J. Biochem. 124, 778-783 (1998)

Pyrrolidone Carboxyl Peptidase from the Hyperthermophilic Archaeon *Pyrococcus furiosus*: Cloning and Overexpression in *Escherichia coli* of the Gene, and Its Application to Protein Sequence Analysis[1]

Susumu Tsunasawa,[2] Satomi Nakura, Tetsuo Tanigawa, and Ikunoshin Kato

Biotechnology Research Laboratories, Takara Shuzo Co., Ltd., Sunaike 2257, Nojicho, Kusatsu, Shiga 525-0055

Received for publication, May 20, 1998

A gene for a pyrrolidone carboxyl peptidase (Pcp: EC 3.4.19.3, pyroglutamyl peptidase), which removes amino-terminal pyroglutamyl residues from peptides and proteins, has been cloned from the hyperthermophilic Archaeon *Pyrococcus furiosus* using its cosmid protein library, sequenced, and expressed in *Escherichia coli*. The DNA sequence encodes a protein containing 208 amino acid residues with methionine at the N-terminus. Analysis of the recombinant protein expressed in *E. coli*, including amino acid sequence analysis from the N-terminus by automated Edman degradation and ionspray mass spectrometric analysis of the peptides generated by enzymatic digestions with lysylendopeptidase and *Staphylococcus aureus* V8 protease, showed its primary structure to be completely identical with that deduced from its cDNA sequence. Comparison of the amino acid sequence of *P. furiosus* Pcp (*P.f.*Pcp) with those of bacterial Pcps revealed that a high degree of sequence identity (more than 40%) and conservation of the amino acid residues comprising the catalytic triad, Cys142, His166, and Glu79. On the other hand, a unique short stretch sequence (positions around 175-185) that is absent in bacterial Pcps was found in *P.f.*Pcp. A similar stretch has also been reported recently in the amino acid sequence of Pcp from the hyperthermophilic Archaeon *Thermococcus litoralis* [Littlechild *et al.*, in abstracts of the "International Congress on Exthermophiles '98" p. 58 (1998)]. To elucidate their contribution to the hyperthermostability of these enzymes, further structural studies are required.

Key words: amino acid sequence, hyperthermostable enzyme, ionspray mass spectrometry, *Pyrococcus furiosus*, pyrrolidone carboxyl peptidase.

Pyrrolidone carboxyl peptidase (Pcp) [EC 3.4.19.3] removes amino-terminal L-pyroglutamic acids from peptides and proteins. Pcps found so far have been divided into two classes (*1*). Type I enzymes are found in both prokaryotes and eukaryotes, belong to one of the cysteine protease families, and are soluble proteins consisting of two or four identical subunits with molecular weight (MW) ranging from 20,000 to 25,000. Type II enzymes are mainly located in mammalian brain tissue, being membrane-bound metallopeptidases having MWs of between 230,000 and 280,000. The role of Type I enzymes still remains unclear, whereas Type II enzymes appear to play an important physiological role in the degradation of other neuropeptides. In addition, Pcps has practical significance in its application to protein sequence analysis as a means of selective removal of N-terminal pyroglutamic acids from peptides and proteins before Edman degradation (*2*). From this standpoint, we have found in the hyperthermophilic Archaeon *Pyrococcus furiosus* an enzyme that is superior in both thermostability and specific activity to any other Pcps found so far. Here, we show the presence of a gene for Pcp in *P. furiosus*, its cloning, and DNA sequencing. We also report several characteristics of the recombinant protein expressed in *Escherichia coli* and its utilization for protein sequence analysis.

MATERIALS AND METHODS

Materials—*P. furiosus* (DM 3638) was cultured as described previously (*3*). *E. coli* JM109, enzymes for *in vitro* manipulation of DNA and their kits, and vector DNAs were the products of Takara Shuzo. *E. coli* DH5αMCR was from Bethesda Research Laboratories. Gigapack II Gold kit and Triple Helix Cosmid Vector were from Stratagene. Lysylendopeptidase from *Achromobacter lyticus* [EC 3.4.21.50] was obtained from Wako Pure Chemical Industries. *Staphylococcus aureus* V8 protease and porcine liver pyrrolidone carboxyl peptidase were from Boehringer Mannheim. Pyrrolidone carboxyl peptidase from *Bacillus amyloliquefacience* was from Toyobo. Synthetic peptides except the N-terminal octapeptide of ovalbumin were

[1] The nucleotide sequence data reported in this paper will appear in the GSDB, DDJB, EMBL, and NCBI nucleotide sequence data bases with the following accession number AB015291, as a *Pyrococcus furiosus* gene.

