Structural Basis for the Unusual Specificity of *Escherichia coli* Aminopeptidase N

Anthony Addlagatta,‡ Leslie Gay, and Brian W. Matthews*

Institute of Molecular Biology, Howard Hughes Medical Institute, and Department of Physics, 1229 University of Oregon, Eugene, Oregon 97403-1229

Received November 8, 2007; Revised Manuscript Received March 7, 2008

ABSTRACT: Aminopeptidase N from *Escherichia coli* is a M1 class aminopeptidase with the active-site region related to that of thermolysin. The enzyme has unusual specificity, cleaving adjacent to the large, nonpolar amino acids Phe and Tyr but also cleaving next to the polar residues Lys and Arg. To try to understand the structural basis for this pattern of hydrolysis, the structure of the enzyme was determined in complex with the amino acids L-arginine, L-lysine, L-phenylalanine, L-tryptophan, and L-tyrosine. These amino acids all bind with their backbone atoms close to the active-site zinc ion and their side chain occupying the S1 subsite. This subsite is in the form of a cylinder, about 10 Å in cross-section and 12 Å in length. The bottom of the cylinder includes the zinc ion and a number of polar side chains that make multiple hydrogen-bonding and other interactions with the α-amino group and the α-carboxylate of the bound amino acid. The walls of the S1 cylinder are hydrophobic and accommodate the nonpolar or largely nonpolar side chains of Phe and Tyr. The top of the cylinder is polar in character and includes bound water molecules. The ε-amino group of the bound lysine side chain and the guanidinium group of arginine both make multiple hydrogen bonds to this part of the S1 site. At the same time, the hydrocarbon part of the lysine and arginine side chains is accommodated within the nonpolar walls of the S1 cylinder. This combination of hydrophobic and hydrophilic binding surfaces explains the ability of ePepN to cleave Lys, Arg, Phe, and Tyr. Another favored substrate has Ala at the P1 position. The short, nonpolar side chain of this residue can clearly be bound within the hydrophobic part of the S1 cylinder, but the reason for its facile hydrolysis remains uncertain.

Aminopeptidases constitute a variety of proteins with important roles in cell maintenance, growth, and development. M1 class aminopeptidases are a large group of metalloproteases classified as gluzincins with a consensus zinc-binding motif [HEXXH-(X18)-E] and an exopeptidase motif (GXMEN) in the active site (*1*). Aminopeptidase N (EC 3.4.11.2) from *Escherichia coli* (ePepN)[1] is a member of the M1 class peptidases, first identified as an alanine aminopeptidase (*2, 3*). Subsequent studies suggested that it has the highest preference for the basic residues lysine and arginine, followed by hydrophobic residues, such as alanine, leucine, phenylalanine, and tyrosine (*4, 5*). Chapplet-Tordo et al. suggested that the side chain of the P1 residue should have a minimal degree of hydrophobicity even if it involves protonated groups (*6*). They also pointed out that ePepN only cleaves peptides with the L configuration.

Crystal structures of PepN from *E. coli* and several of its homologues, for example, PepN from *Neisseria meningitides*, leukotriene A4 hydrolase (LTA4H) from human, and tricorn-interacting factor F3 (F3) from *Thermoplasma acidophilum*, have been determined recently (*7–11*). Although some of them have very little overall sequence similarity, they display high structural conservation and several of their active-site residues are conserved. LTA4H displays aminopeptidase activity in addition to its physiological leukotriene hydrolase function.

Bestatin, a general metalloprotease inhibitor in complex with ePepN and LTA4H, provided the first glimpse of the possible mode of binding of the P1 residue of the substrate in the S1 pocket (*7, 8, 10*). However, bestatin has a D-phenylalanine at its amino terminus as opposed to L-amino acids in substrates. In addition, this phenylalanine adopts different conformations in the ePepN and LTA4H complexes.

To clarify the mode of binding of "natural" peptides, including those with basic and hydrophobic residues in the P1 position, we have determined crystal structures of ePepN in complex with L-Arg, L-Lys, L-Phe, L-Tyr, and L-Trp. In addition, we have also determined the catalytic efficiency of some dipeptides using capillary electrophoresis. These studies provide a molecular basis for the observed substrate preferences for ePepN as well as several other M1 class bacterial and mammalian aminopeptidase homologues.

MATERIALS AND METHODS

Expression and Purification of ePepN. Cloning, expression, purification, and crystallization were carried out as described (*7*). In short, cDNA was cloned into pET15b vector with the N-terminal His tag, and the protein was overex-

* To whom correspondence should be addressed. Telephone: (541) 346-2572. Fax: (541) 346-5870. E-mail: brian@uoregon.edu.

‡ Present address: Pacific Northwest Laboratory, P.O. Box 999, K8-93, Richland, WA 99352.

[1] Abbreviations: LTA4H, leukotriene A4 hydrolase; F3, tricorn-interacting factor; ePepN, *Escherichia coli* aminopeptidase N; L1PepN, *Lactococcus lactis* aminopeptidase N; StPepN, *Streptococcus thermophilus* aminopeptidase N; APN, human cell-surface membrane-associated aminopeptidase N; PSA, cytosolic puromycin-sensitive aminopeptidase; P-LAP, placental leucine aminopeptidase.

10.1021/bi7022333 CCC: $40.75
Published on Web 04/17/2008

Table 1: X-ray Data Collection and Refinement Statistics

	X-ray Data Collection				
ePepN complex	arginine	lysine	phenylalanine	tyrosine	tryptophan
cell parameters					
a and *b* (Å)	120.20	120.56	120.58	120.32	120.64
c (Å)	169.65	170.27	170.76	170.34	170.64
resolution range (Å) (highest shell)	50–2.0 (2.07–2.0)	50–1.5 (1.55–1.50)	50–1.30 (1.32–1.30)	50–1.70 (1.76–1.70)	50–1.85 (1.92–1.85)
R_{sym} (%)	8.1 (57.7)	7.3 (49.0)	12.3 (49.0)	4.0 (44.0)	5.4 (39.0)
$I/\sigma(I)$	13.2 (2.4)	12.3 (1.6)	8.2 (1.18)	32.3 (2.9)	19.5 (2.9)
completeness (%)	99.9 (99.4)	88.8 (48.6)	99.2 (89.9)	99.9 (99.3)	100 (100)
redundancy	5.6 (5.0)	5.3 (3.5)	8.9 (1.9)	5.3 (4.3)	5.2 (5.2)
	Refinement				
number of reflections (R_{free})	92 657 (2855)	195 558 (6049)	339 183 (10 333)	156 281 (4662)	122 587 (6160)
R_{work} (R_{free}) (%)	15.3 (18.8)	14.6 (16.5)	17.7 (19.4)	16.2 (18.4)	15.1 (17.8)
B factor, protein (Å^2)	17.3	14.6	16.9	13.6	15.7
B-factor, ligand (Å^2)	28.4	11.5	12.2	11.6	19.2
B factor, ions (Å^2)	23.7	16.2	16.2	16.2	20.3
B factor, water (Å^2)	27.9	27.6	30.7	23.2	28.8
protein atoms	7252	7244	7479	7125	7144
ligand atoms	12	10	24 (2 conformations)	13	15
ions	3	3	3	3	3
water molecules	765	1111	1113	752	1008
rms deviation					
bond lengths (Å)	0.016	0.008	0.009	0.011	0.012
bond angles (deg)	1.43	1.12	1.15	1.21	1.23
PDB code	3B2P	3B2X	3B34	3B37	3B3B

