Volume 291, number 1, 75–78 FEBS 10236 October 1991

ADONIS 0014579391009463

Molecular cloning and primary structure of *Thermoactinomyces vulgaris* carboxypeptidase T

A metalloenzyme endowed with dual substrate specificity

Sergey V. Smulevitch, Andrey L. Osterman, Olga V. Galperina, Mikhail V. Matz, Olga P. Zagnitko, Rafail M. Kadyrov, Iraida A. Tsaplina, Nikolai V. Grishin, Galina G. Chestukhina and Valentin M. Stepanov

Institute of Microbial Genetics, Lab. of Protein Chemistry, 1st Dorozhny pr. 1, Moscow, 113545, USSR

Received 17 July 1991

A gene coding for an extracellular Zn-carboxypeptidase of *Thermoactinomyces vulgaris* has been cloned and sequenced (EMBL X56901). This enzyme named carboxypeptidase T reveals simultaneously both types of substrate specificity characteristic of mammalian carboxypeptidases A and B. The carboxypeptidase T gene is primarily expressed in *E. coli* as a non-active preproenzyme with an additional 98 amino acid residues at the N-terminus. Primary structure alignment of mature carboxypeptidase T and mammalian metallocarboxypeptidases demonstrated 25–30% overall identity but a full preservation of presumed catalytically important residues. These observations imply a basic uniformity of the general catalytic mechanism for enzymes of that class produced by evolutionarily remote organisms.

Carboxypeptidase T; Extracellular metalloenzyme; Preproenzyme; Primary structure; Genomic library; *Thermoactinomyces vulgaris*

1. INTRODUCTION

General principles of catalytic action of Zn-carboxypeptidases — classical objects of enzymology — still remain to a great extent controversial. Structure–function relationship studies were being performed mostly on bovine pancreatic carboxypeptidase A as a model. In the last few years a remarkable insight in this field was gained by studies of new metallocarboxypeptidases isolated from various organisms and by application of new experimental approaches based on recombinant DNA technology.

Earlier we reported [1] purification and preliminary characterization of an extracellular Zn-carboxypeptidase of *Thermoactinomyces vulgaris* (named carboxypeptidase T), a remote analog of mammalian metallocarboxypeptidases. Carboxypeptidase T shows an unusual substrate specificity, combining features characteristic of both pancreatic carboxypeptidases A and B. 3A resolution X-ray data revealed an overall structure similarity between carboxypeptidase T and bovine carboxypeptidase A [2]. These traits of carboxypeptidase T make it a promising object for further elucidation of the catalytic action and substrate specificity structural basis common to enzymes of that class.

Here we present cloning, sequence determination and preliminary analysis of functionally relevant features of the carboxypeptidase T coding gene (*cpT* gene).

Abbreviations: IPTG, isopropyl β-D-thiogalactopyranozide; X-Gal, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranozide

Correspondence address: A.L. Osterman, Institute of Microbial Genetics, Lab. of Protein Chemistry, 1st Dorozhny pr. 1, Moscow, 113545, USSR. Fax: (7) (095) 315 0501.

2. METHODS

2.1. CpT *gene cloning*

A genomic library of *Thermoactinomyces vulgaris* carboxypeptidase T-producing strain was constructed in phasmid vector lambda-pSL5 after partial *Eco*RI digestion of total DNA, following the described procedure [3]. Replica-screening of the library in a phage form was performed with a ^{32}P-labeled degenerated oligonucleotide 20-mer probe, 5′-CA(T/C)AA(T/C)TA(T/C)AA(T/C)GA(A/G)ATGGT-3′. The 1.9 kb *Kpn*I–*Eco*RI segment hybridizing with the probe was mapped in a 10 kb insert of the selected recombinant phasmid clone and subcloned into m13mp18 and mp19 vectors (Pharmacia). Localization and orientation of the *cpT* gene in the segment was determined by the PCR-technique with the same 20-mer and standard sequencing primer.

2.2. *Identification of the* cpT *gene expression product*

IPTG-induced *E. coli* TG-1 cells harboring the recombinant pOZ119 plasmid (1.9 kb *Kpn*I–*Eco*RI fragment cloned in the pUC119 vector (Pharmacia)) were analyzed by Western blotting using an antiserum raised in rabbit against pure carboxypeptidase T. A specific carboxypeptidase activity in crude extract was estimated with a colored substrate DNP-Gly-Gly-Arg as described [1], with and without activation of a precursor by simultaneous subtilisin (Serva) treatment.

