Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The β-propeller domain of the trilobed protease from *Pyrococcus furiosus* reveals an open Velcro topology

Jürgen Bosch,[a]‡ Tomohiro Tamura,[a]§ Noriko Tamura,[a]§ Wolfgang Baumeister[a] and Lars-Oliver Essen[b]*

[a]Max-Planck-Institute for Biochemistry, Department of Molecular Structural Biology, Am Klopferspitz 18a, D-82152 Martinsried, Germany, and [b]Department of Chemistry, Philipps-University, Hans-Meerwein-Strasse, D-35032 Marburg, Germany

‡ Present address: Howard Hughes Medical Institute and University of Washington, Department of Biochemistry, K-426, 1705 NE Pacific Street, Seattle, WA 98195, USA.
§ Present address: Research Institute of Genome-based Biofactory, National Institute of Advanced Industrial Science and Technology (AIST), 2-17-2-1 Tsukisamu-Higashi, Toyohira-ku, Sapporo 062-8517, Japan.

Correspondence e-mail:
essen@chemie.uni-marburg.de

Received 25 July 2006
Accepted 31 October 2006

PDB Reference: trilobed protease β-propeller domain, 2gop, r2gopsf.

In the proteolytic pathway of prokaryotic and eukaryotic organisms, proteins tagged for proteolysis are firstly shredded into smaller peptides by compartmentalized proteases such as the proteasome complex. Accordingly, a variety of downstream proteases have evolved to further hydrolyze these peptides to the level of free amino acids. In the search for such downstream proteases, a high-molecular-weight protease complex called trilobed protease (TLP) was recently discovered in the archaeon *Pyroccocus furiosus*. The crystal structure of the N-terminal β-propeller domain of the trilobed protease at 2 Å resolution shows that the trilobed protease utilizes this accessory domain to control substrate access to the active site. Modelling of the intact TLP monomer suggests that this protease has an additional side entrance to its active site as in the DPP-IV or tricorn protease complexes.

1. Introduction

The β-propeller motif is a very common architecture based on four-stranded, antiparallel and twisted β-sheets, which are radially arranged around a central tunnel (Fülöp & Jones, 1999). This fold, which was first recognized by Murzin (1992) in the viral enzyme neuraminidase (PDB code 1nn2; Varghese *et al.*, 1983), can adopt a wide range of different biological functions. Accordingly, many proteins containing β-propeller motifs have been found to play roles in diseases such as cancer, Alzheimer's disease, Huntington's disease or arthritis. Five classes of β-propellers are known to date, exhibiting fourfold, fivefold, sixfold, sevenfold or eightfold pseudo-symmetry, respectively. The twisted β-sheets pack face to face *via* hydrophobic interactions between the sheets. Only in smaller β-propellers with fourfold pseudo-symmetry additional stabilization by disulfide bonds is mandatory. In other β-propellers such disulfide bridges are not obligatory, as there is a sufficient amount of hydrophobic interaction between neighbouring blades contributing to protein stabilization. Another factor that contributes to stability is the mode of ring closure. The last blade is mostly formed by a combination of N- and C-terminal β-sheets, resulting in a closed rigid ring-like structure (Velcro type). Only rarely is the ring closure exclusively based on hydrophobic interactions between neighbouring blades (non-Velcro type). This type of ring closure allows more flexibility in the ring or even a slight widening of the tunnel. Most known β-propeller domains use their central tunnel or the entrance to this tunnel to coordinate a ligand, to

doi:10.1107/S0907444906045471

interact with other molecules or domains or to carry out catalytic functions (Pons *et al.*, 2003).

The unique capacity of β-propeller domains to form central channels is utilized by several proteases to generate specific access and exit pathways for substrate peptides and released products. In topologically simpler cases such as the prolyl oligopeptidase POP (PDB code 1qfm; Fülöp *et al.*, 1998) or the dipeptidyl protease DPP-IV (PDB code 1n1m; Rasmussen *et al.*, 2003), a single seven- or eight-bladed β-propeller domain apparently suffices to exert control over substrate access to the active site. In the highly compartmentalized tricorn protease complex from *Thermoplasma acidophilum* (PDB code 1k32; Brandstetter *et al.*, 2001), two β-propeller domains were supposed to separately control the entrance of substrate peptide to the active-site compartment and the release of proteolytic products.

Recently, a new protease complex containing β-domains, the trilobed protease, was discovered in *Pyrococcus furiosus* when cell lysates of this archaeon were screened for high-molecular-weight complexes which exhibit proteolytic activity against small fluorogenic peptides (Tamura *et al.*, manuscript in preparation). According to size-exclusion chromatography, this protease apparently assembles into a trimer of approximately 220 kDa molecular weight and was named according to its trilobal shape in electron-microscopic images. The activity profile of TLP indicated that this protease complex acts in *P. furiosus* downstream of the proteasome-based protein-degradation machinery. Like its functional homologue from *T. acidophilum*, the hexameric tricorn protease complex, the TLP complex apparently only degrades small proteasome-derived peptides. The reaction products are then further decomposed into single amino acids by other peptidases, so that TLP apparently plays a similar role for the recycling of amino acids in *P. furiosus* as the tricorn protease complex in *T. acidophilum* (Brandstetter *et al.*, 2001; Tamura *et al.*, 1996).

According to the predicted domain arrangement, the archaeal TLP protease differs from the tricorn complex and more closely resembles the monomeric POP protease (Fülöp *et al.*, 1998) or the dimeric DPP-IV complex (Rasmussen *et al.*, 2003). Like these eukaryotic proteases, TLP contains a C-terminal α/β-hydrolase domain (356–620) in addition to an N-terminal β-propeller domain (1–347). In POP, the corresponding β-propeller domain was shown to act as a size-exclusion and recognition system for cognate substrate peptides. Furthermore, it protects the cytosolic proteins from being randomly degraded in the active site, which is located in the α/β-hydrolase domain. Therefore, the β-propeller seems to be suitable for nonspecific cellular proteases to restrict access to their active sites and to allow only certain peptides to enter the pore of the β-propeller domain.

In this paper, we describe the X-ray structure of the N-terminal β-propeller domain of the trilobed protease (TLP-β), which is comprised of residues 1–347 of intact TLP. The structure reveals a seven-bladed β-propeller domain, where access to the central tunnel is controlled by a cluster of charged residues. In contrast to most β-propeller domains, but similar to the other compartmentalized proteases with β-propeller domains, the TLP β-propeller belongs to the non-Velcro ring-closure type.

