

Crystal Structure of a Dodecameric Tetrahedral-shaped Aminopeptidase*

Received for publication, August 17, 2004, and in revised form, September 7, 2004
Published, JBC Papers in Press, September 16, 2004, DOI 10.1074/jbc.M409455200

Santina Russo and Ulrich Baumann‡

From the Departement für Chemie und Biochemie, University of Berne, Freiestrasse 3, CH-3012 Bern, Switzerland

Protein turnover is an essential process in living cells. The degradation of cytosolic polypeptides is mainly carried out by the proteasome, resulting in 7–9-amino acid long peptides. Further degradation is usually carried out by energy-independent proteases like the tricorn protease from *Thermoplasma acidophilum*. Recently, a novel tetrahedral-shaped dodecameric 480-kDa aminopeptidase complex (TET) has been described in *Haloarcula marismortui* that differs from the known ring- or barrel-shaped self-compartmentalizing proteases. This complex is capable of degrading most peptides down to amino acids. We present here the crystal structure of the tetrahedral aminopeptidase homolog FrvX from *Pyrococcus horikoshii*. The monomer has a typical clan MH fold, as found for example in *Aeromonas proteolytica* aminopeptidase, containing a dinuclear zinc active center. The quaternary structure is built by dimers with a length of 100 Å that form the edges of the tetrahedron. All 12 active sites are located on the inside of the tetrahedron. Substrate access is granted by pores with a maximal diameter of 10 Å, allowing only small peptides and unfolded proteins access to the active site.

The protein content of all living cells is constantly renewed through synthesis of new proteins and degradation of unneeded or misfolded proteins. This catabolism of proteins is a key cellular function and must be under spatial and temporal control to avert damage to the cell. Because prokaryotes lack membrane-bound compartments, degradation mostly takes place in large macromolecular self-compartmentalizing assemblies whose active sites are arranged in an inner cavity in order to only allow proteolysis of unfolded substrates.

In all three kingdoms the degradation of cytosolic proteins is carried out predominantly by the ATP-dependent proteasome or similar energy-dependent proteases that generate oligopeptides 7–9 amino acids long (1–3). These products are thought to be further processed by assemblies like the energy-independent tricorn protease found in *Thermoplasma acidophilum* (4–6). Other large energy-independent protease complexes that putatively take part in the degradation of oligopeptides created by the proteasome are the mammalian TPPII protease (7), the

DppA D-aminopeptidase from *Bacillus subtilis* (8), and yeast bleomycin hydrolase (9). All these multimeric complexes are metalloenzymes and are composed of rings or barrels with a single central channel and only two openings.

Recently, a novel protease complex called tetrahedral aminopeptidase (TET)¹ has been isolated from the Archaea *Haloarcula marismortui* (10). Electron microscopy analysis of TET at 17 Å resolution showed that it is a 0.4-MDa homododecameric complex with a novel tetrahedral shape that is made up by association of six antiparallel dimers. Contrary to all other self-compartmentalizing proteases described before, the central cavity of this complex is accessible through four narrow channels and through four wider channels. By spectrophotometric assays TET was shown to have aminopeptidase activity of broad specificity with a preference for neutral and basic amino acids (10). The complex most efficiently processes oligopeptides of 9–12 amino acids length. Longer peptides were digested slowly, and cleavage was completely absent for peptides 40 residues long. Furthermore, TET was shown to share significant sequence homology with assigned dinuclear zinc aminopeptidases from bacterial and Archaeal species. Among them is FrvX from the hyperthermophilic Archaea *P. horikoshii*, which shares 27% sequence identity with TET.

FrvX is a polypeptide chain of 353 amino acids with a molecular mass of 39 kDa and a theoretical pI of 5.69. The protein mainly consists of a peptidase domain that belongs to the clan MH, family M42, according to the MEROPS data base (11). This domain contains conserved binding sites for two catalytic zinc ions in the active site, bound by His, Asp, Glu, Asp, and His. Such dinuclear metal clusters promote catalysis by providing a site for substrate binding, polarizing the carbonyl group for nucleophilic attack, and stabilizing the transition state of the reaction (12). Most of the aminopeptidases with such a co-catalytic active site are thought to be zinc enzymes, but typically activity can also be partially or fully observed with other divalent metal ions like Co²⁺, Mg²⁺, Mn²⁺, and Ni²⁺. Furthermore, activity can also be present with only one metal ion bound in the active site.

For a number of dinuclear aminopeptidases there is structural information available, e.g. for bovine lens leucine aminopeptidase (13), *Escherichia coli* methionine aminopeptidase (14), and *A. proteolytica* aminopeptidase (ApAp) (15), which define three subfamilies of dinuclear aminopeptidases (16). All of them share similar folds of the catalytic domain but employ different active site geometries and/or metal ligands. Sequence similarity suggests that FrvX together with all proteases containing a peptidase M42 domain belongs to the ApAp subfamily.

Here we have presented the crystal structure of the energy-independent tetrahedral-shaped aminopeptidase complex FrvX

* This work was supported by Grant 31–67253.01 from the Swiss National Science Foundation, by the Berner Hochschulstiftung, and by the University of Berne. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1XFO) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

‡ To whom correspondence should be addressed. Tel.: 41-031-631-4320; Fax: 41-031-631-4887; E-mail: ulrich.baumann@ibc.unibe.ch.

