The primary structure of carboxypeptidase S3 from *Penicillium janthinellum* IBT 3991

Ib Svendsen[a,*], Estera Szwajcer Day[b]

[a]*Carlsberg Laboratory, Department of Chemistry, 10 Gl. Carlsbergvej, DK-2500 Copenhagen, Denmark*
[b]*Carlsberg Research Laboratory, 10 Gl. Carlsbergvej, DK-2500 Copenhagen, Denmark*

Received 24 June 1995

Abstract The complete amino acid sequence of penicillopeptidase S3, a serine carboxypeptidase isolated from *Penicillium janthinellum* IBT 3991, has been determined. The enzyme consists of 481 amino acids arranged in a single polypeptide chain. Six glycosylation sites were established in positions 41, 218, 256, 326, 384 and 392. The molecule contains six cysteinyl residues among which disulfide bridges was established between Cys-71–Cys-333 and Cys-233–Cys-289. Carboxypeptidase S3 is homologous to carboxypeptidase PEPF (or carboxypeptidase I) from *Aspergillus niger* (67% identical positions). It is proposed that these enzymes form a separate sub-family among the serine carboxypeptidases.

Key words: Carboxypeptidase; Serine protease; Amino acid sequence; Homology

1. Introduction

The serine carboxypeptidases comprise a group of exopeptidases which use the same catalytic mechanism as the serine endopeptidases. A large number of serine carboxypeptidases have now been characterized with respect to amino acid sequence and they appear to belong to either of two sub-groups [1]: a single chain enzyme of approximately 420 amino acid residues (e.g. carboxypeptidase Y [2,3], carboxypeptidase S1 [4] and carboxypeptidase M3 [5] or enzymes containing two chains with approximately 260 and 160 amino acid residues, respectively (e.g. carboxypeptidases M1 [6] and M2 [7] and carboxypeptidase W2 [8]). These enzymes are synthesized as a single chain and then processed.

The great interest in carboxypeptidases stems from their utility in peptide syntheses as well as in amino acid sequencing [1]. Recently a carboxypeptidase from *Penicillium janthinellum*, CPD-S3, has been purified and its enzymatic characteristics determined [9]. CPD-S3 favours substrates of L-configuration with basic amino acid residues in either P_1 or $P_{1'}$. The present paper describes the elucidation of the amino acid sequence of mature CPD-S3, localization of its glycosylation sites and disulfide bridges.

2. Experimental

Penicillocarboxypeptidase S3 (CPD-S3) was isolated as previously described [9]. Amino acid sequencing was performed using an Applied Biosystems gas phase or pulsed liquid sequencer, models 470A and 477A, respectively. Reduction and vinylpyridination, cleavage with cyanogen bromide, hydroxylamine, EndoLysC protease (Boehringer), Glu-C protease isolated from *B. licheniformis* [10], clostripain and Asp-N protease (Boehringer) were carried out as previously described [11]. Separation of peptides was performed by HPLC on a Vydac C18 column using a linear gradient in acetonitril from 10 to 60% in 0.1% trifluoroacetic acid.

3. Results and discussion

CPD-S3 consists of a single polypeptide chain of 481 amino acid residues (Fig. 1). The complete sequence was obtained from N-terminal sequencing of the reduced and vinylpyridinated enzyme together with sequencing of peptides obtained by cleavage with cyanogen bromide, Glu-C protease, Lys-C protease, Asp-N protease, clostripain, cyanogen bromide and hy-

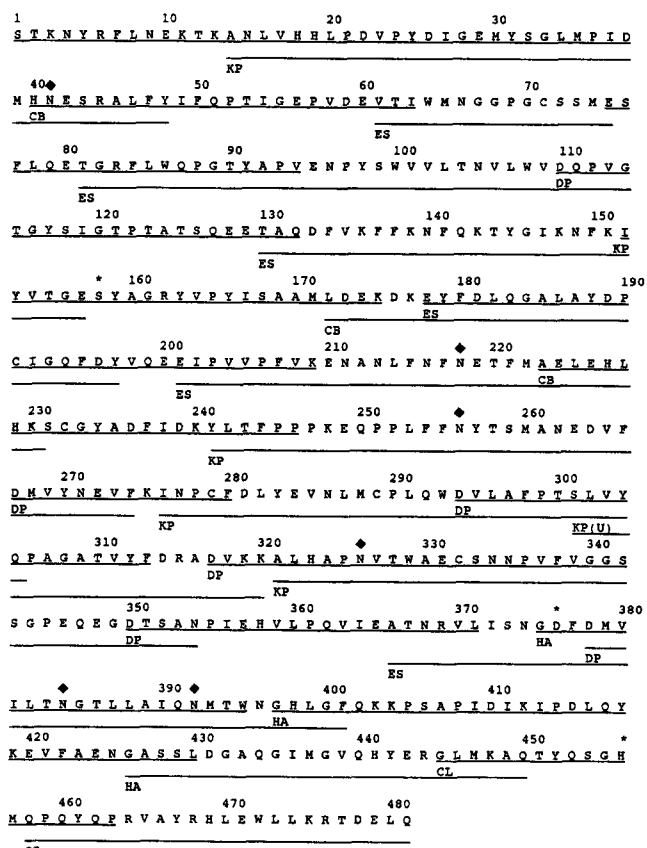

Fig. 1. The amino acid sequence of carboxypeptidase S3 from *Penicillium janthinellum* obtained by N-terminal sequencing and sequencing of peptides obtained from cleavage with cyanogen bromide (CN), hydroxylamine (HA), GluC protease from *B. licheniformis* (ES), LysC protease (KP), AspN protease (DP) and clostripain (CL). KP(U) unspecific cleavage by LysC protease. Diamonds denotes glycosylated Asn residues and * the three residues essential for catalysis.