pressed in BL21 (DE3) cells at 25 °C. The harvested cells from 4 L cell culture were resuspended in +T/G buffer (50 mM Hepes at pH 8.0, 0.5 M KCl, 5 mM imidazole, 10% glycerol, and 0.1% Triton X-100) and passed twice through a French press. The cell lysate was centrifuged at 17000*g* for 30 min. The supernatant was passed though a talon column (cobalt-affinity resin; BD Biosciences, Franklin Lake, NJ), which was pre-equilibrated with +T/G buffer. The column was washed further after loading the lysate with the +T/G buffer until the absorption OD_{280} reached the baseline. +T/G buffer was replaced by −T/G buffer (50 mM Hepes at pH 8.0, 0.5 M KCl, and 5 mM imidazole), and the washing continued until the absorption reached the baseline. Pure protein was eluted with 100 mM imidazole in −T/G buffer at 15 mg/mL concentration. This eluent was dialyzed twice against 4 L of storage buffer (25 mM Hepes at pH 7.5 and 150 mM KCl), aliquoted, and stored at −80 °C until further use.

Crystallization. The amino acids L-Lys, L-Arg, L-Phe, L-Tyr, and L-Trp were dissolved in water (all compounds were obtained from Sigma-Aldrich, St. Louis, MO). This solution was mixed with the enzyme at 11 mg/mL in the storage buffer to obtain a final amino acid concentration of 2 mM and incubated at 25 °C for 30 min. Cocrystals were obtained by mixing 5 μL of the enzyme/amino acid mixture with 5 μL of well solution (2.0 M sodium malonate at pH 7.5) in a hanging drop and equilibrating against the well solution for 3–5 days. Crystals were harvested by equilibrating them in cryoprotectant solution (22% glycerol, 2.0 M sodium malonate at pH 7.5, and 2 mM concentration of the appropriate amino acid) for 5 min and freezing directly in liquid nitrogen in cryo-loops (Hampton Research, Aliso Viejo, CA).

Data Collection, Structure Determination, and Refinement. X-ray data were collected on beam line 8.2.2 at the Advanced Light Source (Lawrence Berkeley National Laboratory, Berkeley, CA) using a Quantum 315 ADSC Area Detector (Area Detector Systems Corp., Poway, CA). X-ray data were processed and scaled using HKL2000 (*12*). All structures were isomorphous with the apo structure in space group $P3_121$ [Protein Data Bank (PDB) code 2HPO]. Structure refinement (*13, 14*) was performed using the CNS and CCP4 program suites, and COOT was used to visualize the model (*15*). Patches of diffuse but strong density were noted outside the S1 subsite and were modeled by a disordered glycerol and a partially occupied malonate. The X-ray data refinement statistics are summarized for each crystal structure in Table 1. Figure 1 shows the electron density for the bound amino acid in the respective ePepN complexes.

Aminopeptidase Activity Assay. Enzymatic activity was assayed by incubating dipeptide substrates of the form XXX-Phe (Table 2) with ePepN for 60 min. The 3 mL reaction mixture contained 500 ng (about 5 pmol) of ePepN and 15 μmol of substrate peptide (3 μmol in the case of Tyr-Phe and Phe-Phe peptides) in 25 mM sodium phosphate at pH 7.5. The reaction was terminated by adding 5 μL of the reaction mixture to 45 μL of 50 mM sodium phosphate at pH 2.5. The extent of hydrolysis was monitored by quantifying the remaining peptide using micellar electrokinetic chromatography on a Hewlett-Packard HP3D capillary electrophoresis system. The electropherogram run parameters were 50 μm × 40 cm extended light path fused-silica capillary; buffer, 50 mM sodium phosphate at pH 2.5; run conditions, 30 kV, 25 °C, 50 mbar, and 10 s injection; peak detection at 200 nm. The results are summarized in Table 2.

Sequence Alignment. ClustalW was used for pairwise sequence alignment (*16*). The homology-derived secondary structure of proteins (HSSP) database was used to analyze the multiple sequence alignments of proteins, which have at least 30% sequence identity with a protein of known structure (*17*).

RESULTS AND DISCUSSION

Hydrophobic Cylindrical S1 Pocket. The overall structure of ePepN consists of the four domains as shown in parts a

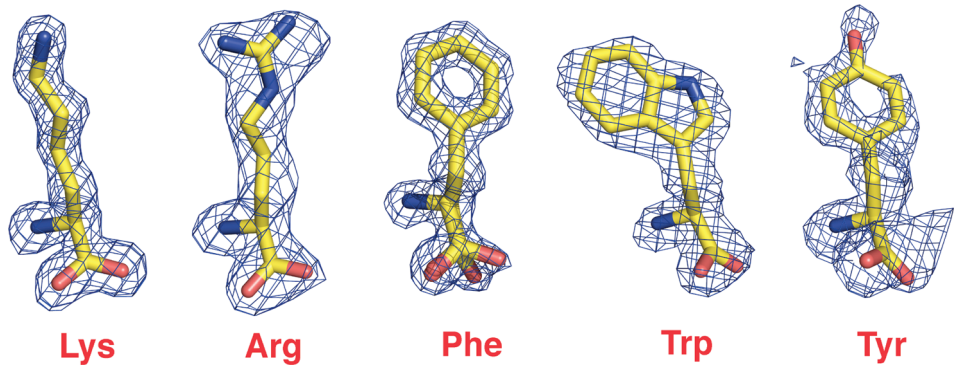

FIGURE 1: Omit electron-density maps for each amino acid in complex with ePepN (Table 1). Coefficients are $(F_o - F_c)$, where F_o values are the observed amplitudes. The calculated amplitudes F_c and the phases were obtained from the refined models with the ligand in the S1 pocket removed. The maps are contoured at 5.0σ for Lys, 4.2σ for Arg, 5.0σ for Phe, 4.0σ for Trp, and 2.8σ for Tyr.