¹ The abbreviations used are: TET, tetrahedral aminopeptidase; ApAp, *Aeromonas proteolytica* aminopeptidase; r.m.s.d., root mean square deviation; contig, group of overlapping clones.

TABLE I
 Data collection and refinement statistics

	Native	Zinc peak	Trigonal native
Crystal parameters	P6 ₃ ; $a = b = 157.9 \text{ \AA}$, $c = 114.2 \text{ \AA}$ 4 molecules /a.s.u. ^a	P6 ₃ ; $a = b = 160.1 \text{ \AA}$, $c = 115.6 \text{ \AA}$ 4 molecules /a.s.u. ^a	P3 _{1/2} 21; $a = b = 160.0$, $c = 418.2$ 12 molecules /a.s.u. ^a
Data collection (XDS)			
Beamline	X06SA, SLS	BW7A, DESY	X06SA, SLS
Wavelength (Å)	0.9791	1.2855	0.92003
Resolution range ^b	40–2.09 (2.21–2.09)	40–2.74 (2.9–2.74)	40–5.0 (5.28–5.0)
No. observations	423,856	698,722	149,006
No. unique reflections	92,985	85,740 ^c	49,903
Completeness (%)	97.2 (82.3)	98.3 (89.7)	96.1 (85.8)
R_{sym}^d (%)	5.9 (43.6)	5.7 (49.8)	8.5 (74.4)
$I/\sigma(I)$	16.7 (3.2)	33.6 (4.6)	6.96 (1.04)
Refinement (REFMAC)			
Resolution range (Å)	25–2.25		
No. reflections working set	74,367		
No. reflections test set	2,300		
No. non-hydrogen atoms	10,844		
Solvent water molecules	305		
R/R_{free} (%)	18.27 / 22.12		
R.m.s.d. bond length (Å)	0.011		
R.m.s.d. bond angles (deg.)	1.243		

^a Asymmetric unit.

^b The values in parentheses of resolution range, completeness, R_{sym} , and $I/\sigma(I)$ correspond to the outermost resolution shell.

^c Friedel pairs were treated as different reflections.

^d $R_{\text{sym}} = \sum_{hkl} \sum_j |I(hkl;j) - \bar{I}(hkl)| / (\sum_{hkl} \sum_j I(hkl;j))$, where $I(hkl;j)$ is the j th measurement of the intensity of the unique reflection (hkl) and $\bar{I}(hkl)$ is mean overall symmetry-related measurements.

from *P. horikoshii*. This dodecameric protein has the same macromolecular shape as the previously described TET from *H. marismortui*. The structure of the monomers and the active site strongly resembles the *A. proteolytica* aminopeptidase.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification—The gene encoding FrvX was amplified by PCR (17) from *P. horikoshii* genomic DNA and cloned into a pET-28a vector (Novagen) introducing a N-terminal His₆ tag. The protein was expressed in *E. coli* RosettaTM (DE3) cells (Novagen). Cells were cultured in Luria Bertani medium at 37 °C to an A₆₀₀ of 0.8 before expression was induced by 1 mM isopropyl-β-D-thiogalactopyranoside. After 14 h of induction, cells were harvested, resuspended in 20 mM Tris-HCl, pH 8.0, 0.02% sodium azide, and disrupted by sonication. After centrifugation at 41,000 × *g* for 30 min, the supernatant was heat-treated at 88–90 °C for 3 min and centrifuged again at 41,000 × *g* for 40 min. The supernatant was applied to a DEAE-Sepharose ion exchange column (Amersham Biosciences) and eluted with a linear gradient from 0–1 M NaCl. The fractions containing FrvX were pooled, concentrated using a Centriprep-30 device (Millipore), and dialyzed against 20 mM Tris-HCl, pH 8.0. The protein solution was then loaded on a MonoQ column (Amersham Biosciences) and eluted by a 0–500 mM NaCl gradient. FrvX was further purified by size exclusion chromatography using a Superdex 75 column (Amersham Biosciences) equilibrated in 20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 2 mM dithiothreitol, 0.02% sodium azide. The purification steps were monitored by SDS-PAGE analysis (18). It was found to be essential for the quality of the crystals that the first purification step was an ion exchange column. Protein initially purified by nickel-nitrilotriacetic acid affinity chromatography only yielded poorly diffracting crystals. Cleavage of the His₆ tag using thrombin (Amersham Biosciences) also yielded crystals of much inferior quality. Typical yields were 15 mg/liter pure protein as measured by Bradford assay (Bio-Rad).

Crystallization—Hexagonal crystals of FrvX were obtained after 2 days by the sitting drop vapor diffusion method. Drops were set up by mixing 2 μl of protein solution (8 mg/ml) with 1 μl of reservoir solution (0.1 M sodium-citrate, pH 6.1, 18% (w/v) polyethylene glycol 400, 0.1 M NaCl) and equilibrated against 100 μl of reservoir solution at 18 °C. Typical crystals have an average size of 200 × 50 × 50 μm³. The spacegroup was determined as P6₃ with cell constants of $a = b = 157.9 \text{ \AA}$, $c = 114.2 \text{ \AA}$, with four monomers in the asymmetric unit.

