Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of triosephosphate isomerase (TIM) from *Methanocaldococcus jannaschii*

P. Gayathri,[a] Mousumi Banerjee,[a] A. Vijayalakshmi,[a] Shamina Azeez,[b] Hemalatha Balaram,[b] P. Balaram[a] and M. R. N. Murthy[a]*

[a]Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India, and [b]Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore 560 084, India

Correspondence e-mail: mrn@mbu.iisc.ernet.in

Received 11 September 2006
Accepted 4 November 2006

PDB Reference: triosephosphate isomerase, 2h6r, r2h6rsf.

The crystal structure of a recombinant triosephosphate isomerase (TIM) from the archaeabacterium *Methanocaldococcus jannaschii* has been determined at a resolution of 2.3 Å using X-ray diffraction data from a tetartohedrally twinned crystal. *M. jannaschii* TIM (MjTIM) is tetrameric, as suggested by solution studies and from the crystal structure, as is the case for two other structurally characterized archaeal TIMs. The archaeabacterial TIMs are shorter compared with the dimeric TIMs; the insertions in the dimeric TIMs occur in the vicinity of the tetramer interface, resulting in a hindrance to tetramerization in the dimeric TIMs. The charge distribution on the surface of the archaeal TIMs also facilitates tetramerization. Analysis of the barrel interactions in TIMs suggests that these interactions are unlikely to account for the thermal stability of the archaeal TIMs. A novelty of the unliganded structure of MjTIM is the complete absence of electron density for the loop 6 residues. The disorder of this loop could be ascribed to a missing salt bridge between residues at the N- and C-terminal ends of the loop in MjTIM.

1. Introduction

Extensive studies on triosephosphate isomerase (TIM) have provided many insights into the folding and assembly of the $(\beta/\alpha)_8$-barrel proteins. TIM, a central enzyme in the glycolytic pathway, catalyzes the interconversion of glyceraldehyde 3-phosphate and dihydroxyacetone phosphate and has been structurally characterized from a large number of organisms. Currently, the structures of TIMs from 17 different species are available. Of these, nine are from eukaryotic sources, six are from prokaryotes and two are from archaeabacteria. Although TIMs are usually dimeric enzymes, the archaeabacterial proteins from *Pyrococcus woesei* and *Thermoproteus tenax* are tetramers. However, these two archaeabacterial TIMs differ in their solution properties. While the *P. woesei* TIM (PwTIM) forms a stable tetramer (Walden *et al.*, 2001), the corresponding *T. tenax* (TtTIM) enzyme appears to dissociate into dimers (Walden *et al.*, 2004), suggesting a significantly weaker set of interactions at the tetramer interface. Both the archaeabacterial TIMs are highly thermostable, with melting temperatures above 358 K. Interestingly, TIM from *Thermotoga maritima*, a thermophilic bacterium, has been shown to be a dimer in solution (Beaucamp *et al.*, 1997), but occurs as a disulfide-bonded tetramer in crystals (Maes *et al.*, 1999). Archaeabacterial TIMs have significantly shorter polypeptide chains (215–230 residues) than the TIMs from other sources (240–250 residues).

Previous studies in this laboratory have focused on folding and stability in the TIM-barrel enzymes using the dimeric *Plasmodium falciparum* TIM (PfTIM) as a model (Maithal,

Ravindra, Nagaraj *et al.*, 2002; Maithal, Ravindra, Balaram *et al.*, 2002; Ray *et al.*, 1999; Gopal *et al.*, 1999; Gokhale *et al.*, 1999). In order to probe the relationship between quaternary structure and stability, we have determined the crystal structure of a tetrameric TIM from the thermophilic archaeon *Methanocaldococcus jannaschii* (MjTIM) at a resolution of 2.3 Å. This structure is a valuable addition to the already existing TIM structures, as only two tetrameric TIM structures are available at present. Comparison of the dimeric and tetrameric TIMs provides an opportunity to examine the roles of the intra-subunit and inter-subunit interactions in determining protein stability.

With the availability of a large number of crystal structures of TIM, an analysis of the interactions that hold the barrel in place and contribute to the robustness of the fold gains significance. In the current analysis, the interactions within the core of the barrel and those between the barrel and the helix layers have been compared for all the TIM structures. The analysis has been carried out in detail for a few representative TIMs and this has been extended to the rest of the TIMs on the basis of a structure-based sequence alignment. The results suggest that these interactions are insufficient to account for the higher thermal stability of archaeal TIMs. Therefore, the thermal stability appears to be a consequence of reduced loop lengths and tetramerization.

Another novelty in the structure reported is the complete disorder of the loop 6 region. Loop 6 of TIM has been a topic of extensive study because of its conformational flexibility and the probability of the loop being in either the open or closed forms. In the large number of TIM structures solved to date with various ligands, loop 6 has been observed either in the closed or open state. In the unliganded structure of MjTIM, the disorder of loop 6 draws attention towards differences in the loop 6 region between the archaeal and other TIM sequences.

2. Materials and methods

2.1. Bacterial strains and culture condition

Escherichia coli strain AA200 was obtained from the *E. coli* Genetic Stock Center, Yale University. *E. coli* AA200 (*gar B10*, *fluA22*, *ompF627*, *fadL701*, *relA1*, *pit-10*, *spot1*, *tpi-1*, *phoM510*, *merB1*) is a null mutant of the inherent TPI gene. Cells carrying the recombinant vector were grown at 310 K in Terrific broth containing 100 mg l^{-1} ampicillin for protein purification.

2.2. Chemicals and reagents

All chemicals used in the assay were purchased from Sigma Aldrich Company (St Louis, MO, USA) and media components were from HiMedia Laboratories Ltd (Mumbai, India).

2.3. Cloning, expression and purification of MjTIM

A plasmid carrying the MjTIM gene was obtained from Soumya S. Ray and S. K. Burley, Rockefeller University, USA. The full-length MjTIM DNA was amplified using the end primers MJF (5′-CACCATGGTGATAGTAATTAACTATA-AAAC-3′) and MJR (5′-CAGGATCCTTAGATGAACTT-TATTAATTCTC-3′). This amplicon was ligated to *Nco*I- and *Bam*HI-digested pTrc99A (Ranie *et al.*, 1993). Clones were confirmed by sequencing. Constraints of primer design required the mutation of Val at position 2 to Leu. *E. coli* AA200 cells transformed with pTrc99A-Mj were induced using 300 µ*M* isopropyl *β*-D-thiogalactopyranoside at an $OD_{600\,nm}$ of 0.6 and were harvested by centrifugation at 6000 rev min^{-1} for 15 min. The pellet obtained was resuspended in lysis buffer containing 20 m*M* Tris–HCl pH 8.0, 1 m*M* EDTA, 0.01 m*M* PMSF, 2 m*M* DTT and 10% glycerol before being disrupted using a probe sonicator (Sonifier B-12, Branson Sonic Power) for 20 min. The lysate was centrifuged at 12 000 rev min^{-1} for 45 min to remove cell debris. After precipitation with 0.1% PEP (polyethyleneimine P), the supernatant was heated at 343 K for 20 min and then centrifuged at 12 000 rev min^{-1} for 30 min. The supernatant was fractionated at 75% ammonium sulfate saturation to precipitate MjTIM. The TIM-containing pellet was dissolved in 20 m*M* Tris–HCl pH 8.0, 2 m*M* DTT and 10% glycerol.

2.4. Purification

The protein solution was dialyzed extensively against 20 m*M* Tris–HCl pH 8.0, 2 m*M* DTT and 10% glycerol at 277 K overnight and purified using a preparative Q-Sepharose anion-exchange column connected to an AKTA-Basic (GE Healthcare) FPLC system. The protein was eluted with a linear salt gradient of 0–1 *M* NaCl in the same buffer. Fractions containing the protein were pooled and precipitated with ammonium sulfate to a concentration of 75%. The pellet was dissolved in 20 m*M* Tris–HCl pH 8.0, 2 m*M* DTT and 10% glycerol and subjected to gel filtration on a preparative Sephacryl-200 column equilibrated with the same buffer. The protein purity was checked by 12% SDS–PAGE and the exact molecular weight was verified by ESI–MS (Esquire3000+, Bruker Daltonics).

2.5. Analytical gel filtration

Analytical gel filtration was performed on a Superdex-200 column (300 × 10 mm) attached to an AKTA Basic FPLC system at a flow rate of 0.5 ml min^{-1} and protein elution was monitored at a wavelength of 220 nm owing to the absence of tryptophan residues. The column was calibrated with *β*-amylase (200 kDa), alcohol dehydrogenase (150 kDa), bovine serum albumin (66 kDa), carbonic anhydrase (29 kDa) and cytochrome *c* (12.4 kDa).

2.6. Enzyme assays

Enzyme activity was determined by the conversion of GAP (glyceraldehyde-3-phosphate) to DHAP (dihydroxyacetone phosphate) in the presence of TIM and a coupling enzyme, *α*-glycerol phosphate dehydrogenase (Oesper & Meyerhof, 1950; Plaut & Knowles, 1972). The reaction mixture (1 ml) contained 100 m*M* triethanolamine, 5 m*M* EDTA, 0.5 m*M* NADH, 20 µg ml^{-1} *α*-glycerol phosphate dehydrogenase and

2.0 m*M* GAP. The enzyme activity was determined by monitoring the decrease in absorbance at 340 nm using a JASCO UV-530 spectrophotometer equipped with a Peltier cell holder. The specific activity was monitored over the temperature range 293–333 K.

