The Acylaminoacyl Peptidase from *Aeropyrum pernix* K1 Thought to Be an Exopeptidase Displays Endopeptidase Activity

András L. Kiss¹, Balázs Hornung², Krisztna Rádi², Zsolt Gengeliczi² Bálint Sztáray², Tünde Juhász¹, Zoltán Szeltner¹, Veronika Harmat³ and László Polgár¹*

¹Institute of Enzymology, Biological Research Center, Hungarian Academy of Sciences, H-1518 Budapest 112, P.O. Box 7, Hungary
²Eötvös Loránd University, Institute of Chemistry, H-1117, Budapest, Pázmány P. stny. 1/A, Hungary
³Hungarian Academy of Sciences- Eötvös Loránd University, Protein Modeling Research Group, H-1117 Budapest, Pázmány P. stny. 1/A, Hungary

Mammalian acylaminoacyl peptidase, a member of the prolyl oligopeptidase family of serine peptidases, is an exopeptidase, which removes acylated amino acid residues from the N terminus of oligopeptides. We have investigated the kinetics and inhibitor binding of the orthologous acylaminoacyl peptidase from the thermophile *Aeropyrum pernix* K1 (ApAAP). Complex pH-rate profiles were found with charged substrates, indicating a strong electrostatic effect in the surroundings of the active site. Unexpectedly, we have found that oligopeptides can be hydrolysed beyond the N-terminal peptide bond, demonstrating that ApAAP exhibits endopeptidase activity. It was thought that the enzyme is specific for hydrophobic amino acids, in particular phenylalanine, in accord with the non-polar S1 subsite of ApAAP. However, cleavage after an Ala residue contradicted this notion and demonstrated that P1 residues of different nature may bind to the S1 subsite depending on the remaining peptide residues. The crystal structures of the complexes formed between the enzyme and product-like inhibitors identified the oxyanion-binding site unambiguously and demonstrated that the phenylalanine ring of the P1 peptide residue assumes a position different from that established in a previous study, using 4-nitrophenylphosphate. We have found that the substrate-binding site extends beyond the S2 subsite, being capable of binding peptides with a longer N terminus. The S2 subsite displays a non-polar character, which is unique among the enzymes of this family. The S3 site was identified as a hydrophobic region that does not form hydrogen bonds with the inhibitor P3 residue. The enzyme–inhibitor complexes revealed that, upon ligand-binding, the S1 subsite undergoes significant conformational changes, demonstrating the plasticity of the specificity site.

© 2007 Elsevier Ltd. All rights reserved.

*Corresponding author

Keywords: thermophilic oligopeptidase; kinetics; substrate specificity; inhibitor binding; X-ray crystallography

Introduction

Acylaminoacyl peptidase (AAP) catalyses the removal of N-terminally blocked amino acid residues from peptides. The products of the reaction are an acyl amino acid and a peptide with a free N terminus shortened by one amino acid residue. The enzyme acts on peptides with different N-terminal acyl groups, like formyl, acetyl, chloroacetyl, carbamyl and other groups. However, N-terminally blocked proteins are not substrates for AAP.¹ AAP

Abbreviations used: AAP, acylaminoacyl peptidase; ApAAP, *Aeropyrum pernix* K1 acylaminoacyl peptidase; Ac, acetyl; Z, benzyloxy carbonyl; Abz, 2-aminobenzoyl; F(NO₂) or Phe(NO₂), 4-nitrophenylalanine; FAPα, fibroblast activation protein α; Nap, 2-naphthylamide; Nan, 4-nitroanilide; Amc, 7-amino-4-methylcoumarin; Nle, norleucine.

E-mail address of the corresponding author: polgar@enzim.hu

0022-2836/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
is a member of the prolyl oligopeptidase family of serine peptidases (clan SC, family S9). The family includes dipeptidyl-peptidase IV, fibroblast activation protein α, oligopeptidase B and the archetype prolyl oligopeptidase, and are all of pharmaceutical importance. The evolutionary relationship of the enzymes of the prolyl oligopeptidase family has been demonstrated recently using a phylogenetic tree. The crystal structure of prolyl oligopeptidase was the first in the family to be solved. It was crystallized, but their structures remain to be solved. The crystal structure of AAP from the thermophile *Aeropyrum pernix* has been demonstrated recently using a phylogenetic tree.

The crystal structure of prolyl oligopeptidase was the first in the family to be solved. It was shown that the enzyme contains a peptidase domain with an α/β hydrolase fold, and that its catalytic triad (Ser554, His680, Asp641) is covered by the central tunnel of an unusual β-propeller. This domain makes prolyl oligopeptidase an oligopeptidase by excluding large, structured peptides from the active site. In this way, the propeller protects large peptides and proteins from proteolysis in the cytosol.

Human and porcine AAP have been crystallized, but their structures remain to be solved. The crystal structure of AAP from the thermophile *Aeropyrum pernix* K1 (ApAAP) has been reported recently. Unlike the tetramer mammalian enzyme, ApAAP is a symmetrical homodimer with each subunit containing two domains: the N-terminal seven-bladed β-propeller and the peptidase domain with an α/β-hydrolase fold, characteristic of this enzyme family. The catalytic triad of ApAAP consists of Ser445, Asp524 and His556. The structure of the enzyme complexed with 4-nitrophenylphosphate delineated the hydrophobic S1 substrate-binding pocket that accepts large non-polar residues, like phenylalanine and leucine. This binding site is different from that of the porcine enzyme, which is specific for the small alanine side-chain. To reveal the binding mode of true peptide substrates, we used product-like inhibitors for structure determination and several oligopeptides for kinetic analysis. Surprisingly, the enzyme, which has been regarded as an exopeptidase, proved to be an endopeptidase as well.