Table 2: Specific Activity of ePepN toward Representative Dipeptide Substrates

dipeptide	rate of hydrolysis [nmol min^{-1} (μg of protein)$^{-1}$]
Arg-Phe	214 ± 24
Ala-Phe	154 ± 10
Lys-Phe	142 ± 36
Tyr-Phe	46 ± 6
Phe-Phe	44 ± 4
Asp-Phe	not detectable

and b of Figure 2. The catalytic domain (domain II) has a structure similar to that of the bacterial endopeptidase thermolysin (*18*). As for thermolysin, the catalytic domain of ePepN is bilobal with a pronounced active-site cleft at the junction of the two lobes. Domain I of ePepN blocks this cleft in the region where the P2 residue of the substrate would bind to thermolysin. This explains why ePepN is an exo- rather than an endopeptidase.

As discussed earlier (*7*), the active site of ePepN is enclosed within a large cavity of volume 2200 Å^3 and is inaccessible to substrates, except for a small opening of about 8–10 Å that may admit small peptides. The S1 subsite has the approximate shape of an elliptical cylinder and opens into the cavity. For the ease of description, we describe the S1 pocket in terms of six locations as seen when viewed (Figure 3) approximately from the direction of the C terminus of the substrate: left, right, top, bottom, front, and back (far end of the cylinder). The cylindrically shaped S1 pocket is made of residues from four parts of the protein. First, there are amino acids within or adjacent to the GXM^{263}EN exopeptidase motif, including Met260 and Met263. Second, there are residues near 120, including Gln119 and Glu121. Third and fourth, the side chains of Asn373 and Tyr376 and Gln821 also contribute. The hydroxyl of Tyr376 forms hydrogen bonds with the carboxylate of Glu121. The same oxygen of Glu121 also interacts with Lys319 that in turn interacts with the Glu264 and Glu320. The second oxygen of the Glu121 interacts with the amino terminus of the substrate.

The walls of the S1 cylinder are essentially hydrophobic, but there are polar regions at each end. The zinc-binding region, at which hydrolysis occurs, is located at the bottom of the cylinder. The top is formed by one residue from the catalytic domain (Asn373) and one from domain IV (Gln821). The longest dimension of the elliptical cylinder is between the left and right faces (10 Å). The distance from the front to back faces is about 7 Å, and that from the top to bottom

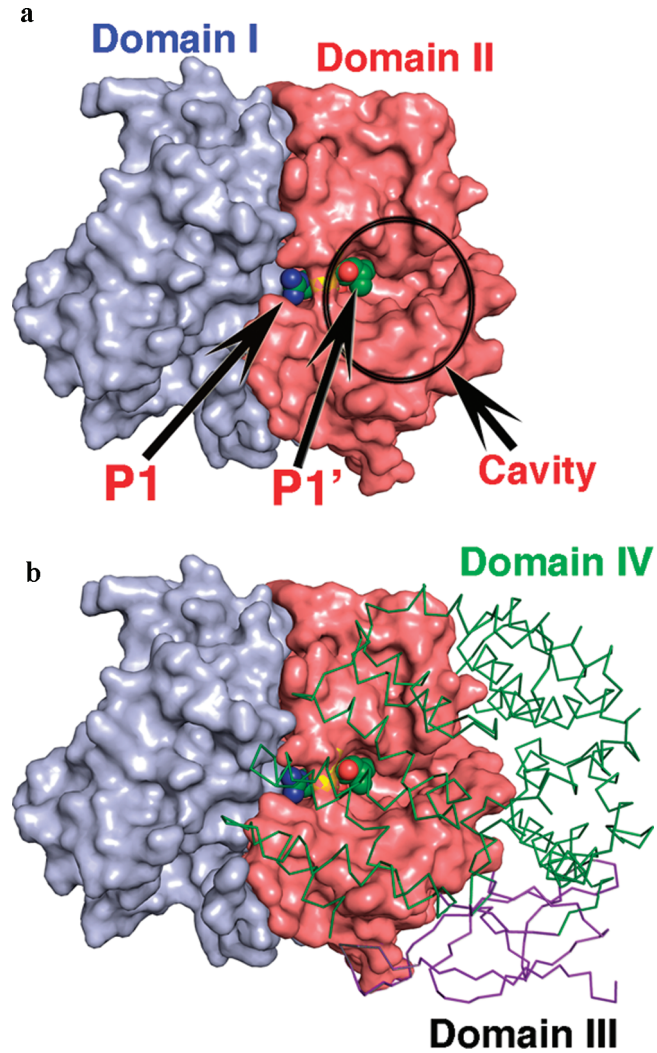

FIGURE 2: (a) Surface diagram of the domain I (blue) and domain II (catalytic, in red) with a peptide (small spheres) occupying the S1 and S1′ subsites. Domain I blocks the subsite where the P2 residue of a thermolysin substrate would normally bind, rendering ePepN as an aminopeptidase. A cavity of about 2200 Å^3 indicated by the circle is formed between domains I, II, and IV (*7*). (b) Overall structure of ePepN. Domains I and II are shown in the same surface representation as in a. For domains III (blue) and IV (green), the Cα backbone is shown.

is about 12 Å. In all of the complexes, the Cα carbon of the amino acid bound in the S1 subsite is 3.2 Å from the main-chain carbonyl oxygen of Ala262, possibly forming a C−H· · ·O interaction.