2.7. Crystallization and data collection

MjTIM was crystallized using the microbatch method under an oil layer containing equal volumes of silicone and paraffin oils (Hampton Research). The protein, at a concentration of $\sim$7 mg ml^{-1}, was incubated with an excess of 3-phosphoglyceric acid (1:50 molar ratio) for 5 h prior to setting up crystallization trials. Initial trials for crystallization using Crystal Screen from Hampton Research gave thin needle-shaped crystals in condition No. 6 (0.2 *M* $MgCl_2$, 0.1 *M* Tris pH 8.5 and 30% PEG 4000). Further optimization resulted in diffraction-quality crystals of approximately 0.3 $\times$ 0.06 $\times$ 0.06 mm in size in the presence of 0.2 *M* $MgCl_2$, 0.1 *M* Tris pH 8.5 and 25–30% PEG 3350.

The crystals were transferred to crystallization buffer containing 20% glycerol as cryoprotectant for a few seconds before flash-freezing. A data set was collected at 100 K at a wavelength of 0.93 Å at the synchrotron beamline BL44XU at SPring-8, Hyogo, Japan using a DIP-2040b image-plate detector. The data set was processed using the *HKL*-2000 program suite (Otwinowski & Minor, 1997). The data could be scaled in space group *P*622 with an $R_{\rm merge}$ of 11.3%.

2.8. Detection of twinning

The intensity statistics, including the cumulative intensity plot and the moments of normalized intensities, were obtained using the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994). Statistical analysis of the intensities clearly indicated that the crystal used for data collection was twinned. The data set was scaled in space group *P*3 and estimation of the twinning fraction from the Yeates twinning tests (Yeates, 1997) was carried out using the *DETWIN* program from *CCP*4 (Collaborative Computational Project, Number 4, 1994). Based on the observations, it was concluded that the data had been collected from a perfect tetartohedrally twinned crystal and although it could be scaled in *P*622, the actual space group was *P*3.

2.9. Structure solution and refinement

From Matthews coefficient calculations (Matthews, 1968) assuming space group *P*3, there were eight polypeptides in the asymmetric unit. MjTIM shares 66% identity with PwTIM. Hence, a tetramer of the polyalanine chain of the PwTIM structure (PDB code 1hg3) was used as the model for molecular-replacement trials. Molecular replacement was carried out using the program *MOLREP* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). As no axial reflections were recorded, the presence of a screw axis could not be detected from the systematic absences and hence molecular replacement was tried out in space groups *P*3, $P3_1$ and $P3_2$. This gave a solution with an *R* factor of 53% and a correlation coefficient of 36% in space group $P3_1$. The high *R* factor for the MR solution in spite of the 66% identity between the two molecules again pointed out that the data could be twinned.

Owing to the presence of tetartohedral twinning, refinement could not be performed using the standard programs available. The refinement was carried out using the *CNS* modules and input files adapted for tetartohedral twinning (Barends *et al.*, 2005) obtained from the IUCr electronic archive (reference: WD5026). The map calculated after 20 cycles of rigid-body refinement and 50 cycles of minimization, using the programs modified for tetartohedral twinning, had clear electron density for most of the side chains of the residues of MjTIM. 5% of the reflections were kept aside as test reflections to monitor the progress of refinement through $R_{\rm free}$ calculations. The test reflections were chosen such that the twin-related reflections were maintained together in the same set. Owing to the presence of eight polypeptides in the asymmetric unit, noncrystallographic symmetry constraints were applied, excluding the loop regions, during a few initial rounds of refinement. After a few rounds of model building using *Coot* v.0.0.33 (Emsley & Cowtan, 2004) and refinement alternately, simulated annealing along with the NCS restraints helped in lowering the $R_{\rm work}$ and $R_{\rm free}$ to 24% and 31%, respectively. This was followed by addition of water molecules and *B*-factor refinement. NCS restraints were maintained during the remaining refinement cycles.

2.10. Structure analysis

PROCHECK (Laskowski, 2003) was used to calculate the parameters defining the quality of the model. Pairwise structural alignments were performed using *ALIGN* (Cohen, 1997), while multiple structural alignment was carried out using *MUSTANG* (v.0.3; Konagurthu *et al.*, 2006) and represented using *ESPript* (Gouet *et al.*, 1999). Only the C^{α} atoms which are common to all the 18 structures were considered in the calculation of r.m.s.d. in the pairwise comparisons. The residues at the interfaces were identified for TIMs from all organisms using the Protein Interfaces, Surfaces and Assemblies server (*PISA*; Krissinel & Henrick, 2005) at the European Bioinformatics Institute (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html). Surface-accessibility calculations were carried out using *NACCESS* (Hubbard & Thornton, 1996). The multiple sequence alignment of TIM sequences from archaea was obtained using *ClustalW* (Thompson *et al.*, 1994) and represented using the *AMAS* server (Livingstone & Barton, 1993). Figures were prepared using *PyMOL* (DeLano, 2002), while the electrostatic surface potential representation was produced using *GRASP* (Nicholls *et al.*, 1991).

3. Results and discussion

3.1. Expression and characterization

Recombinant MjTIM in vector pTrc99A was overexpressed in *E. coli* (strain AA200) and the protein was purified by

thermal precipitation followed by 75% ammonium sulfate fractionation, ion-exchange and gel-filtration chromatography. A 1 l batch of culture yielded 25–30 mg purified protein. The purified protein was seen as a 23 kDa band on SDS–PAGE and the ESI mass spectrum yielded a molecular weight of 23 245 Da (Fig. 1*a*). The molecular weights and the quaternary structure of the purified enzyme in solution were determined by analytical gel-filtration chromatography (Fig. 1*b*). The 23 kDa observed for the subunit molecular weight and the 92 kDa observed for the native enzyme suggested a tetrameric quaternary structure.

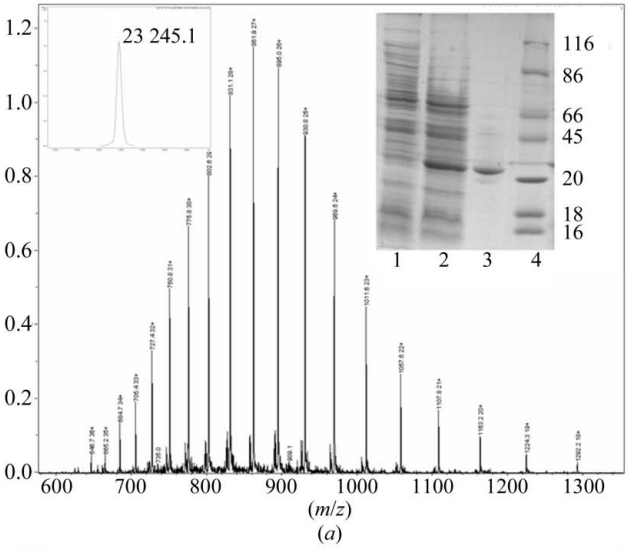
The published sequence of MjTIM corresponds to a molecular weight of 23 319 Da. The difference between the observed and calculated weights (74 Da) is significant. Sequencing of MjTIM cDNA revealed two differences, at positions 2 and 134, from the sequence deposited in the SWISS-PROT database (accession No. Q58923). Leu2 of the published sequence was replaced by a Val to accommodate a restriction-enzyme site for cloning purposes. Residue 134 was found to be Cys (codon UGC) instead of Tyr (codon UAC). The introduction of mutations is not uncommon during PCR amplification. This mutation is unlikely to influence the properties of the enzyme as the residue occurs in a solvent-accessible region. Cysteine-modification experiments have also demonstrated that Cys134 is accessible to the solvent (unpublished work).

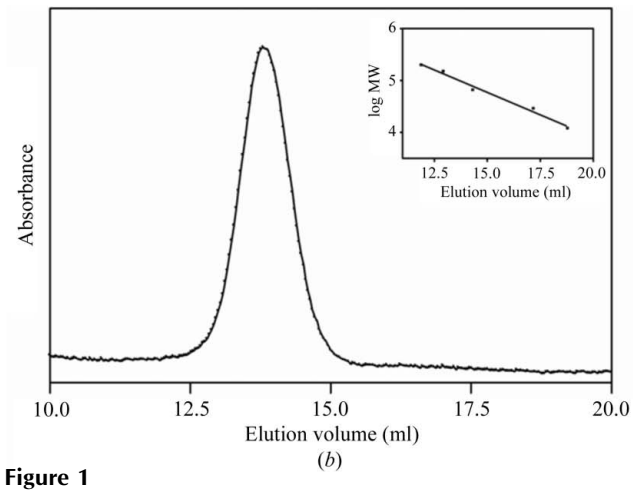
3.2. Kinetic parameters

The purified enzyme is active over the temperature range 293–333 K, showing an increase in specific activity with temperature (Fig. 1*c*). The specific activity of MjTIM is estimated to be around 12 000 U mg^{-1} at 323 K. Unit enzyme activity (U) is defined as the amount of the substrate GAP in micromoles converted to the product DHAP by 1 mg of enzyme per minute.