### Results and Discussion

**The pH-rate profile is dependent on the charge of the substrate**

The study of the pH-dependence of rate constants can furnish useful mechanistic information about the participation of ionizing groups in catalysis. The classical example is the involvement of a histidine residue in the catalysis by chymotrypsin. Here, we determined the pH dependence of $k_{cat}/K_m$, which is characteristic of the ionizing groups of the substrate and the free enzyme. Figure 1 illustrates the reactions of ApAAP with Ac-Phe-Nap and Gly-Phe-Nap, the simplest substrates of ApAAP. This shows that both profiles deviate slightly from the bell-shaped curve (dotted lines, equation (1)), the alkaline limbs being steeper than expected from the dissociation of a single acid. The points fit better to a curve that delineates the dissociation of two-protons in the alkaline pH range (equation (2)), although the differences are almost negligible. The groups with $pK_a$ and $pK_b$ (Table 1) may reflect changes in the active site conformation of the enzyme. The inactivation of the enzyme at high pH is a reversible process. The enzyme was incubated at pH 10 for longer than the activity measurement, but still recovered full activity at neutral pH.

Although the two substrates shown in Figure 1 display rate constants of similar magnitude, their pH-dependence differs, as the active pH-dependent range for the Gly-Phe-Nap substrate significantly narrower. The difference between the $pK_a$ values (1.14 units) is particularly remarkable (Table 1). The $pK_a$ of 6.09 is likely consistent with the ionization of the catalytic His556. This value determined at 70 °C is somewhat lower than usual for a serine peptidase at 25 °C, but the $pK_b$ of the histidine residue diminishes with the increase in temperature. As $d(pK_a)/dT = -0.020$ for the imidazole, the $pK_a$ would be 0.9 unit greater at 25 °C, i.e. 6.99, a reasonable value for a catalytic histidine residue. For the porcine AAP, the $pK_a$ is 7.01. In the case of Gly-Phe-Nap, the ionization of the glycine amino group overrules that of the histidine. The $pK_a$ of 7.2 (Table 1) is an apparent $pK_a$, which may result from overlapping ionizations of His556 and the substrate

**Figure 1.** pH-rate profiles for ApAAP. The reactions were performed with Ac-Phe-Nap (○) and Gly-Phe-Nap (△) in the presence of 0.3 M NaCl. The continuous lines represent simple bell-shaped curves.

**Table 1.** Kinetic parameters for the reactions of ApAAP

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Ac-Phe-Nap</th>
<th>Gly-Phe-Nap</th>
<th>Abz-GFEX</th>
<th>Abz-EFSX</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{cat}/K_m$ (limit)</td>
<td>1.52±0.05</td>
<td>0.766±0.039</td>
<td>2.24±0.05</td>
<td>0.133±0.002</td>
</tr>
<tr>
<td>$k_{cat}/K_m$ (limit)</td>
<td>6.09±0.06</td>
<td>7.23±0.05</td>
<td>5.75±0.04</td>
<td>5.55±0.06</td>
</tr>
<tr>
<td>$pK_a$</td>
<td>9.18±0.18</td>
<td>9.08±0.14</td>
<td>8.58±0.06</td>
<td>9.15±0.10</td>
</tr>
<tr>
<td>$pK_b$</td>
<td>8.77±0.29</td>
<td>9.07±0.23</td>
<td>9.01±0.15</td>
<td>8.52±0.15</td>
</tr>
</tbody>
</table>

$^a$ X=PF(NO₃)³RA.

$^b$ Rate-limiting value in the acidic range.
amino group. Obviously, the binding is better with the neutral form than with the charged amino group, and this may be due to an unfavourable electrostatic environment. In the case of the related prolyl oligopeptidase, the electrostatic environment of the active site region also considerably affected the pH-rate profiles determined with various charged substrates.13

ApAAP displays both exo- and endopeptidase activities, as well as sequence-dependent S1 specificity

Perusal of the crystal structure of ApAAP has shown that there is enough space for the binding of peptides with more N-terminal residues than the Gly of the Gly-Phe-Nap substrate. This observation was confirmed by the hydrolysis of two peptides with Abz-Gly or Abz-Glu preceding the P1 phenylalanine (Figure 2(a)). Substrate Abz-GFEPF(NO2)RA displayed greater kinetic specificity than Ac-Phe-Nap, the classical substrate of ApAAP, whereas the pH-rate profiles were similar and exhibited identical pK1 values (Table 1). This indicates that the enzyme can hydrolyse a peptide with an extended N terminus, which is inconsistent with the classification of ApAAP as an exopeptidase. The kcat/Km for the Abz-EFSPF(NO2)RA peptide is lower by one order of magnitude and does not follow the ionization of a single base in the acidic region. The relatively higher rate constant in the acidic region is very likely due to the change in ionization of the P2 glutamic acid. The negative charge on the side-chain apparently counteracts the catalysis, but this adverse effect reduces with the decrease in pH. No such effect was observed with the P1’ glutamic acid of the Abz-GFEPF(NO2)RA substrate.

The electrostatic effects between enzyme and substrate can be reduced significantly or abolished completely at high concentrations of salt. Such a change was observed in the pH-rate profile for the reaction of prolyl oligopeptidase with succinyl-Gly-Pro-Nan.14 The pH-rate profiles for the two charged substrates shown in Figure 2(a) were determined in the presence of 0.3 M NaCl, in order to partially quench the electrostatic effects. In the absence of the salt, the electrostatic effects were still discernible and the pH-rate profiles were altered substantially (Figure 2(b)). The rate constants for the Abz-EFSPF(NO2)RA increased considerably in the acidic range, whereas those at the maximum activity decreased. When the ionic strength was further reduced by using a tenfold-diluted buffer (3 mM), the change was even more remarkable (data not shown). Different results were obtained with Abz-GFEPF(NO2)RA, inasmuch as the rate constants did not increase in the acidic pH range. It can be concluded that the ionized glutamic acid at position P2 is rather unfavourable, while its neutral form at low pH proved to be more specific. In contrast, the P1’ glutamic acid did not exert such an effect. Our crystal complexes containing Gly at the P2 site indicate that a hydrogen bond is formed between the main chain carbonyl oxygen of the P2 residue and the side-chain of Arg326. (See Crystal structures of ApAAP–inhibitor complexes reveal the oxyanion-binding site and the specificities of the S1, S2 and S3 subsites). This stabilization of the substrate main chain constrains the P2 side-chain to a hydrophobic patch lined by Phe residues. Such an environment is unfavourable for charged residues, which is consistent with the electrostatic effects on the activity (Figure 2(b)). In accordance with the above results, there is only a small effect of salt on the pH-rate profile of the neutral Ac-Phe-Nap hydrolysis.

