Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Molecular Catalysis B: Enzymatic 23 (2003) 405–409
JOURNAL OF MOLECULAR CATALYSIS B: ENZYMATIC
www.elsevier.com/locate/molcatb

Structure and function of an isozyme of earthworm proteases as a new biocatalyst

Manabu Sugimoto [a,*], Kohji Ishihara [b], Nobuyoshi Nakajima [c]

[a] *Research Institute for Bioresources, Okayama University, Kurashiki, Okayama 710-0046, Japan*
[b] *Department of Chemistry, Kyoto University of Education, Fushimi-ku, Kyoto 612-8522, Japan*
[c] *Graduate School of Health and Welfare Science, Okayama Prefectural University, Soja, Okayama 719-1197, Japan*

Received 27 November 2002; accepted 6 January 2003

Dedicated to Professor Dr. Kenji Soda in honor of his 70th birthday

Abstract

The amino acid sequence of the earthworm-serine protease, isozyme C, which shows not only elastase-like activity but also trypsin-like activity, was determined. The catalytic triad of the trypsin family, His, Asp, Ser, was conserved in isozyme C, but the primary substrate determinant of trypsin, Asp, was missing in isozyme C, the same as in elastase. One of the two Gly at the entrance of the substrate-binding pocket of trypsin was replaced by Val as in elastase, however, the other was replaced by Ser whereas Thr is present in elastase. Furthermore, isozyme C also showed esterase-like activity, which was applicable for the synthesis of useful substances.

Keywords: Earthworm; Serine protease; Elastase; Esterase-like activity

1. Introduction

Proteolytic enzymes are ubiquitous in occurrence, being found in all living organisms, and essential for cell growth and differentiation. There is renewed interest in the study of proteolytic enzymes, mainly due to the recognition that these enzymes not only play an important role in the cellular metabolic processes but have also gained considerable attention in the industrial community as a biocatalyst for the preparation of useful compounds [1–4]. Serine proteases are members of proteolytic enzymes which require the amino acid serine at their active site and appear to use the same mechanism for catalysis [5]. They are widely distributed in nature and have diverse substrate specificity. Trypsin, chymotrypsin, and elastase are major serine proteases which catalyze the hydrolysis of peptide bonds on the carboxyl side of the Lys and Arg residues, aromatic amino acid residues, and uncharged nonaromatic amino acid residues [6–8]. Alignment analysis of the amino acid sequences and three-dimensional structure analysis indicated that the catalytic activity is mediated by the conserved His, Asp, and Ser residues and the substrate specificity is determined by the different sets of amino acid residues which are composed of the substrate-binding pocket [5,9,10]. Amino acid mutagenesis showed the possibility of limiting or altering the substrate specificity of the native enzymes, however, the activity of the mutated enzymes was low compared to that of the

* Corresponding author. Tel.: +81-86-424-1661; fax: +81-86-434-1249.
E-mail address: manabus@rib.okayama-u.ac.jp (M. Sugimoto).

1381-1177/$ – see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S1381-1177(03)00105-X

native enzymes [11–13]. These results suggest that a number of changes in the amino acid residue in the structure of the binding pocket should be required to alter the substrate specificity and it is advantageous to compare the protein structure of the native enzymes having various substrate specificities in order to clarify the relationship between the substrate specificity and protein structure.

Earthworms secrete serine proteases that degrade a wide variety of proteins including fibrin, collagen, and elastin, and the therapeutic effects of the earthworm-fibrinolytic enzymes have recently been reported [14–16]. To date, we have characterized alkaline serine proteases with a potent fibrinolytic activity from the earthworm, *Lumbricus rubellus*, and demonstrated that the proteases are composed of six isozymes (isozymes A, B, C, D, E, and F) [17–22]. Isozymes A, B, D, E, and F represent trypsin- and chymotrypsin-like activities, however, isozyme C shows not only elastase-like activity but also trypsin-like activity.

