Molecular Cloning, Nucleotide Sequence, and Expression of a Carboxypeptidase-Encoding Gene from the Archaeobacterium *Sulfolobus solfataricus*

SONIA COLOMBO, GABRIELE TOIETTA, LAURA ZECCA, MARCO VANONI,
AND PAOLO TORTORA*

Dipartimento di Fisiologia e Biochimica Generali, Università degli
Studi di Milano, 20133, Milan, Italy

Received 24 March 1995/Accepted 18 July 1995

Mammalian metallocarboxypeptidases play key roles in major biological processes, such as digestive-protein degradation and specific proteolytic processing. A *Sulfolobus solfataricus* gene (*cpsA*) encoding a recently described zinc carboxypeptidase with an unusually broad substrate specificity was cloned, sequenced, and expressed in *Escherichia coli*. Despite the lack of overall sequence homology with known carboxypeptidases, seven homology blocks, including the Zn-coordinating and catalytic residues, were identified by multiple alignment with carboxypeptidases A, B, and T. *S. solfataricus* carboxypeptidase expressed in *E. coli* was found to be enzymatically active, and both its substrate specificity and thermostability were comparable to those of the purified *S. solfataricus* enzyme.

Mammalian metallocarboxypeptidases play key roles in major biological processes, ranging from digestive-protein degradation, as effected by carboxypeptidases A and B (9, 13), to specific proteolytic processing, as in the maturation of biologically active peptides. This latter is the role, for instance, of carboxypeptidase N and enkephalin convertase (10, 17).

Little is known, however, regarding microbial carboxypeptidases. Recent studies led to the cloning and crystallization of a metallocarboxypeptidase from *Thermoactinomyces vulgaris* (27, 29), which was shown to possess a dual substrate specificity, namely, the capacity to cleave both hydrophobic and basic amino acid residues, combining the activities of carboxypeptidases A and B, respectively. Sequence comparison revealed a sequence similarity of approximately 30% and a three-dimensional structure very similar to those of both of these mammalian enzymes (27, 29). Another microbial carboxypeptidase with broad substrate specificity was also isolated recently from the thermophilic eubacterium *Thermus aquaticus* (16). However, no sequence data are so far available for this enzyme.

The lack of any knowledge regarding archaeobacterial carboxypeptidases led us to purify and characterize one such enzyme from *Sulfolobus solfataricus* (3), an extreme thermoacidophilic archaeobacterium, isolated from volcanic hot springs, that grows optimally at 87°C (7). We found that this protein is a zinc metalloprotease endowed with a unique substrate specificity in that it could release basic, acidic, aromatic, and, to a lesser extent, aliphatic amino acids from artificial substrates. Also, it could withstand temperatures of up to 85°C and some organic solvents at up to 40°C. Furthermore, we found that salt bridges and the zinc ion play major roles in the kinetic thermal stabilization of this molecule (30). Here, we report the molecular cloning and complete nucleotide sequence of the carboxypeptidase-encoding gene. We show that despite the lack of overall sequence homology with known carboxypeptidases,

seven blocks of homology with known metallocarboxypeptidases, spanning the Zn-coordinating histidines, can be identified. We also show that *S. solfataricus* carboxypeptidase expressed in *Escherichia coli* is enzymatically active and that both its substrate specificity and thermostability are comparable to those of the purified *S. solfataricus* enzyme.

MATERIALS AND METHODS

Enzymology. *S. solfataricus* carboxypeptidase was purified as previously described (3). Carboxypeptidase activity was assayed as previously reported (3) by a cadmium-ninhydrin colorimetric method (8). The enzyme was incubated at 60°C in 120 μ l of 0.1 M potassium MES (potassium morpholineethanesulfonic acid) (pH 6.5) with, unless otherwise stated, 2 mM benzyloxycarbonyl-Arg as a substrate. After suitable incubation times, 0.9 ml of cadmium-ninhydrin reagent was added, the mixture was heated for 5 min at 84°C, and the A_{505} was read. One unit of enzyme activity is defined as the amount which hydrolyzes 1 μ mol of substrate per min under the above-mentioned conditions. For thermal stability

FIG. 1. Immunoblot analysis of Y1089/ΔCP14A lysogen, grown either in the absence (–) or in the presence (+) of 1 mM IPTG, probed with either anti-β-galactosidase (A) or anticarboxypeptidase (B) antibodies. (C) Specificity of anticarboxypeptidase antibodies purified on ΔCP14A protein and probed on an *S. solfataricus* crude extract (5 μ g) (lane 7). Lanes 5 and 6, purified carboxypeptidase (50 ng) and an *S. solfataricus* crude extract (5 μ g), respectively, probed with the affinity-purified antibodies used in the immunological screening.