[2] To whom correspondence should be addressed. E-mail: s-tsunas@mx.biwa.or.jp

Abbreviations: Pcp, pyrrolidone carboxyl peptidase; TFA, trifluoroacetic acid; pyr-, pyroglutamyl-; *p*NA, *p*-nitroanilide; *P.f.*Pcp, *Pyrococcus furiosus* pyrrolidone carboxyl peptidase; PTH-, phenylthiohidantoin-.

products of Peptide Institute Inc. The octapeptide was synthesized by the *t*-butoxycarbonic method on a peptide synthesizer (Perkin-Elmer 430A). Hen egg white riboflavin-binding protein and low molecular weight trypsin inhibitor from *Brassia oleracea*, were kindly provided by Drs. T. Mega (Faculty of Science, Osaka University) and S. Watanabe (Seitoku Jr. College of Nutrition), respectively. The source of other specific chemicals and reagents is shown in the text.

Preparation of P. furiosus Cosmid Protein Library—The *P. furiosus* genome DNA (400 μg) was partially digested with *Sau*3AI in a buffer [50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 1 mM dithiothreitol, 100 mM NaCl] and fractionated according to fragment size by sucrose density gradient centrifugation. Triple Helix Cosmid Vector (1 μg) was cleaved with *Bam*HI and mixed with the genome DNA fragments (140 μg) of 35–50 kb fractionated as described above. After ligation by use of a Ligation kit, the *P. furiosus* genome DNA fragments in the cosmid vector were packaged into λ-phage particles by the *in vitro* packaging method using Gigapack II Gold. By using a part of the phage solution thus obtained, *E. coli* DH5αMCR was transformed to give a cosmid library. Then 500 colonies were selected and cultured independently in an L-broth containing 100 μg/ml ampicillin at 37°C with shaking for 16 h. The culture was centrifuged and cells were harvested. These cells were ultrasonicated and further treated at 100°C for 10 min. After centrifugation, the supernatants were used as a protein library to screen for the target activity.

Measurement of Pcp Activity—Substrate solution [Na-phosphate buffer (pH 7.0) containing 1 mM L-pyroglutamyl *p*-nitroanilide (Pyr-*p*NA) and 10 mM DTT; usually 500 μl] was added to an appropriate amount of each sample (usually 5 μl), and the mixture was incubated at 75°C for 15 min. The reaction was stopped by adding 30% acetic acid (usually 100 μl), then the release of the *p*NA was quantified by measuring the absorbancy at 410 nm. One unit of the activity was defined as the amount of enzyme that released 1 μmol of *p*NA per min under the conditions described above. The same procedures were used for screening of the cosmids and clones for Pcp activity.

Screening of the Pcp Gene of P. furiosus—The Pcp activity in the *P. furiosus* cosmid protein library was measured. From two transformants showing the Pcp activity, cosmid DNA was prepared and digested with *Bam*HI, inserted into the *Bam*HI site of the plasmid vector pUC118, and then introduced into *E. coli* JM109. From the transformants, a colony showing the Pcp activity in the supernatant of cells prepared in a similar manner to the cosmid protein library was further selected, and the resultant plasmid was designated pPCP1. pPCP1 was digested with *Sac*I and *Hinc*II, the resulting DNA fragments were inserted into pUC118, and the plasmid pPCP2 was obtained by measuring the Pcp activity in *E. coli* transformants in the same way as described above. A DNA fragment of about 1.1 kbp containing the Pcp gene was subjected to DNA sequencing. To construct a more effective expression system of *P.f.*Pcp, the Pcp gene was amplified from the plasmid pPCP2 by PCR together with an oligonucleotide, GGGTTAAATTTAAAGTATTAGTTACCGGG, into which a *Dra*I site was introduced by replacing a putative initiation codon, ATG, of the Pcp gene with TTT, and M13 primer M4. After amplification, the PCR product was treated with proteinase K, extracted with phenol-chloroform, and precipitated with ethanol. The PCR product was digested with *Dra*I and *Hinc*II, then purified by agarose gel electrophoresis, inserted into the *Nco*I-*Hinc*II site of the plasmid vector pTV118N, and used to transform *E. coli* JM109. *E. coli* JM109/pPCP3 thus obtained showed strong and hyperthermostable Pcp activity.

DNA Sequencing—The DNA fragment of about 1.1 kbp containing the Pcp gene inserted into the plasmid pPCP2 was subjected to DNA sequencing. The nucleotide sequences of the DNA fragments formed by digestion of the parent DNA fragment with several restriction enzymes were analyzed by the dideoxy chain termination method basically as described by Sanger *et al.* (*4*). BcaBEST DNA polymerase (*5*) was used for the chain elongation reaction with fluorescent dideoxy terminators, and the products were analyzed with an autosequencer (Perkin-Elmer-370A).