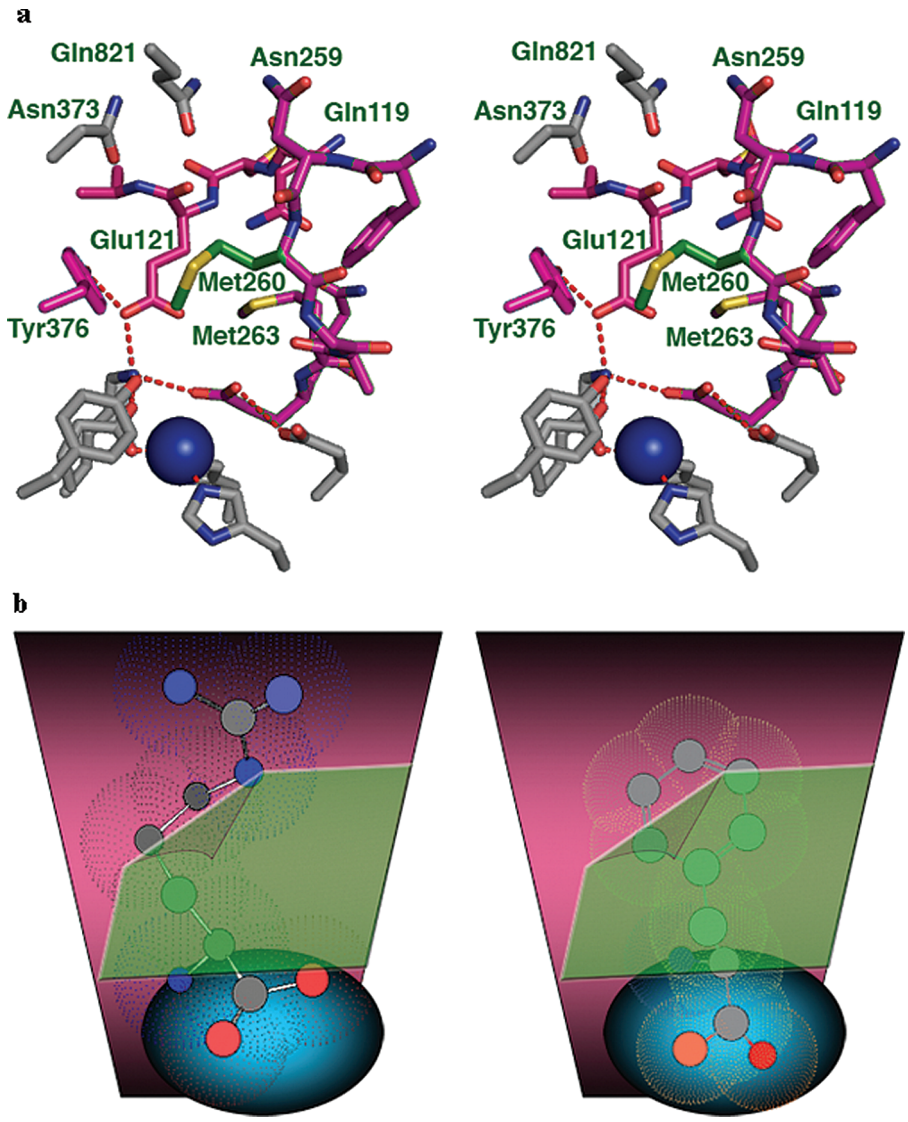

FIGURE 3: S1 pocket of ePepN adopts a cylindrical shape. (a) Stereo diagram of residues that make up the S1 pocket of ePepN. Residues shown in magenta form the hydrophobic cylinder. Met260 is shown in green and acts as a flexible gate to accommodate different residues in the S1 pocket. Residues in gray form the bottom and top caps of the cylinder. The zinc ion at the bottom of the cylinder is shown as a blue sphere. (b) Schematic diagram of the S1 pocket occupied by the polar residue arginine (left) and the hydrophobic residue phenylalanine (right) (shown in a ball-and-stick representation). The light magenta in the middle indicates the hydrophobic region, while the dark top and bottom suggest polar caps. The green "flap" indicates Met260 as shown in the a. The blue oval indicates the region at the bottom of the S1 pocket occupied by the carboxylate of Phe or Lys. This region extends into the adjacent S1′ subsite as shown in Figure 2a.

Complexes of ePepN with Lysine and Arginine. To help follow the discussion, Figure 4 shows simplified sketches of the active-site region in the absence and presence of bound ligand. The structure of ePepN in complex with L-lysine (Lys) was determined at 1.5 Å resolution (Table 1 and Figure 1). The backbone atoms of the lysine interact with the bottom of the S1 cylinder; the side-chain ε-amino group interacts with the top; and the hydrocarbon part of the side chain is enclosed within the hydrophobic walls of the cylinder (Figure 5a). The α-amino terminus of the lysine forms a three-centered interaction with the carboxylates of Glu121 (2.7 Å), Glu264 (2.7 Å), and Glu320 (3.2 Å). One of the carboxylate oxygen atoms of the lysine interacts with both the active-site zinc (2.0 Å) and Tyr381 (2.8 Å). The other carboxylate oxygen hydrogen bonds with the carboxylate of Glu298 (2.6 Å). Glu298 is presumed to activate a water molecule during catalysis, and the carboxylate of the lysine replaces this water molecule. The orientation of the lysine carboxylate is approximately perpendicular (72°) to that of the corresponding peptide of bestatin. The ε-amino group of the lysine side chain interacts with the side-chain oxygen of Asn373 (2.9 Å) and two water molecules W1 and W2 (2.8 Å each) at the top of the S1 cylinder. One of these water molecules (W1) interacts with the main-chain carbonyl of Asn373 (2.8 Å), while the other (W2) interacts with the side chain of Gln821 (2.8 Å). Near W2, there is a triad of three additional water molecules that are present in all known structures of ePepN, including those in this paper.

One of the major differences associated with the binding of lysine is in the conformation of the side chain of Met260. In the holo structure, it is disordered within the middle of the S1 cylinder. When lysine is bound, the side chain of Met260 swings out from the S1 hydrophobic pocket, providing space for the lysine side chain (Figure 5a). In the Lys-bound structure, Met260 is well-ordered. In association with the movement of the side chain of Met260, the main chain moves away from the center of the S1 cylinder by about 0.8 Å. The preceding residue (Asn259) also moves

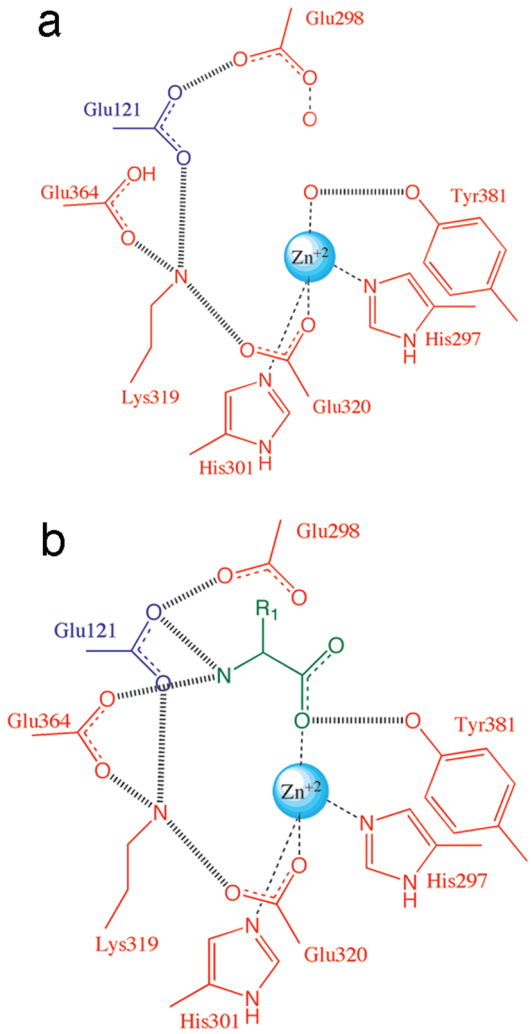

FIGURE 4: (a) Cartoon diagram of the active site of aminopeptidase N in the absence of substrate or product. The active-site zinc is coordinated by His297, His301, Glu320, and a water molecule. For simplicity, a number of protons are not shown explicitly, e.g., on Lys319, Tyr381, and the water molecules. (b) Cartoon diagram of aminopeptidase N with an amino acid bound in the active site. As in a, a number of protons have been omitted for simplicity.

about 0.4 Å, suggesting an expansion of the S1 pocket upon binding of the substrate/product.