2.3. *Sequence analysis*

CpT gene sequencing was performed by the standard chain termination protocol of Sanger et al. [4] and its Sequenase version [5]. Both strands of the 1.9 kb segment were sequenced using the standard

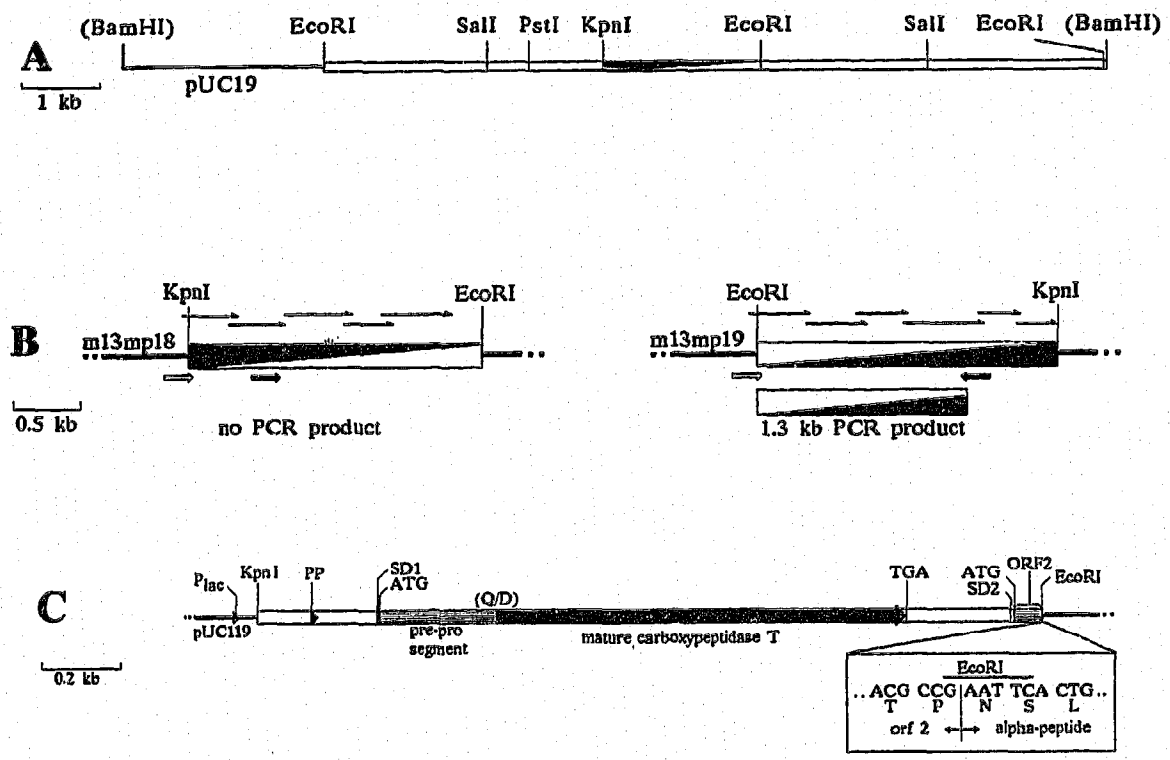

Fig. 1. Strategy of *cpT* gene cloning, localization and sequencing. A. Restrictase map of a 10-kb DNA insert excised by flanking *Bam*HI sites from the selected recombinant phasmid clone together with the pUC19 plasmid which is a constituent part of the phasmid lambda pSL5-vector [3]. The 1.9 kb *Kpn*I–*Eco*RI fragment hybridizing with the gene-specific oligonucleotide probe is marked. B. PCR-based localization of the *cpT* gene within the 1.9 kb *Kpn*I–*Eco*RI fragment. Small arrows below indicate the way of annealing of *cpT*-specific (dark) and standard sequencing (blank) primers. Nucleotide sequence determination of both strands of the fragment is illustrated by long arrows above corresponding to sequencing runs from standard and a set of sequence-derived primers. C. Structural organization of the *cpT* gene and flanking regions (pOZ119 plasmid). The position of open reading frames (ORF1 and ORF2) with corresponding Shine–Dealgarno sequences (SD1 and SD2), start (ATG) and stop (TGA) codons, activation peptide cleavage site (Q/D), the *cpT* gene potential promoter region (PP), as well as the pUC119-vector encoded *lacZ* promoter (P_{lac}) are indicated. An ORF2-*lacZ* α-peptide junction region is zoomed in a box.

sequencing primer and a corresponding set of custom primers. Sequence data were submitted to the EMBL Data Library, accession number X56901.