3.3. Structure solution

3.3.1. Detection of twinning. The cumulative intensity plot for observed acentric reflections was sigmoidal (Fig. 2*a*), indicating the presence of twinning. The moments of intensities in different resolution ranges also suggested the possibility of a perfectly twinned data set. Estimation of the twinning fraction from the Yeates test (Yeates, 1997) for the data scaled in *P*3 indicated twin fractions of 0.42, 0.42 and 0.46 for the three possible twin laws (k, h, $-l$), ($-k$, $-h$, $-l$) and ($-h$, $-k$, l), respectively (Fig. 2*b*). In the case of hemihedral twinning, a significant twin fraction can only be observed for one of the twin laws. Based on the observation of twin fractions greater than 0.4 for all three possible twinning directions, it was concluded that the data set corresponds to a perfect tetartohedrally twinned crystal. Although the data could be scaled in *P*622, the actual space group was *P*3 or one of its





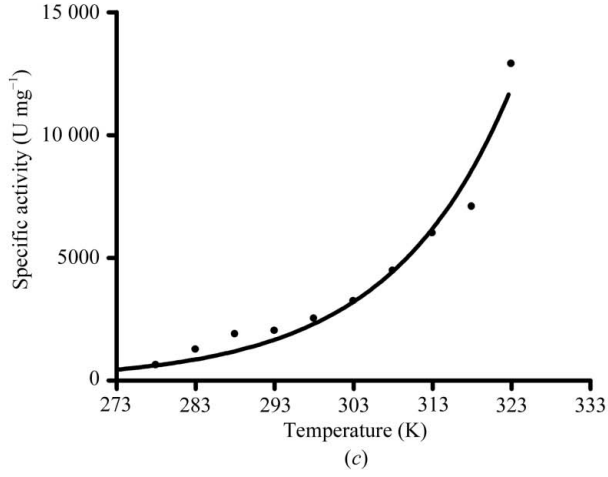


Figure 1
(*a*) The LC–ESI mass spectrum of MjTIM is shown along with its charge-state distribution. Left inset, the deconvoluted weight of MjTIM (23 245 Da); right inset, expression and purification profile of MjTIM on 12% SDS–PAGE (lane 1, uninduced culture; lane 2, induced culture; lane 3, pure protein after all purification steps; lane 4, molecular-weight markers labelled in kDa). (*b*) Analytical gel-filtration profile for MjTIM. The standard marker profile is shown as an inset. The column used for gel filtration was a 10 × 30 cm Superdex-200 (S-200). Buffer containing 50 m*M* phosphate pH 8 with 150 m*M* sodium chloride was used for all runs. (*c*) Dependence of the specific activity of MjTIM on temperature.

Table 1
Data-collection statistics for all space groups compatible with the trigonal lattice.

Values in parentheses are for the last resolution shell.

Space group	$P3_1$	$P3_121$	$P3_112$	$P6_1$	$P6_122$
Unit-cell parameters (Å, °)	$a = b = 139.340$, $c = 80.890$, $\alpha = \beta = 90$, $\gamma = 120$				
Resolution range (Å)	50–2.3	50–2.3	50–2.3	50–2.3	50–2.3
No. of reflections	1087652	1087652	1087652	1087652	1087652
No. of unique reflections	76266	39559	39148	39034	20738
Completeness (%)	97.4 (96.3)	97.4 (96.3)	97.8 (96.8)	98.1 (97.0)	98.0 (97.1)
$R_{\rm merge}$† (%)	9.4 (26.1)	10.6 (29.1)	10.6 (29.3)	8.8 (27.8)	10.7 (29.9)
$\langle I/\sigma(I)\rangle$	20.2 (6.3)	27.9 (8.5)	27.8 (8.3)	28.2 (8.6)	37.4 (11.3)

† $R_{\rm merge} = \sum|I_{hj} - \langle I(h)\rangle|/\sum I_{hj}$, where I_{hj} is the jth observation of I_h and $\langle I_h\rangle$ is its mean intensity.

variants. The scaling statistics in the four possible trigonal space groups are shown in Table 1.

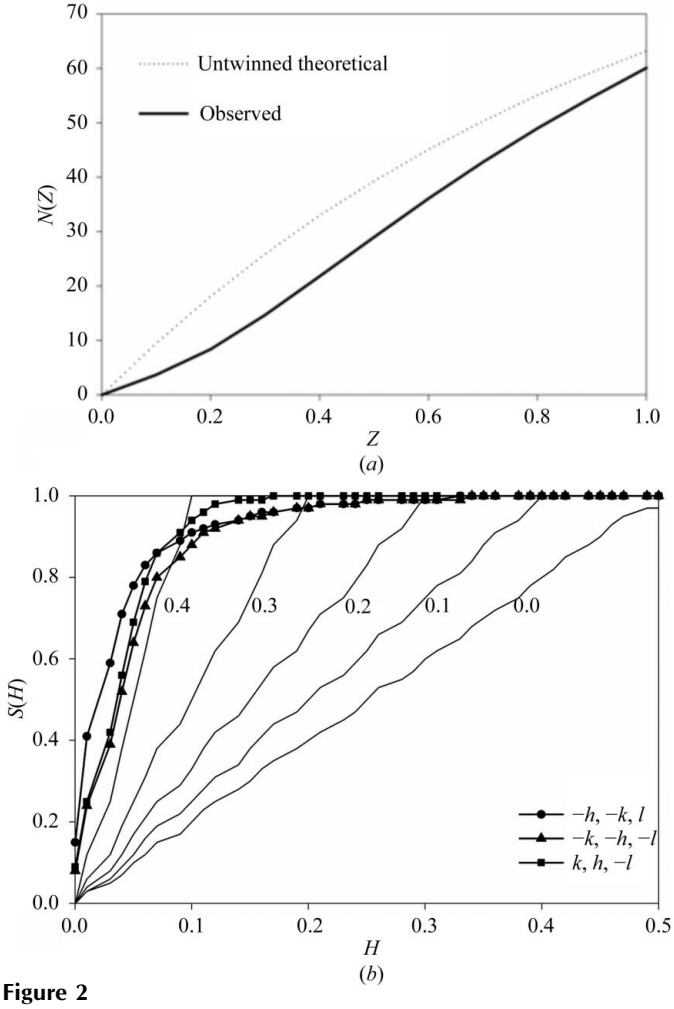


Figure 2
Intensity statistics for the data set scaled in $P3_1$. (*a*) Cumulative intensity plot. (*b*) Yeates twinning-test plots for the three twin operators $-h, -k, l$ (circles), $-k, -h, -l$ (triangles) and $k, h, -l$ (squares). The lines for the expected plots for various twin fractions (labelled 0.0, 0.1, 0.2, 0.3, 0.4) are also shown. H is defined as $H = |I_{\rm obs}(\mathbf{h}_1) - I_{\rm obs}(\mathbf{h}_2)|/[I_{\rm obs}(\mathbf{h}_1) + I_{\rm obs}(\mathbf{h}_2)]$, where $I_{\rm obs}(\mathbf{h}_1)$ and $I_{\rm obs}(\mathbf{h}_2)$ are the observed intensities of two acentric reflections related by the twin law and $S(H)$, the cumulative distribution of H, is given by $S(H) = H/(1 - 2\alpha)$, where α denotes the twin fraction (Yeates, 1997).

Tetartohedral twinning in protein crystals has only been previously reported in two other cases, namely *E. coli* MltA (Barends *et al.*, 2005) and SRP core protein from *Sulfolobus solfataricus* (Rosendal *et al.*, 2004). Of these two, only the structure of MltA has been reported to date. In tetartohedral twinning, the intensity of each reflection is a sum of the diffraction intensities from four of the domains that form the crystal. In such cases, as a result of the fourfold averaging of the intensities, the cumulative intensity statistics will be different from untwinned or hemihedrally twinned data sets. In the present case, the moments of intensities $\langle E\rangle$ and $\langle E^3\rangle$ are 0.933 and 1.208, respectively. These values are closer to the theoretical values for a perfect hemihedral twin than to the estimated values for a perfect tetartohedral twin. Such a departure from the values expected for tetartohedral twinning is known to occur when the noncrystallographic symmetry axes of an oligomeric structure are aligned with the crystallographic axes (Barends *et al.*, 2005). The eight molecules of MjTIM found in the asymmetric unit correspond to two tetramers, each with 222 symmetry. Their packing in the pseudo-$P622$ cell is such that the 222 symmetry of one of the tetramers coincides with the 222 symmetry of the crystal and the two tetramers are related by an approximate 6_1 screw along the c axis. Because of this special packing in the cell, the effects of twinning on the intensity statistics are expected to be reduced, as observed in the case of MltA (Barends *et al.*, 2005).

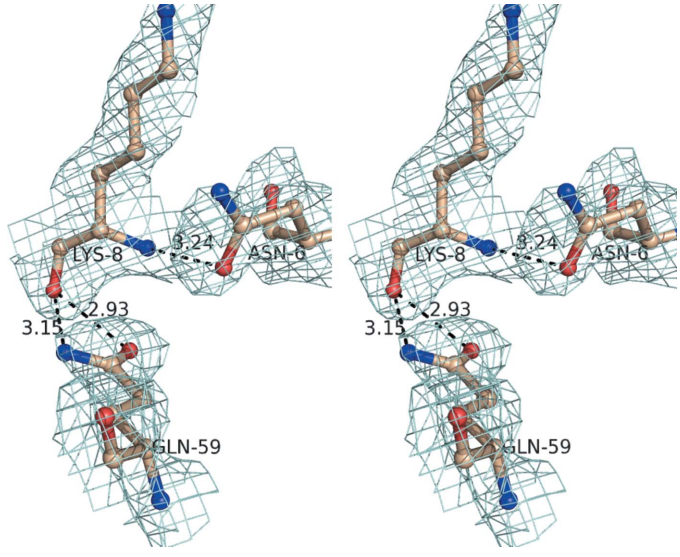


Figure 3
A representative section of the electron density ($2F_{\rm o} - F_{\rm c}$ contoured at 1.2σ) around the active-site Lys8. The residues Asn6 and Gln59 form hydrogen bonds with the main chain of the active-site lysine. The hydrogen bonds are shown as broken lines and the distances between the atoms involved are indicated.