The hydrolyses of substrates having longer N-terminal acyl moieties were examined (Table 2). The glutaryl-GGF-Amc had a rate constant comparable to that of the classic Ac-Phe-Nap, whereas the succinyl-AAPF-Nan substrate was hydrolysed at a significantly slower rate. However, the Abz-SAVLQSGF(NO2)A peptide possessing a potential P1 leucine, proved to be a good substrate. It was suggested that Leu is a specific residue as well as Phe.9,15 However, we have found that the specificity rate constant is lower by one order of magnitude for Ac-Leu-Nan than for Ac-Phe-Nan (Table 2). Also, the
Abz-KARVLf(NO2)EANle was a poor substrate compared with the good substrate Abz-GFEPF(NO2)RA.

Table 2 shows that the ApAAP hydrolyses the Ac-Ala-Amc, which is the classic substrate of the porcine enzyme, very slowly. The rate constant is more than three orders of magnitude lower than the $k_{cat}/K_m$ for Ac-Phe-Nap. The rate constants in Table 2 show the remarkable difference between the mesophilic and the thermophilic enzymes. Thus, Gly-Phe-Nap is a much worse substrate for the porcine AAP compared to the Ac-Phe-Nap, indicating that the porcine enzyme requires a substrate with a blocked N terminus, as expected from a true AAP. In addition, the mammalian AAP does not show an appreciable endopeptidase activity, as demonstrated with four oligopeptides (Table 2).

Indeed, acetylation of the terminal NH2 group of intracellular proteins occurs frequently in eukaryotes but rarely in more ancient cells.16 It appears that AAP originally was less specialized and became a true AAP only at a later stage of evolution. The evolutionary relationship between the thermophile and mesophilic enzymes is supported by Figure 3, showing the amino acid sequence alignment of the two peptidase domains. The identical locations of the catalytic triads and the glycine residues of the oxyanion-binding site are clearly seen (red letters).

The residues of ApAAP involved in substrate binding are also shown (dark green), two of them being identical: Met477 of ApAAP and Met619 of porcine AAP, and Arg 526 of ApAAP and Arg677 of porcine AAP. Because of the different specificities of the two enzymes, identity of the binding sites cannot be expected. A similar result was obtained recently with FAPα, a close relative of the genuine exopeptidase dipeptidyl-peptidase IV of the prolyl oligopeptidase family. The aspartic acid that binds the N-terminal amino group of the substrate in dipeptidyl-peptidase IV corresponds to an alanine in FAPα, thereby reducing the exopeptidase and generating a significant endopeptidase activity.

It follows from the above data that ApAAP is able to hydrolyse oligopeptides as an endopeptidase. This is evident from the hydrolysis of Abz-EALFQGPF(NO2)A, a very good substrate (Table 2), in which the specific Phe residue occupies the fifth position from the N terminus. We have examined a tridecapeptide, Abz-RPIITTAGPSF(NO2)A, which is a good substrate (Table 2).

Table 2. Specificity rate constants for AAPs

<table>
<thead>
<tr>
<th></th>
<th>$k_{cat}/K_m$ (μM$^{-1}$s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ApAAP</td>
<td>Porcine AAP</td>
</tr>
<tr>
<td>Abz-GFEPF(NO2)RA</td>
<td>2.96</td>
</tr>
<tr>
<td>Abz-EFSPF(NO2)RA</td>
<td>0.21</td>
</tr>
<tr>
<td>Abz-GFSPF(NO2)RA</td>
<td>1.92</td>
</tr>
<tr>
<td>Abz-SAYLQSGF(NO2)A</td>
<td>1.24</td>
</tr>
<tr>
<td>Abz-KARVLf(NO2)EANle</td>
<td>0.048</td>
</tr>
<tr>
<td>Abz-RIPTTAGPSF(NO2)A</td>
<td>0.031</td>
</tr>
<tr>
<td>Abz-EALFQGPF(NO2)A</td>
<td>3.22</td>
</tr>
<tr>
<td>Glutaryl-GGF-Amc</td>
<td>0.42</td>
</tr>
<tr>
<td>Succinyl-AAPF-Nan</td>
<td>0.023</td>
</tr>
<tr>
<td>Ac-Ala-Amc</td>
<td>0.00054</td>
</tr>
<tr>
<td>Ac-Phe-Nap</td>
<td>1.50</td>
</tr>
<tr>
<td>Gly-Phe-Nap</td>
<td>0.86</td>
</tr>
<tr>
<td>Ac-Phe-Nan</td>
<td>1.20</td>
</tr>
<tr>
<td>Ac-Leu-Nan</td>
<td>0.14</td>
</tr>
</tbody>
</table>

The reactions were measured at 0.4 μM substrate in 50 mM phosphate buffer (pH 8.0) containing 1 mM EDTA and 0.3 M NaCl.

- In 1% DMSO, 70 °C.
- In 0.02% DMSO, 25 °C.
- Measured photometrically at 14 μM substrate.

Abz-KARVLf(NO2)EANle was a poor substrate compared with the good substrate Abz-GFEPF(NO2)RA.

Table 2 shows that the ApAAP hydrolyses the Ac-Ala-Amc, which is the classic substrate of the porcine enzyme, very slowly. The rate constant is more than three orders of magnitude lower than the $k_{cat}/K_m$ for Ac-Phe-Nap. The rate constants in Table 2 show the remarkable difference between the mesophilic and the thermophilic enzymes. Thus, Gly-Phe-Nap is a much worse substrate for the porcine AAP compared to the Ac-Phe-Nap, indicating that the porcine enzyme requires a substrate with a blocked N terminus, as expected from a true AAP. In addition, the mammalian AAP does not show an appreciable endopeptidase activity, as demonstrated with four oligopeptides (Table 2).

