The Exquisite Structure and Reaction Mechanism of Bacterial Pz-peptidase A toward Collagenous Peptides

X-RAY CRYSTALLOGRAPHIC STRUCTURE ANALYSIS OF PZ-PEPTIDASE A REVEALS DIFFERENCES FROM MAMMALIAN THIMET OLIGOPEPTIDASE*

Received for publication, May 5, 2010, and in revised form, June 30, 2010. Published, JBC Papers in Press, September 3, 2010, DOI 10.1074/jbc.M110.141838

Akio Kawasaki†, Hiroaki Nakano†‡, Allin Hosokawa§, Toru Nakatsu$, Hiroaki Kato$, and Kunihiko Watanabe†‡

From the †Division of Applied Life Sciences, Graduate School of Life and Environmental Sciences, Kyoto Prefectural University, Shimogamo, Sakyō, Kyoto 606-8522, Japan, the ‡Department of Structural Biology, Graduate School of Pharmaceutical Sciences, Kyoto University, 46-29 Yoshida Shimoadachi-cho, Sakyō, Kyoto 606-8501, Japan, and the $Department of Pharmacy, Hyogo University of Health Sciences, 1-3-6 Minatojima, Chuo, Kobe 650-8530, Japan

Pz-peptidase A, from the thermophilic bacterium Geobacillus collagenovorans MO-1, hydrolyzes a synthetic peptide substrate, 4-phenylazobenzoyloxy carbonyl-Pro-Leu-Gly-Pro-d-Arg (Pz-PLGPR), which contains a collagen-specific tripeptide sequence, -Gly-Pro-X-, but does not act on collagen proteins themselves. The mammalian enzyme, thimet oligopeptidase (TOP), which has comparable functions with bacterial Pz-peptidases but limited identity at the primary sequence level, has recently been subjected to x-ray crystallographic analysis; however, no crystal structure has yet been reported for complexes of TOP with substrate analogues. Here, we report crystallization of recombinant Pz-peptidase A in complex with two phosphinic peptide inhibitors (PPIs) that also function as inhibitors of TOP and determination of the crystal structure of these complexes at 1.80–2.00 Å resolution. The most striking difference between Pz-peptidase A and TOP is that there is no channel running the length of bacterial protein. Whereas the structure of TOP resembles an open bivalve, that of Pz-peptidase A is closed and globular. This suggests that collagenous peptide substrates enter the tunnel at the top gateway of the closed Pz-peptidase A molecule, and reactant peptides are released from the bottom gateway after cleavage at the active site located in the center of the tunnel. One of the two PPIs, PPI-2, which contains the collagen-specific sequence, helped to clarify the exquisite structure and reaction mechanism of Pz-peptidase A toward collagenous peptides. This study describes the mode of substrate binding and its implication for the mammalian enzymes.

Hard-to-degrade animal proteins (collagen, keratin, and elastin) are ubiquitously present throughout all animal bodies and are responsible for a wide variety of critical roles in cells (1–4). On the other hand, vast amounts of these proteins in industrial wastes are currently disposed of by incineration; accompanying problems, such as the obvious energy loss and the production of carbon dioxide, urgently need to be solved (2). Thermophilic microbes have been closely studied as a means of handling hard-to-degrade proteins, because thermophiles exhibit no pathogenicity in humans or other animals (5–7). However, little is known about how thermophilic microbes degrade those proteins.

Collagens are the most abundant of the aforementioned hard-to-degrade proteins (1, 3). They consist of three helically wound fibrils containing numerous repeats of a tripeptide unit, -Gly-Pro-X-. In animal bodies, initial degradation of collagen is performed by collagenases; complimentary peptidases to further decompose degraded collagen fragments. A mammalian enzyme, thimet oligopeptidase (TOP)3 or endopeptidase 24–15 (EC 3.4.24.15), can hydrolyze a synthetic oligopeptide substrate, 4-phenylazobenzoyloxy carbonyl-Pro-Leu-Gly-Pro-d-Arg (Pz-peptide), containing a collagen-specific tripeptide sequence (8). As a result of this activity, the TOP enzyme was historically called Pz-peptidase and was thought to be involved in the collagen metabolism in mammalian cells (9). More recently, it has been established that TOP is a metallopeptidase that has no relation to the collagenolytic activity, but instead hydrolyzes bioactive neuropeptides and also is involved in histocompatibility (10). In this situation, we have studied a thermophilic and collagen-degrading bacterium, Geobacillus collagenovorans MO-1, and its collagen-degrading enzymes (11, 12). The MO-1 strain produces two distinct Pz-peptide-hydrolyzing enzymes, Pz-peptidases A and B, which hydrolyze the Pz-peptide at the same sites as TOP but do not act on collagen itself (13). Although Pz-peptidases A and B have many similarities in their catalytic properties, they have distinct molecular masses and do not share any antigenic groups (13). It is noteworthy that they have limited primary structural identity (22%), although both Pz-peptidases belong to the M3 family of proteolytic enzymes, in particular, the M3B subfamily. TOPs, which have comparable functions, have even lower identities with Pz-

* This work was supported in part by the Hyogo University of Health Sciences Research grant for 2102 (to H. N.).