Trigonal crystals of lower quality were obtained from protein batches purified by nickel-nitrilotriacetic acid chromatography as the first purification step and subsequent cleavage of the His₆ tag. 2 μl of protein solution (8 mg/ml) were mixed with 1 μl of reservoir solution (50 mM glycine, pH 3.1, 2% (w/v) polyethylene glycol 8000, 1.0 M lithium sulfate) and equilibrated against 100 μl of reservoir solution at 18 °C. Crystals were observed after 3 days and grew to an average size of 200 × 100 ×

100 μm³. They belong to spacegroup P3_{1/2}21 with cell constants of $a = b = 160.0 \text{ \AA}$, $c = 418.2 \text{ \AA}$. Only the hexagonal crystals were suitable for structure solution.

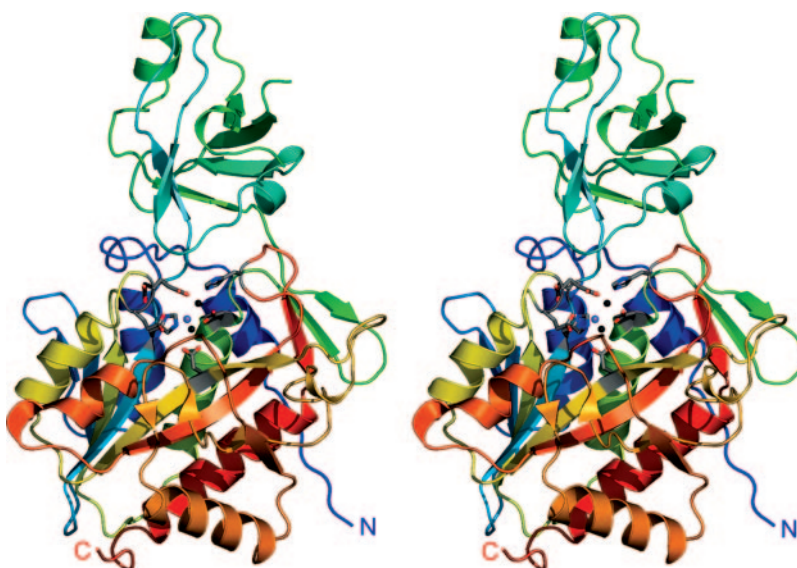
Data Collection—Prior to data collection, crystals were cryoprotected by raising the polyethylene glycol 400 concentration to 35% (v/v) and flash-cooled in a nitrogen stream at 110 K. High resolution (2.25 Å) native data were collected on beamline X06SA at the Swiss Light Source in Villigen using a MAR-165 CCD detector (Marresearch, Hamburg, Germany). A single wavelength anomalous dispersion experiment employing the zinc absorption edge was carried out at beamline BW7A at the European Molecular Biology Laboratory Hamburg Outstation at the Deutsches Elektron Synchrotron in Hamburg using a MAR-165 CCD detector (Marresearch, Hamburg, Germany). Exposure times varied from about 30 s (Deutsches Elektron Synchrotron) to 1 s (Swiss Light Source) for 0.2-degree oscillations. All datasets were integrated and scaled with XDS (19–21). Data collection statistics are summarized in Table I.

Structure Solution, Refinement, and Analysis—Initially, the phase problem was tackled using seleno-methionine-labeled protein. However, the resulting crystals were of much inferior quality, exhibiting split Bragg reflections and signs of twinning. During the efforts to improve those crystals, the structures of the *Thermotoga maritima* (PDB code 1VHO) and *B. subtilis* (PDB code 1VHE) orthologs became available. Initial phases were then determined by molecular replacement using the program MOLREP (22) employing chain A of aminopeptidase 1VHE from *B. subtilis* as a search model. Rigid body refinement was followed by NCS averaging and phase extension using CNS software (23). Automated model building was done by ArpWarp (24). The resulting model was completed using O (25). Refinement was carried out by REFMAC (26). The data in the outmost resolution shell (2.21–2.09 Å) showed large deviations in a Wilson plot and were excluded from refinement. At this point the model contained one Zn⁺² ion/monomer, such as in the original search probe. An anomalous difference Fourier map using the zinc peak dataset revealed a more weakly occupied position of a second Zn²⁺ ion in the active site of three monomers; these zinc ions were subsequently included in all four subunits. Refinement statistics are given in Table I.

The quality of the structure was checked with PROCHECK (27). Comparison of three-dimensional structures was effected using DALI. Structure figures were created using the program PYMOL (www.pymol.org).

Activity Assay—The aminopeptidase activity of FrvX was examined using aminoacyl-*para*-nitroanilides as substrates. The reactions were carried out at 90 °C in 50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM ZnCl₂ with 0.5 mM substrate and 2 or 10 μg/ml enzyme. The concentration of product formed was determined by measuring the absorption of released *para*-nitroaniline at 405 nm ($\epsilon_{405} = 10,400 \text{ M}^{-1} \text{ cm}^{-1}$) every 5 min up to 30 min.