Table 2
Refinement statistics and quality of the model.

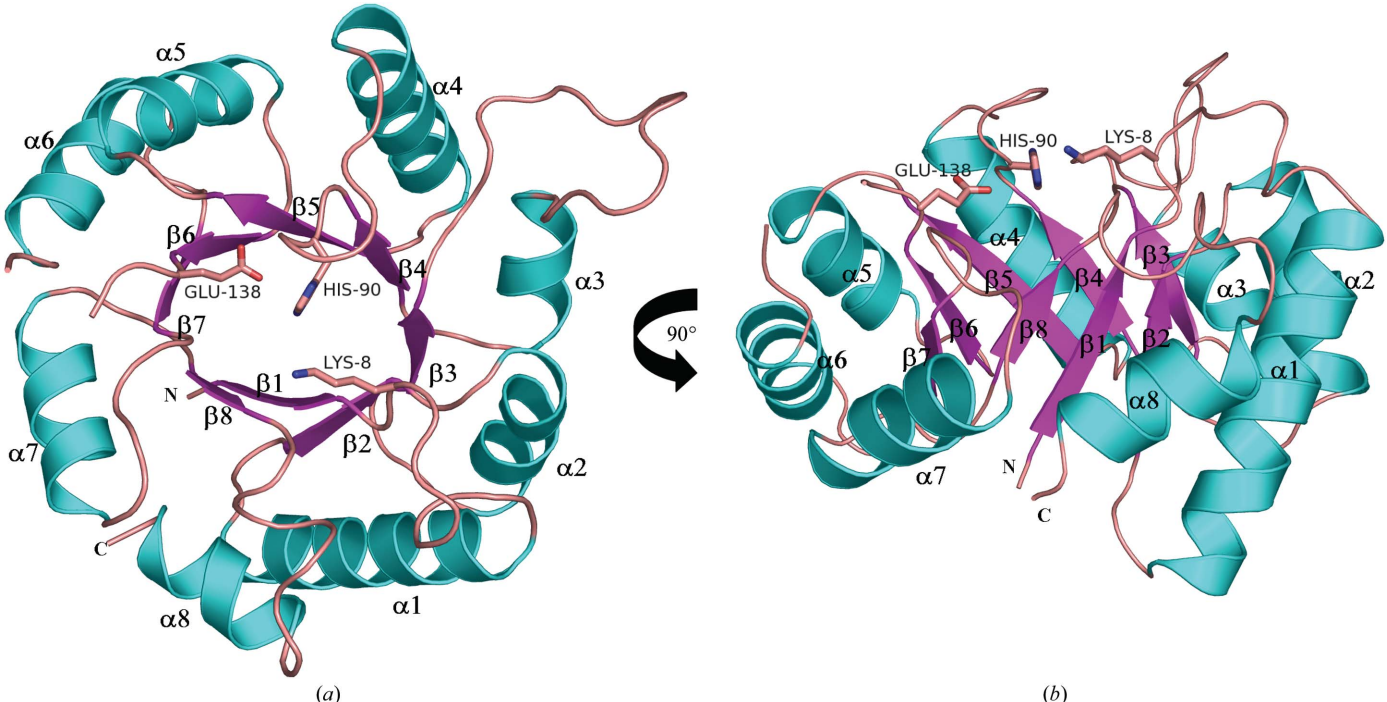
Values in parentheses are for the last resolution shell.

R factor (%)	21.3 (29.9)
R_{free} (%)	27.8 (32.3)
Model quality	
No. of atoms	
Protein atoms	12089
Waters	384
Average *B* factors (Å^2)	
Mean *B* value	36.9
B value from Wilson plot	23.5
R.m.s. deviations from ideal values	
Bond lengths (Å)	0.004
Bond angles (°)	1.091
Residues in Ramachandran plot (%)	
Most allowed region	85.7
Allowed region	12.9
Generously allowed region	0.8
Disallowed region	0.6

Table 3
Sequence and structural comparison of MjTIM (PDB code 2h6r) with other TIMs.

PDB code	Organism	Sequence identity (%)	R.m.s.d. (Å)	No. of aligned C^{α} atoms
1lyx	*Plasmodium falciparum*	12	1.83	172
1amk	*Leishmania mexicana*	24	1.88	172
1tcd	*Trypanosoma cruzi*	25	2.13	174
5tim	*Trypanosoma brucei brucei*	17	2.09	175
1wyi	Human	20	1.93	173
1ypi	Yeast	20	1.96	175
1r2r	Rabbit	20	1.94	177
1tim	*Gallus gallus*	20	1.79	174
1mo0	*Caenorhabditis elegans*	24	1.77	171
1m6j	*Entamoeba histolytica*	24	1.99	175
1tre	*Escherichia coli*	20	1.69	173
1aw1	*Vibrio marinus*	25	1.71	172
1yya	*Thermus thermophilus*	21	1.57	170
1btm	*Bacillus stearothermophilus*	28	1.59	173
1b9b	*Thermotoga maritima*	18	1.72	176
1w0m	*Thermoproteus tenax*	46	0.91	176
1hg3	*Pyrococcus woesei*	66	0.73	186

3.3.2. Refinement. The molecular-replacement solution obtained from *MOLREP* (Collaborative Computational Project, Number 4, 1994) in space group $P3_1$ has been refined, assuming perfect tetartohedral twinning, according to the steps described in §2, to final R_{work} and R_{free} values of 21.3% and 27.8%, respectively. The refinement and data-quality statistics are shown in Table 2. Residues 1–217 could be traced in all eight subunits, except for a stretch of residues in loop 6. There was no clear electron density for the loop in any of the eight subunits and therefore these residues (140–154) could not be modelled. 0.6% of the residues occur in the disallowed region of the Ramachandran plot. Lys8, the active-site residue, occurs near the boundary of the allowed and disallowed regions in all the subunits and hence is in the disallowed region in a few of the subunits. The occurrence of unusual φ, ψ values for the active-site lysine is a characteristic of all TIM structures; the values in MjTIM are ∼51, −122°. The electron density in this region is shown in Fig. 3. The hydrogen bonds formed by Asn6 and Gln59 with the main-chain atoms of Lys8, which hold it in the unusual conformation, are also shown. These interactions may compensate for the energetic penalty imposed by backbone strain (Stites *et al.*, 1994; Gunasekaran *et al.*, 1996). There was no electron density for ligand bound to



(*a*) (*b*)

Figure 4
Two perpendicular views of the overall fold of the MjTIM $(\beta/\alpha)_8$-barrel with the three active-site residues highlighted. The helices and strands are numbered.

the protein, although crystals were obtained in the presence of excess 3-phosphoglyceric acid.

3.4. Overall fold

MjTIM exhibits the classical $(\beta/\alpha)_8$-barrel structure, as expected. It resembles the tetrameric TIMs of archaeal origin, specifically those from *P. woesei* and *T. tenax*. Fig. 4 shows the overall fold of the barrel structure of MjTIM, with the three active-site residues Lys8, His90 and Glu138 highlighted. The MjTIM structure superposes with an r.m.s.d. of 0.75 Å for 198 aligned C^{α} atoms with the PwTIM structure, which was used as the molecular-replacement model and shares 66% sequence identity with MjTIM. Table 3 shows the identities of the MjTIM sequence to the sequences of other known TIM structures. It also lists the r.m.s. deviations observed on superposing the corresponding C^{α} atoms. The closest structures are those from other archaea, with r.m.s. deviations less than 1 Å. The structure-based sequence alignment shown in Fig. 5 permits analysis of the differences between the tetrameric archaeal TIMs and the dimeric TIM structures.

3.5. Oligomeric state

MjTIM occurs as a homotetramer in solution (Fig. 1*b*). The oligomeric state in the crystal structure is in agreement with that in solution. The eight monomers observed in the asymmetric unit of MjTIM correspond to two functional tetramers. In the archaeal TIM structures, PwTIM, TtTIM and MjTIM, the tetramer is formed as a dimer of dimers, the dimer interface being similar to that in mesophilic TIMs. The residues involved in the dimer and tetramer interfaces are highlighted in the structural alignment shown in Fig. 5 and a list of interacting pairs of residues is