In the L-arginine−ePepN (Arg) complex, the backbone of the arginine forms interactions similar to those seen for lysine (Figure 5b). The side chain also mimics that of lysine. At the top of the cylinder, one of the nitrogen atoms of the guanidinium group overlaps the nitrogen of the lysine and forms a hydrogen bond with Asn373 (2.9 Å). This nitrogen is close to Tyr376 (3.8 Å) and has one of its hydrogen atoms directed at the center of the aromatic ring, potentially forming a N−H• • •π interaction. The other nitrogen of the guanidinium group replaces water molecule W2 in the Lys structure and directly interacts with the Gln821 (2.9 Å). In addition, this nitrogen also forms a hydrogen bond with one of the waters of the triad. Similar to the Lys-bound structure, the side chain of Met260 is displaced from the S1 pocket. The main chain in the vicinity of Met260 also undergoes a similar outward movement (1.1 Å). The plane of the guanidinium group packs in the direction of the longest dimension of the cylinder and is stacked between the front and back faces.

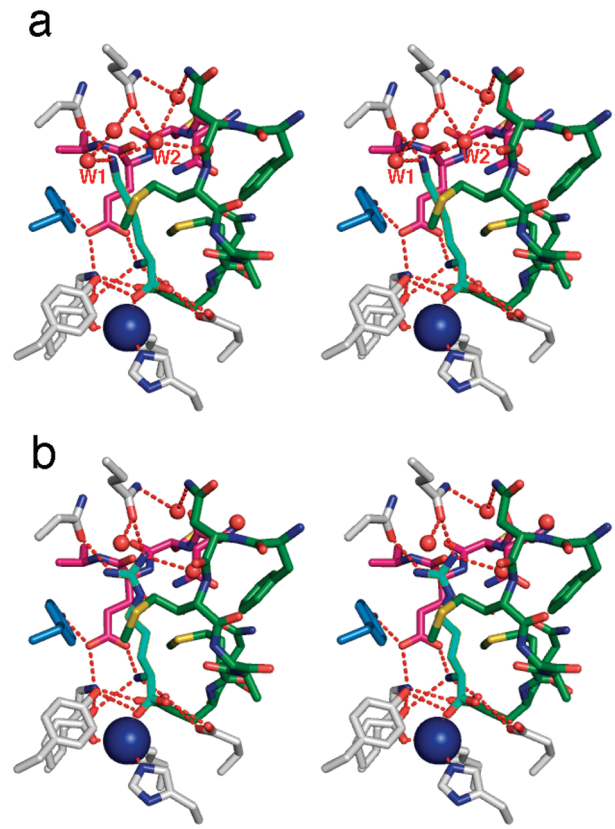

FIGURE 5: Stereo diagram of the active-site region for complexes of ePepN with lysine and arginine. (a) Lysine is shown in blue, with waters as red spheres. Both the bottom and top caps of the S1 subsite make extensive hydrogen-bonded networks with the P1 residue. (b) Arginine is shown in blue.

ePepN in Complex with Phe, Tyr, and Trp. The structure of ePepN in complex with the L-phenylalanine (Phe) was determined at 1.3 Å resolution, the highest among all ePepN structures determined to date. The amino and carboxylate groups of the phenylalanine form interactions with the active-site residues similar to those seen for the Lys and Arg structures, except that the carboxylate group adopts a second alternative conformation at higher occupancy (70%) (Figure 6a). In the new conformation, the carboxylate rotates further away (about 118°) from the direction of the peptide chain and interacts with the second oxygen atom of the carboxylate of Glu298 (2.6 Å). Notwithstanding, this alternative conformation of the carboxylate of the Phe, Glu298, moves away by about 0.2 Å with a slight rotation (15°) of its carboxylate group. The Phe side chain is stacked between Met260 in front and Gln119, Glu121, and Met263 at the back of the S1 pocket. In this conformation, the edges of the phenyl ring forms edge• • •π interactions with Phe258 (right) and Tyr376 (left).

The orientation of the phenyl ring in the Phe complex is perpendicular (about 105°) to that of the D-Phe in the bestatin−ePepN complex (bestatin), even though both occupy the S1 subsite (Figure 6a) (*7, 8*). In bestatin, the aromatic ring of D-Phe stacks over that of Tyr376. In concert with this difference in orientation, the side chain of Met260 remains in the S1 pocket, leaving the P1 side chain exposed to the solvent. In the Phe structure, it swings out and covers the side chain (Figure 6a). In the process, the Met260 side chain becomes disordered in the Phe structure relative to that in the bestatin complex.