Though a large number of serine protease structures have been reported, the structure of the serine proteases which show multiple substrate specificity like isozyme C have not been described. Furthermore, little information is known about the application of the earthworm protease for the organic synthesis. It is of great interest to clarify the molecular basis of these particular properties of isozyme C as a new biocatalyst. We report here the amino acid sequence of isozyme C in comparison to those of trypsin and elastase, and its esterase-like activity which is applicable for the synthesis of useful substances such as optically active compounds.

2. Materials and methods

2.1. Purification of isozyme C

The earthworm-serine protease, isozyme C, was purified to homogeneity according to the previously described procedures [17,20].

2.2. Amino acid sequencing

The purified isozyme C, inactivated by 4-(2-aminoethyl)-benzensulfonate fluoride, was aminoethylated [23] and pyridylethylated [24]. The former *S*-alkylated protein was digested by endoproteinase Lys-C (Roche Diagnostics, Switzerland) and the latter by endoproteinase Asp-N (Roche Diagnostics). The peptides obtained by proteolytic digestion were separated by reverse-phase HPLC on a Sephasil C18 column (Amersham Pharmacia Biotech, Japan) or Wakopak Navi C22-5 column (Wako, Japan) by monitoring at 214 nm. The sequences of the purified peptides were analyzed by the automated Edman degradation using a model 476A protein sequencer (Applied Biosystems, USA).

2.3. Enzymatic synthesis of aromatic acid vinyl ester

To 20 ml of EtOH was added the acetylated aromatic acid vinyl ester (**1**, 10 mmol), isozyme C (20 mg), and molecular sieves 4A (250 mg, heat-dried) [25]. The reaction mixture was stirred for 48 h at 37 °C. The mixture was filtered and washed with EtOH, then concentrated under reduced pressure. The crude mixture was purified by silica-gel chromatography (*n*-hexane: EtOAc) to afford a white powder.

3. Results and discussion

3.1. Structure and function of isozyme C

The amino acid sequence of the earthworm-serine protease, isozyme C, obtained by arranging the amino acid sequences of 18 peptide fragments, is shown in Fig. 1. The sequence data appear in the SWISS-PROT data bank under the accession number P83298 (N. Nakajima and M. Sugimoto). Isozyme C comprises 242 amino acid residues with a calculated molecular mass of 24,836 Da which is very close to that of the native enzyme, 24,664 Da, estimated by ionspray-MS [17].

The sequence similarities of isozyme C with porcine elastase [26] and bovine trypsin [27] were 33 and 31%, respectively. Fig. 2 illustrates the alignment of isozyme C, the active form of trypsin and elastase. The catalytic triad, His57*, Asp102*, and Ser195* (asterisks show chymotrypsinogen numbering [28]) of trypsin corresponding to His44, Asp93, and Ser191 of isozyme C, respectively, showing that isozyme C belongs to the

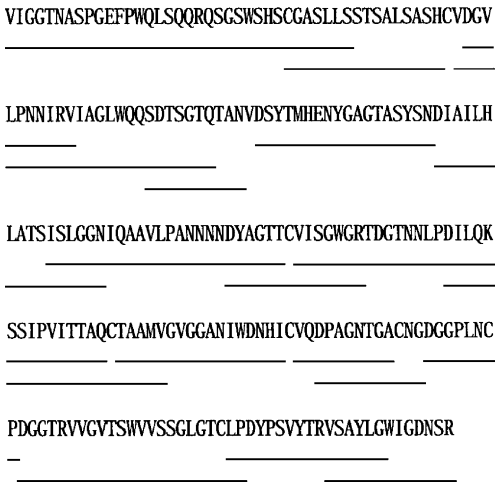


Fig. 1. Amino acid sequence of the earthworm protease (isozyme C) from *L. rubellus*. The lines indicate the N-terminal and internal amino acid sequences of the protein by Edman degradation.

trypsin family. A comparison of the substrate-binding pocket structure of elastase and trypsin suggests that Gly216* and Gly226* of trypsin, which lie at the entrance of the substrate-binding pocket, are replaced by Val209 and Thr221 in elastase [29,30]. These extra side chains in elastase interfere with the entry of bulky aromatic side chains into the substrate-binding pocket. Furthermore, Asp189* of trypsin, which lies at the bottom of the substrate-binding pocket, forms hydrogen bonds with the Lys and Arg residues of the substrate, resulting in the specific binding of trypsin to positively charged amino acid residues, whereas the corresponding residue is replaced by Ser182 in elastase.