* Corresponding author. Mailing address: Dipartimento di Fisiologia e Biochimica Generali, Sezione di Chimica Biologica, Università degli Studi di Milano, via Celoria 26, I-20133 Milano, Italy. Phone: 39-2-70644508. Fax: 39-2-2362451. Electronic mail address: curti@imiucca.csi.unimi.it.

					Met	Asp	Leu	Val	Glu	Lys	Leu	Lys	Asn	Asp	Val	Arg	Glu	Ile	Glu	15
TTC	GGT	CTT	ATG	GAT	TTA	GTT	GAG	AAG	TTA	AAA	AAT	GAC	GTA	AGA	GAA	ATA	GAG			54
<u>Asp</u>	<u>Trp</u>	<u>Ile</u>	<u>Ile</u>	<u>Gln</u>	<u>Ile</u>	<u>Arg</u>	<u>Arg</u>	<u>Lys</u>	<u>Ile</u>	<u>His</u>	<u>Glu</u>	<u>Tyr</u>	<u>Pro</u>	<u>Glu</u>	<u>Leu</u>	<u>Ser</u>	<u>Tyr</u>	<u>Lys</u>	<u>Glu</u>	35
GAC	TGG	ATA	ATT	CAA	ATT	AGA	AGG	AAA	ATC	CAT	GAG	TAT	CCG	GAA	CTT	TCC	TAC	AAG	GAG	114
Tyr	Asn	Thr	Ser	Lys	Leu	Val	Ala	Glu	Thr	Leu	Arg	Lys	Leu	Gly	Val	Glu	Val	Glu	Glu	55
TAT	AAC	ACC	TCT	AAA	CTA	GTA	GCG	GAA	ACG	TTA	AGG	AAA	TTG	GGA	GTA	GAA	GTG	GAA	GAA	174
Gly	Val	Gly	Leu	Pro	Thr	Ala	Val	Val	Gly	Lys	Ile	Arg	Gly	Ser	Lys	Pro	Gly	Lys	Thr	75
GGC	GTT	GGA	TTA	CCC	ACA	GCA	GTG	GTT	GGT	AAG	ATT	AGG	GGA	AGT	AAA	CCA	GGA	AAG	ACT	234
Val	Ala	Leu	Arg	Ala	Asp	Met	Asp	Ala	Leu	Pro	Val	Glu	Glu	Asn	Thr	Asp	Leu	Glu	Phe	95
GTT	GCT	TTG	AGA	GCT	GAT	ATG	GAT	GCC	CTT	CCG	GTA	GAG	GAG	AAC	ACT	GAT	CTA	GAA	TTT	294
Lys	Ser	Lys	Val	Lys	Gly	Val	Met	His	Ala	Cys	Gly	His	Asp	Thr	His	Val	Ala	Met	Leu	115
AAA	TCC	AAA	GTT	AAG	GGA	GTA	ATG	CAC	GCA	TGT	GGT	CAT	GAT	ACT	CAC	GTA	GCA	ATG	CTC	354
Leu	Gly	Gly	Ala	Tyr	Leu	Leu	Val	Lys	Asn	Lys	Asp	Leu	Ile	Ser	Gly	Glu	Ile	Arg	Leu	135
TTA	GGT	GGA	GCT	TAT	CTG	TTA	GTT	AAG	AAT	AAA	GAT	TTA	ATC	AGT	GGT	GAA	ATT	AGG	TTA	414
Ile	Phe	Gln	Pro	Ala	Glu	Glu	Asp	Gly	Gly	Leu	Gly	Gly	Ala	Lys	Pro	Met	Ile	Glu	Ala	155
ATA	TTC	CAA	CCG	GCA	GAG	GAG	GAT	GGA	GGA	TTA	GGA	GGA	GCA	AAA	CCA	ATG	ATT	GAG	GCT	474
Gly	Val	Met	Asn	Gly	Val	Asp	Tyr	Val	Phe	Gly	Ile	His	Ile	Ser	Ser	Ser	Tyr	Pro	Ser	175
GGA	GTT	ATG	AAC	GGT	GTA	GAT	TAT	GTA	TTT	GGA	ATA	CAT	ATA	TCG	AGT	AGT	TAT	CCT	TCT	534