FIG. 1. **Stereo schematic of the FrvX monomer structure.** The rainbow coloring runs from blue (N terminus) to red (C terminus). Active site residues and zinc ions are shown as sticks and spheres. Zinc ions are shown as black spheres. The catalytic water is shown as a violet ball.



RESULTS

Structure Refinement—The four monomers present in the hexagonal asymmetric unit of FrvX comprise residues 6–121 and 133–353 (polypeptide chains B, C, and D) and residues 1–121 and 133–353, as well as four additional residues of the N-terminal His tag (chain A). In all four monomers, residues 122–132, which are part of a solvent-exposed loop, were not visible in the electron density. Altogether, the final model consists of 1357 amino acid residues, 601 water molecules, and 8 zinc cations. Refinement statistics are given in Table I. A Ramachandran plot as defined by Kleywegt and Jones (28) shows 1.5% outliers.

Monomer Structure—The monomer fold resembles those found in other MH clan aminopeptidases, e.g. from *A. proteolytica* (PDB code 1AMP, (29)), but is also similar to the related clan MF leucine aminopeptidase from *Bos taurus* (PDB code 1BPM, (13)). These two structures have a DALI (30) Z-score/ C^{α} r.m.s. deviations of 23.4/2.6 and 13.2/3.3 Å. The sequence identities are only 17 and 9%, respectively. Much more similar are the deposited but hitherto unpublished structures from *B. subtilis* (PDB code 1VHO) and *T. maritima* (PDB code 1VHE), which exhibit C^{α} -r.m.s. deviations of about 1 Å.

The FrvX monomer is composed of two globular α/β domains and has approximate dimensions of $62 \times 60 \times 47$ Å (Fig. 1). The bigger domain (domain I) contains the N-terminal residues 1–72 and the C-terminal amino acids 165–353. The second domain (domain II) is smaller (amino acids 74–164) and less structured. The main feature of domain I is a central eight-stranded β -sheet that is surrounded by α -helices. The concluding two strands on each side of the sheet are antiparallel and rather short, whereas the central four strands are parallel and much longer. The sheet is sandwiched between two groups of α -helices. The first group consists of five helices, including the N-terminal two; the second group is located on the other side of the β -sheet and comprises two short α -helices. In addition, a smaller, three-stranded antiparallel β -sheet is located at the surface of the molecule.

Domain two is built of a tilted, six-stranded antiparallel β -sheet flanked by two very short α -helices. The two domains are connected by two loop regions. Of these, the first extends from a central β -strand of domain I to the second strand of the β -sheet of domain II; the second domain extends from the third strand of the sheet of domain II to the small surface-located β -sheet of domain I. The first link comprises two ligands of one active site zinc ion.

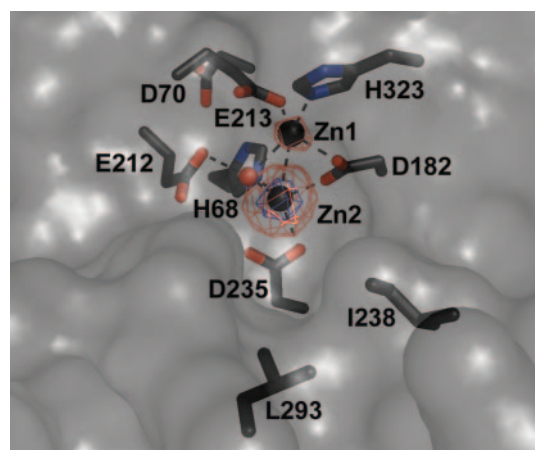


FIG. 2. **Close-up view of the active site.** Amino acid side chains are shown as sticks. Atomic colors are as follows: oxygen, red; nitrogen, blue; carbon, gray. Zinc ions and the water molecule are shown as black or red spheres, respectively. The imaginary Fourier map depicted in chicken wire representation is contoured at 18 (blue) and 9 (orange) S.D. above the mean.

The active site is located in the cleft between the two domains that is mainly generated by a large, solvent-exposed loop stabilized by crystal contacts. All catalytic residues are part of domain I.

Dinuclear Active Site—Contrary to the search probe 1VHE, the active site contains in the refined FrvX model a second zinc ion (denoted as Zn1) identified by imaginary Fourier maps using reflection data measured at the zinc edge. The catalytic water molecule was identified in the $2F_o - F_c$ and $F_o - F_c$ electron density maps in one subunit and subsequently added to all four crystallographically independent monomers. A close-up view of the active site of FrvX is shown in Fig. 2. Both zinc ions are tetrahedrally coordinated and are bridged by Asp-182 and a water molecule. The ion denoted as Zn1 is additionally coordinated by His-323 and Glu-213 and Zn2 by His-68 and Asp-235. It is possible that the two terminal acidic residues Glu-213 and Asp-235 provide a fifth ligand to both zinc ions. Glu-212 forms a hydrogen bond with the water molecule. The distance between the two metal ions is about 3.4 Å. Overall, the active site of FrvX is very similar to the one of the ApAp. The dinuclear metal center has the same symmetric geometry and the same amino acid ligands. An overlay of the metal clusters of FrvX and *Aeromonas* enzyme reveals only