These results restrict the substrate specificity of elastase to uncharged nonaromatic amino acids residues. From the alignment analysis, the amino acid residue of isozyme C corresponding to Asp189* is Gly185, and those corresponding to Gly216* and Gly226* were Val212 and Ser226, both of which have extra side chains. These structural similarities of isozyme C with elastase should show elastase-like activity. The unique characteristic of isozyme C is to show the trypsin-like activity of which the specific activity is 33% of the elastase-like activity. At the entrance of the substrate-binding pocket, Val212 and Ser226, which have large side chains, lie in isozyme C, however, Ser226 is replaced by Thr221 in elastase. The difference in an extra methyl group between Ser226 and Thr221 might cause a space at the entrance of the substrate-binding pocket in isozyme C that might allow entry of Lys and Arg into the pocket. The replacement of Asp189* by Gly182 in isozyme C should decrease the activity for Lys or Arg residues of peptide bonds, resulting in the substrate specificity of isozyme C which shows multiple substrate specificity [17,19,20].

3.2. Application of esterase-like activity of isozyme C

Isozyme C, potentially degraded various proteins such as elastin, collagen, and fibrin, and some peptides

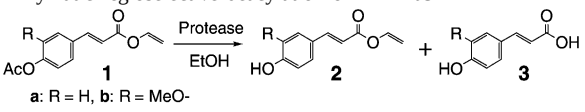
```
          20        30        40        50        60        70        80        90        100
BT    IVGGYTCGANTVPYQVSL---N-S-GYHFCGGSLINSQWVVSAAHCYKSGIQVRL----G-EDNINVVEGNEQFISASKSIVHPSYNS--NTLN
C     VIGGTNASPGEFPWQLSQQRQSGS-WSHSCGASLLSSTSALSASHCVDGVLPNNIRVIAG-LWQQSDTSGT-QTANVDSYTMHENYGAGTASYS   91
EL    VVGGTEAQRNSWPSQISLQYRSGSSWAHTCGGTLIRQNWVMTAAHCVDRELTFRVVVGEH-NLNQNDGT-E-QYVGVQKIVVHPYWNTDDVAAG   91

          110       120       130       140       150       160       170       180
BT    NDIMLIKLKSAASLNSRVASISLP-TSC-ASAGTQCLISGWGNTKSSGTSYPDVLKCLKAPILSDSSCKSAYP---GQ-ITSNMFCAGYLEG--
C     NDIAILHLATSISLGGNIQAAVLPANNNNDYAGTTCVISGWGRTD-GTNNLPDILQKSSIPVITTAQCTAAMVGVGGANIWDNHICVQDPAGNT  184
EL    YDIALLRLAQSVTLNSYVQLGVLPRAGTILANNSPCYITGWGLTR-TNGQLAQTLQQAYLPTVDYAISSSSSY--WGSTVKNSMVCAGG-DGVR  181

          190       200       210       220       230       240  245
            *                   *         *
BT    GKDSCQGDSGGPVVCSGK----LQGIVSWGS----GCAQKNKPGVYTKVCNYVSWIKQTIASN
C     --GACNGDSGGPLNCPDGGTR-VVGVTSWVVSSGLGRCLPDYPSVYTRVSAYLGWIGDNSR--  242
EL    --SGCQGDSGGPLHCLVNGQYAVHGVTSFVS--RLGCNVTRKPTVFTRVSAYISWINNVIASN  240
```

Fig. 2. Sequence alignment of isozyme C (C), bovine trypsin (BT), and porcine elastase (EL). The numbering shown above the sequence of BT is based on that according to the chymotrypsinogen numbering system [28]. The amino acid residues of the catalytic triad are represented by reversal letters. The primary substrate specificity determinant of trypsin, Asp189* and Gly216* and Gly226* at the entrance of the substrate-binding pocket is indicated by asterisks.