Gly	Val	Phe	Ala	Thr	Arg	Lys	Gly	Pro	Ile	Met	Ala	Thr	Pro	Asp	Ala	Phe	Lys	Ile	Ile	195
GGA	GTT	TTC	GCA	ACT	AGA	AAA	GGC	CCT	ATA	ATG	GCT	ACG	CCG	GAC	GCA	TTC	AAG	ATA	ATC	594
Val	His	Gly	Lys	Gly	Gly	His	Gly	Ser	Ala	Pro	His	Glu	Thr	Ile	Asp	Pro	Ile	Phe	Ile	215
GTT	CAC	GGG	AAG	GGC	GGT	CAT	GGT	TCT	GCT	CCT	CAT	GAG	ACT	ATT	GAC	CCA	ATT	TTT	ATA	654
Ser	Leu	Gln	Ile	Ala	Asn	Ala	Ile	Tyr	Gly	Ile	Thr	Ala	Arg	Gln	Ile	Asp	Pro	Val	Gln	235
TCC	TTA	CAA	ATA	GCT	AAC	GCA	ATC	TAC	GGC	ATA	ACA	GCA	AGG	CAA	ATT	GAT	CCA	GTT	CAA	714
Pro	Phe	Ile	Ile	Ser	Ile	Thr	Thr	Ile	His	Ser	Gly	Thr	Lys	Asp	Asn	Ile	Ile	Pro	Asp	255
CCC	TTT	ATC	ATA	TCC	ATT	ACT	ACA	ATA	CAT	TCA	GGT	ACA	AAG	GAT	AAC	ATA	ATA	CCA	GAT	774
Asp	Ala	Glu	Met	Gln	Gly	Thr	Ile	Arg	Ser	Leu	Asp	Glu	Asn	Val	Arg	Ser	Lys	Ala	Lys	275
GAT	GCC	GAA	ATG	CAG	GGA	ACA	ATT	AGA	AGT	TTA	GAC	GAG	AAC	GTT	AGA	AGT	AAG	GCT	AAG	834
Asp	Tyr	Met	Arg	Arg	Ile	Val	Ser	Ser	Ile	Cys	Gly	Ile	Tyr	Gly	Ala	Thr	Cys	Glu	Val	295
GAC	TAT	ATG	AGA	AGA	ATA	GTT	TCG	TCA	ATA	TGT	GGA	ATC	TAT	GGT	GCA	ACT	TGT	GAG	GTT	894
Lys	Phe	Met	Glu	Asp	Val	Tyr	Pro	Thr	Thr	Val	Asn	Asn	Pro	Glu	Val	Thr	Asp	Glu	Val	315
AAA	TTC	ATG	GAA	GAC	GTC	TAT	CCA	ACT	ACC	GTA	AAT	AAC	CCT	GAG	GTA	ACT	GAT	GAG	GTA	954
Met	Lys	Ile	Leu	Ser	Ser	Ile	Ser	Thr	Val	Val	Glu	Thr	Glu	Pro	Val	Leu	Gly	Ala	Glu	335
ATG	AAA	ATT	CTA	TCT	TCA	ATA	TCA	ACA	GTT	GTT	GAG	ACA	GAG	CCA	GTG	CTA	GGA	GCG	GAG	1014
Asp	Phe	Ser	Arg	Phe	Leu	Gln	Lys	Ala	Pro	Gly	Thr	Tyr	Phe	Phe	Leu	Gly	Thr	Arg	Asn	355
GAC	TTC	TCC	AGA	TTC	TTA	CAG	AAG	GCT	CCA	GGA	ACG	TAT	TTC	TTT	CTG	GGA	ACC	AGA	AAC	1074
Glu	Lys	Lys	Gly	Cys	Ile	Tyr	Pro	Asn	His	Ser	Ser	Lys	Phe	Cys	Val	Asp	Glu	Asp	Val	375
GAA	AAG	AAA	GGA	TGC	ATA	TAT	CCC	AAT	CAC	AGC	TCT	AAG	TTC	TGT	GTA	GAT	GAG	GAC	GTG	1134
Leu	Lys	Leu	Gly	Ala	Leu	Ala	His	Ala	Leu	Leu	Ala	Val	Lys	Phe	Ser	Asn	Lys	TER		395
CTA	AAA	TTA	GGT	GCC	TTA	GCT	CAC	GCA	TTA	TTG	GCA	GTA	AAG	TTC	AGT	AAT	AAA	TAA	AGG	1194
GCT	TAA	TGC	TAA	ATG	AGT	ATT	GGA	AAT	GGA	GAA	GGA	AAA	GGT	TCT	CAA	TAT	TTT	GAG	GAA	1254
TTC																				1257