† The atomic coordinates and structure factors (codes 3AHM, 3AHN, and 3AHO) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

‡ Both authors contributed equally to this work.

§ To whom correspondence should be addressed: Division of Applied Life Sciences, Graduate School of Life and Environmental Sciences, Kyoto Prefectural University, Shimogamo, Sakyō, Kyoto 606-8522, Japan. Tel. and Fax: 81-75-703-5667; E-mail: kwatanab@kpu.ac.jp.

3 The abbreviations used are: TOP, thimet oligopeptidase; PPI, phosphinic peptide inhibitor.
peptidases (at most 14%) despite belonging to the M3A subfamily. Thus, it is of great interest to compare the structure and function of Pz-peptidases with those of TOPs.

The molecular structure of TOP recently was revealed by x-ray crystallographic analysis at 2.0 Å resolution using structural data from neurolysin, a highly homologous neuropeptidase (14). No crystal structure analysis for the complexes of TOP and neurolysin with the substrate analogues has been reported yet, although the dynamic movements of the domains of TOP and neurolysin participating in peptide hydrolysis are suggested. Previously, we succeeded in crystallizing recombinant Pz-peptidase A in complex with phosphinic peptide inhibitors (PPIs), which also inhibit TOP and neurolysin, and completed the preliminary x-ray analysis (15). One of PPIs contains the collagen-specific tripeptide sequence, Gly-Pro-X; therefore, we expect that this structure will help to clarify the recognition and metabolism of the collagen-specific sequence. In this study, we report the entire structure of Pz-peptidase A itself at 2.00 Å resolution and reveal the structure of the enzyme in complex with two PPIs at 1.80 and 1.88 Å resolution, as well as other new findings.

**EXPERIMENTAL PROCEDURES**

**Protein Purification and Crystallization**—For crystallization, recombinant Pz-peptidase A was purified from an *Escherichia coli* strain BL21(DE3) harboring plasmid pETA-1 according to methods previously described (16). The purified protein solution, concentrated to ~20 mg/ml in 50 mM Tris-HCl (pH 7.5), was incubated in the absence or presence of one of the two PPIs in 12% (w/v) PEG 4000, 0.5 M magnesium acetate, and 0.1 M Tris-HCl (pH 7.0) for 5 days by the hanging-drop vapor diffusion method at 293 K. The PPIs used were benzoxycarbonyl-Phe-(PO2CH2)-Gly-Pro-Nle (PPI-1) and Gly-Pro-Phe-(PO2CH2)-Gly-Pro-Nle (PPI-2) (gifts from Dr. Vincent Dive), at final concentrations were 0.5 mM (15).

**Diffraction Data Collection**—The crystals of recombinant Pz-peptidase A used for data collection had dimensions of ~1.20 × 0.50 × 0.10 mm. The crystal, in a cryoprotectant consisting of 14% (w/v) PEG 4000, 0.5 M magnesium acetate, 0.1 M Tris-HCl (pH 7.0), and 10% (v/v) isopropanol, was scooped up in a cryoloop, frozen in liquid nitrogen, and then mounted on a goniometer in a nitrogen stream at 93 K. X-ray diffraction was detected on an R-AXIS VII imaging plate system attached to a Rigaku CuKα radiation rotating-anode generator (FR-E) with a crystal-to-detector distance of 120 mm. Data were collected to 1.80 Å resolution (0.5° frames) with an exposure time of 1 min, indexed, and integrated with the MOSFLM program (17) and scaled using the SCALA program from the CCP4 suite (18). The crystal of Pz-peptidase A belongs to the monoclinic space group *P*2₁ with unit-cell parameters *a* = 56.63, *b* = 193.84, *c* = 60.24 Å, and β = 106.54°. Assuming two molecules per asymmetric unit, the calculated Matthews coefficient VM value is 2.73 Å³/Da (19). The solvent content of the crystal was, therefore, calculated to be 48.8%. Data collection statistics are given in Table 1.