Table 1
Enzymatic regioselective deacylation of **1** in EtOH[a]



a: R = H, **b**: R = MeO-

Protease	**1a**		**1b**	
	2a (%)[b]	**3a** (%)[b]	**2b** (%)[b]	**3b** (%)[c]
Isozyme C	55	0	48	0

[a] Incubated in EtOH at 37 °C for 48 h.
[b] Isolated yield.
[c] The isolated yield of the by-product (also hydrolysis of vinyl ester moiety).

such as β-amyloid [19,20]. The isozyme preferentially cleaved the peptide bonds consisting of neutral or hydrophobic amino acids, besides the hydrolysis of the bonds on the carboxyl side of the Lys and Arg residues [19].

To further investigate the catalytic functions of the protease, the ester-bond hydrolyzing activity of isozyme C was determined. As a result, the hydrolyzing activity for ethyl acetate, glucose acetate, *p*-nitrophenyl acetate, *p*-nitrophenyl butyrate, and tosyl-L-Lys-*O*Me were found to be the same as isozyme A [20]. Isozyme C was also applicable for the regioselective deacylation of **1** [25]. The hydrolysis by isozyme C exclusively produced the deacetylated vinyl ester **2** (aromatic acid vinyl ester) as shown in Table 1. Isozyme A and B also showed the similar esterase activity. Recently, the application of the earthworm-protease isozymes for the stereoselective hydrolysis of a wide varieties of ester compounds such as the *N*-Boc and *N*-Cbz proline esters [34] was reported.

Therefore, isozyme C showed not only proteolytic activity (trypsin- and elastase-like activities) but also esterase-like activity, which were considered to be applicable for the synthesis of useful substances such as optically active compounds [31,32], as well as the other earthworm-protease isozymes such as isozyme A [33].

Acknowledgements

This work was supported by a Grant-in-Aid for Scientific Research (B)(1) No.10556023 (1998–2000) from the Ministry of Education, Science, Sports, and Culture, Japan. We would like to acknowledge the financial supports of The Asahi Glass Foundation (2000–2001 and 2003–2005), Japan.