FIG. 2. Nucleotide and deduced amino acid sequences of *cpsA*. The in-frame ATG and encoded Met are shown in boldface. The experimentally determined N-terminal sequence is underlined; nonmatching residues are italicized. The nucleotide sequence is numbered in the 5'-to-3' direction beginning with the first *S. solfataricus* nucleotide sequenced.

FIG. 3. Expression of *S. solfataricus* carboxypeptidase in *E. coli*. (A) Immunoblot analysis of *E. coli* BL21(DE3)[pLysE] transformed with plasmid pT7-7 (lane 1) or plasmid pT7-7/CpsA (lane 2). Lane 3, purified *S. solfataricus* carboxypeptidase (250 ng). (B) Specific (bars) and residual (squares) activities of carboxypeptidase after different heat treatments of a BL21(DE3)[pLysE, pT7-7/CpsA] crude extract in 50 mM Tris-HCl (pH 7.4)–5 mM 2-mercaptoethanol.

tests, extracts from BL21(DE3)[pLysE, pT7-7/CP14A] were heated for 3 min at either 70 or 80°C, and denatured proteins were eliminated by centrifugation. Total proteins were assayed with the Pierce reagent. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed by the method of Laemmli (14). Immunoblotting was performed by standard techniques. Antibodies were raised in rabbits according to standard protocols (12). Specific anticarboxypeptidase antibodies were affinity purified essentially as described previously (26) with carboxypeptidase immobilized on nitrocellulose.

Enzymes, radioactive biochemicals, and synthetic oligonucleotides. Restriction enzymes were purchased from Promega (Madison, Wis.) and Boehringer Mannheim (Mannheim, Federal Republic of Germany). Radioactive biochemicals were obtained from Amersham (Amersham, United Kingdom). Oligonucleotides were obtained from Primm s.r.l. (Milan, Italy).

Bacteria strains, genomic library, and cloning vectors. *S. solfataricus* MT-4 (ATCC 49155) was kindly supplied by A. Gambacorta, Istituto per la Chimica di Molecole di Interesse Biologico del Consiglio Nazionale delle Ricerche, Arco Felice, Naples, Italy. *S. solfataricus* cells were grown aerobically in a minimal medium containing 2 g of yeast extract per liter as a carbon source and were collected in stationary phase as described previously (7). The expression library constructed in the λ gt11 vector from *S. solfataricus* genomic DNA was kindly provided by M. Rossi, University of Napoli, Naples, Italy. *E. coli* Y1090 (Δ lacU169 *proA*⁺ Δ lon *araD139* *strA* *supF* *trpC22::Tn10* *mcrA*/pMC9) and Y1089 (Δ lacU169 *proA*⁺ Δ lon *araD139* *strA* *hflA150* *chr::Tn10*/pMC9) were used for plating λ gt11 phages and obtaining lysogens, respectively (31). Plasmid pGEM-3Z (Promega) was used for subcloning and DNA sequencing, while plasmid pT7-7 (1) was used as an expression vector. *E. coli* JM101 [*F'* *traD36* *proA*⁺ *proB*⁺ *lacZ* Δ M15/*supE* *thi* Δ (*lac-proAB*) (23)] and BL21(DE3) [pLysE] [*hsdS* *gal*(λ Its857 *ind1* [*Sam7* *nin5* *lacUV5-T7* gene 1) (28)], respectively, were used as hosts for these plasmids. The carboxypeptidase expression plasmid was constructed as follows. The CP14A fragment containing the *cpsA* gene was obtained by *Eco*RI digestion of plasmid pGEM-3Z/CP14A (see Results and Discussion). After filling of recessed 3' termini, the CP14A fragment was inserted into *Sma*I-cut pT7-7, generating plasmid pT7-7/CP14A. The CP14A fragment contains the complete sequence of the gene encoding carboxypeptidase and about 50 bp downstream of the translational stop codon; the 5' terminus of this fragment contains the translational start codon, so that translation starts at the plasmid ATG.

Isolation and characterization of phage clones. Standard recombinant DNA techniques were performed as described by Sambrook et al. (23). The library was screened with affinity-purified antibodies by an adaptation of the method of Young and Davis (31). The insert of the positive clone was analyzed by restriction mapping, subcloned, and sequenced by the chain termination method with the Pharmacia T7 sequencing kit as suggested by the manufacturer.

Computer analysis. The amino acid sequence of carboxypeptidase was compared with those of other proteins in the SwissProt protein sequence database by using the FASTA program (18) as implemented in the Genetics Computer Group package of programs. Homology analyses with proteins of the carboxypeptidase family were conducted with the Gibbs sampling algorithm (15) as implemented in MACAW 2 (25). The algorithm locates relatively short patterns that are shared by otherwise dissimilar sequences and that reflect structural and functional constraints that arise from the energetic interaction among residues or between residue and ligand, irrespective of evolutionary history.

Nucleotide sequence accession number. The *cpsA* sequence shown in Fig. 2 has been deposited in the EMBL data library under accession number Z48497.