Purification of the Recombinant P. furiosus Pcp Expressed in E. coli—*E. coli* JM109 carrying pPCP3 was grown at 37°C for 16 h with shaking in 250 ml of an L-broth medium containing 0.01% ampicillin. After centrifuging the culture (6,000 rpm, 10 min), the cells thus collected (4.6 g wet weight) were suspended in 22 ml of 50 mM Tris-HCl buffer (pH 7.5), ultrasonicated and treated at 100°C for 10 min. The suspension was centrifuged (12,000 rpm, 30 min), and the supernatant was used as a crude enzyme source. To the supernatant, 50 mM Tris-HCl buffer (pH 7.5) containing 1% protamine sulfate was added to give a final concentration of 0.1%. The solution was kept at 5°C for 30 min, then centrifuged (12,000 rpm, 30 min), and the supernatant (28 ml) was fractionated with ammonium sulfate at 40 to 80% saturation. The precipitate was dissolved in 5 ml of 1.2 M ammonium sulfate, 20 mM K-phosphate buffer containing 5 mM EDTA and 1 mM 2-mercaptoethanol, and the resultant clear solution was applied to a column of Phenyl Sepharose CL-4B (Pharmacia, 2.5×18 cm) equilibrated with the same buffer. The column was washed with the buffer, the adsorbed enzyme was eluted with a decreasing linear gradient of ammonium sulfate from 1.2 to 0 M. The active fractions of the eluate were combined, concentrated by ultrafiltration, and dialyzed against 20 mM K-phosphate buffer (pH 7.0) containing 5 mM EDTA and 1 mM 2-mercaptoethanol. The dialysate (26 ml) was loaded on a DEAE-Sepharose CL-6B column (Pharmacia, 1.5×9.6 cm) equilibrated with the same buffer. The column was developed with a linear gradient from 0 to 1 M of KCl in the same buffer, and active fractions were collected (88 ml), concentrated by ultrafiltration and dialyzed against 10 mM Na-phosphate buffer (pH 7.5). The enzyme thus obtained showed a single band on SDS-PAGE. Protein was measured by the method of Bradford (*6*).

Protein Analysis—The N-terminal sequences of the intact recombinant *P.f.*Pcp and the substrate peptides and proteins generated by digestion with *P.f.*Pcp were analyzed by direct sequencing with a pulse-liquid phase protein sequencer (Perkin-Elmer 477A). A triple quadrupole equipped with an ionspray ion source mass spectrometer (Perkin-Elmer-Sciex, API-III) was used to determine the molecular masses of the both intact and *S*-carboxymethylated recombinant *P.f.*Pcps by introducing each protein sample (200 pmol) dissolved in 100 μl of 0.1% formic acid–50% acetonitrile into the mass spectrometer through a

fused silica tube (100 μm i.d.) at a flow rate of 2 μl/min. The quadrupole was scanned from 350 to 2,000 Da using a step size of 0.1 Da and a dwell time of 0.5 ms per step. Ionspray voltage was set at 5 kV, and the orifice voltage was 80 V. *S*-Carboxymethylation of the recombinant *P.f.*Pcp was basically carried out according to Crestfield *et al.* (*7*). In advance of LC-MS analysis, the *S*-carboxymethylated *P.f.*Pcp was digested with each protease at 30°C for 24 h under the conditions indicated in parentheses (buffer; molar substrate-to-enzyme ratio), lysylendopeptidase [0.1 M Tris-HCl buffer (pH 9.0) containing 2 M urea; 200:1], *S. aureus* V8 protease [0.05 M ammonium bicarbonate (pH 7.8) containing 2 M urea; 200:1]. Aliquots of each digest (25 pmol) were loaded onto a packed capillary reversed phase column (Poros IIR/H, 0.3×100 mm, LC-Packing) connected to the ionspray mass spectrometer. A solvent flow of 10 μl/min was generated by use of a Waters 625 LC system. Solvent A was 0.05% TFA in H_2O, and solvent B was 0.05% TFA in acetonitrile. Separation of the peptides obtained was effected with a gradient of 0–60% B over 45 min. The column effluent, except for the unadsorbed fraction, was passed into an ionspray probe of the mass spectrometer. The total ion chromatograms were recorded in the single quadrupole mode. Amino acid analyses of the both intact and *S*-carboxymethylated recombinant *P.f.*Pcps were performed with an amino acid analyzer (Hitachi L-8500S) for the hydrolyzate with 5.7 M HCl containing 0.2% phenol or with 4 M methanesulfonic acid containing 0.2% 3-(2-aminoethyl)indole at 100°C for 24 h (*8*).