Indeed, acetylation of the terminal NH2 group of intracellular proteins occurs frequently in eukaryotes but rarely in more ancient cells.16 It appears that AAP originally was less specialized and became a true AAP only at a later stage of evolution. The evolutionary relationship between the thermophile and mesophilic enzymes is supported by Figure 3, showing the amino acid sequence alignment of the two peptidase domains. The identical locations of the catalytic triads and the glycine residues of the oxyanion-binding site are clearly seen (red letters). The residues of ApAAP involved in substrate binding are also shown (dark green), two of them being identical: Met477 of ApAAP and Met619 of porcine AAP, and Arg 526 of ApAAP and Arg677 of porcine AAP. Because of the different specificities of the two enzymes, identity of the binding sites cannot be expected. A similar result was obtained recently with FAPα, a close relative of the genuine exopeptidase dipeptidyl-peptidase IV of the prolyl oligopeptidase family. The aspartic acid that binds the N-terminal amino group of the substrate in dipeptidyl-peptidase IV corresponds to an alanine in FAPα, thereby reducing the exopeptidase and generating a significant endopeptidase activity.

It follows from the above data that ApAAP is able to hydrolyse oligopeptides as an endopeptidase. This is evident from the hydrolysis of Abz-EALFQGPF(NO2)A, a very good substrate (Table 2), in which the specific Phe residue occupies the fifth position from the N terminus. We have examined a tridecapeptide, Abz-RPIITTAGPSF(NO2)A, which is a good substrate (Table 2).
(NO₂)A, which was hydrolysed by ApAAP but practically not by the porcine enzyme, although the substrate contains a specific alanine (residue 8). This indicates that the porcine AAP is indeed an exopeptidase, which is further supported by the study on the hydrolysis of α melanocyte-stimulating hormone, which was cleaved only after the N-terminal Ac-Ser by rat liver AAP.¹⁸

When internally quenched fluorescent substrates are used, the detectable cleavage(s) must occur between the fluorophore (Abz) and the quencher (F(NO₂)) of its fluorescence. However, ApAAP having endopeptidase activity may also hydrolyse the peptide bond after the hydrophobic Phe(NO₂) residue. The occurrence of such a cleavage was probed with a spectrophotometric method measuring the increase in absorbance at 300 nm, related to the hydrolysis of the scissile peptide bond.¹⁹ If the cleavage takes place at the NH group of Phe(NO₂), the absorbance decreases during the reaction, and increases if the hydrolysis occurs at the carboxyl group.²⁰,²¹ We did not observe a significant absorbance change with the Abz-GFEPF(NO₂)RA, the Abz-EFSPF(NO₂)RA or the Abz-SAVLQSGF(NO₂)A substrates, while the fluorometric reactions proceeded to completion. This indicated that neither the F(NO₂)–X nor the X–F(NO₂) bond was hydrolysed to a noticeable extent. However, the absorbance increased during the hydrolysis of Abz-KARVLF(NO₂)JEANle, which yielded a rate constant of 1.3 times higher than that found with the fluorometric method. This revealed that at least two parallel reactions took place, and the cleavage at the Phe (NO₂) residue undoubtedly demonstrated the existence of the endopeptidase activity of ApAAP.

Abz-KARVLF(NO₂)JEANle has several potential hydrophobic cleavage sites: (i) splitting after Phe (NO₂) follows from the increase in the absorbance at 300 nm; (ii) the cleavage after Leu is a possibility; and (iii) the hydrophobic Val may also be a candidate. Using amino acid sequence analysis of the hydrolysis products, we have determined the true scissile bonds. After 50% hydrolysis of the substrate detected by fluorimetric measurement, the following fragments were obtained: EANle, RVLF(NO₂)EANle, which yielded a rate constant of 1.3 times higher than that found with the fluorometric method. This revealed that at least two parallel reactions took place, and the cleavage at the Phe (NO₂) residue undoubtedly demonstrated the existence of the endopeptidase activity of ApAAP.

To reveal the binding modes of substrates and inhibitors, we have crystallized ApAAP with the product-like inhibitors Ac-Phe-OH and Gly-Phe-OH, as well as with the substrate Abz-GFEPF(NO₂)RA. The crystal structures contained the dimer of ApAAP. The monomers A and B displayed similar conformations and accommodated the ligands at the S1 and S2 binding sites in similar positions and conformations.

Figure 4 illustrates the binding modes of Ac-Phe-OH and Gly-Phe-OH in their complexes with ApAAP. Several hydrogen bonds and polar contacts are formed between the enzyme and the inhibitors. Hydrogen bond distances are shown in Table 3. The O³ atom of the active site serine is in close proximity to the carboxylate C atom of the bound inhibitor. The structures of the complexes revealed the oxyanion binding site, which stabilizes the transition state in all serine peptidases.²² The oxyanion-binding site of ApAAP is composed of the main chain NH group of Tyr446 and Gly369, which donate the hydrogen bonds to the ligands in both complexes. In the ApAAP complex published earlier,⁹ the phosphate group of 4-nitrophenylphosphate is located in a different orientation in the oxyanion hole, being the phosphate oxygen atoms far from the Tyr446 main chain NH group (PDB code 1VE7).

The main chain of the inhibitor is further stabilized by hydrogen bonds formed between the NH group of the P1 residue and the carbonyl group of Gly369 (Table 3), as well as the carbonyl oxygen of the P2 residue and the side-chain of Arg526 (Figures 4 and 5). Arg526 is conserved in all other AAP amino acid sequences.⁵ The hydrogen bond between the P2 carbonyl oxygen and the Arg526 guanidinium group, and the Arg526–Glu88 salt-bridge strongly resembles the complexes of prolyl-oligopeptidase, where the corresponding residues are Arg643 and Asp149. The above hydrogen bond is present in the complex with Gly-Phe-OH,
stabilizing the main chain of the inhibitor, but with Ac-Phe-OH the guanidinium group is connected to the Ac-Phe-OH inhibitor via a hydrogen-bonded water molecule (Table 3 and Figure 4).