RESULTS AND DISCUSSION

Molecular cloning and nucleotide sequencing of *cpsA*, a carboxypeptidase-encoding gene from *S. solfataricus*. A carboxypeptidase-encoding gene was cloned by using an expression library constructed from *S. solfataricus* genomic DNA in the λ gt11 vector. Phage clones were screened with affinity-purified anticarboxypeptidase antibodies; a positive phage, named λ CP14, was isolated by repeated plaque purification. A lysogen for the positive phage was obtained by using *E. coli* Y1089 as the host, and the proteins expressed were analyzed by immunoblotting with anticarboxypeptidase (Fig. 1B) and anti- β -galactosidase (Fig. 1A) antibodies; an IPTG (isopropyl- β -D-thiogalactopyranoside)-inducible fusion protein of ca. 158 kDa was recognized by both antibodies (Fig. 1, lanes 2 and 4).

Antibodies affinity purified from the protein expressed by the λ CP14A recombinant lysogen recognized a protein comigrating with carboxypeptidase in an *S. solfataricus* crude extract (Fig. 1C, lane 7). Moreover, in crude extracts of *E. coli* Y1089/ λ CP14, the carboxypeptidase specific activity was fourfold higher than it was in the absence of the inducer or in the control strain Y1089/ λ gt11 (data not shown). These findings suggest that the isolated *S. solfataricus* DNA insert in phage

TABLE 1. Substrate specificities of recombinant and native CpsA^a

Substrate	Rate of hydrolysis (%) ^b with:	
	Recombinant CpsA	Native CpsA
CbzArg	100	100
CbzAsp	60	53
CbzGlyGlyPhe	196	237
CbzPhe	11	25
CbzAla	3	1

^a Rates of hydrolysis were determined with the indicated benzyloxycarbonyl (Cbz) substrates as indicated in Materials and Methods. The equivalent of 100 ng of pure enzyme was used in each test.

^b All values are percentages expressed relative to the rate of hydrolysis of CbzArg (100%) for each protein.

B

		Block 1		Block 2	
CpsA	16	DWIIQIRRKIHEYPELSYKEY	36 ... 93	LEFKSKVKGVMHACGHDTHVAMLL	116
CPA	33	VSKLQIGRSYEGRPYVLKFS	53 ... 58	NRPAIWIDLGHSREWITQATGVW	81
CPB	30	VTQSVIGTTFEGRNMYVLKIG	50 ... 55	NKPAIFIDCGFHAREWISPAFCQW	78
CPT	30	VKKFSIGKSYEGRELWAVKIS	50 ... 58	NEPEVLYTALHHAREHLTVEMALY	81
2 ^{ary}	Str	ββββββββ tt βββββββ		βββββββ tt αααααααα	
		Block 3		Block 4	
CpsA	162	DYVFGIHSSSYPSGVFATRKG	183 ... 197	HGKGGHGSAPHETIDP	212
CPA	126	WRKTRSVTSSSLCVGVDANRNW	147 ... 152	GKAGASSSPCSEYHG	167
CPB	123	WRKTRSTMAGSSCLGVRPNRNF	144 ... 149	CEVGASRSPCSEYTCG	164
CPT	128	WRKNRQPNSSSYVGTDLNRNY	149 ... 155	CCGGSSGSPSSEYTRG	170
2 ^{ary}	Str		ααα	β β tt ttβ	
		Block 5		Block 6	
CpsA	238	IISITTIHSGTKDNIIPDDA	257 ... 297	FMEDVYPTTVNNP	309
CPA	189	FKAFLSIHYSQLLLYPYGY	208 ... 243	IITTIYQASGGSI	255
CPB	187	IKAYLTIHYSQMMLYPYSY	206 ... 241	GATTIYPAAGGSD	253
CPT	197	IKTLITFHITYSELILYPYGY	216 ... 250	QASDLYITDGDMT	262
2 ^{ary}	Str	ββββββββ ββββ		ββββt αα	
		Block 7			
CpsA	359	GCIYPNHSSKFCVDE	373		
CPA	256	DWSYNQGIKYSFTFE	270		
CPB	254	DWSYDQGIKYSFTFE	268		
CPT	263	DWAYGQHKIFAFTFE	277		
2 ^{ary}	Str	αααααtt βββββββ			

FIG. 4. Multiple alignment of carboxypeptidase from *S. solfataricus* (CpsA), mammalian carboxypeptidases A (CPA) and B (CPB), and carboxypeptidase from *T. vulgaris* (CPT). Proteins were aligned by using the Gibbs algorithm as implemented in MACAW with the BLOSUM62 substitution matrix. (A) Scheme highlighting the positions of the homology blocks within each protein. (B) Close-up view of homology blocks. Residues involved in zinc chelation are shown in boldface. Residues implicated in catalysis and conserved in CpsA are italicized. Secondary-structure assignments (2^{ary} Str) from the published three-dimensional structure of carboxypeptidase A are shown (α, α helix; β, β-strand; t, turn).