2. Material and methods

2.1. Purification and crystallization of TLP-β

The TLP-β domain containing a C-terminal His$_6$-tag was overexpressed in BL21(DE3)-RIL cells (Stratagene). Cells were cultured at 310 K and 250 rev min^{-1} in 2 l baffled flasks filled with 500 ml TB medium and supplemented with 100 µg ml^{-1} ampicillin, 34 µg ml^{-1} chloramphenicol and 10 µl antifoam Se-15 (Sigma) for improved oxygenation. The cells were induced with 1 m*M* IPTG at an OD$_{600}$ of 2. After 8 h induction, cells were harvested by centrifugation at a final OD$_{600}$ of ~12. The cell pellet was washed once in buffer *A* (300 m*M* NaCl, 50 m*M* NaH$_2$PO$_4$, 4 m*M* MgSO$_4$ pH 6.0) and resuspended in two volumes of buffer *A*. The homogenized cell suspension was lysed by two passes using a cell disrupter (Avestin) at 277 K with a maximum pressure of 200 MPa. After the first pass, 15 U Benzonase (Merck) was added to the lysate. The lysate was cleared of cell debris by centrifugation for 1 h at 75 600*g* and 277 K (Beckman, Avanti, J-25) and incubated with Ni–NTA resin (Qiagen) for 1 h with slow stirring at 277 K. The Ni–NTA resin was transferred into an XK16 column (Amersham Pharmacia) and washed with buffer *A* until a stable OD$_{280}$ baseline was reached. The resin was then washed with 20 column volumes of buffer *B* [300 m*M* NaCl, 50 m*M* NaH$_2$PO$_4$, 10%(*v*/*v*) glycerol pH 8.0]. Recombinant TLP-β was eluted with a linear gradient of 0–300 m*M* imidazole and collected in 1 ml fractions. Fractions containing the β-propeller domain were pooled after SDS–PAGE analysis and desalted with a HiPrep 26/10 desalting column (Amersham Pharmacia) in 50 m*M* Tris–HCl pH 8.0. The sample was loaded onto a Resource Q column (Amersham Pharmacia) and eluted with a linear gradient of 0–2 *M* NaCl. Fractions were inspected by SDS–PAGE analysis, pooled and concentrated to a final volume of 1.5 ml, not exceeding a protein concentration of 40 mg ml^{-1}. A final size-exclusion chromatography step using a Superdex 75 column (Amersham Pharmacia) led to a homogeneous TLP-β preparation, which was concentrated to a final concentration of 30 mg ml^{-1}. Aliquots were flash-frozen with liquid nitrogen and stored at 77 K.

Sparse-matrix crystallization screens (Hampton Research) were set up as hanging-drop experiments in VDX plates at 277, 289 and 293 K and with an initial protein concentration of 10 mg ml^{-1}. 2 µl protein solution was mixed with 2 µl reservoir solution. The first needles were observed after 24 h in condition No. 40 of Crystal Screen I [100 m*M* sodium citrate pH 6.5, 20%(*v*/*v*) 2-propanol, 20%(*w*/*v*) PEG 4000]. After 72 h, well shaped crystals were observed in condition No. 37 of Crystal Screen 1 [100 m*M* sodium acetate pH 4.6, 8%(*w*/*v*) PEG 4000]. Further screening led to optimal crystal growth at 293 K using 100 m*M* sodium acetate pH 4.4, 20 m*M* MgCl$_2$, 300 m*M* 1,6-hexanediol and 10%(*w*/*v*) PEG 4000. TLP-β crystals grew

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

	Native	$IrCl_3$	K_2PtBr_4	K_2PtCl_6	K_2ReCl_6	$KReO_4$	$UO_2(OAc)_2$
Unit-cell parameters (Å)							
a	54.7	54.9	54.8	54.9	54.9	55.1	55.3
b	88.5	88.1	88.5	88.2	87.6	88.7	88.4
c	153.9	154.0	153.9	153.9	154.1	153.9	154.0
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Wavelength (Å)	1.000	1.105	1.071	1.071	1.176	1.176	0.933
Resolution (Å)	2.0	2.9	2.5	2.5	2.9	2.9	2.0
No. of unique reflections	39936	17268	26594	26350	16780	18386	48797
Mosaicity (°)	0.49	0.55	0.49	0.45	0.88	0.76	0.51
$I/\sigma(I)$	44.7 (4.5)	21.5 (6.8)	18.6 (2.8)	27.6 (4.6)	18.9 (4.8)	15.9 (7.0)	6.4 (2.0)
$R_{\rm merge}$ (%)	2.9 (35.4)	3.3 (13.1)	4.4 (31.8)	2.4 (16.2)	2.9 (12.9)	4.3 (10.7)	8.2 (34.9)
Completeness (%)	88.1 (95.6)	99.8 (99.7)	99.5 (98.6)	92.9 (98.3)	94.3 (85.5)	97.2 (75.6)	83.5 (64.3)
No. of heavy-atom sites	—	6	9	6	5	—	10
Phasing power	—	0.29	0.67	0.78	0.83	—	1.10

from precipitate and reached overall dimensions of approximately 0.4 × 0.4 × 0.25 mm.

2.2. Data collection

TLP-β crystals were sequentially soaked in crystallization solutions containing increasing amounts of glycerol in 2%(*w*/*v*) steps and 1,6-hexanediol in 100 m*M* steps. Each crystal was incubated for at least 30 min before being transferred to a condition with higher cryoprotectant concentration. Nevertheless, ice formation upon flash-freezing could be reduced but not completely avoided, as TLP-β crystals suffered a severe loss of diffraction at glycerol conditions exceeding 10%(*w*/*v*). A compromise between ice formation and obtainable resolution was a cryocondition with 100 m*M* sodium acetate pH 4.4, 20 m*M* $MgCl_2$, 600 m*M* 1,6-hexanediol, 10%(*w*/*v*) PEG 4000, 8% glycerol, in which all data sets were collected. Several native and heavy-metal derivative data sets were collected at beamline BW6, DESY Hamburg on a MAR CCD 130 mm detector. One data set was collected at ID14-1, ESRF Grenoble. All data were processed and reduced with *MOSFLM* v.6.1 (Leslie, 2006) and *SCALA* from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). Data in the 2.31–2.35 and 2.17–2.21 Å resolution shells had to be excised from the native data set owing to the overlap of spots with intense ice rings in these regions. Accordingly, the native data set has a low overall completeness of 88.1% (95.6% in the outer shell) despite its fourfold redundancy.