**Structure Analysis and Refinement**—Molecular replacement calculations were performed on Pz-peptidase A and the complexes of Pz-peptidase A and either of the two PPIs using the MOLREP program (20). The whole structure of putative oligoendopeptidase F from *Geobacillus stearothermophilus* (Protein Data Bank code 2H1N), which shares 77% amino acid identity with Pz-peptidase A, was employed for phase determination. In generating the search model, the residues that were not identical between sequences were replaced by Ala residues. Employing x-ray diffraction data from the complex of Pz-peptidase A with PPI-1, we found a single solution with a correlation coefficient of 0.367 and *R*-factor of 48.3% (8.0 – 4.0 Å) after the translation-function calculation. To improve the accuracy of the solution, the resultant structure was subjected to 40 cycles of rigid body refinement using data from 8.0 to 3.0 Å resolution. The *R*-factor was refined to 47.2%.

Although the first electron density map obtained by rigid body refinement was insufficient to trace the loop regions of Pz-peptidase A with PPI-1, repeated cycles of model fitting and refinement using the programs REFMAC5 (21) and COOT (22) allowed us to trace the entire Pz-peptidase A. In this step, the alanine-replaced residues were restored to the original amino acid residues. Later refinement steps included the refinement of grouped or individual temperature factors. Omit maps were used to check the model. Water molecules were identified using the *waterpick* procedure in the CNS program (23).

The final model for the complex of Pz-peptidase A with PPI-1 included all 1,128 amino acids and 1,445 water molecules in an asymmetric unit containing two Pz-peptidase A molecules. The final *R*<sub>free</sub> and *R*<sub>f</sub> were 17.4 and 21.4%, respectively. This model had excellent stereochemistry when it was evaluated with the PROCHECK program from the CCP4 suite (24). On the Ramachandran plot, 94.1% of the nonglycine residues are in the most favored region, and none of the residues lie in the generously allowed or disallowed regions. The root mean square deviations of the bond lengths and angles are 0.014 Å and 1.350°, respectively. The structures of Pz-peptidase A without any inhibitors and in complex PPI-2 were solved by employing molecular replacement methods using the structure of Pz-peptidase A with PPI-1 with the MOLREP program from the CCP4 suite (18). Table 1 summarizes the refinement statistics. The final coordinates and structure factors have been deposited in the Protein Data Bank (3AHH for Pz-peptidase, 3AHN for the complex of Pz-peptidase A with inhibitor PPI-1, and 3AHO for the complex of Pz-peptidase A with inhibitor PPI-2).

**RESULTS AND DISCUSSION**

**Overall Structure of Recombinant Pz-peptidase A and Comparison with Thimet Oligopeptidase**—X-ray crystallographic analyses were performed with three species of crystals, recombinant Pz-peptidase A only, and the complexes of the recombinant enzyme with either of the two inhibitors (PPI-1 or -2). Recombinant Pz-peptidase A could crystallize without any genetic modification, even in the presence or absence of PPIs, and their shapes were essentially indistinguishable. The data for the final models are summarized in Table 1.

Pz-peptidase A adopts a globular shape covered with predominant components, 25 α helices and only five short β strands (Fig. 1). This overall fold is classified to an M3 metallopeptidase family containing neurolysin (25) and TOP (14). The overall peptide folds of Pz-peptidase A and TOP are signif-
Recognition of Collagen-specific Sequence by Pz-peptidase A

TABLE 1

X-ray diffraction data statistics
Values in parentheses are for the highest resolution shell. r.m.s.d., root mean square deviation.