λCP14 encodes a carboxypeptidase. After removal of the phage arms by *KpnI-SacI* digestion, the λCP14 *KpnI-SacI* fragment was subcloned in pGEM-3Z, generating plasmid pGEM-3Z/λCP14. The λCP14 insert was then cleaved into three fragments that were subcloned in pGEM-3Z, generating plasmid pGEM-3Z/CP14A, pGEM-3Z/CP14B, and pGEM-3Z/CP14C. The complete sequence of the CP14A fragment was determined with synthetic oligonucleotides as primers. Within fragment CP14A, an 1,179-bp open reading frame encoding a protein of 393 amino acids was found (Fig. 2). The predicted protein would have a molecular mass of 43,068 Da, in good agreement with the enzyme molecular mass assessed by SDS-PAGE (42 kDa per subunit). One ATG codon (in boldface in Fig. 2) at position 10 was in frame with λgt11 *lacZ*. The amino acid sequence encoded between nucleotides 10 and 102 (underlined in Fig. 2) matched the experimentally determined N-terminal amino acid sequence of pure *S. solfataricus* carboxypeptidase, with four exceptions (shown in italics in Fig. 2). This finding indicates that translation starts at position 10 and confirms the identity of the cloned gene. The observed discrep-

ancy might be due to mistakes in the amino acid sequence determination, to strain variation, or to the presence of different carboxypeptidase isoforms. We propose to call this gene *cpsA*.

As reported for other *S. solfataricus* protein-encoding genes (4–6, 19–21, 24), the codon usage of the *cpsA* gene shows a strong preference towards A and T, reflecting the low G+C content of the DNA from this organism. The overall A+T content of the *cpsA* gene is 60%, rising to 66% when only the third position is considered, with Trp and Met excluded. Only AGG and AGA codons, which are very rare in *E. coli* but common in eukaryotes, are used for arginine, in agreement with data previously reported (4–6).

No putative terminator signal matching the consensus (TTT TTTY) proposed by Reiter et al. (22) was found downstream of the termination codon.

Expression of *S. solfataricus* carboxypeptidase in *E. coli*. *E. coli* carboxypeptidase expression plasmids were constructed as described in Materials and Methods and transformed into strain BL21(DE3)[pLysE]. Carboxypeptidase activity was as-

sayed in *E. coli* BL21(DE3)[pLysE] transformed with plasmid pT7-7/CP14A, encoding carboxypeptidase, and the control plasmid pT7-7. A 42-kDa protein was specifically detected by anticarboxypeptidase immunopurified antibodies (Fig. 3A) in BL21(DE3)[pLysE, pT7-7/CP14A] and not in BL21(DE3)[pLysE, pT7-7]. Accordingly, carboxypeptidase activity was detected in BL21(DE3)[pLysE, pT7-7/CP14A] but not in BL21(DE3)[pLysE, pT7-7]. Assuming that the specific activity of the expressed enzyme is the same as that of the carboxypeptidase from *S. solfataricus*, we calculated that under these conditions active carboxypeptidase accumulated to about 0.3% of total *E. coli* proteins.

In order to better characterize the activity of the recombinant protein, thermal stability and substrate specificity were tested. Under the assay conditions used, no activity was detectable in crude extracts of the control strain BL21(DE3)[pLysE, pT7-7] (data not shown). Figure 3B shows that after 3 min at 70 and 80°C, 100 and 80%, respectively, of the initial activity was retained, with over a 10-fold increase in specific activity. The *Sulfolobus*-produced enzyme has been shown to have a similar thermal stability (reference 30 and data not shown).

E. coli-expressed CpsA showed a broad substrate specificity, as shown by using an enzyme which was partially purified (70% purity) according to a modification (29a) of a published procedure (3). A substrate specificity comparable to that of the purified *S. solfataricus* enzyme was displayed, as demonstrated by the release of basic, aromatic, and acid amino acids from benzyloxycarbonyl-amino acids (Table 1 and reference 3).

Sequence homology. Comparison of the amino acid sequence of the *S. solfataricus* carboxypeptidase with those of proteins in the SwissProt data bank did not show significant similarities to carboxypeptidases. Global alignment programs like FASTA often lack the sensitivity necessary to detect low-level, local homologies. We thus used a recently described algorithm to test for local homologies between the protein encoded by the cloned *S. solfataricus* gene and other carboxypeptidases, namely, mammalian carboxypeptidases A and B and *T. vulgaris* carboxypeptidase T. This analysis identified seven major homology blocks, which are schematically shown as white boxes in Fig. 4A.

Biochemical and genetic analyses have identified several key residues in metallo-carboxypeptidases (see reference 2 for a comprehensive review). In bovine carboxypeptidase A, His-69, Glu-72, and His-196 are involved in zinc chelation. Figure 4B shows that two histidines (His-104 and His-245) are aligned to His-69 and His-196, respectively, of the bovine enzyme. No equivalent of Glu-72 is found; however, Asp-109 might be a viable substitute. Residues involved in catalysis in carboxypeptidase A include Arg-127, Asn-144, Arg-145, Tyr-248, and Glu-270. Homologous residues for the last three have been identified at positions 181, 302, and 373, respectively.