Interestingly, the imidazole group of the catalytic His556 is too far from the carboxylate group of Ac-Phe-OH to establish a hydrogen bond (Table 3). Though in the complex with Gly-Phe-OH the corresponding distance is shorter, the orientation of the carboxylate group is unfavourable for hydrogen bond formation (Figure 4). This is different from the analogous prolyl oligopeptidase complex, where the carboxylate group of the product-like inhibitor (Z-Gly-Pro-OH) can form a hydrogen bond with the catalytic histidine residue (3.02 Å).23 Moreover, the association constant ($1/K_i$) for Z-Gly-Pro-OH increased with the protonation of the catalytic histidine in prolyl oligopeptidase. Because of the structural difference between the inhibitor complexes formed with the two enzymes, we have determined the pH-dependence of $1/K_i$ for Ac-Phe-OH and ApAAP. Figure 6 illustrates the titration of ApAAP with Ac-Phe-OH, which provided a $pK_a$ of

Table 3. Hydrogen bonds and polar contacts between the A and B molecules of ApAAP and the bound inhibitors

<table>
<thead>
<tr>
<th></th>
<th>Ac-Phe-OH (A) (Å)</th>
<th>Ac-Phe-OH (B) (Å)</th>
<th>Gly-Phe-OH (A) (Å)</th>
<th>Gly-Phe-OH (B) (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S445 O$^{−}$-P1 O1</td>
<td>2.50</td>
<td>2.62</td>
<td>2.58</td>
<td>2.47</td>
</tr>
<tr>
<td>Y446 N-P1 O1</td>
<td>2.91$^a$</td>
<td>2.93$^a$</td>
<td>2.92$^a$</td>
<td>2.85$^a$</td>
</tr>
<tr>
<td>G369 N-P1 O1</td>
<td>2.79$^a$</td>
<td>2.74$^a$</td>
<td>2.75$^a$</td>
<td>2.69$^a$</td>
</tr>
<tr>
<td>G369 N-P1 O2</td>
<td>3.06</td>
<td>3.14</td>
<td>3.54</td>
<td>3.36</td>
</tr>
<tr>
<td>G369 O-P2 N</td>
<td>2.79$^a$</td>
<td>2.93$^a$</td>
<td>2.80$^a$</td>
<td>2.83$^a$</td>
</tr>
<tr>
<td>R526 N$^{−}$-P2 O</td>
<td>3.68$^b$</td>
<td>3.72$^b$</td>
<td>3.21$^b$</td>
<td>2.99$^b$</td>
</tr>
<tr>
<td>H356 N$^{−}$-P1 O2</td>
<td>5.04$^a$</td>
<td>5.12$^b$</td>
<td>3.46</td>
<td>4.34</td>
</tr>
</tbody>
</table>

P1 and P2 represent the inhibitor moieties at the S1 and S2 subsites, respectively.

$^a$ Hydrogen bond.

$^b$ Water-mediated hydrogen bond.

Figure 4. Stereo view of the active site of molecule A of the ApAAP dimer. Superposition of the active site region of the ApAAP–Ac-Phe-OH complex on the ApAAP–Gly-Phe-OH structure is shown. The amino acid residues are shown in ball- and -stick representation. The carbon atoms of the ApAAP–Ac-Phe-OH and the ApAAP–Gly-Phe-OH complexes are coloured orange and green, respectively. The other atoms are illustrated in standard colours. The main chain NH groups of Gly369 and Tyr446 constitute the oxyanion-binding site. Hydrogen bonds of the ApAAP–Gly-Phe-OH complex are illustrated as black broken lines. The S2–P2 interactions are different in the two complexes. While the Gly residue of Gly-Phe-OH is hydrogen bonded directly to Arg526, the acetyl group of Ac-Phe-OH is connected to Arg526 via a water molecule (w) with hydrogen bonds represented as yellow broken lines. The distances are compiled in Table 3.

Figure 5. Scheme of the binding of Abz-Gly-Phe-OH. The binding interactions for the P1 and P2 residues are identical with those of Gly-Phe-OH. The catalytic triad is shown on the right-hand side of the scheme. While the catalytic serine residue was present in the complex formed with Gly-Phe-OH, it was replaced by alanine in the complex formed with Abz-Gly-Phe-OH. The hydrogen bonds are represented by broken lines. The catalytic serine side-chain (Ser445) is reconstructed using the crystal structure of the Gly-Phe-OH complex. Phe485 contacts P1 and P3, and Phe488 interacts with the P1 and P2 residues of the inhibitor.
6.19±0.02 and a limiting value of 3.90(±0.06) mM in the presence of 0.3 M NaCl. The pK\textsubscript{a} value is similar to the pK\textsubscript{a} of 6.09±0.06 extracted from the pH–k\textsubscript{cat}/K\textsubscript{m} profile of the Ac-Phe-Nap reaction (Figure 1). This indicates that the ionization of histidine influences the binding of inhibitor even if an intimate ion pair cannot be formed. We have examined the salt effects on the formation of the enzyme–inhibitor complex, because high ionic strength depresses the electrostatic effects. Indeed, 1/K\textsubscript{i} was considerably higher in the absence of 0.3 M NaCl (18.8(±1.0) mM\textsuperscript{−1}), and the pK\textsubscript{a} extracted from the dissociation curve also changed (pK\textsubscript{a} =5.36). It may be noted that we did not collect more points at low pH because the 1/K\textsubscript{i} values tended to decrease with the decrease in pH. Nevertheless, the errors of the data were relatively low, apparently because the points were suitable for extrapolation to the low pH region. This finding differed from that found with prolyl oligopeptidase, where the salt effect elicited only a minor change in 1/K\textsubscript{i}, and there was practically no change in the pK\textsubscript{a} value. Such differences between the binding modes of the two enzymes suggest that the electrostatic environments and the structures of the active sites are somewhat divergent. It is worthy of note that product-like inhibitors form a salt-bridge with the catalytic histidine of the less related trypsin-like peptidases.24

It is interesting to ask why the inhibitors to ApAAP do not form hydrogen bonds with protonated imidazole, as in other serine peptidases. The answer is not simple. (i) The titration of ApAAP was performed at 70 °C, where the protein is more flexible than in the crystal structure determined at 100 K, and the high temperature may permit the carboxylate ion to form a hydrogen bond with the imidazole group. (ii) Upon protonation, the catalytic histidine modifies the electrostatic environment that facilitates the binding event, even if the hydrogen bond fails to form. Such a salt-bridge-like electrostatic interaction can explain the results of titration with Ac-Phe-OH (Figure 6).