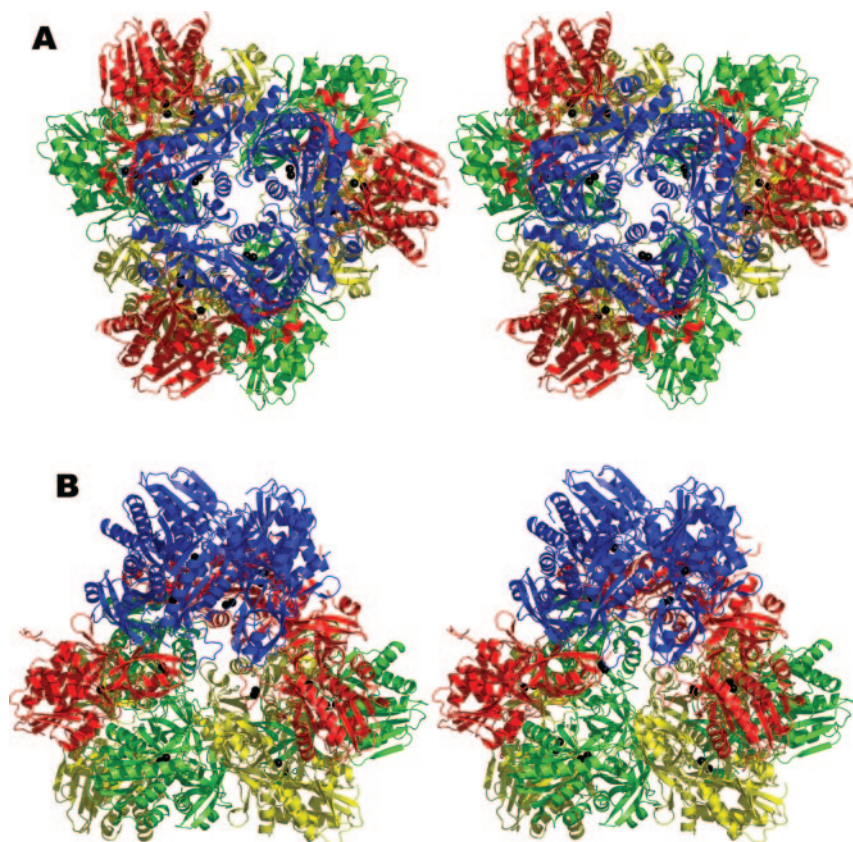


FIG. 3. **The dodecameric tetrahedron.** Schematic representation, zinc ions are shown as *black spheres*. *A*, top stereo view down the crystallographic 3-fold axis. *B*, side stereo view, approximately along a non-crystallographic 3-fold axis. The edge dimer subunits A and D and their symmetry equivalents are colored in *red* and *blue*, respectively. The B and C subunits of the asymmetric unit and their symmetry equivalents are *green* and *yellow*.

very little change in the position of the ligands and metal ions.

Because of the similarity of the active sites of FrvX and ApAp, we suggest that the reaction mechanism of FrvX is similar to the one proposed for ApAp (31), which is mainly based on electron paramagnetic resonance spectroscopic and crystallographic data (15, 32, 33). In this mechanism, the carbonyl group of the scissile peptide bond first binds to Zn1, followed by binding of the N-terminal amine group to Zn2. This coordination of the nitrogen atom to Zn2 putatively leads to the break of the interaction between Zn2 and the catalytic water molecule and, hence, transfer of the latter to Zn1. Glu-212 (Glu-151 in ApAp) is thought to act as general base, inducing the formation of the nucleophilic hydroxide ion, which attacks the carbonyl group leading to the formation of a tetrahedral intermediate. The last and rate-limiting step is consequently the break of the peptide bond involving a proton transfer from Glu-212 to the new N-terminal amine group of the leaving peptide.

It is well known and accepted that some dinuclear peptidases are at least partially active with only one metal bound in the active site. For the *Aeromonas* enzyme it was shown that the two zinc ions possess different affinities for their binding site and also a different effect on enzymatic activity (34, 35). If only the high affinity site (Zn1) is occupied, the protein still retains about 80% of its activity. It has been proposed that in the absence of a metal ion bound in the low affinity site (Zn2), the N-terminal amine group is bound by Asp-179 (Asp-235 in FrvX).

For FrvX, an anomalous difference Fourier calculation of data measured at the zinc absorption edge provides insight into the occupancy level of both metal ions (Fig. 2). Interestingly, in all four monomers embodied in the asymmetric unit the Zn2 site possesses the higher occupancy. This indicates that contrary to ApAp, in FrvX the Zn2 site exhibits higher affinity toward a zinc ion and consequently is more important for enzyme activity than the Zn1 site. This finding may suggest

that the roles of Zn1 and Zn2 in FrvX are inverted. Whether FrvX is active with only one metal ion bound and, if this is the case, which zinc binding site is vital for the reaction to occur needs further investigation.

Dodecamer Structure—Gel filtration experiments indicated that FrvX is a dodecamer in solution. Similar experiments (data not shown) employing homologs from *Lactococcus lactis* (Swiss-Prot accession number Q48677) and *Thermoanaerobacter tengcongensis* (Swiss-Prot accession number Q8RCI7) hinted at dimeric and monomeric quaternary structures, respectively. The self rotation function of FrvX in the trigonal and hexagonal crystal forms revealed the presence of tetrahedral particles, which was confirmed by the high resolution structure of the hexagonal space group on which we will focus in the following.