The absence of Arg and Asn at positions 127 and 144 (carboxypeptidase A numbering) is intriguing. Both Arg-127 and Asn-144 have been proposed to act as hydrogen bond donors; the former would act in stabilizing the oxyanion in the transition state, and the latter would cooperate with Arg-145 and Tyr-248 in hydrogen bonding the C-terminal carboxylate of the substrate. Arg-127 has also been proposed to act as the non-metal location of the intact carbonyl just prior to catalysis (reviewed in reference 2). Despite the wealth of structural information available on carboxypeptidase A, few mutagenesis experiments have been designed to directly probe the roles of residues thought to be involved in catalysis and binding. The first of these experiments addressed the role of Tyr-248, which, on the basis of structural and chemical modification experiments, was suggested to be the general acid catalyst required

for peptide bond hydrolysis. The experiments showed that the phenolic hydroxyl of Tyr-248 was not essential for catalysis, although it was participating in substrate binding, since the mutated enzyme had the same k_{cat} as the wild-type enzyme had but a higher K_m for some, but not all, substrates (11). Tyr-163 and Thr-180 are found aligned with Arg-127 and Asn-144 and might functionally replace these residues. Site-directed mutagenesis and structural studies are under way in order to define the roles of the above-mentioned residues in the CpsA catalytic mechanism and to clarify the molecular basis of the unusually broad substrate specificity and chemio- and thermostability of this interesting enzyme.

ACKNOWLEDGMENTS

We thank M. Rossi for the generous gift of the *S. solfataricus* gene bank and A. Gambacorta for supplying the MT-4 strain of *S. solfataricus*.

This work has been supported by grants from the Ministero dell'Università e della Ricerca Scientifica e Tecnologica (40% program) and the Consiglio Nazionale delle Ricerche.

REFERENCES

1. Bianchi, M. E. 1991. Production of functional rat HMG1 protein in *Escherichia coli*. *Gene* **104**:271–275.
2. Christianson, D. W., and W. N. Lipscomb. 1989. Carboxypeptidase A. *Acc. Chem. Res.* **22**:62–69.
3. Colombo, S., S. D'Auria, P. Fusi, L. Zecca, C. A. Raia, and P. Tortora. 1992. Purification and characterization of a thermostable carboxypeptidase from the extreme thermophilic archaeobacterium *Sulfolobus solfataricus*. *Eur. J. Biochem.* **206**:349–357.
4. Colombo, S., M. Grisa, P. Tortora, and M. Vanoni. 1994. Molecular cloning, nucleotide sequence and expression of a *Sulfolobus solfataricus* gene encoding a class II fumarase. *FEBS Lett.* **337**:93–98.
5. Cubellis, M. V., C. Rozzo, P. Montecucci, and M. Rossi. 1990. Isolation and sequencing of a new β -galactosidase-encoding archaeobacterial gene. *Gene* **94**:89–94.
6. Cubellis, M. V., C. Rozzo, G. Nitti, M. I. Arnone, G. Marino, and G. Sannia. 1989. Cloning and sequencing of the gene coding for aspartate aminotransferase from the thermoacidophilic archaeobacterium *Sulfolobus solfataricus*. *Eur. J. Biochem.* **186**:375–381.
7. De Rosa, M., A. Gambacorta, B. Nicolaus, P. Giardina, E. Poerio, and V. Buonocore. 1984. Glucose metabolism in the extreme thermoacidophilic archaeobacterium *Sulfolobus solfataricus*. *Biochem. J.* **224**:407–414.
8. Doi, E., D. Shibata, and T. Matoba. 1981. Modified colorimetric ninhydrin methods for peptidase assays. *Anal. Biochem.* **118**:173–184.
9. Folk, J. E. 1971. Carboxypeptidase B, p. 57–79. In P. D. Boyer (ed.), *The enzymes*, 3rd ed. vol. 3. Academic Press, New York.
10. Fricker, L. D., and S. H. Snyder. 1982. Enkephalin convertase: purification and characterization of a specific enkephalin-synthesizing carboxypeptidase localized to adrenal chromaffin granules. *Proc. Natl. Acad. Sci. USA* **79**:3886–3890.
11. Gardell, S. J., C. S. Craick, D. Hilvert, M. S. Urdea, and W. Rutter. 1985. Site-directed mutagenesis shows that tyrosine 248 of carboxypeptidase A does not play a crucial role in catalysis. *Nature (London)* **317**:551–555.
12. Harlow, E., and D. Lane. 1988. *Antibodies: a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
13. Hartsuck, J. A., and W. N. Lipscomb. 1971. Carboxypeptidase A, p. 1–56. In P. D. Boyer (ed.), *The enzymes*, 3rd ed. vol. 3. Academic Press, New York.
14. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680–685.
15. Lawrence, C. E., S. F. Altschul, M. S. Boguski, J. S. Liu, A. F. Neuwald, and J. C. Wotton. 1993. Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. *Science* **262**:208–214.
16. Lee, S.-H., E. Minagawa, H. Taguchi, H. Matsuzawa, T. Ohta, S. Kamino-gawa, and K. Yamauchi. 1992. Purification and characterization of a thermostable carboxypeptidase (carboxypeptidase Taq) from *Thermus aquaticus* YT-1. *Biosci. Biotechnol. Biochem.* **56**:1839–1844.
17. Levin, Y., R. A. Skidgel, and E. G. Erdős. 1982. Isolation and characterization of the subunits of human plasma carboxypeptidase N (kininase I). *Proc. Natl. Acad. Sci. USA* **79**:4618–4622.
18. Lipman, W. J., and W. R. Pearson. 1985. Rapid and sensitive protein similarity searches. *Science* **227**:1435–1441.
19. Olsen, G. J., N. R. Pace, M. Nuell, B. P. Kaine, R. Gupta, and C. R. Woese. 1985. Sequence of 16S rRNA gene from the thermoacidophilic archaeobacterium *Sulfolobus solfataricus* and its evolutionary implications. *J. Mol. Evol.* **22**:301–307.
20. Pisani, F. M., C. De Martino, and M. Rossi. 1992. A DNA polymerase from