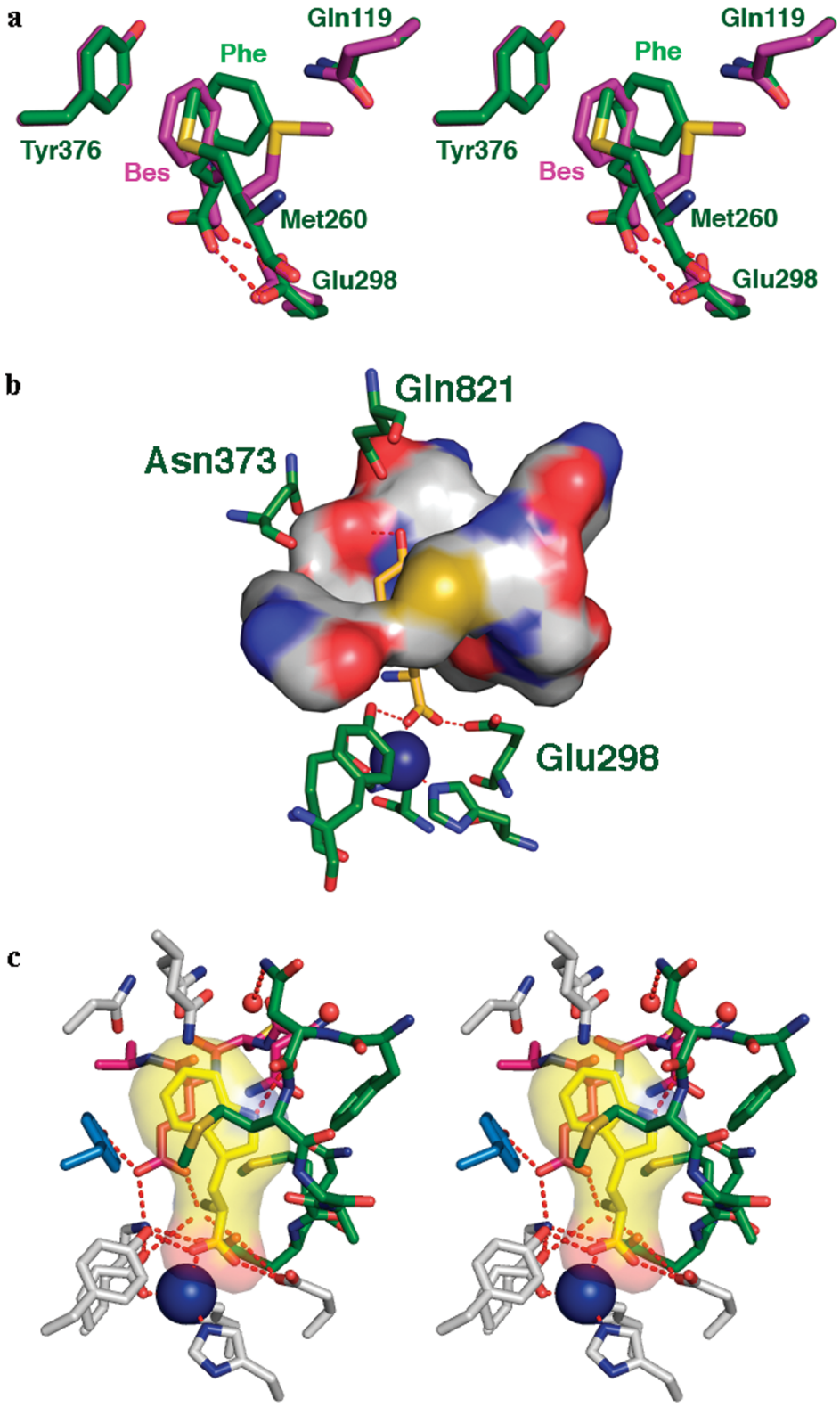

FIGURE 6: (a) Stereo diagram of the overlay of ePepN in the S1 pocket region in complex with Phe (green) and bestatin (magenta). The aromatic groups of bestatin and Phe are almost perpendicular. For the Phe complex, this causes the side chain of Met260 to move from inside to the surface of the hydrophobic cylinder. (b) Structure of Tyr (yellow) in complex with ePepN, depicted near the S1 pocket. The hydrophobic cylinder of the S1 pocket is represented by a surface diagram. Much of the hydrophobic region of the Tyr is enclosed within the hydrophobic cylinder. (c) Tryptophan (yellow) fills the entire pocket, making close contacts with the S1 subsite residues. The surface of the tryptophan is also shown in yellow.

Tryptophan and tyrosine bind in a fashion almost identical to that of phenylalanine, except that the tyrosine has only about 50% occupancy. The side chains of Tyr and Trp have an orientation similar to that of Phe (parts b and c of Figure 6). The carboxylate groups of these amino acids adopt the second conformation seen in the Phe structure. The tryptophan molecule fills the entire S1 pocket, making tight contacts with the protein. However, there is no appreciable difference between the holo and Trp structure, except in the region of Met260, as noted in the other complexes. The six-

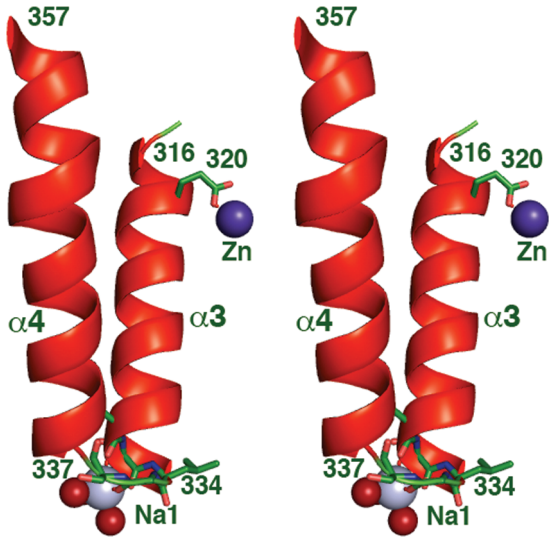

FIGURE 7: Sodium ion (gray) bound within the helix−turn−helix region. One of these α helices contains Glu320, which coordinates the zinc ions (blue sphere). Backbone carbonyls of Ser332, Asp334, and Gly335 in combination with three water molecules complete the octahedral geometry around the sodium ion.

membered ring of the indole side chain of the Trp makes contacts as short as 3.2 Å with the aromatic ring of Tyr376, and the indole nitrogen forms a hydrogen bond with a water molecule (3.0 Å) that is part of the triad. In the Tyr structure, the hydroxyl group hydrogen bonds to the main-chain carbonyl of the Asn373 (3.1 Å) and a water molecule that in turn interacts with the side chains of Asn373 and Gln821. Met260 is ordered in the Trp complex but not in that for Tyr.

Two Sodium Ions Bound to ePepN. The atomic resolution (1.3 Å) structure of the Phe complex made it possible to identify two sites on the surface of the ePepN that have octahedral geometry and appear to bind sodium ions. One of these presumed sodium ions (Na1) binds within the helix−turn−helix region between the α3 (Leu316−Leu334) and α4 (Arg337−Ala357) helices (Figure 7). One of these helices contains Glu320, which coordinates the active-site zinc ion. The main-chain carbonyls of Ser332 (2.4 Å), Asp334 (2.9 Å), Gly335 (2.3 Å), and three water molecules (2.2, 2.4, and 2.5 Å) provide octahedral coordination. The second sodium ion (Na2) binds on the surface of domain III. In addition to four water molecules (2.5, 2.5, 2.6, and 3.2 Å), the main-chain carbonyl (2.5 Å) and one of the side-chain oxygen atoms of Asp452 (2.4 Å) complete the octahedral ligation. All of the other structures in this paper, as well as the holo and bestatin-bound structures described earlier (*7*), also suggest the presence of these metal ions. In our studies, the activity of the ePepN increases by about 1.3 times in the presence of 100 mM NaCl relative to that of the sodium free. However, at concentrations of 200 mM or higher, the activity of the enzyme decreases to about 70% or lower (data not shown). The activity of one of the homologues of ePepN from *Streptococcus thermophilus* (StPepN) was shown to increase 2.3-fold in the presence of 100–280 mM NaCl (*3*). We previously suggested (*7*) that residues within the helix−turn−helix region that bind Na1 may undergo conformational changes between accepting and hydrolyzing peptides. It is possible that sodium ions play a role in stabilizing one or both of these states.