2.3. Structure determination and refinement

Molecular replacement using the coordinates of the homologous β-propeller domain of POP was unsuccessful in our hands, as the sequence identity to TLP-β was only 10.7%. Accordingly, the structure was solved *via* multiple isomorphous replacement (MIR) using heavy-metal soaks. The space group was determined to be $P2_12_12_1$, with unit-cell parameters a = 54.75, b = 88.50, c = 153.97 Å. Heavy-metal sites were automatically identified by *SOLVE* (Terwilliger & Berendzen, 1999) at a resolution of 2.5 Å. Phases were extended to 2 Å using 30 cycles of *RESOLVE* (Terwilliger, 1999). An initial molecular model that was about 50% complete was built using the program *MAID* (Levitt, 2001). The NCS matrix that related the two partially defined TLP-β molecules in the asymmetric unit was delineated using the program *LSQKAB* (Collaborative Computational Project, Number 4, 1994). An initial mask based on the C^α backbone was used in 100 cycles of density-modification averaging with phase extension from 3 to 2 Å, resulting in a considerably improved electron-density map. As the tracing proceeded and the polyalanine model was extended, new masks were iteratively calculated and used in phase improvement. Manual model rebuilding was performed in *O* (Jones *et al.*, 1991).

The model was initially refined by simulated annealing using loosely defined NCS restraints and data to 2.5 Å resolution. A test set of 3% of the unique reflections was used to calculate a free R factor (Brünger, 1992). During automatic refinement and manual rebuilding, the resolution was increased stepwise to 2.0 Å. The structure was refined with *CNS* v.1.1 (Brünger *et al.*, 1998) to an $R_{\rm work}$ of 25.9% and an $R_{\rm free}$ of 31.5%. A significant decrease in the R factor was obtained by 20 cycles of TLS refinement (Winn *et al.*, 2001) using the program *REFMAC* v.5.1 (Murshudov *et al.*, 1997), where each blade was assigned to a single TLS group. After 30 cycles of restrained refinement, a final $R_{\rm work}$ and $R_{\rm free}$ of 20.9% and 25.8% were obtained, respectively.

3. Results and discussion

3.1. Structure determination

The isolated TLP-β domain could be efficiently expressed and purified with yields of about 200 mg per litre of culture. Large orthorhombic crystals grew from precipitate within a few days of incubation. The three-dimensional structure of TLP-β was subsequently solved at 2.0 Å resolution by multiple isomorphous replacement (MIR). A heavy-atom search identified five useful heavy-atom derivatives (Table 1). Five heavy-atom sites were initially found by *SOLVE* (Terwilliger & Berendzen, 1999). Using the Pt^{2+} data set and a second round of *SOLVE*, a total of 36 sites were identified in the derivative data sets (Table 1). After heavy-atom parameter and phase refinement, an overall figure of merit of 0.54 was obtained for the MIR phases (Table 2), which was improved to 0.72 by phase extension to 2.0 Å using the program *RESOLVE* (Terwilliger, 1999). The resulting electron-density map revealed two TLP-β molecules in the asymmetric unit. At this stage, automatic chain tracing with the program *MAID* (Levitt, 2001) yielded an unambiguous trace for only 181 of the 356 residues per monomer. Noncrystallographic symmetry averaging using *DM* (Cowtan & Main, 1993) further improved the phases so that 260 residues could be assigned for molecule *A* and 220 for molecule *B*. The model was completed by alternating cycles of refinement with *REFMAC* v.5.1 (Murshudov *et al.*, 1997) and manual rebuilding using the program *O* (Jones *et al.*, 1991). Finally, a total of 308 residues in molecule *A* and 301 residues in molecule *B* of the 356 residues per monomer were defined by electron density.

Table 2
Statistics of structure refinement and quality of the model.

Resolution range (Å)	18.8–2.00 (2.08–2.00)
No. of reflections (unique)	39936
No. of reflections (test set)	1235
$R_{\rm work}$† (%)	20.9
$R_{\rm free}$‡ (%)	25.8
No. of non-H atoms in protein model	5285
No. of water molecules	251
R.m.s. deviation, bond lengths (Å)	0.008
R.m.s. deviation, bond angles (°)	1.13

† $R_{\rm work} = \sum_{hkl} \big||F_{\rm obs,work}| - k|F_{\rm calc}|\big| / \sum_{hkl} |F_{\rm obs,work}| \times 100$. ‡ $R_{\rm free} = \sum_{hkl} \big||F_{\rm obs,test}| - k|F_{\rm calc}|\big| / \sum_{hkl} |F_{\rm obs,test}| \times 100$.

The final crystallographic ($R_{\rm work}$) and free ($R_{\rm free}$) *R*-factor values were 20.9% and 25.8%, respectively, for data between 18.8 and 2.0 Å resolution. The model contains two TLP-β monomers and 251 water molecules. Of the 356 residues making up a complete monomer, residues 1–10, 125–150 and 348–356 were missing in both molecules owing to a lack of electron density. Likewise, residues 189–198 in molecule *A* and residues 34–40 and 341–347 in molecule *B* were not defined. Data statistics and the quality of the model are summarized in Tables 1 and 2. A Ramachandran plot (Ramakrishnan & Ramachandran, 1965) shows that 86.9% and 10.9% of the