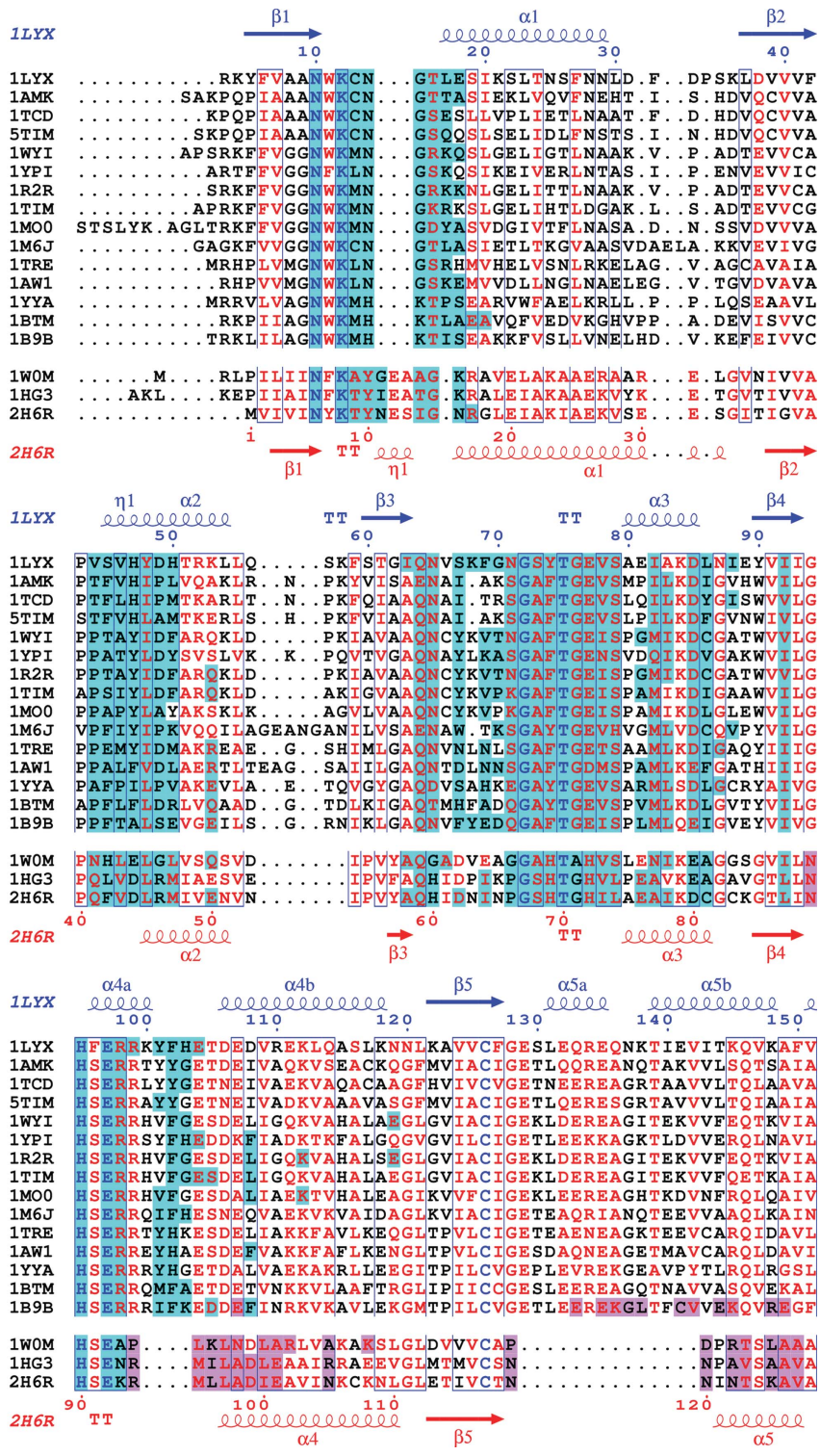


Figure 5
Structure-based sequence alignment of TIMs from various organisms. The structures used to obtain the alignment are denoted by their PDB codes and the details are shown in Table 3. The sequence numbering and secondary-structure assignments at the top (shown in blue) and bottom (in red) correspond to the PfTIM and MjTIM sequences, respectively. The residues involved in dimer and tetramer interfaces are highlighted in cyan and pink, respectively. The three tetrameric TIMs from archaea are grouped together as a separate block from the rest of the TIM sequences. The loop 6 residues in MjTIM, which are missing in the structure, have been included in the alignment according to the pairwise sequence alignment with PwTIM and are highlighted in cyan letters. The structure-based alignment was performed using *MUSTANG* (Konagurthu *et al.*, 2006) and represented using *ESPript* (Gouet *et al.*, 1999). The residues involved in the interfaces were identified using the *PISA* server (Krissinel & Henrick, 2005) and are highlighted in the representation.

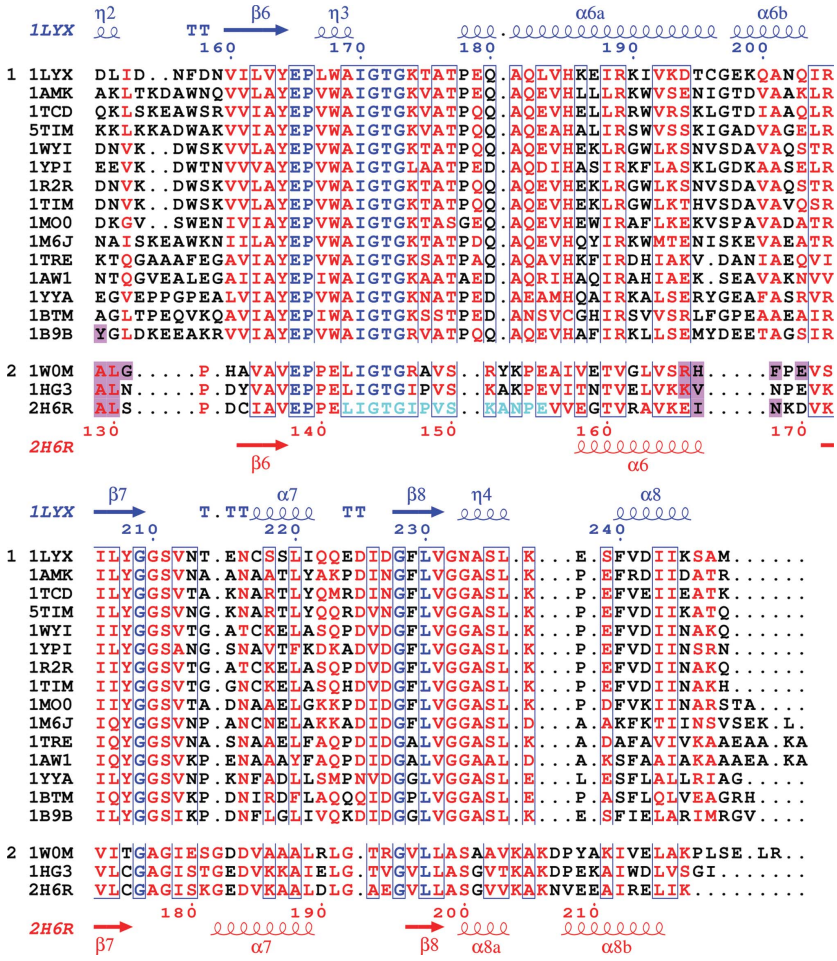


Figure 5 (continued)

given in Tables 4 and 5. A detailed comparison between the structures of tetrameric and dimeric TIMs has been carried out with the aim of understanding the structural adaptations required for tetramerization.

The characteristics of the MjTIM tetramer are similar to those of PwTIM, whereas TtTIM differs from these two in that it exists as a weak tetramer exhibiting an equilibrium between active tetramers and inactive dimers in solution (Walden *et al.*, 2004). Another example of a tetrameric TIM is that from *Thermotoga maritima* (TmTIM), which is formed from monomers of length similar to those of the dimeric TIMs, unlike archaeal TIMs, which are about 30 residues shorter. This TIM is an exception as it forms a covalently bonded tetramer by disulfide linkage involving Cys142 of twofold-related dimers (Maes *et al.*, 1999). The tetramer of TmTIM is not functionally significant; the enzyme is a dimer in solution (Beaucamp *et al.*, 1997) and tetramerization appears to be a crystallization artifact.

3.5.1. Dimer interface. The dimer interface consists mainly of hydrogen-bond interactions involving loop 3, loop 1 and helix 2. The buried surface area on dimerization corresponds to 14.2% of the total surface area of the *A* and *B* subunits. 37.3% of the buried surface area is polar. There are two salt bridges at the interface: one between Arg46 and Asp44 and the other involving Asp80 and Arg17. The salt bridge between Arg46 and Asp44 is also conserved in PwTIM. Most of the hydrogen bonds at the dimer interface are conserved in the MjTIM and PwTIM structures.

The main chain at the dimer interface of MjTIM superposes well with those of the dimeric TIMs, showing that the dimer interfaces are similar. Superposition of the dimer interfaces of MjTIM and a dimeric TIM (PfTIM), considering 40 C^{α} atoms at the interface, results in a low r.m.s.d. of 0.46 Å. Although there are some sequence changes, features of the interface such as the percentage of the buried surface area compared with the total (17.9%) and the percentage of polar residues in the buried surface area (34.5%) remain similar in the two TIMs.

3.5.2. Tetramer interface. The tetramer interface of MjTIM reveals many more nonpolar interactions than the dimer interface. Table 5 shows that the interactions at the tetramer interface are mainly between residues from loop 4 and helix 5. 9.2% of the total surface area of the *AB* and *CD* dimers is buried at the tetramer interface. Met95, Leu96 and Leu97 are the major residues involved in nonpolar contacts, while Glu101, Asn120, Asn122 and Thr123 form hydrogen bonds at the tetramer interface in MjTIM (Table 5). The interactions between the hydrophobic residues of helix 5, Ala126, Ala129 and Leu130, also contribute to the tetramer interface.