Role of Met260 and Orientation of the Carboxylate of the Product in the Active Site. The only residue that undergoes a substantial conformational change upon binding the various amino acids is Met260. In the holo structure, it is disordered and occupies the middle of the S1 pocket. In the complexes, it shifts outward and encloses the bound ligand within the S1 cylinder. The outward shift of Met260 occurs in all of the current structures, although there are some differences. For the Lys, Arg, and Trp complexes, Met260 adopts the same well-ordered conformation. In contrast, in the Phe and (partially occupied) Tyr structures, it is disordered. In all cases, the carboxylate group of the amino acid bound in the active site is roughly perpendicular to the overall peptide chain direction. In this configuration, one of the carboxylate oxygens replaces the water molecule that is present in the holo enzyme and is presumed to act as a nucleophile. In the Phe structure, the carboxylate adopts two conformations that respectively hydrogen bond to one or the other oxygen atoms of Glu298 (Figure 6a). In the Lys and Arg complexes, the plane of the carboxylate is about 80° away from the plane through N, Cα, and Cβ (Figures 1 and 5). In the Trp and Tyr complexes and one of the conformers of the Phe, the carboxylate rotates about 40° to interact with the second oxygen atom of the carboxylate of Glu298 (Figures 1 and 6). However, there is no correlation between the carboxylate orientation of the product amino acid and the behavior of Met260 in the S1 pocket. On the basis of the current structures, it is not clear how the product of catalysis leaves the active site. At a minimum, some sort of conformational change would be required. Presumably, the release of the product allows the nucleophilic water molecule to be replaced.

Specificity of ePepN. To avoid the possible influence of hetero groups, such as *para*-nitroanilide or amido-4-methylcoumarin, on enzymatic activity, we used the natural peptides Lys-Phe, Arg-Phe, Tyr-Phe, Phe-Phe, Ala-Phe, and Asp-Phe as substrates. Because we are interested in understanding the specificity of the S1-binding pocket, we fixed the P1′ amino acid as phenylalanine based on an earlier study (*3*). Consumption of the peptide was monitored using capillary electrophoresis. The kinetic data suggest that the preferred amino acid at the P1 position is, in decreasing order, Arg, Lys, Ala, Tyr, and Phe (Table 2). This is consistent with other studies (*4, 5*).

In the absence of structural information, it would be puzzling as to how the same enzyme could cleave large hydrophobic amino acids, such as Phe and Tyr, but also cleave very polar ones, such as Lys and Arg. The present complexes show how this occurs. The S1-binding site consists of a relatively long hydrophobic cylinder with polar groups at the distal end. The nonpolar walls of the cylinder readily accommodate the side chains of Phe and Tyr. This nonpolar region can also bind the hydrocarbon part of a lysine or arginine side chain. In addition, the polar cap at the end of the S1 cylinder can equally make hydrogen-bonding interactions that accommodate both the ε-amino group of lysine or the guanidinium group of arginine. It is also clear that the nonpolar walls of the S1 cylinder will not favor the binding of a short polar side chain, such as that of Asp, explaining its low catalytic preference (Table 2).

It is of interest that Ala is one of the most favored residues at the P1 position (Table 2). On one hand, the small,

Table 3: Conservation of Some of the Residues That Make Up the S1 Pocket in the ePepN, F3, and LTA4H Families[a]

ePepN (307 homologues)		F3 (157 homologues)		LTA4H (136 homologues)	
residue	conservation (%)	residue	conservation (%)	residue	conservation (%)
Gln119	99	His, Gln, Lys	8, 82, 4	Gln	99
Glu121	100	Glu, Gln	93, 6	Glu, Gln	14, 86
Met260, Phe, Ala*	90, 4, 3	Ala, Ser, Met	63, 12, 8	Tyr, Phe	38, 53
Ala262	100	Ala	99	Gly	96
Tyr376	99	Phe, Tyr	97, 2	Tyr, Phe	21, 62

[a] The names underlined indicate the residues present in the ePepN, F3, and LTA4H structures, respectively.

hydrophobic side chain can readily be accommodated within the S1 subsite. On the other hand, because of its smaller size, Ala would not be expected to bind as tightly as Phe or Tyr. Ito et al. (*8*) have suggested that the conformation of the side chain of Met260 in the holo form may play a role in the stability of peptides with alanine on the amino terminus.

Specificity of ePepN Homologues. A comparison of the crystal structures of ePepN, F3, and LTA4H indicates that they share an almost identically shaped S1 pocket. However, there are variations in the residues that make up the S1 pocket not only between these three structures but also between their homologues, as summarized in Table 3. Met260 is conserved in about 90% of the 367 homologues of ePepN, most of which are prokaryotic. In contrast, 62% of the 187 F3 homologues have alanine in this position. Similarly, 171 LTA4H homologues, which are predominantly eukaryotic, have either Phe (53%) or Tyr (38%). Another strong correlation in this region is at the "X" position in the GXM^{263}EN exopeptidase motif (Ala262 in ePepN). An alanine is conserved in 100% of ePepN and 99% of the F3 homologues, while it is a glycine for 96% of the LTA4H family.

Kinetic studies on LTA4H suggest that it prefers either basic or aromatic residues in the S1 pocket (*19*), similar to ePepN. This is consistent with the conservation of the S1 site. The substrate specificity of the F3 enzyme has not been examined. The presence of a histidine in place of Gln119 of ePepN that points into the S1 pocket and also the absence of a methionine to seal off the hydrophobic cylinder suggest that it may prefer an acidic residue at the P1 position. Although the shape of the S1 pocket is similar between ePepN and LTA4H, the latter acts as bifunctional enzyme by hydrolyzing leukotriene A4 to leukotriene B4 (natural substrate) in addition to its aminopeptidase activity. This suggests that small differences in the identity of residues in the similar shaped S1 subsite play a pivotal role in modulating the affinity of the enzyme. Together, it is conceivable that M^{260}GAMEN (as in ePepN), AGAMEN (as in F3), and F/YGGMEN (as in LTA4H homologues) act as three distinct functional exopeptidase motifs that play specific roles in the substrate recognition. On the basis of this, these three motifs can be classified as three distinct subclasses of thermolysin-based M1 class exopeptidases, namely, the Met-M1-gluzincins, Ala-M1-gluzincins and Phe/Tyr-M1-gluzincins, respectively.

Bacterial Homologues that Reduce the Bitterness in Cheese. Enzymatic hydrolysis of β-casein in milk results in a bitter taste, because of the formation of low-molecular-weight peptides composed of mainly hydrophobic amino acids. Several of these have basic or hydrophobic residues at the P1 position (*20*). Aminopeptidase N from lactic acid bacteria, such as *Lactococcus lactis* (LlPepN) (*21*) and *Streptococcus thermophilus* (StPepN) (*3*), that are shown to have a preference for peptides with basic and hydrophobic residues in the P1 position reduce the bitterness of cheese. Sequence alignment and homology modeling (data not shown) suggest that LlPepN and StPepN have the MGAMEN exopeptidase motif and a cylindrically shaped hydrophobic S1 pocket similar to ePepN. On the basis of these comparisons, we suggest that these aminopeptidases adopt a similar mechanism as that of the ePepN in selection and hydrolysis of the peptides that are responsible for the bitterness of cheese.