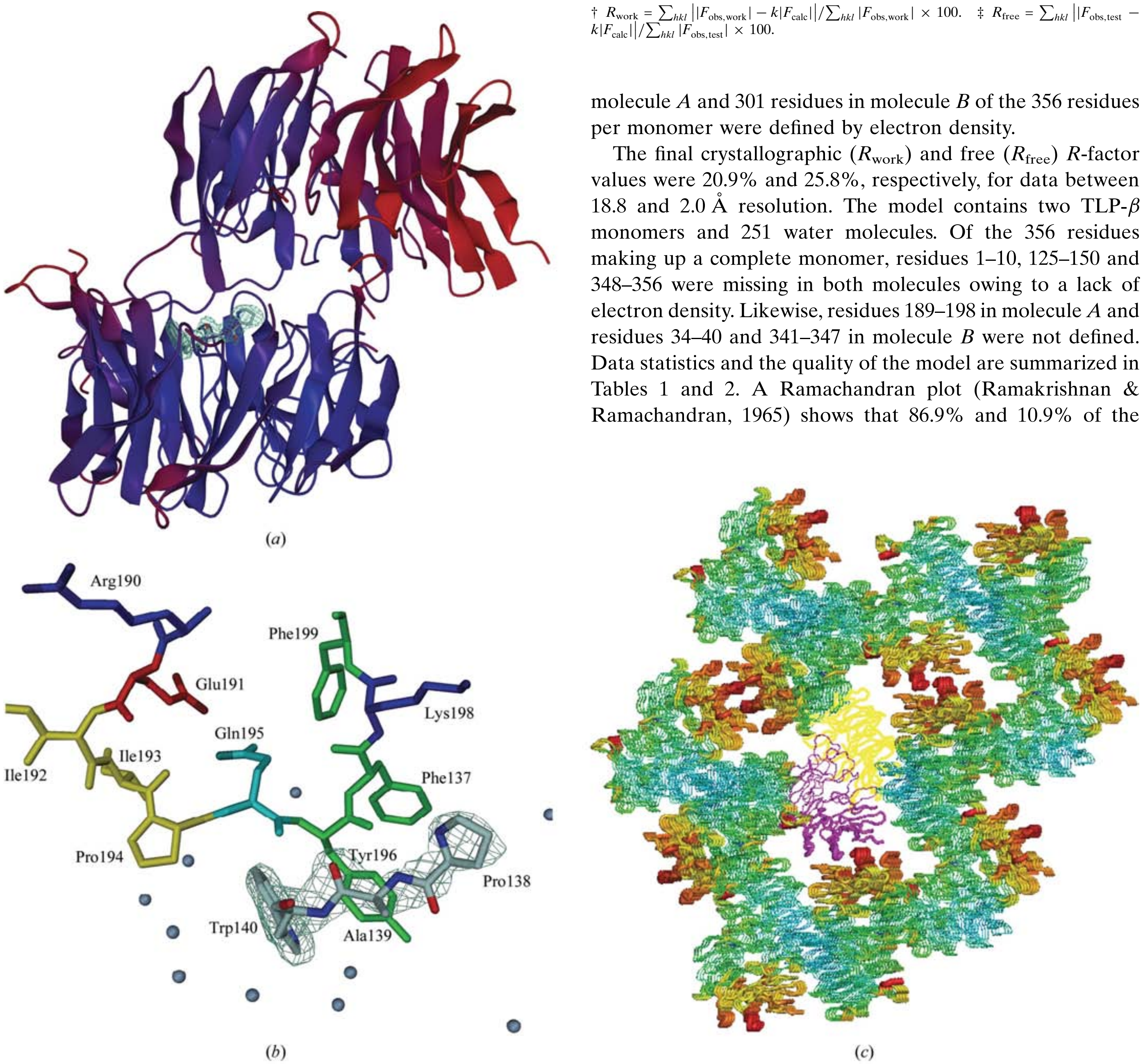
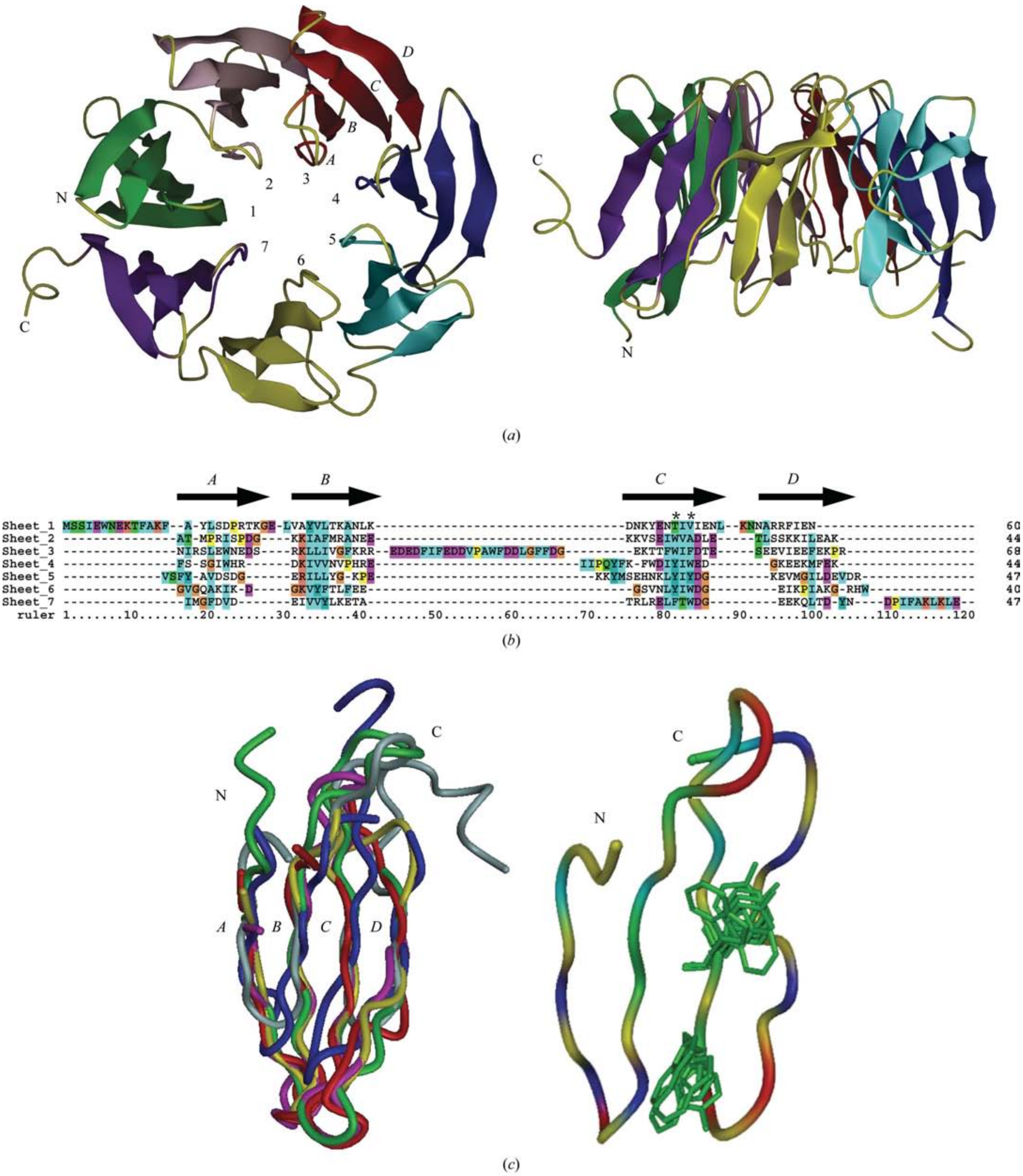