3. RESULTS AND DISCUSSION

We have found phasmid vector lambda-pSL5 designed by Yankovsky et al. [3] to be an effective tool for the construction and screening of genomic libraries (for example see [6]). To clone the carboxypeptidase T gene (*cpT* gene) a genomic library of *T. vulgaris* has been constructed represented by more than 10^4 recombinant phasmid clones with 10–20 kb insertion fragments. Using carboxypeptidase T, the N-terminal sequence-derived [1] oligonucleotide probe with 32-fold degeneracy:

...-His - Asn - Tyr - Asn - Glu - Met - Val -...

5′-CA^T_C AA^T_C TA^T_C AA^T_C GA^A_G ATG GT-3′

for library screening we have isolated a recombinant phasmid carrying a 10 kb DNA-insert. After removal of the phage arms by *Bam*HI digestion and religation we obtained the insert cloned in pUC19 (see Fig. 1A) and analyzed it by restrictase mapping and Southern blotting. PCR-based analysis of a hybridization-positive 1.9 kb subfragment cloned in both orientations in m13 sequencing vectors, as depicted in Fig. 1B, has shown it to contain the presumably full-size *cpT* gene in an orientation from *Kpn*I to *Eco*RI. This has been confirmed by further sequence determination. The complete DNA sequence of the 1.9 kb fragment and the deduced amino acid sequence of preprocarboxypeptidase T are presented in Fig. 2.

The *cpT* gene coding region is presented by the unique large open reading frame (ORF1) of 1269 bp starting from the ATG codon with a typical Shine–Dalgarno sequence (SD1) AGGAGG 13 bp upstream. Due to the lack of information on transcription regulatory signals in *Thermoactinomyces* we cannot identify unambiguously the promoter region of the *cpT* gene. Still near 150 bp uptream from the translation start we could find two overlapping segments:

(137) TAGAAA...(18 bp)...TATATT
(148) TAGAAA...(18 bp)...TAATAT

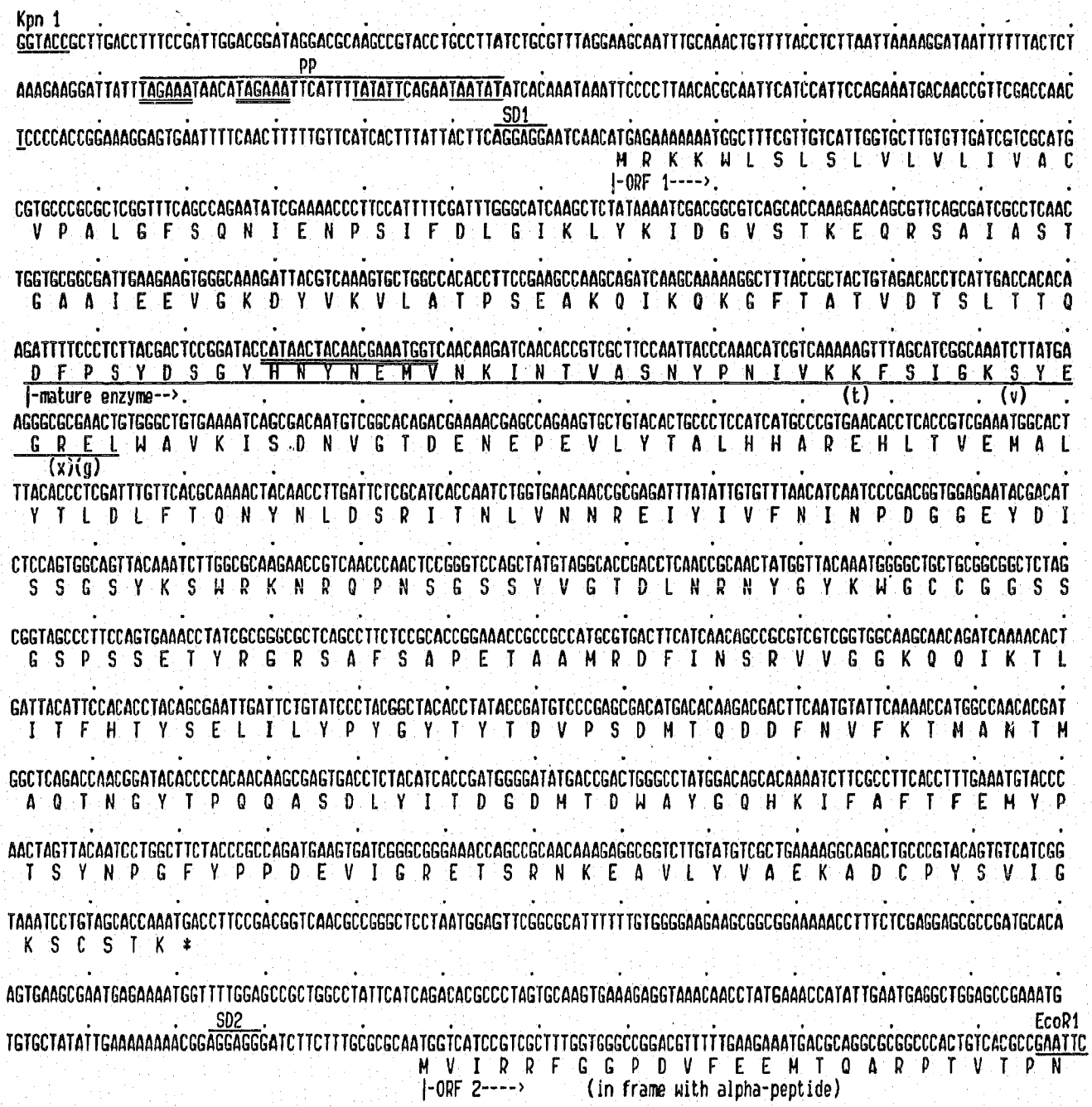

Fig. 2. The nucleotide sequence of the 1.9 kb *Kpn*I–*Eco*RI DNA fragment and the deduced amino acid sequence of preprocarboxypeptidase T. Amino acid residues of mature carboxypeptidase T matching N-terminal sequencing data [1] are underlined, 4 mismatches are shown in parentheses below. Some functionally relevant features are indicated as in Fig. 1C. A segment corresponding to the N-terminal sequence derived probe is double underlined.

preceded by an AT-enriched region, similar to some bacilli promoters. The 3′-flanking region failed to reveal a definite stem-and-loop structure characteristic for procaryotic transcription terminators. At the very 3′-end of the 1.9 kb fragment we have localized the beginning of another open reading frame (ORF2), preceded by an appropriate Shine–Dalgarno sequence (SD2). In the recombinant plasmid pOZ119 (pUC119 with 1.9 kb *Kpn*I–*Eco*RI insert), ORF2 happened to be merged directly in frame with a corresponding segment of the *lacZ* structural gene derived from the vector (see Fig. 1C). A fusion protein composed of amino acids 1–25 of ORF2 and residues 25–127 of the *lacZ* α-peptide was expressed with partial restoration of capacity to complement galactosidase activity. This was manifested by a definite blue color of *E. coli* TG-1 (pOZ119) colonies on IPTG/XGal medium. Therefore the absence of a 3′-palindromic terminator structure together with a demonstrated IPTG-induced run-through transcription of ORF1 and chimeric ORF2-*lacZ* indicate that the *cpT*

gene might be involved into a sort of operon structure and transcribed as part of polycistronic RNA.

ORF1 is coding for a precursor of carboxypeptidase T with a calculated M_r value 47 500 which is in good agreement with a position of an additional immunopositive band on Western blotting. The molecular organization of carboxypeptidase T precursor is quite typical for secretory proteolytic preproenzymes. The putative pre-pro region is composed of 98 amino acid residues, starting with a characteristic signal peptide motif, 3 positively charged residues, followed by a prolonged hydrophobic stretch (see Fig. 2). Nevertheless, solely on the basis of comparison with various microbial signal peptides that are usually 28–31-residues long, we failed to predict the precize position of the signal peptide cleavage site in the carboxypeptidase T precursor. This may be due to the distinctive substrate specificity of a *Thermoactinomyces* signal peptidase.

