlin.
[3] James, M.N.G., Sielecki, A.R. and Hofmann, T. (1985) in: Aspartic Proteinases and their Inhibitors (Kostka, V. ed.) pp.163–177, De Gruyter, Berlin.
[4] Suguna, K., Bott, R.R., Padlan, E.A., Subramanian, E., Sheriff, S., Cohen, G.H. and Davies, D.R. (1987) J. Mol. Biol. 196, 877–900.
[5] Iwasaki, S., Tamura, G. and Arima, K. (1967) Agric. Biol. Chem. 31, 546–551.
[6] Ottesen, M. and Rickert, W.S. (1970) Compt. Rend. Trav. Lab. Carlsberg 37, 301–325.
[7] Rickert, W.S. and McBride-Warren, P.A. (1977) Biochim. Biophys. Acta 480, 262–268.
[8] Etoh, Y., Shoun, H., Beppu, T. and Arima, K. (1979) Agric. Biol. Chem. 43, 209–215.
[9] Bech, A.-M. and Foltmann, B. (1981) Neth. Milk Dairy J. 35, 275–280.
[10] Boel, E., Bech, A.-M., Randrup, K., Draeger, B., Fill, N.P. and Foltmann, B. (1986) Proteins Struct. Funct. Genet. 1, 363–369.
[11] Gray, G.L., Hayenga, K., Cullen, D., Wilson, L.J. and Norton, S. (1986) Gene 48, 41–53.
[12] Tonouchi, N., Shoun, H., Uozomi, T. and Beppu, T. (1986) Nucleic Acids Res. 14, 7557–7568.
[13] Hirs, C.H.W. (1967) Methods Enzymol. 11, 199–203.
[14] Gross, E. (1967) Methods Enzymol. 11, 238–255.
[15] Dixon, H.B.F. and Perham, R.N. (1968) Biochem. J. 109, 312–314.
[16] Baudyš, M. and Kostka, V. (1983) Eur. J. Biochem. 136, 89–99.
[17] Weber, K. and Osborn, M.C. (1969) J. Biol. Chem. 244, 4406–4412.
[18] Thannhauser, T.W., McWherter, C.A. and Scheraga, H.A. (1985) Anal. Biochem. 149, 322–330.
[19] Baudyš, M., Kostka, V., Gruner, K. and Pohl, J. (1982) Coll. Czech. Chem. Commun. 47, 709–718.
[20] Edman, P. and Begg, J. (1967) Eur. J. Biochem. 1, 80–91.
[21] Pavlík, M. and Kostka, V. (1985) Anal. Biochem. 151, 520–525.
[22] Chang, J.-Y., Brauer, D. and Wittmann-Liebold, B. (1978) FEBS Lett. 93, 205–214.
[23] Hayashi, R. (1977) Methods Enzymol. 47, 84–93.
[24] Kornfeld, R. and Kornfeld, S. (1985) Annu. Rev. Biochem. 54, 631–664.
[25] Dreyer, T., Halkier, B., Svendsen, I. and Ottesen, M. (1986) Carlsberg Res. Commun. 51, 27–41.
[26] Foltmann, B. and Pedersen, V.B. (1977) in: Acid Proteases, Structure, Function and Biology (Tang, J. ed.) pp.3–22, Plenum, New York.