Figure 1
Crystal packing and disorder in TLP-β crystals. (*a*) Ribbon diagram of the TLP-β dimer present in the asymmetric unit; colours represent *B* values, ranging from 15 Å^2 (blue) to 80 Å^2 (red). Between the two TLP-β molecules, the short segment PAW is depicted as a stick model. This three-amino-acid segment is part of the missing loop containing residues 125–150, namely residues Pro138–Trp140. The $\sigma_{\rm A}$-weighted $2F_{\rm obs} - F_{\rm calc}$ electron-density map (cyan) is contoured at 1.5σ. (*b*) Interactions between the PAW tripeptide and the loop region of molecule *B*. The electron-density map is contoured at 1.5σ; the side chains of the loop region are highlighted according to overall residue properties (red, acidic; blue, basic; green, aromatic; cyan, polar; yellow, aliphatic). Water molecules in close proximity are coloured light blue. (*c*) Cross section through the crystal lattice (colours represent *B* values); only the C$^\alpha$ traces are shown. Two subunits representing molecule *A* (yellow) and molecule *B* (magenta) are highlighted within the lattice.



```
                                                                                        * *
Sheet_1 MSSIEWNEKTFAKF--A-YLSDPRTKGE-LVAYVLTKANLK---------------------------------DNKYENTIVIENL--KNNARRFIEN---------------------   60
Sheet_2 ---------------AT-MPRISPDG----KKIAFMRANEE---------------------------------KKVSEIWVADLE-----TLSSKKILEAK------------------   44
Sheet_3 ---------------NIRSLEWNEDS----RKLLIVGFKRR--EDEDFIFEDDVPAWFDDLGFFDG---------EKTTFWIFDTE-----SEEVIEEFEKPR-----------------   68
Sheet_4 ---------------FS-SGIWHR------DKIVVNVPHRE---------------------------IIPQYFK-FWDIYIWED--------GKEEKMFEK------------------   44
Sheet_5 -------------VSFY-AVDSDG------ERILLYG-KPE-----------------------------KKYMSEHNKLYIYDG--------KEVMGILDEVDR---------------   47
Sheet_6 ---------------GVGQAKIK-D-----GKVYFTLFEE-----------------------------------GSVNLYIWDG---------EIKPIAKG-RHW--------------   40
Sheet_7 ----------------IMGFDVD-------EIVVYLKETA----------------------------------TRLRELFTWDG---------EEKQLTD-YN----DPIFAKLKLE--   47
  ruler 1.......10........20........30........40........50........60........70........80........90.......100.......110.......120
```

Figure 2
(*a*) Ribbon diagram of a TLP-β monomer as viewed along the sevenfold pseudo-symmetry axis and in a side view. Blades are coloured differently: starting clockwise from the N-terminus, blade 1 (green, 11–59), blade 2 (grey, 63–103), blade 3 (red, 107–170), blade 4 (blue, 174–216), blade 5 (cyan, 219–256), blade 6 (yellow, 266–297) and blade 7 (violet, 304–347). (*b*) Structure-based sequence alignment for each blade of TLP-β. Arrows above the alignment represent β-strands. Residues marked with an asterisk are highlighted as side chains in (*c*). (*c*) Superposition of the seven β-sheets of the three-dimensional structure of TLP-β on each other. The angle between strands *A* and *D* in each blade is close to 90° if viewed from the side. Blade 6 is coloured by residue properties as in Fig. 1(*b*). Two highly conserved positions occurring in strand *C* of each blade are depicted with their side chains.

residues in molecule *A*, and 84.4% and 14.1% in molecule *B* occupy the most favoured and additionally allowed regions as determined by the program *PROCHECK* (Laskowski *et al.*, 1993).

3.2. Subunit interactions and crystal contacts

In the asymmetric unit of the TLP-β crystals the two monomers are packed together in a face-to-face manner, burying a large hydrophobic surface of 1538 Å^2 from solvent access (Fig. 1*a*). As seen in the cross section through the crystal lattice of the orthorhombic TLP-β crystals (Fig. 1*c*), the two molecules are not equivalent because of different degrees of conformational disorder. Each molecule *A* has three interacting regions with neighbouring molecules, whereas molecules *B* make only two crystal contacts. The looser packing of molecule *B* is reflected by the distribution of the *B* values in the two molecules (Fig. 1*a*), which shows a progressive increase of thermal disorder in molecule *B* and rationalizes the significant improvement by final TLS refinement. Accordingly, molecule *B* exhibits a significantly larger overall *B* factor (54 Å^2) than molecule *A* (34 Å^2).

Within the enclosure of the TLP-β dimer, a segment of three residues with the amino-acid sequence PAW was trapped, mainly by hydrophobic interactions with the loop region Pro194–Phe197 of molecule *B*. The residues of the segment were identified as Pro138-Ala139-Trp140 of the otherwise unassigned region 125–150 by the shape of electron density as shown in Fig. 1(*b*). However, owing to its location on the side of the β-propeller domain close to the active site (see below), the observed interaction is unlikely to mimic a subunit interaction of an intact TLP oligomer.