*Corresponding author. Fax: (45) (33) 27 47 08.

SSDI 0014-5793(95)00840-3

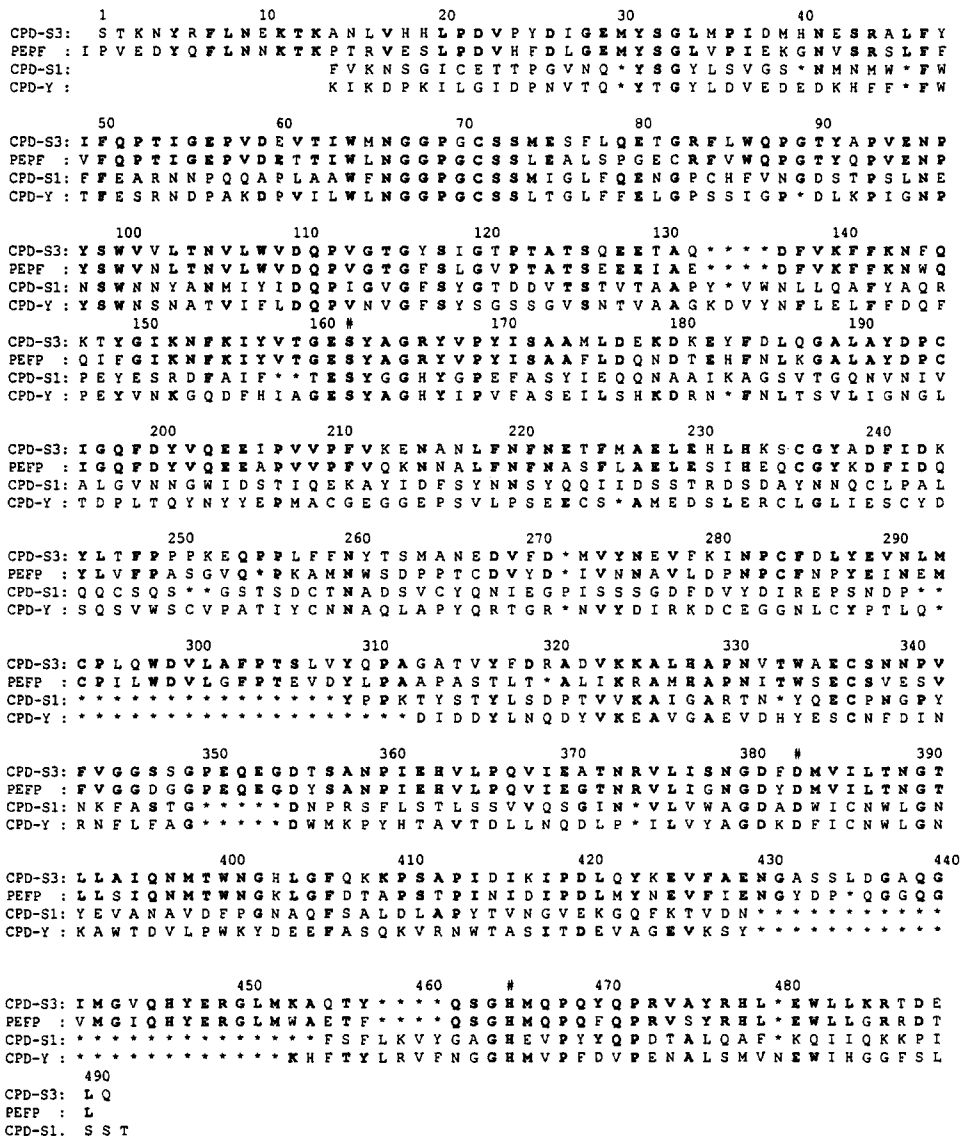

Fig. 2. Comparison of carboxypeptidase CPD-S3, PEPF from *A. niger* [12], CPD-S1 from *Penicillium janthinellum* [10] and CPD-Y from yeast [2,3]. Identical residues are shown in bold letters. *denotes gaps. # shows the three essential amino acids.

droxylamine. The peptides sequenced necessary to demonstrate the overlaps are underlined. All amino acids have been sequenced at least twice.

The molecular mass of the enzyme was calculated from the determined sequence to be 54.011 Da with respect to the protein part of the molecule. However, the enzyme is a glycoprotein with an estimated carbohydrate content of 32% [9]. This gives a mass of approximately 72,000 Da in reasonable agreement with the 62,000 Da determined by SDS-gel electrophoresis [9].