The structure-based sequence alignment shows that in the dimeric TIMs there are large insertions at sites corresponding to the region of the tetramer interface. The major insertions include four residues (Lys100–His104 in PfTIM numbering) in loop 4, a long stretch from Glu129 to Val142 and Asp155–Asn159 (PfTIM numbering) in the vicinity of helix 5 and Cys196–Gln200 on helix 6. The insertion in the loop 4 region in the dimeric TIMs is comprised of bulky residues such as Tyr/Phe. Helix 5 and loop 5 are shortened in the tetrameric TIMs. The shortening of the loop results in a more compact structure as the helix is pulled towards the barrel. The length of the insertions is fairly constant in all the dimers, as can be observed from the structure-based alignment in Fig. 5. Fig. 6(*a*) shows the MjTIM tetramer, while Fig. 6(*b*) shows a superposition of dimers of tetrameric and dimeric TIMs, with MjTIM and PfTIM as the representatives. The insertions at the putative tetramer interface are highlighted in the case of the PfTIM dimer. From the relative orientations of the dimers in Fig. 6, it appears that the insertions can act as a hindrance to tetramerization.

An electrostatic surface representation of the dimers in the dimeric and tetrameric TIMs is shown in Figs. 7(*a*) and 7(*b*). In the dimeric TIMs, the region corresponding to the tetramer interface of archaeal TIMs is strongly negatively charged,

Table 4
List of interacting pairs of residues at the dimer interface.

VdW denotes van der Waals interactions.

Subunit *A*	Secondary structure	Subunit *B*	Secondary structure	Type of interaction
Asn6	β1	Thr70	L3	Hydrogen bond
Thr9	L1	Gly67	L3	Hydrogen bond
Tyr10	L1	Gly67	L3	Hydrogen bond
		Ser68	L3	Hydrogen bond
Ile14	L1	Ala77	L3	Hydrogen bond
		Asp80	α3	VdW
Gly15	L1	Asp80	α3	Hydrogen bond
Arg17	α1	Asp80	α3	Salt bridge
Phe42	α2	Phe42	α2	VdW
Val43	α2	Leu45	α2	VdW
Asp44	α2	Arg46	α2	Salt bridge
Leu45	α2	Val43	α2	Hydrogen bond
Arg46	α2	Asp44	α2	Salt bridge
Gln59	L3	Gly71	L3	Hydrogen bond
His60	L3	Gly71	L3	Hydrogen bond
Pro66	L3	Asn11	L1	Hydrogen bond
Gly67	L3	Thr9	L1	Hydrogen bond
Ser68	L3	Tyr10	L1	Hydrogen bond
		Glu92	L4	Hydrogen bond
His69	L3	Thr9	L1	Hydrogen bond
		Glu92	L4	Hydrogen bond
Thr70	L3	Asn6	L1	Hydrogen bond
		Glu92	L4	Hydrogen bond
Gly71	L3	Phe42	L3	Hydrogen bond
		Gln59	α2	Hydrogen bond
		His60	L3	Hydrogen bond
Ile73	L3	Phe42	α2	Hydrogen bond
Asp80	α3	Ile14	L1	Hydrogen bond
His90	L4	Thr70	L3	Hydrogen bond
Glu92	L4	Ser68	L3	Hydrogen bond
		His69	L3	Hydrogen bond
		Thr79	L3	Hydrogen bond

Table 5
List of interacting pairs of residues at the tetramer interface.

VdW denotes van der Waals interactions.

Subunit *A*	Secondary structure	Subunit *C*	Secondary structure	Type of interaction
Arg94	L4	Ala98	α4	VdW
Met95	L4	Leu96	L4	VdW
		Leu97	L4	Hydrogen bond
Leu96	L4	Arg94	L4	VdW
		Met95	L4	VdW
		Leu96	L4	VdW
Leu97	L4	Met95	L4	Hydrogen bond
		Ile100	L4	VdW
		Val125	α5	VdW
Ala98	L4	Arg94	L4	VdW
		Asn119	L4	VdW
Ile100	L4	Leu97	L4	VdW
Glu101	α4	Asn120	α5	VdW
		Asn122	α5	Hydrogen bond
		Thr123	α5	Hydrogen bond
Asn120	α5	Glu101	α5	Hydrogen bond
Asn122	α5	Glu101	α5	Hydrogen bond
Thr123	α5	Glu101	α5	VdW
Ala126	α5	Ala129	α5	VdW
		Leu130	α5	VdW
Val127	α5	Leu97	L4	VdW
Ala129	α5	Ala126	α5	VdW
Leu130	α5	Ala126	α5	VdW

suggesting that repulsive interactions will hinder tetramerization. In contrast, hydrophobic patches are observed for MjTIM at the site of tetramerization. Fig. 7(*d*) shows a structure-based alignment of the sequences in the tetramer-interface region, where a large insertion is observed in the dimers. Inspection of the inserted region, which encompasses a segment of helix 5, shows the presence of several aspartate and glutamate residues, resulting in a strongly negatively charged region. A critical hydrophobic segment, Met95–Leu97 in MjTIM, is replaced by polar residues in the dimeric TIMs. The patch of nonpolar surface in MjTIM is fully buried during tetramerization, as is evident from the electrostatic surface-potential representation of the tetrameric MjTIM (Fig. 7*c*) shown in the same orientation. Thus, the surface in tetrameric TIMs is modified such that it facilitates tetramerization.

Shortening the sequence without a significant change in the overall fold requires several key mutational changes, which accommodate the shorter connecting loops. Asn119, Asn120 and Asn168 of MjTIM occur in the region corresponding to the insertion in helix 5 in dimeric TIMs. The direction of the main chain changes at these residues. A conserved glycine is present in all the dimeric TIMs at the position corresponding to Asn89 (MjTIM numbering). The side chain of Asn89 is involved in a hydrogen bond with the main-chain atom of Met95 of the tetramer interface.

Helix 5 of tetrameric TIMs is rich in alanines (Ala126, Ala128 and Ala129 in MjTIM). These alanines are conserved in the tetrameric TIMs. There is greater variability in the dimeric TIMs at these positions, particularly the residue corresponding to Ala129 of MjTIM. It has been suggested that the increased alanine content in archaeal TIMs is essential for stabilization of the truncated helix 5 (Walden *et al.*, 2001).

In contrast to MjTIM and PwTIM, TtTIM exists in a dimer–tetramer equilibrium. The nature of the tetramer interface is not conserved in different archaeal TIMs. The Met95-Leu96 segment at the tetramer interface of MjTIM is replaced by Leu98-Lys99 in TtTIM. Other tetramer-interface residues that are unique to the TtTIM are Pro97, Asn101, Pro122, Asp123, Ala130 and Phe171 (TtTIM numbering). This observation suggests that there can be different classes of tetramers among the archaeal TIMs: some are held by strong tetramer interfaces as in the case of MjTIM and PwTIM, while others such as TtTIM may exist as weak tetramers.

3.6. Thermostability

The crystal structures of TIMs from the thermophilic bacteria *Bacillus stearothermophilus* (BsTIM), *Thermus thermophilus* and *Thermotoga maritima* (TmTIM), all three of which are dimeric in solution, have been determined (Delboni *et al.*, 1995; Maes *et al.*, 1999). The tetrameric TIMs from archaea form a separate class of thermophilic TIMs. There is no clear consensus on the structural basis for thermostability in proteins. Several factors including oligomerization, enhanced polar interactions, tighter packing in the core of the protein and accessible surface area have been advanced to account for thermal stability (Robinson-Rechavi *et al.*, 2006).

In the case of BsTIM, the thermostability of the enzyme has been attributed to extensive hydrophobic interactions, reduced cavity volume, the presence of proline residues and the replacement of a conserved Asn-Gly pair of mesophilic TIMs (Asn14-Gly15 in PfTIM) at loop 1 with His-Lys (Delboni *et al.*, 1995). The replacement of Asn-Gly by His-Lys appears to be an adaptation to prevent deamidation, which can occur at high temperatures, thereby indirectly facilitating thermostability. This replacement is also found in the other two thermophilic bacterial TIMs (Fig. 5). In the case of the archaeal TIMs there is an insertion of three residues in this region. Interestingly, the MjTIM sequence has as many as 16 Asn residues out of a total of 219 residues, suggesting that deamidation may not be a significant contributor to lowering thermal stability.

A detailed comparison of the structure of the *T. maritima* enzyme with nine other dimeric TIM structures suggested that the thermostable protein possesses a larger number of salt bridges and has a larger buried hydrophobic surface (Maes *et al.*, 1999). The difficulties in defining precisely the factors that govern the thermostability of proteins, despite the availability of a large body of structural information, suggests that thermal stability can be achieved by entirely different approaches which have evolved independently. In the case of the archaeal TIMs PwTIM and TtTIM, oligomerization has been suggested to be a major factor in contributing to thermal stability (Walden *et al.*, 2001, 2004). A detailed analysis of the MjTIM structure has been carried out in order to assess whether structural correlations can be derived to rationalize thermal stability.