|                    | Recombinant enzyme | Complex with PPI-1 | Complex with PPI-2 |
|--------------------|--------------------|--------------------|--------------------|
| **Crystallographic data** |                    |                    |                    |
| Space group        | $P_2_1$            | $P_2_1$            | $P_2_1$            |
| Wavelength (Å)     | 1.5418             | 1.5418             | 1.5418             |
| Unit cell parameters (Å) | $a = 56.63, b = 193.84, c = 60.24; \alpha = \gamma = 90^\circ, \beta = 106.54^\circ$ | $a = 56.38, b = 194.15, c = 59.93; \alpha = \gamma = 90^\circ, \beta = 106.22^\circ$ | $a = 56.81, b = 194.26, c = 60.09; \alpha = \gamma = 90^\circ, \beta = 106.28^\circ$ |
| Completeness (%)   | 98.5 (96.5)        | 96.2 (93.5)        | 99.9 (100.0)       |
| Resolution (Å)     | 20.00-2.00 (2.11-2.00) | 20.00-1.80 (1.90-1.80) | 20.00-1.88 (1.98-1.88) |
| Multiplicity       | 4.5 (4.3)          | 6.0 (5.9)          | 3.0 (2.9)          |
| Rmerge (%)         | 7.6 (36.6)         | 6.2 (17.2)         | 9.1 (22.1)         |
| I/\sigma(I)        | 16.4 (3.0)         | 22.8 (7.0)         | 10.9 (2.9)         |
| **Refinement**     |                    |                    |                    |
| No. of reflections | 78,182             | 104,042            | 95,887             |
| No. of protein atoms | 9383              | 9383              | 9383              |
| No. of inhibitor atoms | 0                | 84                | 86                |
| No. of water molecules | 1281             | 1445              | 1230              |
| Rmerge/Rfree (%)   | 16.9/22.9          | 17.4/21.4          | 17.6/22.1          |
| r.m.s.d. bond length (Å) | 0.020           | 0.014             | 0.018             |
| r.m.s.d. bond angle | 1.70°             | 1.35°             | 1.56°             |
| No. of zinc ions   | 1                  | 1                 | 1                 |

FIGURE 1. Overall structure of Pz-peptidase A with inhibitor PPI-1. A and B, schematic representation of the structure in two different directions. The directions in A and B are perpendicular to each other. These figures, as well as Figs. 2–7, except for Fig. 3, were prepared with the PyMOL program. Twenty five α helices and five β strands are depicted in blue and red, respectively, whereas the loops are in green. The inhibitor PPI-1 molecule, drawn in a stick representation, is in pink. The gray arrow is a direction for Fig. 2.

The most striking difference of Pz-peptidase A from TOP is that there is no channel that runs the length of the molecule, dividing it into two distinct domains (Fig. 2). Whereas the structure of TOP is just like an open bivalve composed of two domains, presumably with a hinge-like motion, the corresponding domain of Pz-peptidase A is totally closed and globular. In contrast, it is remarkable that the overall folds of the secondary elements for Pz-peptidase A and TOP exhibit such similarity (Fig. 3). Whereas TOP contains 24 α helices and nine short β strands, Pz-peptidase A has 25 α helices and five short β strands. Most of the secondary structure elements correspond to each other. It is noteworthy that many deletions occur in Pz-peptidase A relative to TOP, specifically in the loops between α1 and α2, α2 and α3, α3 and α4, α5 and α6, β1 and β2, β2 and α7, α13 and α14, α14 and β3, β4 and β5, β5 and α15, α15 and α16, α17 and α18, α21 and α22, and α22 and α23. These deletions suggest that the loop regions in Pz-peptidase A are much shorter than those in TOP, whereas the secondary elements are comparable in length. Pz-peptidase A from a thermophilic strain G. collagenovorans MO-1, can maintain activity over 30 min at 60 °C, which is far more thermo-
Recognition of Collagen-specific Sequence by Pz-peptidase A

As thermostable proteins tend to have shorter loops to gain more thermostability than their thermostable counterparts, shorter loops in Pz-peptidase A likely function to elevate its thermostability (26).

Comparison with Structural Homologues—In addition to TOP and neurolysin, we searched homologous enzymes that assume similar folds by using DALI (27). Consequently, G. stearothermophilus oligoendopeptidase F (Protein Data Bank code 2H1N, M3B family), human angiotensin-convert- ing enzyme (Protein Data Bank code 1O8A, M2 family) (28, 29), and E. coli peptidyl dipeptidase (Protein Data Bank code 1Y79, M3A family) (30) were screened. The sequence identity of Pz-peptidase A with those enzymes is as high as 77% for G. stearothermophilus oligoendopeptidase F but <5% for the two latter proteins. Their Z-scores are 64.9 for G. stearothermophilus oligoendopeptidase F, 26.9 for human angiotensin-convert- ing enzyme, and 22.6 for E. coli pepsidipidase. The human angiotensin-convert- ing enzyme has a closed globular structure, which is similar to that of Pz- peptidase A even though the proteins belong to different families. The M3A and M3B subfamilies are classified according to their primary sequences. M3A subfamily members, including TOP, neurolysin, and E. coli pepsidipidase, have an active site in the cleft between bivalve struc- tures, consisting of two distinct domains; in contrast, the globular structures are observed among M3B subfamily members, including Pz-peptidase A and G. stearothermophilus oligoendopeptidase F as well as the human angiotensin-converting enzyme of the M2 family. This difference in structure can be employed as a new factor in classification of the subfamilies and analysis of the similarity between the M2 and M3 families. However, it is unlikely that the hinge-like motion of the two domains affects either substrate specific- ity or catalytic activity. By contrast, only the study of an angiotensin-converting enzyme-related carboxypeptidase belonging to the M2 family revealed a large hinge-bending motion of 16° in response to the entry of a substrate analogue inhibitor into the active site (31). On the other hand, Tryp- anosoma cruzi metallocarboxypeptidase, belonging to the M14 family (32), also is reported to have similar behavior; therefore, the bivalve-structured proteins are proposed to be classified into the cowrin family, named after cowry shells (33).