No specific enzymatic activity of the carboxypeptidase T precursor could be registered in the crude extract of *E. coli* TG-1 (pOZ119) cells. Subtilisin treatment of the same extract in the presence of a carboxypeptidase T substrate — dNP-Gly-Gly-Arg — caused a zymogen activation manifested by substrate cleavage. No activity towards the same substrate appeared in the course of parallel treatment of a control TG-1 (pUC119) extract.

The carboxypeptidase T activation peptide cleavage site, Gln-Asp (see Fig. 2), deduced from the N-terminal sequence of the mature enzyme [1] is quite dissimilar to those of known bacilli zymogens. Still, the amino acid sequence of *Streptomyces griseus* extracellular mature metallocarboxypeptidase [7] — the only microbial analog of carboxypeptidase T (64% sequence identity) studied so far — possesses the N-terminal Asp residue as well. Therefore a specific proteolytic activation pathway might be characteristic for both *Streptomyces* and *Thermoactinomyces*.

No primary structure identity between the activation peptides and 25–30% between the mature enzymes can be registered, comparing carboxypeptidase T to any of the A- and B-type pancreatic metallocarboxypeptidases of mammals. At the same time all catalytically important residues [8] are strictly preserved in carboxypeptidase T as denoted in Table I. This strongly confirms their functional importance for that class of enzymes and implies the basic uniformity of the general catalytic mechanism of carboxypeptidase T and evolutionary remote metallocarboxypeptidases of mammals. Detailed consideration of the carboxypeptidase T structural organization will be published elsewhere.

Table I

Active site forming residues in bovine carboxypeptidase A and *T. vulgaris* carboxypeptidase T

Supposed functional role [8]	Residue in:	
	CpA	CpT
Zn-coordination	His-69	His-69
Zn-coordination	Glu-72	Glu-72
Polarization of the scissile bond carbonyl	Arg-127	Arg-129
Fixation of the substrate carboxylate	Asn-144	Asn-146
Fixation of the substrate carboxylate	Arg-145	Arg-146
Zn-coordination	His-196	His-204
Fixation of the substrate carboxylate	Tyr-248	Tyr-255
H_2O-promotion, H-donor for leaving group	Glu-270	Glu-277

Carboxypeptidase T was shown to possess an unusual 'dual' substrate specificity; an ability to split off with comparable efficiency hydrophobic (like carboxypeptidase A) as well as positively charged (like carboxypeptidase B) C-terminal amino acids [1]. We have found amino acid residues forming the primary specificity binding pocket to be much more variable than those directly responsible for catalysis. Site-directed mutagenesis experiments are in progress in order to establish the precise role of the individual residues governing substrate specificity of carboxypeptidase T.

REFERENCES

[1] Osterman, A.L., Stepanov, V.M., Rudenskaya, G.N., Khodova, O.M., Tsaplina, I.A., Yakovleva, M.B. and Loginova, L.G. (1984) Biochemistry (USSR) 49, 292–301.
[2] Polyakov, K.M., Obmolova, G.V., Kuranova, I.P., Strokopytov, B.V., Vainshtein, B.K., Mosolova, O.V., Stepanov, V.M. and Rudenskaya, G.N. (1989) Molecular Biology (USSR) 23, 266–272.
[3] Yankovsky, N.K., Fonstein, M.Y., Lashina, S.Y. (1989) Gene 81, 203–207.
[4] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[5] Tabor, S. and Richardson, C.C. (1987) Proc. Natl. Acad. Sci. USA 84, 4767–4771.
[6] Osterman, A.L., Karasin, A.I., Zagnitko, O.P., Kaluger, S.V., Chestukhina, G.G., Fonstein, M.Yu., Yankovsky, N.K. and Stepanov, V.M. (1989) Molecular Biology (USSR) 23, 463–472.
[7] Narahashi, Y. (1990) J. Biochem. 107, 879–886.
[8] Christianson, D.W. and Lipscomb, W.N. (1989) Acc. Chem. Res. 22, 62–69.