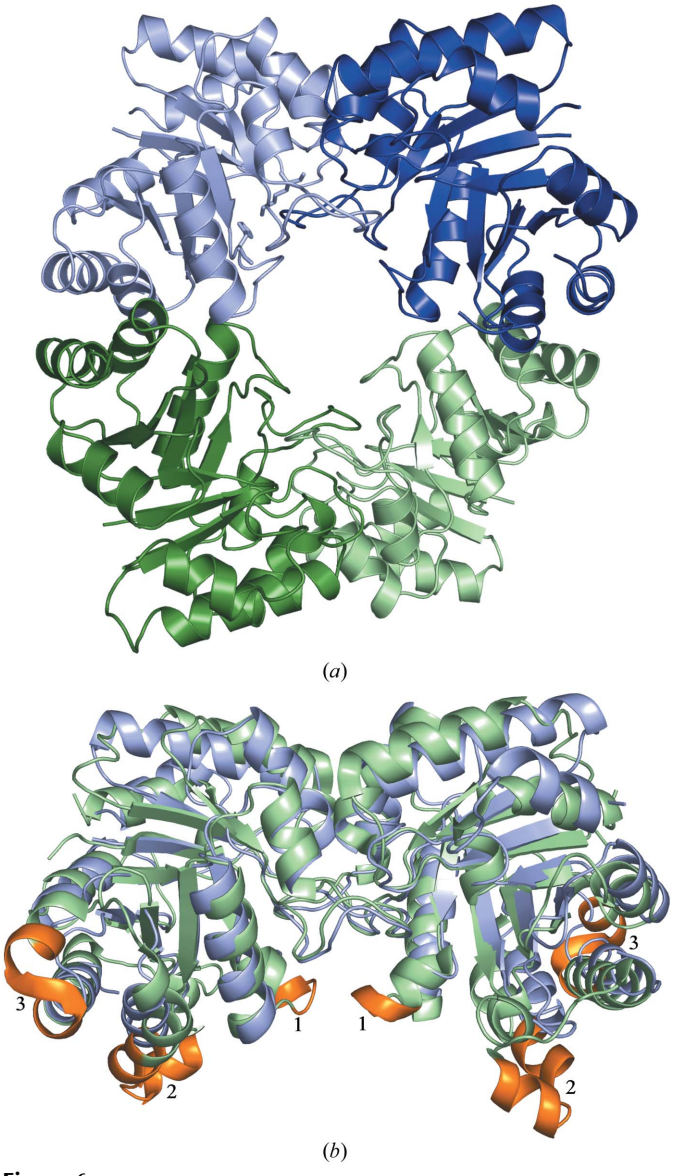


Figure 6
(*a*) Tetramer of MjTIM represented as a dimer of *AB* (blue) and *CD* (green) dimers. (*b*) Superposition of dimers of MjTIM (blue) and PfTIM (PDB code 1lyx, green), with the regions of insertions in dimeric TIMs highlighted in orange. The insertion regions in loop 4, helix 5 and helix 6 in both subunits of PfTIM are labelled 1, 2 and 3, respectively.

3.7. Barrel interactions in TIMs

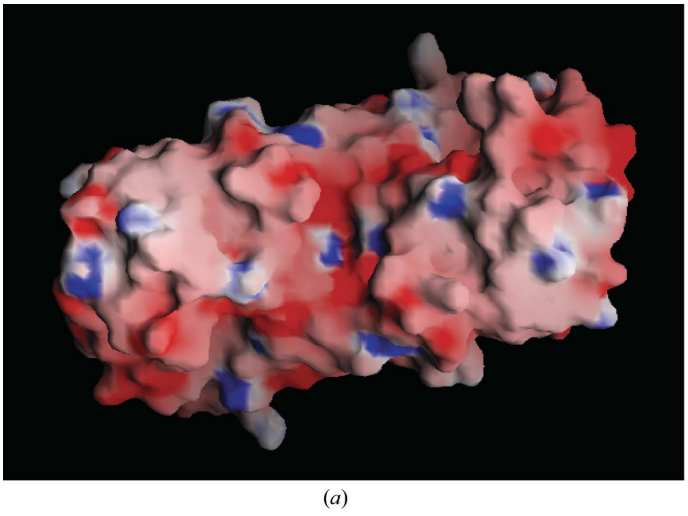
The TIM barrel, one of the most common protein folds, has been suggested to be one of the most stable protein structural motifs. The analysis of the interactions forming the core of the barrel and their conservation across various organisms might provide insights regarding the stability of the protein. The interactions that stabilize the TIM-barrel structure may be classified as those involving residues with side chains pointing towards the core of the barrel (inner barrel residues) and those involving residues with side chains pointing towards the outer surface of the barrel (outer barrel residues) and their interactions with the helix and loop residues and finally the interactions in the layer of helices surrounding the barrel.

3.7.1. Inner barrel interactions. The innermost core of the barrel is mostly held together by nonpolar interactions. Fig. 8 shows the interactions in the barrel core in the tetrameric MjTIM and the dimeric PfTIM. In all TIM barrels, hydrophobic residues with bulky side chains such as isoleucine and leucine and aromatic residues mediate contacts across the strands. A conserved aromatic residue pointing into the barrel, corresponding to Tyr90 of PfTIM, is found in all the dimeric TIMs. However, this is replaced by a glycine in the archaeal TIMs and a histidine in the TIM of the psychrophilic bacterium *Vibrio marinus* (VmTIM). In archaeal TIMs, the substitution is compensated for by the presence of an aromatic residue (Tyr57 in MjTIM) corresponding to a Gly/Ser/Ala in dimeric TIMs. The observed mutations lead to compensatory interactions, which may be essential for the stability of the archaeal TIM barrel. Tyr57, on the third strand of the barrel in MjTIM, is more centrally placed with respect to the barrel than Tyr90 (PfTIM), which occurs on strand 4 (Fig. 8). Thus, Tyr57 participates in a larger number of interactions within the barrel. The presence of a histidine with fewer interactions in the position corresponding to Tyr90 of PfTIM might be one of the reasons for the low thermostability of VmTIM.

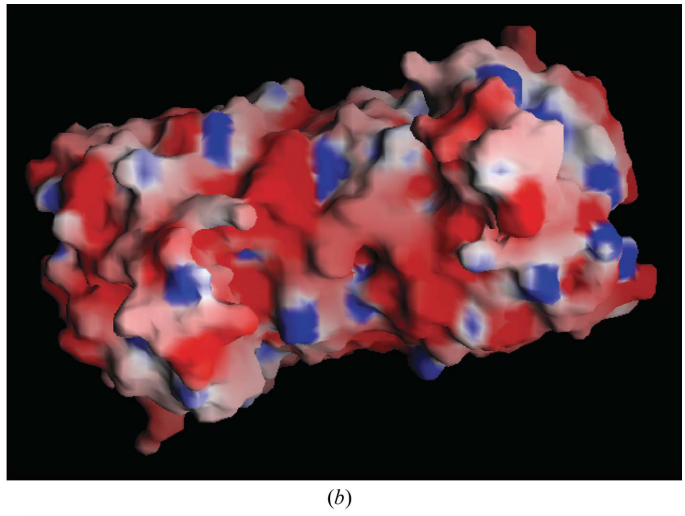
An analysis of the conservation of residues in the core of the barrel shows that Leu197 (MjTIM) is strictly conserved. This residue interacts with non-adjacent strands of the TIM

barrel. Cys117 (MjTIM) is also a strictly conserved residue, although it does not make many interactions with the barrel residues and occurs at the carboxy end of strand 5 of the barrel. It has been shown that this cysteine is essential for stability and efficient folding in studies carried out using yeast TIM (Gonzalez-Mondragon *et al.*, 2004). Another residue that is conserved in all TIMs is Ala136 (MjTIM). The remainder of the inner barrel residues are nonpolar across TIMs from all organisms, except in the case of the residue corresponding to Leu174 (MjTIM). In a few cases (1m6j, 1tre, 1aw1 and 1btm; see Fig. 5) glutamine replaces Leu174. The glutamine side chain forms hydrogen bonds with the main chain of the barrel residues.

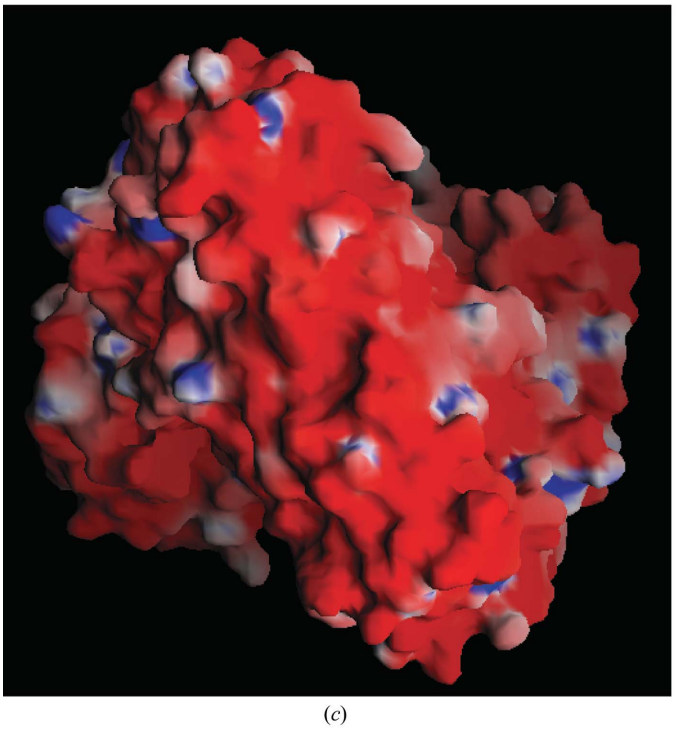
The conserved residues Asn6 (Asn10 of PfTIM) and Gln59 (Gln64 of PfTIM) occur at the tip of strand 1 and strand 3, respectively. The glutamine forms a hydrogen bond with the main-chain atom of the last residue of strand 2. These two residues are also involved in hydrogen bonds with the main-chain atoms of the active-site lysine and are crucial for the activity and stability of TIMs. This glutamine residue is replaced by a glutamate only in *Leishmania* TIM. Previous studies have shown that mutation of Glu65 in *Leishmania* TIM



(*a*)



(*b*)



(*c*)

(*d*)

Figure 7
Electrostatic surface-potential representation of dimeric and tetrameric TIMs. (*a*) MjTIM dimer, (*b*) PfTIM dimer, (*c*) MjTIM tetramer. MjTIM and PfTIM dimers are shown in the same orientation, with the tetrameric interface of MjTIM facing towards the viewer. The tetramer is also represented in the same orientation. The figures were generated using *GRASP* (Nicholls *et al.*, 1991). Blue, red and white regions represent positively charged, negatively charged and hydrophobic regions respectively. (*d*) A section of the structure-based sequence alignment of all TIMs, showing the insertion region in the vicinity of helix 5. Acidic residues are coloured blue, hydrophobic red, basic magenta and polar amino acids with hydroxyl and/or amide groups green.

to glutamine leads to an increase in the stability of the enzyme (Williams *et al.*, 1999).