With the product-like inhibitors, the location of the P1 phenylalanine side-chain is different from that observed with the p-nitrophenylphosphate.9 The latter compound extends to the bottom of the S1 pocket and interacts with Ile489 through the bulky nitro group (Figure 7). Comparison of the structures of the enzyme without bound ligand and with bound p-nitrophenylphosphate or phenylalanine residue in the S1 pocket reveals that some of the side-chains of the residues lining the S1 pocket
change position significantly, depending on the bound ligand (Figure 7). These changes affect the sides of the S1 pocket (Phe488 and Met477 widening, Leu492 narrowing it), while the bottom of the pocket changes only slightly (RMS deviations are shown in Table 4).

An attempt to map both the S and S’ subsites was made by determining the crystal structure of the practically inactive ApAAP S445A variant with the Abz-GFEPF(NO2)RA substrate. A similar approach was employed successfully for prolyl oligopeptidase,23 but in another study the “inactive” enzyme hydrolysed the substrate.13 Likewise, the substrate was cleaved during the lengthy crystallization period in the present investigation, so that only its acyl portion was seen in the crystal structure (Figure 8). In the case of subtilisin modified similarly at the catalytic serine residue, an ∼106-fold decrease in activity was demonstrated, indicating a small remaining activity, but with different mechanism not involving an acyl-enzyme intermediate.25 With the analogous variant of porcine prolyl oligopeptidase, we obtained an ∼107-fold rate decrease.13 As in these cases, the remaining activity of the S445A ApAAP variant could cause the cleavage in the peptide. The hydrogen bond network at the S1 and S2 subsites of the ApAAP–Abz-Gly-Phe-OH complex, as well as the orientation of the P1 carboxylate group, resembles that of the structure of the ApAAP–Gly-Phe-OH complex. The differences observed in Ac-Phe binding can be related to its less polar character.

Residues forming the S2 and S3 sites are shown in Figure 8(a). The S2 site, lined by the hydrophobic side-chains of Phe153, Phe169, Phe371 and Phe488 is capable of accommodating bulkier residues of non-polar character. In the present structures, the glycyl and acetyl moieties of the inhibitors are bound in roughly similar positions (Figure 4). Though the main chain atoms of the P2 residue are bound in similar positions in ApAAP and prolyl oligopeptidase, the environment of the P2 side-chain would be less polar in ApAAP. In contrast to dipeptidyl-peptidase IV,26 ApAAP does not form salt-bridges or even hydrogen bonds with the amino group of the Gly-Phe-OH inhibitor. Instead, the amino group turns out into the solvent of the central cavity. This is in accordance with the observations showing that ApAAP is not specific for splitting dipeptides with a charged N terminus from the substrate.

The N-terminal Abz moiety resembles a hydrophobic amino acid, which makes it a candidate for exploring the S3 binding site. The ApAAP residues contacting the Abz moiety, Leu115, Phe153, Phe155...
and Phe485, define a hydrophobic patch in the complex (Figure 8). The aromatic ring of the Abz group is partially buried by the side-chain of Arg526, while its amino and carbonyl groups are exposed to the solvent in the central cavity of the enzyme. As shown in Figure 8(b), the depression of the enzyme surface extends beyond the S3 site, suggesting that it can accommodate further peptide residues.

### Conclusion

We have demonstrated that ApAAP exhibits endopeptidase activity besides the well-known exopeptidase activity, while the porcine enzyme is, indeed, an exopeptidase as correctly classified. ApAAP is able to hydrolize the peptide bond after an Ala residue, which contradicts the notion that the enzyme is specific for hydrophobic amino acids, in particular to phenylalanine, in accord with the non-polar S1 subsite of ApAAP. The crystal structures of the complexes formed between the enzyme and product-like inhibitors identified the oxyanion hole and the S1, S2 and S3 binding sites. These complexes revealed also that upon ligand binding the S1 subsite undergoes significant conformational changes, demonstrating the plasticity of the specificity site.

### Materials and Methods

#### Cloning the ApAAP gene

The ApAAP gene APE1547 (obtained from the National Institute of Technology and Evaluation, Tokyo) was amplified with the polymerase chain reaction (PCR) method, using the primers

- 5′-GTAGACCATAAGGCTATTATGATGTA-3' (5′-primer; NdeI cleavage site underlined)
- 5′-AGTTTGATCCCTCTCTCTCTCTGGT-3' (3′-primer; BamHI cleavage site underlined).

The gene product was digested with BamHI and NdeI restriction endonucleases, and ligated to a pET22b vector opened with the same endonucleases.

#### Constructing the S445A mutant

The mutation was introduced into ApAAP by the two-step PCR method as described for the H507A variant of the porcine AAP.2 The following primers were used

- 5′-TACATCATGGGGCTACGGTACGGGCGGCTAC-3′ (sense)
- 5′-GTAGCGCGCCGTAAGGTACGCCCATGTATGTA-3′ (antisense).

The mutated nucleotides are designated by lower case letters. An extra recognition site was created for the MluI restriction enzyme (underlined) with silent mutation to verify the incorporation of the mutant oligonucleotides into the PCR product.