Enzymatic Properties of the Pcp—The specificity for several synthetic peptides and proteins with N-terminal pyroglutamic acids was examined. The N-terminal amino acid sequences of the resultant peptides and proteins were analyzed with a protein sequencer as above. The pH optimum and stability were determined at 75°C in a buffer containing 10 mM DTT over the pH range 4.0–10.6. The buffers used were 50 mM Na-acetate for pH 4.0–5.0, 20 mM Na-phosphate for pH 5.8–8.0, 50 mM Na-borate for pH 8.0–9.6, and disodium hydrogen phosphate-NaOH for pH 9.9–10.6. The pH stability was measured with an aliquot of each solution with drawn after incubation at a given pH at 75°C for 30 min by using the standard assay system. The effect of temperature on the activity was determined by incubating an appropriate amount of the enzyme in 50 mM Na-phosphate buffer (pH 7.0) containing 10 mM DTT at temperatures varying from 20 to 100°C for 5 min. Thermostability of the enzyme was determined by measuring the remaining activity in the standard assay system after incubation in the above buffer at 75°C for 1 h. The effects of denaturants on the enzyme activity were also examined.

Computer Analysis—A homology search for the amino acid sequence of *P.f.*Pcp was done by GenBank using FASTA search program.

RESULTS AND DISCUSSION

Cloning of the Pcp Gene from P. furiosus—To screen the Pcp gene in *P. furiosus*, the cosmid protein library described in "MATERIALS AND METHODS" was used as previously described (*9*). From the cosmid protein library made of 500 different clones, the supernatants of two transformants showed high Pcp activity. Cosmid DNAs were prepared from these two transformants, and the restriction patterns of each insert DNA fragment were compared. Both inserts had a *Bam*HI fragment of about 4.2 kbp long. A cosmid DNA was thus selected and subjected to screening and sequencing of the Pcp gene. From subcloning and assay of its heat-stable Pcp activity as described in "MATERIALS AND METHODS," the Pcp gene was located within the 1.1-kbp *Sac*I-*Hin*cII fragment. The plasmid containing the *Sac*I-*Hin*cII restriction fragment was designated pPCP2. The restriction map of the 1.1-kbp insert and its complete nucleotide sequence are shown in Figs. 1 and 2, respectively. The first ATG codon, located at nucleotides 117–119 downstream from the in-frame terminator TAG (nucleotides 69–71) in the determined sequence, was concluded to be the initiation codon from the N-terminal sequence analysis of the Pcp protein. The Pcp gene thus encompasses 624 nucleotides, with the initiation codon ATG and the termination codon TGA. From analysis of the nucleotide sequence flanking the Pcp gene, a Shine-Dalgarno sequence, GGAG, was found at nucleotides 80–83, but, as the consensus Archaeal promoter sequences, only one conserved box A (*10*) sequence, TTAAA, was found immediately preceding the initiation codon.

Expression of the Pcp Gene in E. coli—The recombinant *E. coli* JM109/pPCP3 produced 33.4 units/mg protein of *P. furiosus* Pcp in the crude extract heated at 100°C for 10 min with induction by isopropyl-1-thio-β-D-galactopyranoside. The enzyme was purified to homogeneity by heat treatment, treatment with protamine sulfate, fractionation with ammonium sulfate, and Phenyl-Sepharose CL-4B and DEAE-Sepharose CL-6B column chromatographies (Table I).

Characterization of the Recombinant P. furiosus Pcp—The N-terminal 25 amino acid residues of the recombinant *P.f.*Pcp were determined by direct sequencing of the intact protein. The result indicates that the initiation codon of the *P.f.*Pcp gene corresponds to nucleotides 117–119 (ATG) in Fig. 2. The molecular masses of the native and carboxymethylated Pcps were determined by ionspray triple quadrupole mass spectrometry to be 45,643.0±2.0 Da and 22,936.7±1.9 Da, respectively. The value obtained for the carboxymethylated Pcps was in close agreement with the value (22,939 Da) calculated from the translated nucleotide sequence from the ATG described above as the initiation codon. These results suggest that at least one of the two cysteine residues of *P.f.*Pcp forms a disulfide bridge. If Cys142 of *P.f.*Pcp takes part in a

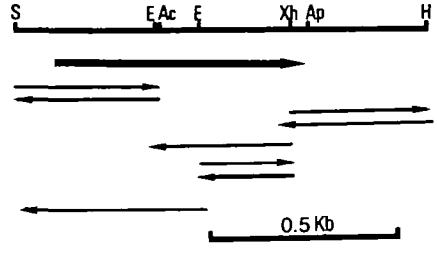

Fig. 1. **Restriction map of the 1.1-kb DNA fragment containing the Pcp gene from *P. furiosus* and the sequencing strategy.** The open reading frame is indicated by the thick arrow. Arrows below the DNA fragment show the direction of sequencing and the region sequenced. Abbreviations: Ac, *Acc*III; Ap, *Apa*LI; E, *Eco*RI; H, *Hin*cII; S, *Sac*I; Xh, *Xho*I.