References

[1] R. Gupta, Q.K. Beg, P. Lorenz, Appl. Microbiol. Biotechnol. 15 (2002) 59.
[2] A.V. Bacheva, F.M. Plieva, E.N. Lysogorskaya, I.Yu. Filippova, Y. Lozinsky, Bioorg. Med. Chem. Lett. 11 (2001) 1005.
[3] J.U. Klein, A. Prykhodzka, V. Cerovsky, J. Pept. Sci. 6 (2000) 541.
[4] T. Nagashima, A. Watanabe, H. Kise, Enzyme Microb. Technol. 14 (1992) 842.
[5] J. Kraut, Annu. Rev. Biochem. 46 (1977) 331.
[6] C.S. Craik, Q.-L. Choo, G.H. Swift, C. Quinto, R.J. MacDonald, W.J. Rutter, J. Biol. Chem. 259 (1984) 14255.
[7] P.E. Wilcox, Meth. Enzymol. 19 (1970) 64.
[8] E.G. Del Mar, G. Largman, J.W. Brodric, M. Fasset, M.C. Geokas, Biochemistry 19 (1980) 468.
[9] T.A. Steitz, R.G. Shulman, Annu. Rev. Biophys. Bioeng. 11 (1982) 419.
[10] K. Wang, L. Gan, I. Lee, L. Hood, Biochem. J. 307 (1995) 471.
[11] L. Graf, C.S. Craik, A. Patthy, S. Roczniak, R.J. Fletterick, W.J. Rutter, Biochemistry 26 (1987) 2616.
[12] L. Graf, A. Jancso, L. Szilagyi, G. Hegyi, K. Pinter, G. Naray-Szabo, J. Hepp, K. Medzihradszky, W.J. Rutter, Proc. Natl. Acad. Sci. U.S.A. 85 (1988) 4961.
[13] L.B. Evnin, J.R. Vasquez, C.S. Craik, Proc. Natl. Acad. Sci. U.S.A. 87 (1990) 6659.
[14] Y.S. Kim, M.K. Pyo, K.M. Park, B.S. Hahn, K.Y. Yang, H.S. Yun-Choi, Arch. Pharm. Res. 21 (1998) 374.
[15] T.M. Hrzenjak, M. Popovic, L. Tiska-Rudman, Patrol. Oncol. Res. 4 (1998) 206.
[16] Q. Fan, C. Wu, L. Li, R. Fan, C. Wu, Q. Hou, R. He, Biochim. Biophys. Acta 1526 (2001) 286.
[17] N. Nakajima, H. Mihara, H. Sumi, Biosci. Biotechnol. Biochem. 57 (1993) 1726.
[18] N. Nakajima, K. Ishihara, M. Sugimoto, H. Sumi, K. Mikuni, H. Hamada, Biosci. Biotechnol. Biochem. 60 (1996) 293.
[19] N. Nakajima, M. Sugimoto, K. Ishihara, K. Nakamura, H. Hamada, Biosci. Biotechnol. Biochem. 63 (1999) 2031.
[20] N. Nakajima, M. Sugimoto, K. Ishihara, J. Biosci. Bioeng. 90 (2000) 174.
[21] M. Sugimoto, N. Nakajima, Biosci. Biotech. Biochem. 65 (2001) 1575.
[22] N. Nakajima, K. Ishihara, M. Sugimoto, T. Nakahara, H. Tsuji, Biosci. Biotech. Biochem. 66 (2002) 2739.
[23] W.E. Schwartz, P.K. Smith, G.P. Ryer, Anal. Biochem. 106 (1960) 43.
[24] M. Friedman, L.H. Krull, J.F. Cavins, J. Biol. Chem. 245 (1970) 3868.

[25] K. Ishihara, N. Nakajima, T. Itoh, H. Yamaguchi, K. Nakamura, T. Furuya, H. Hamada, J. Mol. Catal. B: Enzym. 76 (1999) 307.
[26] D.M. Shotton, B.S. Lartley, Nature 225 (1970) 802.
[27] I.L. Huerou, C. Wicker, P. Guilloteau, R. Toullec, A. Puigserver, Eur. J. Biochem. 193 (1990) 767.
[28] B.S. Hartley, D.L. Kauffman, Biochem. J. 101 (1966) 229.
[29] D.M. Shotton, H.C. Watson, Nature 225 (1970) 811.
[30] C.S. Craik, C. Largman, T. Fletcher, S. Roczniak, P.J. Barr, R. Fletterick, W.J. Rutter, Science 19 (1985) 291.
[31] T. Suzuki, T. Nakayama, T. Kurihara, T. Nishino, N. Esaki, J. Mol. Catal. B: Enzym. 16 (2002) 255.
[32] G.-J. Kim, G.-S. Choi, J.-Y. Kim, J.-B. Lee, D.-H. Jo, Y.-W. Ryu, J. Mol. Catal. B: Enzym. 17 (2002) 29.
[33] N. Nakajima, M. Sugimoto, K. Ishihara, J. Mol. Catal. B: Enzym. 23 (2003) 191.
[34] M. Kurokawa, T. Shindo, M. Suzuki, N. Nakajima, K. Ishihara, T. Sufai, Tetrahedron: asymmetry (2003), in press.