Active Site and Gateway for Substrates—From the crystal structure of recombinant Pz-peptidase A, we found that the active site of Pz-peptidase A is located in the center of the mol- ecule surrounded by secondary elements and that the three residues of a typical zinc metallopeptidase motif HEXXH (His\textsuperscript{356}, Glu\textsuperscript{357}, and His\textsuperscript{360}) (34) reside in the middle of helix α15, interacting with one zinc atom and two water molecules (Fig. 4). The third ligand, Glu\textsuperscript{384}, protruding from helix α16, also was observed to coordinate with the zinc atom. One water molecule forming an H-bond with Glu\textsuperscript{357} (Fig. 4, Water 1) most likely participates in the process of the reaction, whereas the other water molecule (Fig. 4, Water 2) is near the active site. The gateways for substrates in Pz-peptidase A are found at the top and bottom of the protein molecule, shown in Fig. 1A and Fig. 5. The top gateway is an open space surrounded by two helices, α2 and α15, and the loop between strands β4 and β5, whereas the bottom gateway is a space formed by the loop between α12 and α13 and the loop spanning α19, α20, and helix α12. The two gateways are connected in a tunnel, wherein the active site equipped with the HEXXH motif, one zinc atom, and the third Glu ligand is ready for hydrolysis at its inner surface (Fig. 5). The top gateway is larger than the bottom one; thus, the top gateway works as an entrance for longer oligopeptide sub- strates, whereas the bottom gateway works as an exit for shorter oligopeptides after cleavage at the active site. In the M3A family, a gating mechanism of the bivalve structure induced by the substrate uptake would be advantageous for substrate binding as well as hydrolysis efficiency. Thus, the tunnel structure of Pz-peptidase A, which belongs to the M3B family, seems more disadvantageous for substrate uptake than the bivalve struc- ture. However, the $K_{\text{m}}$ value for Pz-peptidase A is 8.2 μM for the substrate 7-methoxycoumarin-3-carboxyl-pro-Leu-Gly-Pro- t-Lys(2,4-dinitrophenyl), which is compatible with the figures (8.6 μM) for TOP (16, 35). The tunnel structure of Pz-peptidase A can function as a gateway for substrates in a comparable manner with that of M3A family enzymes such as TOP. The gateway most likely allows linear oligopeptides to enter the active site. In addition, the gateway and the tunnel are too small for intact proteins to get in, which is in good agreement with the fact that Pz-peptidase A never shows activity toward whole col- lagens (13).

Enzyme Inhibitor and Collagen-specific Sequence Recognition in the Active Site—In place of crystallographic analysis with substrates, an unstructured peptide, neurotensin, which is a 13-residue substrate, was employed to model the substrate recog- nition in the active site of TOP by theoretical fitting and energy minimization (14). On the other hand, we have found that Pz-peptidase A is sensitive to PPIs, which specifically inhibit TOP and neurolysin (16, 36, 37) by a similar mechanism and are applicable to x-ray analyses (38). Inhibitors PPI-1 and PPI-2 potently inhibit Pz-peptidase A with $K_i = 90.1$ and 88.9 nM, respectively (16), whereas PPI-2 inhibits neurolysin with $K_i = 0.4$ nM, more strictly and specifically than TOP ($K_i = 21.0$ nM) (36, 37). It should be noted that PPI-2 is a better molecule for

![Pz peptidase A and TOP](image-url)
FIGURE 3. Secondary structure alignment of Pz-peptidase A and human TOP. Conserved sequences are highlighted in black, and semiconservative ones are boxed. Deleted sequences are shown with dots. Secondary structure elements are depicted schematically (coil, α helix; arrow, β strand) above for Pz-peptidase A and below for human TOP.
understanding the enzyme-substrate interaction and hydrolysis of collagenous peptides because it contains two instances of the collagen-specific sequence, -Gly-Pro-X-. We succeeded in obtaining crystals of Pz-peptidase A in complex with PPI-1 and PPI-2 under the same crystallization conditions reported previously (15).