similar manner to those of bacterial Type I Pcps, then it must be Cys188 that provides the sulfhydryl group through which *P.f.*Pcp forms a dimer. On the other hand, the molecular weight determined by the sedimentation equilibrium with an analytical ultracentrifugation (Beckmann Optima mode XL-A) was approximately 96,300 on the assumption that the partial specific volume of the *P.f.*Pcp was 0.754 cm^3/g from its amino acid composition (K. Yutani *et al.*, in preparation). The native *P.f.*Pcp is therefore estimated to form a tetramer composed of two identical dimers as described above. Next, a more detailed structural analysis of the protein was performed by LC-MS analysis of each peptide produced by either lysylendopeptidase or *Staphylococcus aureus* V8 protease digestion. The observed masses of the resulting peptides are listed in Table II with the expected average masses of amino acid residues deduced from the nucleotide sequence of the Pcp gene. The observed mass in each case is within 0.5 Da of the expected mass. Several small peptides of less than 6 residues from both digests, which might be eluted in the unadsorbed fraction on the LC column used, could not be analyzed, but by mass analyses of the peptides corresponding to the overlaps of each digest, the entire amino acid sequence of the recombinant *P.f.*Pcp was elucidated to be identical with that deduced from the nucleotide sequence of its gene.

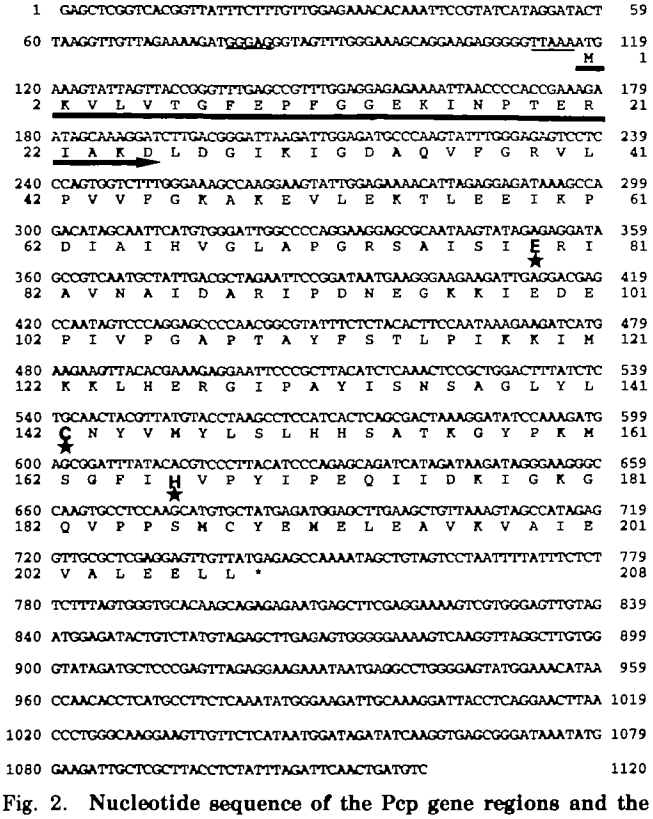

Fig. 2. **Nucleotide sequence of the Pcp gene regions and the deduced amino acid sequence of the *P.f.*Pcp.** The nucleotide numbering starts with the 5′-untranslated region of the DNA fragment. Amino acid numbering is shown beneath nucleotide numbers. The amino acid residues directly identified by sequence analysis of the N-terminus of the intact *P.f.* Pcp are underlined. Amino acid residues estimated to be involved in the catalytic triad are indicated by bold letters and asterisks (★).

TABLE I. **Purification of the recombinant *P.f.*Pcp from *E. coli*.** Values are from 4.6 g (wet weight) of *E. coli* JM109/pPCP3.

Step	Protein (mg)	Activity[a] (U)	Purification (fold)	Recovery (%)
Crude extract (with heat treatment)	78.1	2,606	1	100
Supernatant (with protamine sulfate treatment)	127.6	2,579	0.6	98.9
Precipitates (at 40-80% ammonium sulfate)	55.2	2,706	1.47	103.8
Phenyl-Sepharose CL-4B	10.4	1,583	4.45	60.7
DEAE-Sepharose CL-6B	2.97	796	8.02	30.5

[a]One unit of the activity is defined as the amount of the enzyme that releases 1 μmol of *p*NA from the synthetic peptide Pyr-*p*NA per min at pH 7.0 at 75°C.