3.3. Overall structure of the TLP-β monomer

The TLP-β fold is a four-stranded antiparallel β-sheet which is repeated seven times. All β-sheets are aligned around a sevenfold axis of pseudo-symmetry that runs through the

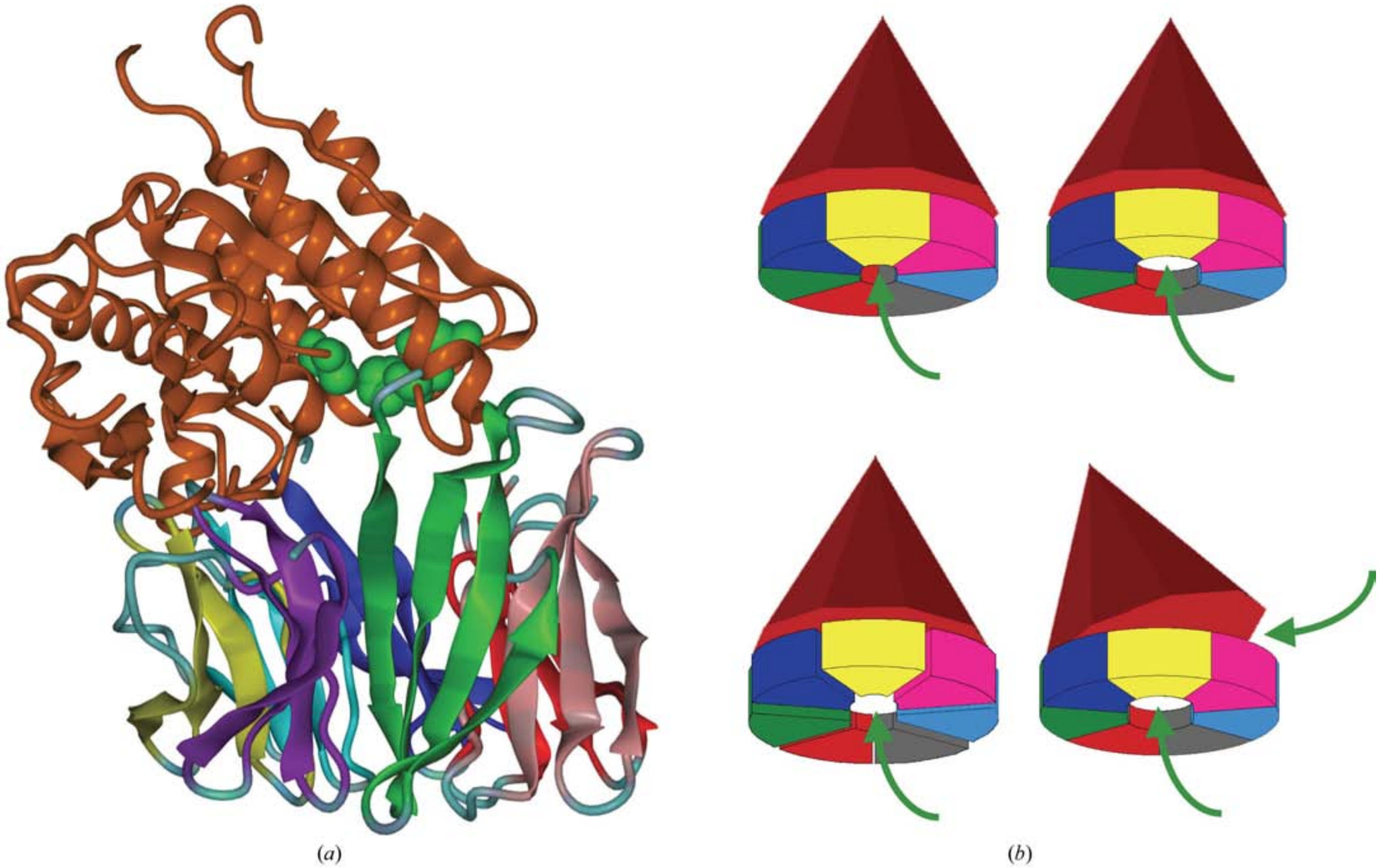


Figure 3
(*a*) Three-dimensional representation of a hypothetical model of the intact trilobed protease. Structure prediction of the α/β-hydrolase domain was performed using the *LIBELLULA* server (Juan *et al.*, 2003) and is shown in brown. The β-propeller domain is depicted as in the previous figures. Both domains were fitted using the program *TOP* (Lu, 2000) according to the overall structure of prolyl oligopeptidase (Fülöp *et al.*, 1998). Residues of the active site are shown as green spheres. (*b*) Schematic overview of possible substrate pathways in the monomeric trilobed enzyme. Green arrows represent substrate access to the active site; the catalytic domain is represented as a cone-shaped domain and the β-propeller domain is colour coded by individual blades. The first model represents a diffuse only model, where the pore cannot be widened, and the second model allows a limited widening of the central pore; both cases are applicable to the Velcro-type closures of β-propeller domains. The third and the fourth model result in larger movements of the β-propeller domain and are limited to non-Velcro-type β-propellers, allowing a larger breathing mechanism for the uptake of larger substrates. The last model is a combination of the breathing mechanism and an additional domain movement, allowing much larger substrates to access the catalytic site *via* the side entrance.

centre of the molecule, forming the circular arrangement of a β-propeller-like topology (Fig. 2*a*). The structure of the TLP-β monomer consists of one roughly doughnut-shaped domain with a diameter of 60 Å and a height of 50 Å.

Following the common terminology of β-propellers (Faber *et al.*, 1995; Ghosh *et al.*, 1995; Renault *et al.*, 1998; Wall *et al.*, 1995; Xia *et al.*, 1992), the β-strands are labelled according to the β-sheet in which they occur (1–7). Within one four-