### Preparation of ApAAPs

The same preparation method was used for the wild-type enzyme and the S445A variant. The PET22b plasmid harbouring the ApAAP gene was transformed into a Rosetta DE3 Escherichia coli strain. LB medium (50 ml) containing 100 mg/1 of ampicillin and 12 mg/1 of chloramphenicol was inoculated with a single-cell colony developed on an agar plate. When the cell suspension had an absorbance at 600 nm of 0.6, 5 ml of cell suspension was added to each of four LB media (4 × 600 ml) containing 100 mg/1 of ampicillin and 12 mg/1 of chloramphenicol. The cells were incubated at 37 °C for 16 h and subsequently centrifuged. The cell paste (15 g wet weight, obtained from 4 × 600 ml) was sonicated in 80 ml of PS1 buffer (50 mM sodium phosphate (pH 6.5), 1 M (NH4)2SO4, 10 mM EDTA). The supernatant of the sonicate was heated at 85 °C for 60 min, and the precipitate was removed by centrifugation. The supernatant was applied to a Phenyl-Sepharose column (9 mm × 85 mm) equilibrated with PS1 buffer. The unbound proteins were removed by washing with 30 ml of PS2 buffer (50 mM phosphate (pH 6.0), 1 M (NH4)2SO4, 10 mM EDTA). The ApAAP was eluted from the column with 20 ml of Q51 buffer (20 mM phosphate (pH 8.0), 1 mM EDTA, 0.1% (v/v) Triton X-100). The enzyme solution was then concentrated to 5 ml by ultrafiltration on an Amicon PM 30 membrane, and finally 2 ml of the solution was applied to a Q-Sepharose column (10 mm × 308 mm) equilibrated with Q51 buffer. The pure enzyme was eluted with a linear gradient of 0–1 M NaCl in Q52 buffer (20 mM phosphate (pH 8.0), 1 mM EDTA, 0.1% (v/v) Triton X-100).

The fraction containing the active enzyme was homogeneous as judged by SDS-PAGE. The protein concentration of the dimer enzyme was calculated from the absorbance at 280 nm using a value of 126,062 Da and A280(0.1%) = 0.93. The overall yields of ApAAP and its S445A variant were about 50 mg.

The preparation of the porcine AAP was as described.12

### Kinetics

With the exception of 4-nitroanilide substrates, the hydrolyses of all substrates were measured fluorometrically using a Cary Eclipse fluorescence spectrophotometer equipped with a Peltier four-position multichannel holder accessory and a temperature controller. The excitation and emission wavelengths were 340 nm and 410 nm, respectively, for the substrate with Nap, and 370 nm and 440 nm, respectively, for the substrate with Amc leaving groups. The substrates with internally quenched fluorescence were prepared by solid-phase synthesis, and the hydrolytic reactions were measured at 337 nm and 420 nm excitation and emission wavelengths, respectively. The liberation of 4-nitroanilide was monitored spectrophotometrically at 410 nm.

The pseudo-first-order rate constants were determined at substrate concentrations lower than 0.1 Km and were calculated by non-linear regression data analysis, using the GraFit software.27 The specificity rate constants (kcat/Km) were obtained by dividing the first-order rate constant by the total enzyme concentration in the reaction mixture.

The pH-dependence of kcat/Km was measured at the temperature optimum of 70 °C in a complex buffer (30 mM Mes, 30 mM Heps, 30 mM Taps, 30 mM glycine, 1 mM EDTA, 0.3 M NaCl, unless stated otherwise). Theoretical curves for the bell-shaped pH-rate profiles were calculated by non-linear regression analysis, using equation (1) and
the GraFit software. In equation (1), $k_\text{cat}/K_m(\text{limit})$ stands for the pH-independent maximum rate constant, and $K_1$ and $K_2$ are the dissociation constants of a catalytically competent base and acid, respectively. When an additional ionizing group modified the bell-shaped character of the pH-dependence curve, the data were fit to equation (2) (a bell-shaped curve with double ionization at the basic limb), or equation (3) (as equation (2), but the emergence of a new enzyme form was considered at low pHi). Equation (4) represents a double bell-shaped curve with double ionization at the basic limb.

$$
k_\text{cat}/K_m = k_\text{cat}/K_m(\text{limit})[1/(1 + 10^{pH - pK1} + 10^{pH - pK2})]
$$

$$
k_\text{cat}/K_m = k_\text{cat}/K_m(\text{limit})[1/(1 + 10^{pH - pK1} + 10^{pH - pK2} + 10^{pH - pK3})]
$$

$$
k_\text{cat}/K_m = k_\text{cat}/K_m(\text{limit})[1/(1 + 10^{pH - pK1} + 10^{pH - pK2})] + k_\text{cat}/K_m(\text{limit})[1/(1 + 10^{pH - pK3})]
$$

$$
k_\text{cat}/K_m = k_\text{cat}/K_m(\text{limit})(1/(1 + 10^{pH - pK1} + 10^{pH - pK2}) + 10^{pH - pK3})$$(4)

The $K_i$ values, the dissociation constants of the enzyme-inhibitor complex, were calculated from equation (5), where $k_i$ and $k_0$ are pseudo first-order rate constants determined at substrate concentrations $< 0.1$ $K_m$ in the presence and in the absence of inhibitor (I), respectively.

$$
k_i/k_0 = 1/(1 + I/K_i)
$$

**Crystallographic studies on ApAAP complexes**

The native ApAAP and its S445A variant were concentrated to 10 mg/ml in 20 mM Tris–HCl (pH 8.0). Ac-Phe-OH, Gly-Phe-OH or Abz-GFEPF(NO2)RA were dissolved in the protein solutions at concentrations of 66.5 μM, 11.9 μM and 1.9 μM, respectively. The crystals were grown at room temperature in hanging drops composed of protein solution and reservoir buffer mixed in 1:1 (v/v) ratio. Crystallization conditions published earlier were modified for the co-crystallization. Crystals of ApAAP complexed with Ac-Phe-OH were obtained from 78 mM sodium acetate buffer (pH 5.0) containing 0.44 mM EDTA, 2.2% (w/v) PEG 400 and 0.56% (w/v) β-octyl-glucoside. The ApAAP complex with Gly-Phe-OH was crystallized from the same reservoir buffer containing 10% (v/v) dimethyl sulphoxide (DMSO). The ApAAP S445A mutant and Abz-GFEPF(NO2)RA were co-crystallized from reservoir containing 20 mM Tris–HCl (pH 8.0) and 20% (v/v) ethanol. The crystals of the native enzyme were cryoprotected by being dipped in mother liquor containing 35% (v/v) glycerol. The cryoprotectant solution for the S445A mutant contained 5% ethanol and 30% PEG 400.