Enzymatic Properties of the Recombinant P. furiosus Pcp—The effects of temperature and pH on the recombinant *P.f.*Pcp were examined. The optimum pH was 6-9, and the optimum temperature was around 90°C (Fig. 3). More than 80% of Pcp activity was retained after heating at 75°C for 150 min in the range of pH 5.0 to 9.0. The Pcp partially purified from crude extract of *P. furiosus* by the same procedures as for the recombinant protein exhibited the same properties with respect to temperature and pH as the recombinant enzyme. After dialysis against 50 mM Na-phosphate buffer (pH 7.0) containing a thiol group blocking reagent such as *N*-ethylmaleimide or *p*-chloro-

TABLE II. **Comparison of observed mass values with theoretical ones for the generated peptides by digestion of the *S*-carboxymethylated recombinant *P.f.*Pcp with lysylendopeptidase (K−) or *S. aureus* V8 protease (SP−).**

Peptide No.[a]		Position		Theoretical mass (average mass)	Observed mass
	SP1-2[b]	M1	-E14	1,511.8	1,510.5
	SP1-3	M1	-E20	2,194.6	2,194.4
	SP1-4′	M1	-D27	3,005.5	3,005.1
K2		V3	-K15	1,379.6	1,379.5
K3		I16	-K24	1,041.2	1,041.0
K4		D25	-K30	659.7	659.5
K5		I31	-K47	1,802.2	1,802.1
	SP4	R21	-E50	3,240.9	3,241.6
	SP4′	D28	-E50	2,428.9	2,429.4
	SP4-5[b]	R21	-E53	3,582.3	3,581.0
	SP5-10[b]	V51	-E99	5,304.1	5,303.5
	SP6-10	K54	-E99	4,962.7	4,961.1
K8-9[b]		T55	-K96	4,464.1	4,464.8
K9		P61	-K96	3,750.3	3,751.3
	SP8	D59	-E79	2,157.6	2,157.4
	SP9	R80	-E87	1,666.9	1,667.1
K11		I98	-K118	2,258.6	2,258.2
K11-12		I98	-K119	2,386.8	2,387.1
K15		L124	-K156	3,723.3	3,723.4
	SP11-12[b]	D100	-E126	3,024.6	3,024.9
	SP13	R127	-E172	5,161.0	5,160.3
K17		M161	-K177	1,911.4	1,911.2
	SP14	Q173	-E190	2,006.4	2,007.0
K19		G181	-K197	1,969.3	1,969.0
	SP17	A195	-E201	728.9	728.8
K20		V198	-K208	1,198.4	1,198.7

[a]Peptides are named by using letters and numbers. Letters indicate the enzyme used for fragmentation: K, lysylendopeptidase; SP, *S. aureus* V8 protease, and numbers indicate the sequential order of the peptides predicted to be produced from the *S*-carboxymethylated recombinant *P.f.*Pcp by each protease. [b]Peptides produced by either partial (−) or non-specific cleavages (′).

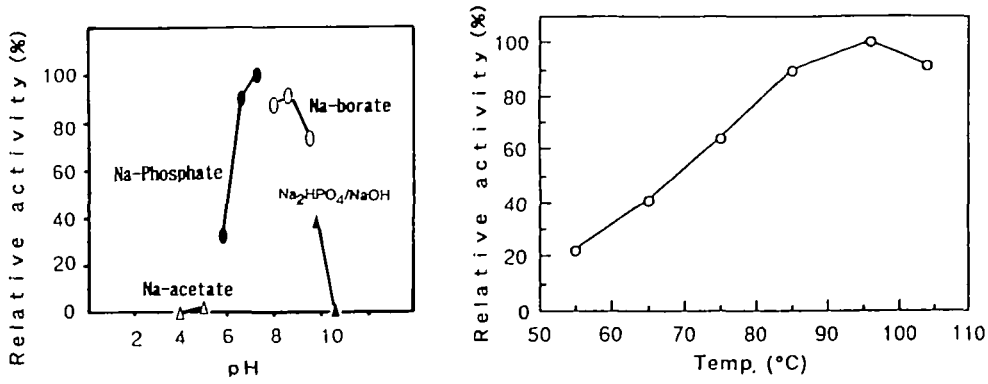

Fig. 3. **Effects of pH (a) and temperature (b) on the activity of *P.f.*Pcp.** Detail are described in the text.

TABLE III. **N-terminal sequence analyses of N-blocked peptides and proteins after treatment with *P.f.*Pcp.**

Peptides/proteins	N-terminal amino acid sequences[a]
Physaleamin	Pyr-A-D-P-N-K-F-Y-G-L-M-NH_2
Neurotensin	Pyr-L-Y-E-N-K-P-R-R-P-Y-I-L
α-MSH	Ac-S-Y-S-M-E-H-F-R-W-G-K-P-V-NH_2
N-terminal peptide of ovalbumin	Ac-G-S-I-S-A-A-S-M
Riboflavin binding protein (hen egg-white)	Pyr-Q-Y-G-C-L-E-G-D-T-H- - -
Low-molecular weight trypsin inhibitor (*B. oleracea*)	Pyr-D-S-E-C-L-K-E-Y-G-G - - - -

[a]Amino acids shown in underlined letters were identified with a protein sequencer.