The two inhibitors were trapped equivalently at the active site through hydrophobic and H-bonds but not through covalent bonds. As shown in Fig. 6, the N terminus of PPI-2 lies at the active site, oriented toward the top. In the crystal of Pz-peptidase A with the PPI-2 molecule, His356, Glu357, His360, and Glu384 reside around a zinc atom in the active site within H-bonding distance (Fig. 6). In addition to their participation, H-bonds link the side chains of the other residues in the active site of Pz-peptidase A with the PPI-2 molecule. An oxygen atom of the phenol ring of Tyr486 forms an H-bond with a nitrogen atom of the phenylalanine residue of PPI-2. Two oxygen atoms of the phosphoryl group in the inhibitor connect with the phenol ring of Tyr490 and the carboxylic group of Glu357, respectively, via H-bonds. Glu357 binds to the inhibitor by displacing a water molecule present in the recombinant Pz-peptidase A structure (Fig. 4). The carbonyl oxygen of a glycine residue of the inhibitor near the C terminus has two H-bonds with the imidazole ring of His479 and the phenol ring of Tyr490. Furthermore, the carbonyl oxygen atom of a proline residue following the second glycine donates one H-bond to the phenol ring of Tyr487.

It is noteworthy that three Tyr residues (Tyr486, Tyr487, and Tyr490) in the active site are responsible for binding the inhibitor PPI-2. Their equivalent residues (Tyr609, Tyr610, and Tyr612) are conserved in TOP (Fig. 3). The interactions of Tyr residues with the substrate were predicted on the basis of modeling the complex of TOP with neurotensin and primary sequence comparison of Pz-peptidases with TOP (13, 14). In addition, very recently, the roles of two Tyr residues (Tyr605 and Tyr612) in TOP were investigated and suggested to play key roles in stabilizing the transition state, as well as accommodating various substrate structures in collaboration with adjacent Gly residues, using H-bonds (39). Unfortunately, there is no residue corresponding to Tyr605 in Pz-peptidase A. However, in the structure of Pz-peptidase A, we found that the Tyr residues exhibit remarkable interactions with glycine and proline residues following a phosphinic group (PO2CH2) contained in the inhibitor. The phosphinic group occurs at the closest position to a zinc atom in the active site. This location agrees with the case involving TOP, in which the group corresponds to the scissile site and the P1’ position of peptide substrates with the following glycine residue (35, 36). The phenol rings of Tyr487 and Tyr490 of Pz-peptidase A interact with the main chains of proline and glycine residues at the S2′ and S1′.
subsites, respectively, through H-bonds, whereas the ring of Tyr486 makes contact with the nitrogen atom of phenylalanine of PPI-2 at the S1 subsite. The three Tyr residues probably contribute to the elevation in affinity of the collagen-specific sequence Gly-Pro for these subsites through their hydrophobic side chains. In particular, there is a preference for Gly at the P1/H1 position, judging from two H-bonds between the phenol ring of Tyr490 and the main chain of the inhibitors surrounding a residue at the P1/H1 position. It is possible that the two H-bonds are formed due to lower steric hindrance only when a residue with a smaller side chain, such as Gly or Ala, occurs at the P1/H1 position, as seen in PPI-1 and PPI-2. In addition, there is no H-bond between the nitrogen atom of proline at the P2/H2 position and active site residues, because the proline residue is an “imino” acid and has no hydrogen available on the nitrogen atom. Interestingly, in the crystal structure of the complex with PPI-1, there is no H-bond between the corresponding nitrogen atom of PPI-1, as though the subsite could fit a proline residue better served in TOP (Fig. 3). Residue Trp377 is located in the vicinity of the N terminus of the two inhibitors and plays a role in the restriction of substrate chain length as well as accommodating the repetitive Gly-Pro sequences on the active site. Fig. 7 shows that the indole ring of Trp377 rotates ~90° toward the N terminus Gly-Pro residues in inhibitor PPI-2, whereas the orientation of the ring in the complex with PPI-1 is the same as that in recombinant Pz-peptidase A without any inhibitors. The indole ring of Trp377 is directed apart from inhibitor PPI-1, which has a noncollagenous sequence, but it gets closer to the first Gly-Pro residues of PPI-2, which contain a collagenous sequence. The change of the indole ring of Trp377 in position does not result in any movement by other residues around the active site. Considering all of these findings, Trp377 specifically hydrolyzes collagenous peptides of appropriate length.