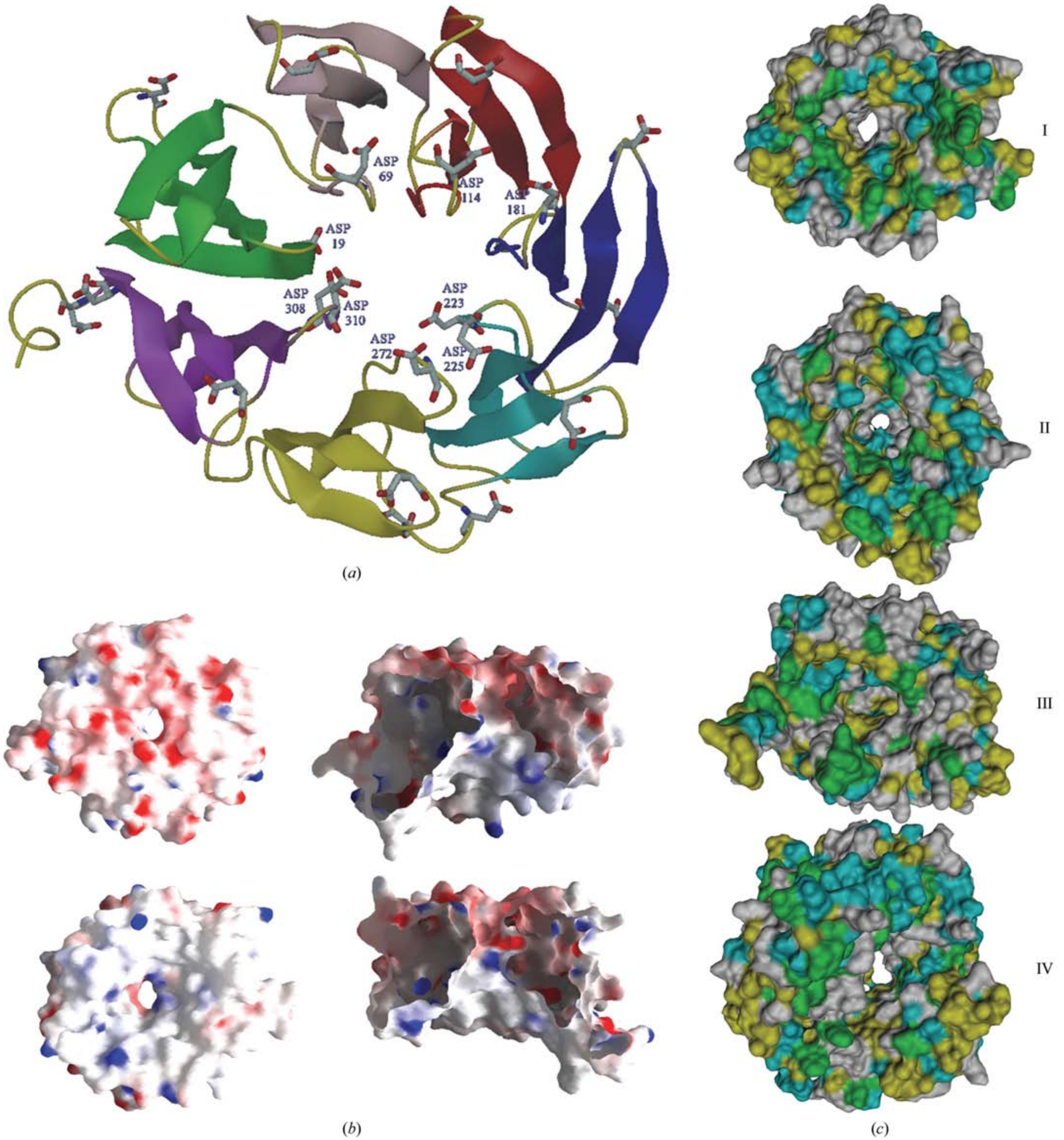


Figure 4
Comparison of TLP-β with other β-propeller domains from compartmentalized proteases. (*a*) Ribbon diagram of TLP-β as viewed along the sevenfold pseudo-symmetry axis. The nine aspartate residues forming the negative ring at the entrance to the channel are highlighted as a ball-and-stick model. (*b*) Electrostatic surface representation of TLP-β in top, bottom and cut-open side views. Negative surface charges are coloured red and positive charges blue. (*c*) Surface representation (colour codes: aromatic, green; aliphatic, yellow; polar, cyan; charged, grey) showing the interior side of the channel of the superimposed structures of TLP-β (I), POP (II), TRI-β7 (III) and DPP-IV (IV) from the bottom view.

Table 3
Structural alignment of individual TLP-β blades to blade 5 of TLP-β.

A low sequence identity results in the same overall topology with low r.m.s.d. values between the compared blades.

Blade No.	Residues	Matching C^α positions	Sequence identity (%)	R.m.s.d. (Å)
1	50	18	16.7	0.96
2	44	24	16.7	0.83
3	44	21	19.0	1.52
4	40	19	15.8	0.87
6	40	25	28.0	0.99
7	44	25	32.0	1.05

stranded antiparallel β-sheet, the strands are labelled from the inside to the outside of the molecule (*A*–*D*). The structure strictly follows the 'W' structural topology, describing the down–up–down–up of the β-sheets from the inside to the outside (Fig. 2*a*). In contrast to other β-propeller structures, the TLP-β structure does not follow the Velcro closure mechanism (Neer & Smith, 1996). Like the β-propeller domains of the prolyl oligopeptidase POP (Fülöp *et al.*, 1998), tricorn protease (Brandstetter *et al.*, 2001) and dipeptidyl peptidase IV (Rasmussen *et al.*, 2003), the TLP-β domain belongs to the group of β-propellers that stabilize their rings by hydrophobic interactions between the N-terminal and C-terminal blades. This open Velcro topology allows increased conformational freedom of the β-propeller domain (Harris *et al.*, 2001), where a breathing mechanism was postulated for the recognition and uptake of substrates by the POP protease.

The TLP-β domain is highly regular and resembles the POP-β domain more than the slightly distorted TRI-β7 domain of tricorn (Brandstetter *et al.*, 2001) or the eight-bladed β-propeller domain of DPP-IV (Rasmussen *et al.*, 2003). A superposition of the individual TLP-β blades (Fig. 2*c*) shows the high structural conservation within the TLP-β domain. Although the sequence identity is low and varies between 15.8% and 32%, the overall structure of four antiparallel β-sheets is retained, as can be derived from the rather low r.m.s.d. values for the C^α positions (Table 3). Interestingly, all loops on the N-terminal side of the β-propeller domain are well defined, whereas several loops on the C-terminal side are not. Presumably, these missing loops are involved in physical interactions with the catalytic domain or contribute to the stabilization of the TLP oligomer.