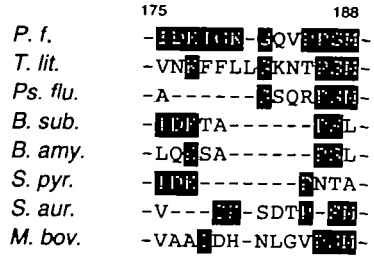

Fig. 4. **Amino acid sequence alignment of *P.f.*Pcp with those of Pcps from other bacteria at the region where a unique short stretch sequence (positions at 175-188) is observed in *P.f.*Pcp.** Shaded residues are identical with those of *P.f.*Pcp. Abbreviations are as follows: *T. lit, Thermococcus litoralis* (*17*); *Ps.flu, Pseudomonas fluorescens* (*15*); *B. sub, Bacillus subtilis* (*12*); *B. amy, Bacillus amyloliquefaciens* (*14*); *S. Pyr, Streptococcus pyrogens* (*13*); *S. aur, Staphylococcus aureus* (*16*); *M. bov, Mycobacterium bovis* (Kim, J.K. and Choe, Y.K., unpublished results).

mercurybenzoate (1 mM), the enzyme activity was completely lost. But addition of DTT (10 mM) in the reaction buffer [50 mM Na-phosphate buffer (pH 7.0)] completely restored the activity, even after heating at 75°C overnight. The enzyme was inhibited by Hg^{2+} ions, and to a lesser extent, by Zn^{2+} and Co^{2+} ions, but other metal ions examined had no effect. These properties suggest that *P.f.*Pcp is a thiol protease similar to other Type I Pcps. In fact, substitution of Cys142 with Ser by site-directed mutagenesis resulted in a complete loss of the activity (K. Yutani *et al.*, in preparation), and this strongly suggests that Cys142 is involved in the catalytic triad of the enzyme. On the other hand, the fact that only about one of the activity was lost on substitution of Ser for Cys188, which seems to be involved in dimer formation through disulfide bridging suggests that Cys188 is not essential for the activity, but contributes to the catalytic action of the enzyme by formation of a suitable oligomeric structure. The resistance of the enzyme activity to several denaturants was examined. Approximately 70% of the original activity was retained after preincubation with 10 mM DTT, 50 mM Na-phosphate buffer (pH 7.0) containing less than either 0.01% SDS, 1 M urea, or 1 M guanidine-HCl at 37°C for 15 min.

Application of the P.f.Pcp to Sequence Analysis—The specificity of the enzyme for several peptides with or without an N-terminal pyroglutamyl group was tested, and it was found that the enzyme specifically cleaved only N-terminal pyroglutamyl residues from these substrates (Table III). The activity of the enzyme was also examined for two proteins, hen egg white riboflavin-binding protein (*11*) and low-molecular weight trypsin inhibitor from *Brassica oleracea* (S. Watanabe *et al.*, in preparation), which have pyroglutamyl residues at their N-termini. Table III also shows the results of N-terminal amino acid sequence analyses of these two proteins blotted on PVDF membranes. It should be noted that N-terminal pyroglutamyl residues were released from the proteins upon incubation in 50 mM Na-phosphate buffer (pH 7.0) containing 10 mM DTT at 50°C for 5 h at a molar substrate-to-enzyme ratio of 25 to 1 without previous denaturation of the substrate proteins. Under the same conditions except for the molar substrate-to-enzyme ratio of 300 to 1, the enzyme released N-terminal pyroglutamyl residues to the same extent.

Sequence Comparison with Pcps from Other Species—Several structural features of *P.f.*Pcp are apparent from its amino acid sequence. In comparison with known amino acid sequences of bacterial type I Pcps (*12-16*), *P.f.*Pcp displays a high degree of sequence identity (more than 40%) and conservation of amino acid residues comprising the catalytic triad, Cys, His, and Glu, together with a unique short stretch sequence (positions around 175-185) that is absent in bacterial Pcps, as shown in Fig. 4. A similar stretch has been also recently reported in the amino acid sequence of Pcp from the hyperthermophilic Archaeon, *Thermococcus litoralis* (*17*). These facts indicate that *P.f.*Pcp is basically a type I Pcp similar to those from bacteria, and that its hyperthermostability must be derived from features of its 3D structure. To specify the structural features responsible

for hyperthermostability, further structural studies including the contribution of the stretch sequence observed in Pcps from thermophilic Archaeon are required. Such studies are now in progress and the results will be published elsewhere.