Multifunctional Mammalian Aminopeptidases. Mammalian homologues of ePepN, such as human cell-surface membrane-associated aminopeptidase N (APN), cytosolic puromycin-sensitive aminopeptidase (PSA), placental leucine aminopeptidase (also called as oxytocinase) (P-LAP), and LTA4H, are members of the M1 class gluzincin family. *In vivo*, each of these enzymes has specific substrates, while all of them display aminopeptidase activity with wide substrate preference *in vitro*. APN, *in vivo*, specifically converts the neuropeptide from angiotensin II to III by removing arginine from the P1 position (*22*). PSA, an important cell-cycle-regulating enzyme, functions by trimming antigen peptides (*23*). While, *in vivo*, P-LAP controls the concentration of peptides, such as oxytocin, that are important in preventing premature delivery, *in vitro*, it cleaves peptides with hydrophobic and basic residues at the P1 position (*24*). LTA4H, in addition to its epoxy hydrolase function on leukotriene A4, also acts as an aminopeptidase (*25*). Of the known members of the M1 class mammalian aminopeptidases, only the crystal structure of LTA4H has been determined (*9*). As noted above, its active site is very similar to that of ePepN. On the basis of sequence homology and structural comparison, it seems clear that all of the M1 class aminopeptidases share a common mechanism of action.

ACKNOWLEDGMENT

We acknowledge the use of beamline BL8.2.2 at the Advanced Light Source for X-ray data collection and thank Jean Delaney for help in preparing the manuscript.

REFERENCES

1. Hooper, N. M. (1994) Families of zinc metalloproteases. *FEBS Lett. 354*, 1–6.
2. Luoma, S., Peltoniemi, K., Joutsjoki, V., Rantanen, T., Tamminen, M., Heikkinen, I., and Palva, A. (2001) Expression of six peptidases from *Lactobacillus helveticus* in *Lactococcus lactis*. *Appl. Environ. Microbiol. 67*, 1232–1238.
3. Chavagnat, F., Casey, M. G., and Meyer, J. (1999) Purification, characterization, gene cloning, sequencing, and overexpression of

aminopeptidase N from *Streptococcus thermophilus* A. *Appl. Environ. Microbiol. 65*, 3001–3007.
4. Chandu, D., and Nandi, D. (2003) PepN is the major aminopeptidase in *Escherichia coli*: Insights on substrate specificity and role during sodium-salicylate-induced stress. *Microbiology 149*, 3437–3447.
5. Chandu, D., Kumar, A., and Nandi, D. (2003) PepN, the major Suc-LLVY-AMC-hydrolyzing enzyme in *Escherichia coli*, displays functional similarity with downstream processing enzymes in Archaea and eukarya. Implications in cytosolic protein degradation. *J. Biol. Chem. 278*, 5548–5556.
6. Chappelet-Tordo, D., Lazdunski, C., Murgier, M., and Lazdunski, A. (1977) Aminopeptidase N from *Escherichia coli*: Ionizable active-center groups and substrate specificity. *Eur. J. Biochem. 81*, 299–305.
7. Addlagatta, A., Gay, L., and Matthews, B. W. (2006) Structure of aminopeptidase N from *Escherichia coli* suggests a compartmentalized, gated active site. *Proc. Natl. Acad. Sci. U.S.A. 103*, 13339–13344.
8. Ito, K., Nakajima, Y., Onohara, Y., Takeo, M., Nakashima, K., Matsubara, F., Ito, T., and Yoshimoto, T. (2006) Crystal structure of aminopeptidase N (proteobacteria alanyl aminopeptidase) from *Escherichia coli* and conformational change of methionine 260 involved in substrate recognition. *J. Biol. Chem. 281*, 33664–33676.
9. Kyrieleis, O. J., Goettig, P., Kiefersauer, R., Huber, R., and Brandstetter, H. (2005) Crystal structures of the tricorn interacting factor F3 from *Thermoplasma acidophilum*, a zinc aminopeptidase in three different conformations. *J. Mol. Biol. 349*, 787–800.
10. Thunnissen, M. M., Nordlund, P., and Haeggstrom, J. Z. (2001) Crystal structure of human leukotriene A(4) hydrolase, a bifunctional enzyme in inflammation. *Nat. Struct. Biol. 8*, 131–135.
11. Nocek, B., Mulligan, R., Bargassa, M., Collart, F., and Joachimiak, A. (2007) Crystal structure of aminopeptidase N from human pathogen *Neisseria meningitidis*, *Proteins*, in press.
12. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307–326.
13. Brunger, 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) Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 54*, 905–921.
14. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 53*, 240–255.
15. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 2126–2132.
16. Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G., and Thompson, J. D. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res. 31*, 3497–3500.
17. Sander, C., and Schneider, R. (1991) Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins 9*, 56–68.
18. Matthews, B. W. (1988) Structural basis of the action of thermolysin and related zinc peptidases. *Acc. Chem. Res. 21*, 333–340.
19. Orning, L., Gierse, J. K., and Fitzpatrick, F. A. (1994) The bifunctional enzyme leukotriene-A4 hydrolase is an arginine aminopeptidase of high efficiency and specificity. *J. Biol. Chem. 269*, 11269–11273.
20. Saha, B. C., and Hayashi, K. (2001) Debittering of protein hydrolyzates. *Biotechnol. Adv. 19*, 355–370.
21. Klein, J. R., Klein, U., Schad, M., and Plapp, R. (1993) Cloning, DNA sequence analysis and partial characterization of pepN, a lysyl aminopeptidase from *Lactobacillus delbruckii* ssp. lactis DSM7290. *Eur. J. Biochem. 217*, 105–114.
22. Albiston, A. L., Ye, S., and Chai, S. Y. (2004) Membrane bound members of the M1 family: More than aminopeptidases. *Protein Pept. Lett. 11*, 491–500.
23. Hattori, A., and Tsujimoto, M. (2004) Processing of antigenic peptides by aminopeptidases. *Biol. Pharm. Bull. 27*, 777–780.
24. Tsujimoto, M., and Hattori, A. (2005) The oxytocinase subfamily of M1 aminopeptidases. *Biochim. Biophys. Acta 1751*, 9–18.
25. Haeggstrom, J. Z., Tholander, F., and Wetterholm, A. (2007) Structure and catalytic mechanisms of leukotriene A4 hydrolase. *Prostaglandins Other Lipid Mediators 83*, 198–202.

BI7022333