X-ray data were collected on a Rigaku R-AXIS II diffractometer using Cu Kα radiation at 100 K. Data

**Table 5.** Data collection and refinement statistics of the ApAAP–inhibitor complexes

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P2₁2₁2₁</td>
<td>P2₁2₁2₁</td>
<td>P2₁2₁2₁</td>
</tr>
<tr>
<td>Unit cell parameters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a$ (Å)</td>
<td>63.682</td>
<td>63.400</td>
<td>63.810</td>
</tr>
<tr>
<td>$b$ (Å)</td>
<td>103.989</td>
<td>103.334</td>
<td>103.922</td>
</tr>
<tr>
<td>$c$ (Å)</td>
<td>168.594</td>
<td>168.224</td>
<td>169.255</td>
</tr>
<tr>
<td>Resolution range</td>
<td>19.54–2.00 (2.10–2.00)</td>
<td>19.73–2.00 (2.10–2.00)</td>
<td>19.96–2.40 (2.50–2.40)</td>
</tr>
<tr>
<td>$R_{merge}$</td>
<td>0.115 (0.383)</td>
<td>0.098 (0.560)</td>
<td>0.137 (0.380)</td>
</tr>
<tr>
<td>Completess (%)</td>
<td>94.8 (80.2)</td>
<td>89.1 (53.2)</td>
<td>99.1 (99.9)</td>
</tr>
<tr>
<td>No. observed reflections</td>
<td>442,438 (21,062)</td>
<td>359,946 (20,672)</td>
<td>181,789 (21,129)</td>
</tr>
<tr>
<td>No. unique reflections</td>
<td>76,464 (8216)</td>
<td>67,288 (5399)</td>
<td>45,069 (5151)</td>
</tr>
<tr>
<td>$I/\sigma(I)$</td>
<td>13.7 (3.10)</td>
<td>15.1 (2.70)</td>
<td>11.68 (4.36)</td>
</tr>
<tr>
<td>$R$</td>
<td>0.1776</td>
<td>0.1836</td>
<td>0.2178</td>
</tr>
<tr>
<td>$R_{merge}$</td>
<td>0.2169</td>
<td>0.2244</td>
<td>0.2613</td>
</tr>
<tr>
<td>RMS deviation from ideal Bond lengths (Å)</td>
<td>0.190</td>
<td>0.121</td>
<td>0.005</td>
</tr>
<tr>
<td>Bond angles (deg.)</td>
<td>1.274</td>
<td>1.308</td>
<td>0.878</td>
</tr>
<tr>
<td>General planes (Å)</td>
<td>0.004</td>
<td>0.005</td>
<td>0.002</td>
</tr>
<tr>
<td>Ramachandran plot</td>
<td>Residues in core regions</td>
<td>863</td>
<td>856</td>
</tr>
<tr>
<td></td>
<td>Residues in allowed regions</td>
<td>84</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>Residues in disallowed regions</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No. protein residues</td>
<td>1146</td>
<td>1148</td>
</tr>
<tr>
<td></td>
<td>No. residues in dual conformations</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>No. residues with disordered side-chains</td>
<td>23</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>No. ligands</td>
<td>Two Ac-Phe-OH</td>
<td>Two Gly-Phe-OH</td>
</tr>
<tr>
<td></td>
<td>No. water molecules</td>
<td>349</td>
<td>332</td>
</tr>
<tr>
<td></td>
<td>No. glycerol molecules</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

Data within parentheses are for the last resolution shell.

a. $R_{merge} = \sum_{i=1}^{n} \sum_{j=1}^{N} (I_i - \langle I_i|_{I_{ij}} \rangle) / \sum_{i=1}^{n} \sum_{j=1}^{N} \langle I_i|_{I_{ij}} \rangle \times N_i$, with $I_i = \langle I_i|_{I_{ij}} \rangle / \langle N_i \rangle$.

b. A randomly selected 5.0% of the reflections in a test set for monitoring the refinement process.

c. The residues in the disallowed region are the catalytic serine S445 and A445 in the native and mutant enzyme, respectively.
processing and data reduction were carried out with the XDS and XSCALE programs. The enzyme structures were solved by molecular replacement using the program MOLREP. The monomer of the uncomplexed enzyme was used as search model (PDB code 1VE6). Refinement was carried out with the program REFMAC5, using restrained maximum-likelihood refinement and TLS refinement. During refinement, non-crystallographic restraints were added to the regions of ApAAP molecules possessing similar conformation. Model building was carried out using the program Coot.

The final model of the ApAAP–Ac–Phe–OH complex contains residues 9–581 of ApAAP monomers, the bound inhibitor located at the active site of each monomer, 349 water molecules and two glycerol molecules. The ApAAP–Gly–Phe–OH complex displays residues 9–581 and 8–582 of ApAAP monomers A and B, respectively, the bound inhibitor in each monomer, 332 water molecules and five glycerol molecules. The crystals of the bound inhibitor in each monomer, 8 water and 8 glycerol molecules. The crystals of the uncomplexed enzyme was used as search model (PDB code 1VE6). Refinement was carried out with the program REFMAC5, using restrained maximum-likelihood refinement and TLS refinement. During refinement, non-crystallographic restraints were added to the regions of ApAAP molecules possessing similar conformation. Model building was carried out using the program Coot.

The atomic coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 2HU5, 2HU7 and 2HU8 for the ApAAP–Gly–Phe–OH, ApAAP–Ac–Phe–OH and S445A ApAAP–Abz–Gly–Phe–OH complexes, respectively.

Acknowledgements

Thanks are due to Ms I. Szamosi for excellent technical assistance. This work was supported by the Hungarian Science Fund OTKA T/16 (T046057).

References


† http://www.pymol.sourceforge.net/


*Edited by R. Huber*