The four subunits embodied in the asymmetric unit build up the physiological tetrahedral-shaped dodecamer by application of the crystallographic 3-fold axis located at $(x = 1/3, y = 2/3, z)$. Each edge of the tetrahedron has a length of about 100 Å and is built by a dimer (chains A and D in the asymmetric unit). Chains B and C each form half of an edge; these edges are then completed by the symmetry-equivalent subunits generated by the crystallographic 3-fold axis (Fig. 3). The intersubunit contacts involve rather extended surfaces: formation of the “edge dimer” from monomers buries 3300 Å² of the solvent-accessible surface, and the AC and AB interfaces measure about 2300 Å² each. Contact residues are indicated in Fig. 4. The contacts are of polar and non-polar nature.

Interestingly, the *B. subtilis* ortholog (37% sequence identity to FrvX, TREMBL P95421, PDB entry 1VHE), which was used as the molecular replacement probe, appears to be dodecameric in the crystal lattice. Here, the tetrahedron is generated entirely by the crystallographic symmetry of its spacegroup F432. The resulting tetrahedron is similar to the one observed in the FrvX structure with a slightly different orientation of the sub-

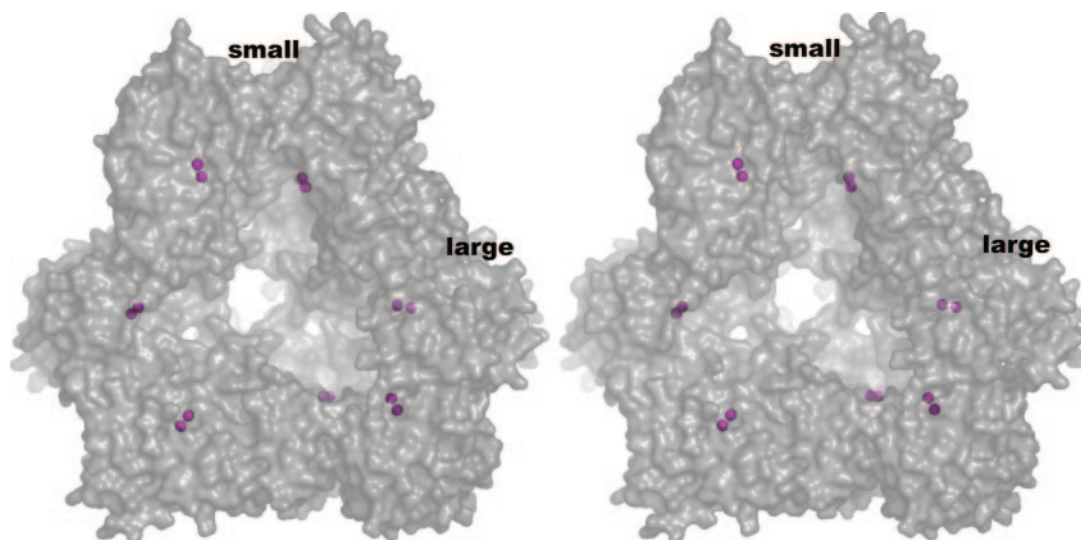


FIG. 5. **Internal view of FrvX in a cut-open surface presentation.** The active site zinc ions are shown as magenta spheres. The view is approximately along one of the wide channels. Openings of large and small channels roughly perpendicular to the view direction are indicated.

B. subtilis enzyme is a dodecamer under physiological conditions as well.

The homologous enzyme from *T. maritima* (TREMBL Q9X0E0, PDB entry 1VHO), which shares 30% sequence identity with FrvX, appears to be a monomer in the crystal lattice. This is surprising because most of the crucial contacts appear to be conserved (Fig. 4). Some more conspicuous changes occur at position 220, where the dodecameric enzymes possess an arginine or lysine and the monomeric enzyme a leucine, and at position 167, where the *Thermotoga* enzyme has a deletion. However, the major reason for the difference between the *T. maritima* and the two dodecameric enzymes appears to reside in the loop between residues 121 and 132, which is invisible in the FrvX structure presented here. This loop is located at the inside of the tetrahedron and close to the edge dimer interface between monomers A and D. Although in the dodecameric *B. subtilis* enzyme this loop is rather well ordered and has a conformation compatible with the visible amino acids in FrvX, the conformation is very different in the monomeric *T. maritima* enzyme and interferes with the edge dimer formation and, consequently, with the dodecameric quaternary structure.

All 12 active sites are located inside the tetrahedron. Access is given essentially by four channels located at the faces of the tetrahedron. The channel, which originates on the center of the triangular face, has a diameter of about 10 Å. Its entrance walls are built by the type II β -turns 80–82 and 242–245. Another channel was described for TET from *H. marismortui*; this channel runs through the vertices of the tetrahedron and is formed by three symmetry-equivalent type II β -turns (residues 306–309). In the case of FrvX, this channel is nearly completely blocked by Phe-224. Even the larger channel is too small to allow properly folded proteins to enter; only small peptides or unfolded protein chains can serve as substrate (Fig. 5).