To break down native collagens, it is necessary to perform at least three steps: binding, unwinding, and hydrolysis (40). Binding of collagenous peptides at the active site has been discussed
above. As one additional element, Cys\textsuperscript{329} forms an H-bond with the carbonyl oxygen atom of a residue at the P\textsubscript{2} position in the crystal structures of two complexes with two respective inhibitors. This residue participates in binding collagenous peptides, as in the case of TOP, in which the equivalent cysteine residue is conserved at position 427 (Fig. 3). On the other hand, the machinery for unwinding collagens is not required for Pz-peptidase because these proteins are involved only in the hydrolysis of shorter collagenous peptides (15-residue length at most). As for the hydrolytic mechanism of Pz-peptidase A acting on collagenous peptide fragments containing a Gly-Pro sequence, a zinc atom and one zinc-bound water molecule play an essential role. The acidic residue Glu\textsuperscript{357} acts as a general base and generates the hydroxide nucleophile from the zinc-bound water (Fig. 4, Water 1); the nucleophilic hydroxide then attacks the carbonyl carbon of the target peptide bond between Gly and its preceding residues in the substrate, leading to the formation of a gem-diolate intermediate. Subsequently, the nitrogen atom of the Gly residue gains a hydrogen atom trapped in Glu\textsuperscript{357} and Tyr\textsuperscript{496}, which is H-bonded to the phosphinic group in the complexes with two inhibitors, can act as the second catalytic residue by forming H-bonds with the substrate carbonyl oxygen and zinc atoms in the active site.

In summary, we suggest a reaction mechanism in which collagenous peptide substrates enter the tunnel from the top gateway of the Pz-peptidase A molecule (Fig. 5), and reactant peptides are subsequently released from the bottom gateway after cleavage at the active site located in the center of the tunnel, following the reaction steps indicated above. As demonstrated in our previous report (13), the Pz-peptidase A is rather specific for shorter peptide substrates and collagenous peptides, but cannot act on longer hormone peptides. The fact that Pz-peptidase A can hydrolyze the peptides containing hydroxyproline and the distinct difference in $K_i$ values suggest that Pz-peptidase A is less specific than TOP.

Another characteristic feature of Pz-peptidase A is that each molecule contains six Cys residues. This is quite peculiar because thermostable enzymes from thermophilic bacteria tend to have a smaller number of Cys residues in the molecules, due to their high sensitivity to oxidation at high temperature (41). In comparison, human TOP has 16 Cys residues and has tend to have a smaller number of Cys residues in the molecules, because thermostable enzymes from thermophilic bacteria