The six glycosylation sites were indirectly determined to be at positions 41, 218, 256, 326, 384 and 392 (Fig. 1). These positions were always empty when sequenced, they are located two residues N-terminal to Ser or Thr and when compared to the homologous carboxypeptidase, PEPF, from Aspergillus niger [12] (Fig. 2), that sequence (translated from the corresponding nucleotide sequence) always has an Asn in the corresponding positions (and one additional site at position 179 (Fig. 2)).

CPD-S3 contains 6 Cys residues in positions 71, 191, 232, 279, 289 and 332 (Fig. 1). In an attempt to determine the disulfide bonds, the native enzyme was digested with Glu-C protease and the resulting peptides separated by HPLC. It was found that the peptides Val-61–Glu-75/Cys-332–Glu-346 (Fig. 1) eluted as one sharp peak as did the peptides His-228–Glu-250/Val-274–Glu-285 which strongly indicates that disulfide bonds are formed between Cys-71–Cys-333 and Cys-233–Cys-279. The postulated disulfide bond between Cys-71 and Cys-333 is equivalent to the corresponding established disulfide bond in CPD-Y. A third disulfide bond between Cys-191 and Cys-289 is a likely possibility although no verification has been obtained.

Fig. 2 compares the amino acid sequences of four mature serine carboxypeptidases: CPD-S1 and CPD-S3 from *Penicil-*

lium janthinellum established by amino acid sequencing (this paper and [4]), CPD-Y from yeast by amino acid sequencing [2] and nucleotide sequencing [3] and PEPF from *Aspergillus niger* by nucleotide sequencing [12]. PEPF is believed to be the same enzyme as CPD-I characterized by Dal Degan et al. and for which the 19 N-terminal amino acid residues have been determined [13]. Within the 19 N-terminal amino acids of CPD-I and PEPF four differneces were noticed. Since these cannot be due to amino acid sequencing errors they probably reflect strain differences.

PEPF and CPD-S3 are very closely related: 67% identical positions distributed throughout the sequences. When CPD-S3 is compared to CPD-S1 and CPD-Y only 15% and 17% identity is observed. However, the identical amino acid is very unevenly distributed centering around the three amino acids essential for catalysis: Gly-160–Pro-169, Gly-380–Asp-384 and Ser-463–Pro-472 (Fig. 2) covering the Ser, Asp and His, respectively, of the catalytic triad. Furthermore, the segment Trp-64–Ser-74 is almost identical in all four enzymes. These observations indicate that all four enzymes have a common origin from which they have diverged. Although all of them, according to the present classification, belong to the sub-family characterized by a single polypeptide chain [1], CPD-S3 and PEPF (CPD-I) are much longer than e.g. CPD-Y (481 amino acid residues versus 421) distributed as an extension of the N-terminal with 13–14 residues and probably 2 extra internal loops as suggested in the comparison in Fig. 2. The three-dimensional structure is naturally necessary for this suggestion to be confirmed. We therefore propose that CPD-S3 and PEPF (CPD-I) be classified as a separate sub-family.

Acknowledgements: Mmes Bodil Corneliussen and Lone Sørensen are thanked for excellent technical assistance.

References

[1] Breddam, K. (1987) Carlsberg Res. Commun. 51, 83–128.
[2] Svendsen, I., Martin, B., Wisvanatha, T. and Johansen, J.T. (1982) Carlsberg Res. Commun 47, 15–28.
[3] Valls, L.A., Hunter, C.P., Rothmann, J.H. and Stevens, T.H. (1987) Cell 48, 887–897.
[4] Svendsen, I., Hofmann, T., Endrizzi, J., Remington, S.J. and Breddam, K. (1993) FEBS Lett. 333, 39–43.
[5] Sørensen, S.B., Svendsen, I. and Breddam, K. (1989) Carlsberg Res. Commun. 54, 193–202.
[6] Sørensen, S.B., Breddam, K. and Svendsen, I. (1986) Carlsberg Res. Commun. 51, 475–485.
[7] Sørensen, S.B., Svendsen, I. and Breddam, K. (1987) Carlsberg Res. Commun. 52, 285–296.
[8] Breddam, K., Sørensen, S.B. and Svendsen, I. (1987) Carlsberg Res. Commun. 52, 297–312.
[9] Dey, E.S. and Aasmul-Olsen, (1993) Enzyme Microb. Technol. 15, 1042–1046.
[10] Svendsen, I., Jensen, M.R. and Breddam, K. (1991) FEBS Lett. 292, 165–167.
[11] Ballance, G.M. and Svendsen, I. (1988) Carlsberg Res. Commun. 53, 411–420.
[12] Homberg, J.P.T.W., van der Jarai, G., Buxton, F.P. and Visser, J. (1994) Gene 151, 73–79.
[13] Degan, F.D., Ribadeau-Dumas, B. and Breddam, K. (1992) Appl. Environ. Microbiol. 58, 2144–2152.