3.4. A structural comparison between TLP-β and POP suggests loose confinement of the active-site in TLP

Using the program *LIBELLULA* (Juan *et al.*, 2003), a homology model for the catalytic α/β-hydrolase domain of TLP was generated using the POP structure as a template. The C^α-backbone model for the catalytic domain of TLP and the structure of the TLP-β domain were then superimposed on the corresponding structural elements of POP using the program *TOP* (Lu, 2000). The three-dimensional model of the intact TLP protease (Fig. 3*a*) indicates a large cleft between the catalytic and the β-propeller domain, with a widening of approximately 10 Å. Such a side opening is not present in POP, where access to the active site is only possible *via* the β-propeller domain. However, such side openings were found in the structures of the DPP-IV (Rasmussen *et al.*, 2003) and tricorn protease (Brandstetter *et al.*, 2001) complexes, where they were suggested to act as putative exhaust pathways for the reaction products. Therefore, it might be proposed that one or both of the putative substrate-access pathways are operational in the TLP complex (Fig. 3*b*). In the first model, the substrate would diffuse to the active site only *via* the pore of TLP-β, which cannot be widened. A second model might additionally allow a limited widening of the β-propeller pore, whereas in the third model a breathing mechanism could open and close the pore to allow substrate access. In the last model, the breathing mechanism might be coupled with an additional domain movement opening a second substrate pathway to the active site.

3.5. The central pore of TLP-β domain in comparison with POP, TRI-β7 and DPP IV

The central pore of TLP-β has an average diameter of 5 × 10 Å (Fig. 4*a*). A highly charged ring formed by nine aspartates and six lysines and arginines constricts the pore entrance. The water-filled channel would hence be freely accessible to polar substrate peptides such as the β-propeller domain of POP (Fülöp *et al.*, 1998). Figs. 4(*b*) and 4(*c*) show a surface representation produced using *GRASP* (Nicholls *et al.*, 1991) for TLP-β and the β-propeller domains POP-β, TRI-β7 and DPP-IV-β. The shape of the channel differs significantly between these examples. TLP-β and POP-β have a funnel-like channel with almost identical width, starting with a 5 × 8 Å entrance for TLP and a 5 × 5 Å constriction for POP-β, which in both cases conically expands to a 20 × 25 Å cavity towards the active site. TRI-β7, however, contains a tunnel-like channel with 8 × 15 Å width but with a blocked access to the channel owing to a basic lid closure formed by residues Arg369, Arg414, Arg645 and Lys646. In contrast, DPP-IV (Rasmussen *et al.*, 2003) is an eight-bladed β-propeller domain. The entrance through the central channel is approximately 10 × 12 Å wide and opens up to 20 × 25 Å (Fig. 4*c*). Obviously, access to the active site benefits from pore widening owing to the additional blade. However, displacement and twisting of blade 1 makes the β-propeller fold of DPP-IV (Rasmussen *et al.*, 2003) highly asymmetric and provides an additional side entrance to the active site that is more than 20 Å wide. In all cases the length of the funnel in the β-propeller domains is similar at ~30 Å. Although all four molecules harbour charged or polar residues around the pore-entrance area, a closer examination of the innermost β-strands forming the channel walls reveals differing properties. The first half of the channel walls of TLP-β are covered by charged residues. Only at the channel end, which approaches the active site, does the number of hydrophobic residues increase. The same is true for the TRI-β7 and the POP-β domains. Interestingly, POP-β contains twice as many aromatic residues in its central pore compared with TLP-β and TRI-β7. An even higher number of aromatic residues compared with POP are found in the channel of DPP-IV. These findings might indicate the preference for certain substrates of each of the listed proteases being based not on size but on overall hydrophobicity and hydrophilicity.

We gratefully thank the staff of beamline BW6 of the MPG Outstation at DESY, Hamburg and the staff of beamline ID14-1 at the ESRF, Grenoble for their help and assistance during data collection.

References

Brandstetter, H., Kim, J. S., Groll, M. & Huber, R. (2001). *Nature (London)*, **414**, 466–470.

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Cowtan, K. D. & Main, P. (1993). *Acta Cryst.* D**49**, 148–157.

Faber, H. R., Groom, C. R., Baker, H. M., Morgan, W. T., Smith, A. & Baker, E. N. (1995). *Structure*, **3**, 551–559.

Fülöp, V., Bocskei, Z. & Polgar, L. (1998). *Cell*, **94**, 161–170.

Fülöp, V. & Jones, D. T. (1999). *Curr. Opin. Struct. Biol.* **9**, 715–721.

Ghosh, M., Anthony, C., Harlos, K., Goodwin, M. G. & Blake, C. (1995). *Structure*, **3**, 177–187.

Harris, M. N., Madura, J. D., Ming, L. J. & Harwood, V. J. (2001). *J. Biol. Chem.* **276**, 19310–19317.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Juan, D., Grana, O., Pazos, F., Fariselli, P., Casadio, R. & Valencia, A. (2003). *Proteins*, **50**, 600–608.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Leslie, A. G. W. (2006). *Acta Cryst.* D**62**, 48–57.

Levitt, D. G. (2001). *Acta Cryst.* D**57**, 1013–1019.

Lu, G. (2000). *J. Appl. Cryst.* **33**, 176–183.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Murzin, A. G. (1992). *Proteins*, **14**, 191–201.

Neer, E. J. & Smith, T. F. (1996). *Cell*, **84**, 175–178.

Nicholls, A., Sharp, K. & Honig, B. (1991). *Proteins*, **11**, 281–296.

Pons, T., Gomez, R., Chinea, G. & Valencia, A. (2003). *Curr. Med. Chem.* **10**, 505–524.

Ramakrishnan, C. & Ramachandran, G. N. (1965). *Biophys. J.* **5**, 909–933.

Rasmussen, H. B., Branner, S., Wiberg, F. C. & Wagtmann, N. (2003). *Nature Struct. Biol.* **10**, 19–25.

Renault, L., Nassar, N., Vetter, I., Becker, J., Klebe, C., Roth, M. & Wittinghofer, A. (1998). *Nature (London)*, **392**, 97–101.

Tamura, T., Tamura, N., Cejka, Z., Hegerl, R., Lottspeich, F. & Baumeister, W. (1996). *Science*, **274**, 1385–1389.

Terwilliger, T. C. (1999). *Acta Cryst.* D**55**, 1863–1871.

Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.

Varghese, J. N., Laver, W. G. & Colman, P. M. (1983). *Nature (London)*, **303**, 35–40.

Wall, M. A., Coleman, D. E., Lee, E., Iniguez-Lluhi, J. A., Posner, B. A., Gilman, A. G. & Sprang, S. R. (1995). *Cell*, **83**, 1047–1058.

Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Cryst.* D**57**, 122–133.

Xia, Z. X., Dai, W. W., Xiong, J. P., Hao, Z. P., Davidson, V. L., White, S. & Mathews, F. S. (1992). *J. Biol. Chem.* **267**, 22289–22297.