The shortest distance between active sites is about 30 Å and occurs between the three subunits located at the vertices of the tetrahedron. The largest distance between two active sites is about 63 Å and spans nearly the whole internal cavity.

Activity Assay—The activity of FrvX was examined employing eight mono-amino-*para*-nitroanilides (Ala, Leu, Val, Arg, Lys, Glu, Met, Pro) and an Ala-Ala-Phe-amino-*para*-nitroanilide. The results are summarized in Fig. 6. Leucine-*para*-nitroanilide was shown to be processed by FrvX at 81.3 $\mu\text{mol}/\text{mg}/\text{hour}$. All other tested amino acids are released in significantly lower rates with preference for acidic and neutral residues.

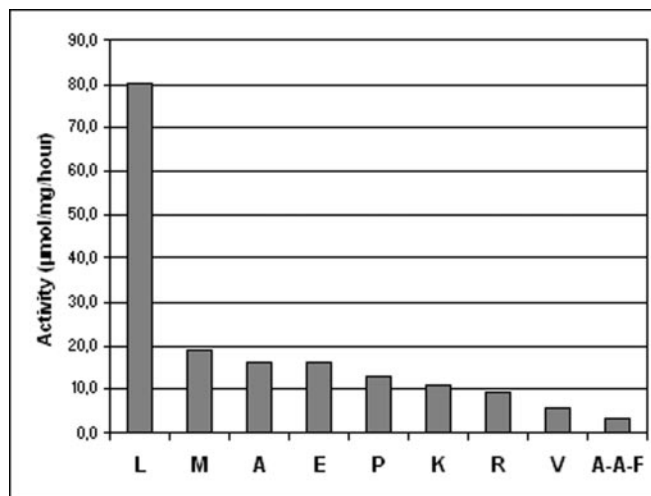


FIG. 6. **Substrate specificity of FrvX.** Assays were performed as described under “Experimental Procedures” employing the indicated amino acid-*para*-nitroanilides.

Interestingly, valine, which of the examined amino acids is most similar to leucine, is processed the slowest. The *para*-nitroaniline group of the tripeptide is released only at a rate of 3.5 $\mu\text{mol}/\text{mg}/\text{hour}$. This can be partially explained as a delay caused by the need for cleaving not one, but three, bonds before *para*-nitroaniline release. The second reason is probably a lower affinity for phenylalanine, which can be explained by the size of the active site pocket. The substrate pocket, which is mainly formed by Leu-293 and Ile-238, seems to be perfected for leucine. In addition, slightly longer side chains like the methionine and glutamic acid side chains can be well fitted in the pocket, the latter being putatively stabilized by H-bonds with Lys-261. It seems clear though that larger side chains, like aromatic residues, collide with Leu-293. The valine side chain probably collides with Thr-298, which would explain the low hydrolyzation rate of Val-*para*-nitroanilide.

DISCUSSION

The crucial role of aminopeptidases has been highlighted in many reviews (16, 36–38). They have important functions in protein maturation, metabolism of biologically active peptides, and protein degradation. In some Archaea, like *T. acidophilum*,

the tricorner protease together with its three monomeric aminopeptidase factors is involved in the further degradation of peptides generated by the proteasome. However, the occurrence of tricorner protease seems not to be ubiquitous. *Halobacterium* and *Pyrococcus*, for example, do not possess obvious homologs of tricorner. The large, self-compartmentalizing aminopeptidase TET from *Haloarcula* possesses the necessary activity for processing most peptides down to amino acids. Contrary to all other large self-compartmentalizing ring- or barrel-shaped protease complexes like the proteasome, TET possesses a tetrahedral-shaped dodecameric architecture. We present here for the first time the crystal structure of the similar dodecameric aminopeptidase FrvX from *P. horikoshii*. This protease belongs to the M42 aminopeptidase family and shares a similar fold and active site as *A. proteolytica* aminopeptidase. The tetrahedron is built by 6 dimers forming the edges in such a way that all 12 active centers are located on the inner side of the tetrahedron. Substrate access is highly restricted via four 10-Å wide channels located at the center of the triangular faces of the tetrahedron. Thus, only small peptides and completely unfolded proteins can serve as substrates.

Contrary to the rather broad specificity of TET, FrvX has a remarkable preference for leucine, whereas lysine is cleaved at a much lower rate. On the other hand, even the less preferred amino acids like lysine or arginine are still processed at rates comparable with those of TET.

During our search for a suitable dodecameric member of the M42 family, we also examined homologs from *T. tengcongensis* and *L. lactis*. Contrary to FrvX, these enzymes migrate as monomers and dimers, respectively, on gel filtration columns. The crystal lattice of the PDB entry 1VHO, the *T. maritima* homolog, reveals a monomer. Obviously, monomers and dimers are not self-compartmentalizing quaternary structures, and special precautions for the regulation of their activity must be provided by the cell. Neither the physiological role nor the localization of the *B. subtilis* or *T. maritima* gene products is known presently, so the explanation for this finding has to await further studies. It seems likely that the dodecameric enzymes are located within the cytoplasm, as has been established for TET.