- the archaeon *Sulfolobus solfataricus* shows sequence similarity to family B DNA polymerases. *Nucleic Acids Res.* **20**:2711–2716.
21. Ramirez, C., and A. T. Matheson. 1991. A gene in the archaeobacterium *Sulfolobus solfataricus* that codes for a protein equivalent to the alpha subunits of the signal recognition particle receptor in eukaryotes. *Mol. Microbiol.* **5**:1687–1693.
 22. Reiter, W. D., P. Palm, A. Henschen, F. Lottspeich, W. Zillig, and B. Grampp. 1987. Identification and characterization of the genes encoding three structural proteins of the *Sulfolobus* virus-like particle SSV1. *Mol. Gen. Genet.* **206**:144–153.
 23. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
 24. Sanangelantoni, A. M., D. Barbarini, G. Di Pasquale, P. Cammarano, and O. Tiboni. 1990. Cloning and nucleotide sequence of an archaeobacterial glutamine synthetase gene: phylogenetic implications. *Mol. Gen. Genet.* **221**:187–194.
 25. Schuler, G. D., S. F. Altschul, and D. J. Lipman. 1991. A workbench for multiple alignment and analysis. *Proteins* **9**:180–190.
 26. Smith, D. E., and P. A. Fisher. 1984. Identification, developmental regulation, and response to heat shock of two antigenically related forms of a major nuclear envelope protein in *Drosophila* embryos: application of an improved method for affinity purification of antibodies using polypeptides immobilized on nitrocellulose blots. *J. Cell Biol.* **99**:20–28.
 27. Smulevitch, S. V., A. L. Osterman, O. V. Galperina, M. V. Matz, O. P. Zagnitko, R. M. Kadyrov, I. A. Tsaplina, N. V. Grishin, G. G. Chestukhina, and V. M. Stepanov. 1991. Molecular cloning and primary structure of *Thermoactinomyces vulgaris* carboxypeptidase T. A metalloenzyme endowed with dual substrate specificity. *FEBS. Lett.* **291**:75–78.
 28. Studier, F. W., and B. A. Moffatt. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**:113–130.
 29. Teplyakov, A., K. Polyakov, G. Obmolova, B. Strokopyton, I. Kuranova, A. Osterman, N. Grishin, S. Smulevitch, O. Zagnitko, O. Galperina, M. Matz, and V. M. Stepanov. 1992. Crystal structure of carboxypeptidase T from *Thermoactinomyces vulgaris*. *Eur. J. Biochem.* **208**:201–288.
 - 29a. Tortora, P. Unpublished work.
 30. Villa, A., L. Zecca, P. Fusi, S. Colombo, G. Tedeschi, and P. Tortora. 1993. Structural features responsible for kinetic thermal stability of a carboxypeptidase from the archaeobacterium *Sulfolobus solfataricus*. *Biochem. J.* **295**:827–831.
 31. Young, R. A., and R. W. Davis. 1983. Efficient isolation of genes using antibody probes. *Proc. Natl. Acad. Sci. USA* **80**:1194–1198.