The recombinant *P.f.*Pcp has higher thermostability, higher resistance to such denaturants as 1 M urea and 0.01% SDS, and higher specific activity even at moderate temperature than other known Pcps. For example, Pcps from *P. furiosus*, *B. amyloliquefaciens*, and porcine liver showed activities of 5.8, 0.90, and 0.53 U/mg, respectively, under the assay conditions described in the text except that enzymatic reaction was done at 37°C. Therefore, *P.f.*Pcp can act on native proteins in buffer including these denaturants as at higher temperature. The recombinant *P.f.*Pcp produced in *E. coli* is presently the most suitable enzyme for the removal of N-terminal pyroglutamic acids of peptides and proteins. It is also expected to find industrial application in the degradation of pyroglutamyl peptides formed during manufacturing process of nutrient peptides. By simultaneous digestion of proteins from foods such as casein and wheat glutein with non-specific endoproteases and *P.f.*Pcp at higher temperature, by-production of pyroglutamyl peptides, which are suspected both to disturb the absorption of other peptides in intestine with their slight solubility and to exhibit unknown physiological activities, will be grately diminished.

REFERENCES

1. Awade, A.C., Cleuziat, Ph., Gonzales, Th., and R.-Baudouy, J. (1994) Pyrrolidone carboxyl peptidase: An enzyme that removes pyroglutamic acid from pGlu-peptides and pGlu-proteins. *Proteins* **20**, 34-51
2. Doolittle, R.F. and Armentroyt, R.W. (1968) Pyrrolidone peptidase: An enzyme for selective removal of pyrrolidone carboxylic acid residues from peptides. *Biochemistry* **7**, 516-521
3. Laderman, K.A., Davis, B.R., Krutzsch, H.C., Lewis, M.C., and Anfinsen, C.B. (1993) The purification and characterization of an extremely thermostable α-amylase from the hyperthermophilic Archaebacterium *Pyrococcus furiosus*. *J. Biol. Chem.* **268**, 24394-24401
4. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467
5. Ishino, Y. (1992) Rapid and reliable DNA sequencing with a dideoxy sequencing kit. *Am. Biotechnol. Lab.* **10**, 47
6. Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254
7. Crestfield, A.M., Moore, S., and Stein. W.H. (1962) The preparation and enzymatic hydrolysis of reduced and S-carboxymethylated proteins. *J. Biol. Chem.* **238**, 622-627
8. Simpson, R.J., Neuberger, M.R., and Liu, T.-Y. (1976) Complete amino acid analysis of proteins from a single hydrolysate. *J. Biol. Chem.* **251**, 1936-1940
9. Tsunasawa, S., Izu, Y., Miyagi, M., and Kato, I. (1997) Methionine aminopeptidase from the hyperthermophilic Archaeon *Pyrococcus furiosus*: Molecular cloning and overexpression in *Escherichia coli* of the gene, and characteristics of the enzyme. *J. Biochem.* **122**, 843-850
10. Reiter, W.-D., Huneepohl, U., and Zilling, W. (1990) Mutational analysis of an archaebacterial promoter: essential role of a TATA box for transcription efficiency and start-site selection *in vitro*. *Proc. Natl. Acad. Sci. USA* **87**, 9509-9513
11. Hamazume, Y., Mega, T., and Ikenaka, T. (1984) Characterization of hen egg white- and yolk-riboflavin binding proteins and amino acid sequence of egg white-riboflavin binding protein. *J. Biochem.* **95**, 1633-1644
12. Awade, A., Cleuziat, Ph., Gonzales, Th., and R.-Baudouy, J. (1992) Characterization of the *pcp* gene encoding the pyrrolidone carboxyl peptidase of *Bacillus subtilis*. *FEBS Lett.* **305**, 67-73
13. Cleuziat, Ph., Awade, A., and R.-Baudouy, J. (1992) Molecular characterization of *pcp*, the structural gene encoding the pyrrolidone carboxyl peptidase from *Streptococcus pyrogenes*. *Mol. Microbiol.* **6**, 2051-2063
14. Yoshimoto, T., Shimoda, T., Kitazato, A., Kabashima, T., Ito, K., and Tsuru, D. (1993) Pyroglutamyl peptidase gene from *Bacillus amyloliquefaciens*: Cloning, sequencing, expression, and crystallization of the expressed enzyme. *J. Biochem.* **113**, 67-73
15. Gonzales, Th. and R.-Baudouy, J. (1994) Characterization of the *pcp* gene of *Pseudomonas fluorescens* and of its product, pyrrolidone carboxyl peptidase (Pcp). *J. Bacteriol.* **179**, 2569-2576
16. Patti, J.M., Shneider, A., Garza, N., and Boles, J.O. (1995) Isolation and characterization of *pcp* gene encoding a pyrrolidone carboxyl peptidase in *Staphylococcus aureus*. *Gene* **166**, 95-99
17. Littlechild, J., Singleton, M.R., and Isupov, M. (1998) Cloning and X-ray structure of pyrrolidone carboxyl peptidase from the Archaeon *Thermococcus litoralis*. In abstracts of the International Congress on Exthermophiles '98, p. 58