Acknowledgments—We thank Dr. Andrea Schmidt from the European Molecular Biology Laboratory outstation at Deutsches Elektron Synchrotron in Hamburg, Dr. Clemens Schulze-Briese and Dr. Takashi Tomizaki from beamline X06SA, Swiss Light Source, Paul Scherrer Institut Villigen, and Dr. Martin Walsh and Dr. Gavin Fox from the beam line BM14, European Radiation Synchrotron Facility, Grenoble, for help with data collection.

REFERENCES

- Wenzel, T., Eckerskorn, C., Lottspeich, F., and Baumeister, W. (1994) *FEBS Lett.* **349**, 205–209
- Lowe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W., and Huber, R. (1995) *Science* **268**, 533–539
- Groll, M., Ditzel, L., Lowe, J., Stock, D., Bochtler, M., Bartunik, H. D., and Huber, R. (1997) *Nature* **386**, 463–471
- Brandstetter, H., Kim, J. S., Groll, M., and Huber, R. (2001) *Nature* **414**, 466–470
- Walz, J., Tamura, T., Tamura, N., Grimm, R., Baumeister, W., and Koster, A. J. (1997) *Mol. Cell* **1**, 59–65
- Tamura, T., Tamura, N., Cejka, Z., Hegerl, R., Lottspeich, F., and Baumeister, W. (1996) *Science* **274**, 1385–1389
- Geier, E., Pfeifer, G., Wilm, M., Lucchiari-Hartz, M., Baumeister, W., Eichmann, K., and Niedermann, G. (1999) *Science* **283**, 978–981
- Remaut, H., Bompard-Gilles, C., Goffin, C., Frere, J. M., and Van Beeumen, J. (2001) *Nat. Struct. Biol.* **8**, 674–678
- Joshua-Tor, L., Xu, H. E., Johnston, S. A., and Rees, D. C. (1995) *Science* **269**, 945–950
- Franzetti, B., Schoehn, G., Hernandez, J. F., Jaquinod, M., Ruigrok, R. W., and Zaccai, G. (2002) *EMBO J.* **21**, 2132–2138
- Rawlings, N. D., and Barrett, A. J. (1999) *Nucleic Acids Res.* **27**, 325–331
- Maher, M. J., Ghosh, M., Grunden, A. M., Menon, A. L., Adams, M. W., Freeman, H. C., and Guss, J. M. (2004) *Biochemistry* **43**, 2771–2783
- Burley, S. K., David, P. R., Taylor, A., and Lipscomb, W. N. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 6878–6882
- Roderick, S. L., and Matthews, B. W. (1993) *Biochemistry* **32**, 3907–3912
- Chevrier, B., Schalk, C., D'Orchymont, H., Rondeau, J. M., Moras, D., and Tarnus, C. (1994) *Structure* **2**, 283–291
- Lowther, W. T., and Matthews, B. W. (2002) *Chem. Rev.* **102**, 4581–4608
- Mullis, K., Faloua, F., Scharf, S., Saiki, R., Horn, G., and Erlich, H. (1986) *Cold Spring Harbor Symp. Quant. Biol.* **51**, Pt. 1, 263–273
- Laemmli, U. K. (1970) *Nature* **227**, 680–685
- Kabsch, W. (1988) *J. Appl. Crystallogr.* **21**, 916–924
- Kabsch, W. (1988) *J. Appl. Crystallogr.* **21**, 67–71
- Kabsch, W. (1993) *J. Appl. Crystallogr.* **26**, 795–800
- Vagin, A., and Teplakov, A. (1997) *J. Appl. Crystallogr.* **30**, 1022–1025
- 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) *Acta Crystallogr. Sect. D* **54**, Pt. 5, 905–921
- Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6**, 458–463
- Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard. (1991) *Acta Crystallogr. Sect. A* **47**, Pt. 2, 110–119
- Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Cryst. Sect. D* **53**, 240–255
- Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26**, 283–291
- Kleywegt, G. J., and Jones, T. A. (1996) *Structure* **4**, 1395–1400
- Chevrier, B., D'Orchymont, H., Schalk, C., Tarnus, C., and Moras, D. (1996) *Eur. J. Biochem.* **237**, 393–398
- Sander, C., and Schneider, R. (1991) *Proteins* **9**, 56–68
- Stamper, C., Bennett, B., Edwards, T., Holz, R. C., Ringe, D., and Petsko, G. (2001) *Biochemistry* **40**, 7035–7046
- Bennett, B., and Holz, R. C. (1997) *Biochemistry* **36**, 9837–9846
- De Paola, C. C., Bennett, B., Holz, R. C., Ringe, D., and Petsko, G. A. (1999) *Biochemistry* **38**, 9048–9053
- Bennett, B., and Holz, R. C. (1997) *J. Am. Chem. Soc.* **119**, 1923–1933
- Prescott, J. M., Wagner, F. W., Holmquist, B., and Vallee, B. L. (1983) *Biochem. Biophys. Res. Commun.* **114**, 646–652
- Gonzales, T., and Robert-Baudouy, J. (1996) *FEMS Microbiol. Rev.* **18**, 319–344
- Taylor, A. (1993) *Trends Biochem. Sci.* **18**, 167–171
- Taylor, A. (1993) *FASEB J.* **7**, 290–298