**REFERENCES**

1. Nimni, M. E. (1983) *Semin. Arthritis Rheum.* **13**, 1–86
2. Suzuki, Y., Tsjijimoto, Y., Matsu, H., and Watanabe, K. (2006) *J. Biosci. Bioeng.* **102**, 73–81
3. Linsenmayer, T. F. (1991) in *Cell Biology of Extracellular Matrix* (Hay, E. D., ed.) pp. 7–44, Plenum Press, New York
4. Myers, L. K., Tang, B., Rosloniee, E. F., Stuart, J. M., Chiang, T. M., and Kang, A. H. (1998) *J. Immunol.* **161**, 3589–3595
5. Harrington, D. J. (1996) *Infect. Immun.* **64**, 1885–1891
6. Peterkofsky, B. (1982) *Methods Enzymol.* **82**, 453–471
7. Watanabe, K. (2004) *Appl. Microbiol. Biotechnol.* **63**, 520–526
8. Barrett, A. J. (1991) *Biochem. J.* **277**, 295–296
9. Barrett, A. J., and Tsiljar, U. (1992) *Matrix Biol.* **1**, 95–96
10. Rock, K. L., York, I. A., and Goldberg, A. L. (2004) *Nat. Immunol.* **5**, 670–677
11. Okamoto, M., Yonejima, Y., Tsjijimoto, Y., Suzuki, Y., and Watanabe, K. (2001) *Appl. Microbiol. Biotechnol.* **57**, 103–108
12. Ito, Y., Horinaka, M., Tsjijimoto, Y., Matsu, H., and Watanabe, K. (2006) *J. Bacteriol.* **188**, 6572–6579
13. Miyake, R., Shigeri, Y., Tatsu, Y., Yamoto, N., Umekawa, M., Tsjijimoto, Y., Matsu, H., and Watanabe, K. (2005) *J. Bacteriol.* **187**, 4140–4148
14. Ray, K., Hines, C. S., Coll-Rodriguez, J., and Rodgers, D. W. (2004) *J. Biol. Chem.* **279**, 20480–20489
15. Kawasaki, A., Nakano, H., Umekawa, M., Tsjijimoto, Y., Matsu, H., Shimizu, T., Nakatsu, T., Kato, H., and Watanabe, K. (2007) *Acta Crystallogr. F Struct. Biol. Cryst. Commun.* **63**, 142–144
16. Sugihara, Y., Kawasaki, A., Tsjijimoto, Y., Matsu, H., and Watanabe, K. (2007) *Biosci. Biotechnol. Biochem.* **71**, 594–597
17. Leslie, A. G. (1999) *Joint CCP4 and ESF-EACMB Newsletter on Protein Crystallography* (SERC Daresbury Laboratory, Warrington, UK)
18. Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. D Biol. Crystallogr.* **50**, D760–763
19. Matthews, B. W. (1968) *J. Mol. Biol.* **33**, 491–497
20. Vagin, A., and Tsyplakov, A. (2000) *Acta Crystallogr. D Biol. Crystallogr.* **56**, 1622–1624
21. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255
22. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132
23. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921
24. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26**, 283–293
25. Brown, C. K., Madaus, K., Lian, W., Beck, M. R., Tolbert, W. D., and Rodgers, D. W. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3127–3132
26. Corazza, A., Rosano, C., Pagano, K., Alverdi, V., Esposito, G., Capanni, C., Bemparad, F., Plakoutsi, G., Stefani, M., Chiti, F., Zuccotti, S., Bolognesi, M., and Vignolo, P. (2006) *Proteins* **62**, 64–79
27. Dali-Holm, L., and Sander, C. (1993) *J. Mol. Biol.* **233**, 123–138
28. Natesh, R., Schwager, S. L., Sturrock, E. D., and Acharya, K. R. (2003)
Nature 421, 551–554
29. Rawlings, N. D., Barrett, A. J., and Bateman, A. (2010) Nucleic Acids Res. 38, D227–233
30. Comellas-Bigler, M., Lang, R., Bode, W., and Maskos, K. (2005) J. Mol. Biol. 349, 99–112
31. Towler, P., Staker, B., Prasad, S. G., Menon, S., Tang, J., Parsons, T., Ryan, D., Fisher, M., Williams, D., Dales, N. A., Patane, M. A., and Pantoliano, M. W. (2004) J. Biol. Chem. 279, 17996–18007
32. Niemirowicz, G., Fernández, D., Solà, M., Cazzulo, J. J., Avilés, F. X., and Gomis-Ruth, F. X. (2008) Mol. Microbiol. 70, 853–866
33. Gomis-Ruth, F. X. (2008) Crit. Rev. Biochem. Mol. Biol. 43, 319–345
34. Rawlings, N. D., and Barrett, A. J. (1995) Methods Enzymol. 248, 183–228
35. Tisdar, U., Knight, C. G., and Barrett, A. J. (1990) Anal. Biochem. 186, 112–115
36. Jirácek, J., Yiotakis, A., Vincent, B., Checler, F., and Dive, V. (1996) J. Biol. Chem. 271, 19606–19611
37. Jirácek, J., Yiotakis, A., Vincent, B., Lecoq, A., Nicolaou, A., Checler, F., and Dive, V. (1995) J. Biol. Chem. 270, 21701–21706
38. Cámara-Artigas, A., Jara-Pérez, V., and Andújar-Sánchez, M. (2009) in Drug Design of Zinc-Enzyme Inhibitors (Surpran, C. T., and Winum, J. Y., eds) pp. 751–787, John Wiley & Sons, Inc., Hoboken, New Jersey
39. Bruce, L. A., Sigman, J. A., Randall, D., Rodriguez, S., Song, M. M., Dai, Y., Elmore, D. E., Pabon, A., Glucksman, M. J., and Wolfson, A. J. (2008) FEBS J. 275, 5607–5617
40. Gioia, M., Monaco, S., Fasciglione, G. F., Coletti, A., Modesti, A., Marini, S., and Coletta, M. (2007) J. Mol. Biol. 368, 1101–1113
41. Vieille, C., and Zeikus, G. J. (2001) Microbiol. Mol. Biol. Rev. 65, 1–43
42. Philominathan, S. T., Koide, T., Hamada, K., Yasui, H., Seifert, S., Matsushita, O., and Sakon, J. (2009) J. Biol. Chem. 284, 10868–10876
43. Petrek, M., Otyepka, M., Banáš, P., Kosinová, P., Koca, J., and Damborský, J. (2006) BMC Bioinformatics 7, 316

Recognition of Collagen-specific Sequence by Pz-peptidase A