3.7.2. Outer barrel interactions. One of the prominent features of the TIM barrels of archaea that differentiates them from the rest of the TIMs is the complete absence of aromatic residues in the outer surface of the barrel (Fig. 8). The aromatic residues Tyr164, Tyr208 and Phe229 (PfTIM numbering) are replaced by Val, Cys/Thr and Ala/Gly, respectively, in archaea. The replacement of the bulky aromatic residues by smaller hydrophobic residues makes the archaeal TIM structures more compact. Phe127 on β5 of the PfTIM barrel is changed to Thr118 in MjTIM. Fig. 9 shows a superposition of the interface of strand 5 and helix 5 of MjTIM and PfTIM. Helix 5 has moved closer to strand 5 in MjTIM owing to the shorter loop connecting them and the presence of a smaller residue (Thr118) at the interface. A similar argument holds for the mutation of Tyr164 (PfTIM) in the next strand, the side chain of which points towards the loop 5. Phe229 also is an instance of mutation to a smaller residue (valine) in MjTIM.

Gln133 of PfTIM, which interacts with Phe127, is replaced by Thr123 in MjTIM. Thr123 (corresponding to glutamine in all dimeric TIMs except *Gallus gallus* TIM) of helix 5 forms two hydrogen bonds with Thr118 (MjTIM numbering). These hydrogen bonds compensate for the hydrophobic interactions between Phe127 and Gln133 of PfTIM. The C^{γ} group of Thr118 forms van der Waals contacts with Ile135 and Val137 (MjTIM numbering), which are outer barrel residues on the adjacent strand. The corresponding residues are also hydrophobic in the other dimeric TIMs.

The TIM barrel is mainly stabilized by hydrophobic interactions and a few hydrogen bonds. The N-terminal residues of the strands have some polar or charged residues that are mostly surface-exposed and can form polar contacts with the loop regions. The interactions between the barrel and the helices are also mainly hydrophobic. The α-helices and loops are reinforced by hydrogen bonds and salt bridges occurring between the side chains of surface-exposed residues. The numbers of salt bridges and hydrogen bonds formed in the dimeric and thermostable tetrameric TIMs are comparable, a feature noted previously for thermophilic TIMs (Maes *et al.*, 1999).

The analysis of the barrel interactions in TIMs points out certain features that are conserved across all the organisms. The thermostability of the archaeal TIMs cannot be explained on the basis of the number of interactions in the inner or outer barrel regions. The major difference between the tetrameric

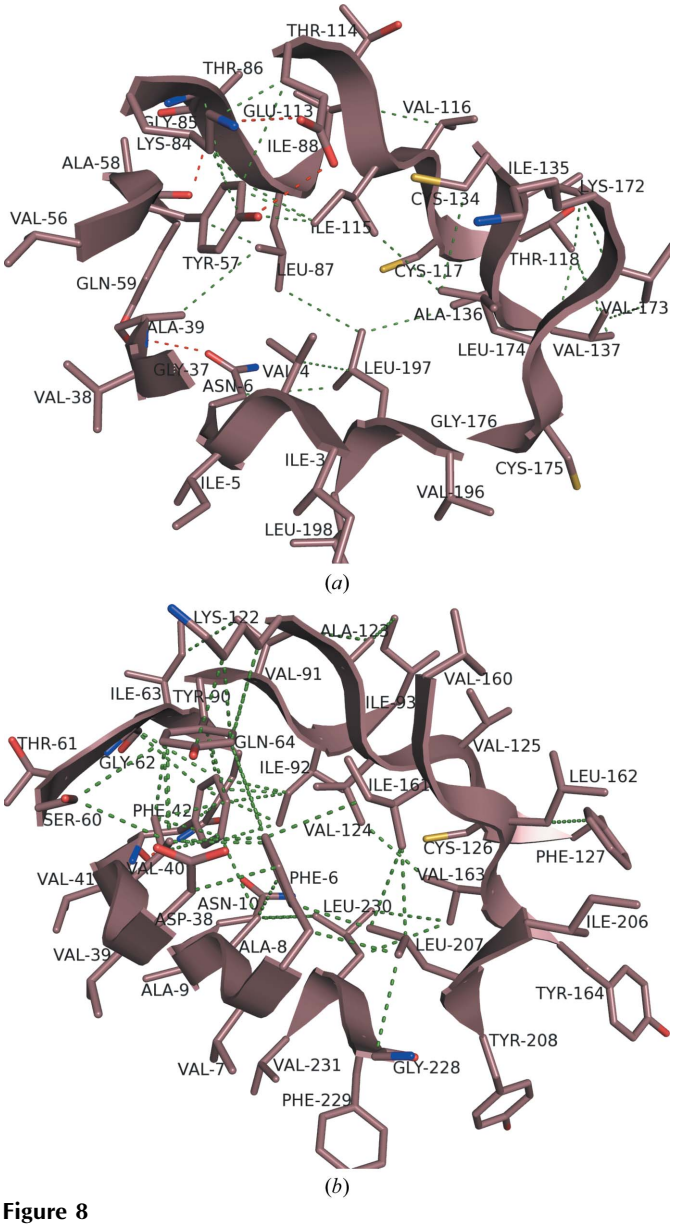


Figure 8
Barrel residues in (*a*) MjTIM and (*b*) PfTIM. Only the residues which form part of the β-strands are shown. The atoms involved in van der Waals interactions are connected by green dashed lines, while hydrogen bonds are shown in red. The hydrogen bonds involving only main-chain atoms, which form the parallel β-sheet of the barrel, are not shown.

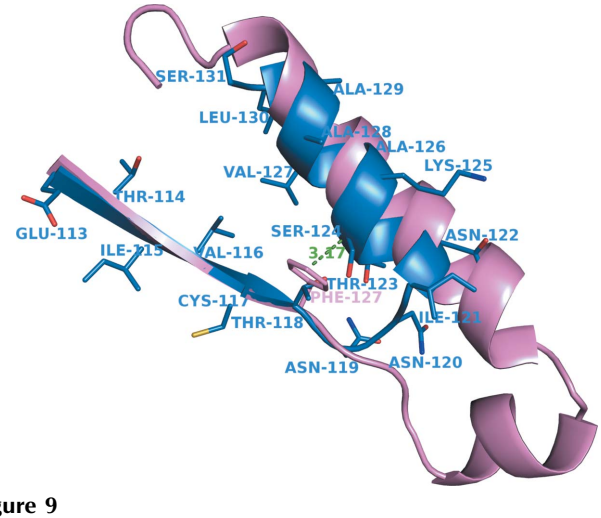


Figure 9
Superposition of β5 and α5 of MjTIM and PfTIM. Residues from MjTIM are shown in blue, while those from PfTIM are shown in pink. The closest distance between the Phe127 side chain of MjTIM and the main-chain C^{α} atom of MjTIM is indicated. This figure was generated by rotating the coordinates of the PfTIM β5 and α5 residues according to the matrix obtained by superposition of only strand β5 of the two TIMs.

and dimeric TIM barrels is the presence of aromatic residues. The mutations of aromatic residues at critical positions of the barrel and the shortening of the loops have contributed to the compactness of the tetrameric TIM structures and these might contribute to the increase in stability.

3.8. Disorder of loop 6

The active-site loop 6 includes the residues connecting strand 6 to helix 6 in all TIMs and extends from Pro166 to Thr177 in PfTIM. A notable feature of the active-site loop is its ability to exist in two distinct conformations: open and closed (Joseph *et al.*, 1990). The closed form is largely found in liganded TIM structures, while the open form is found in unliganded structures. PfTIM is unusual in that open forms have frequently been observed even when the active site is occupied by ligand (Parthasarathy, Balaram *et al.*, 2002; Parthasarathy, Ravindra *et al.*, 2002). The loop dynamics have been a subject of considerable interest and the consensus is that loop 6 moves as a rigid body about hinges that consist of Pro166–Trp168 and Lys174–Ala176 (PfTIM numbering) on either side (Sun & Sampson, 1998, 1999).

Fig. 10 illustrates examples of loop 6 in both the open and closed conformations. In the PwTIM structure complexed with 2-carboxyethyl phosphonic acid the loop exists in a closed conformation (Walden *et al.*, 2001; Fig. 10*a*). In all cases where the loop has been crystallographically characterized, an ordered structure held together by internal interactions is observed. Two distinct loop types may be identified. In the dimeric TIMs, the N-terminal hinge segment is made up of residues P*X*W, of which Trp168 (PfTIM) is strictly conserved. As seen in Fig. 10, this residue stabilizes the loop structure by

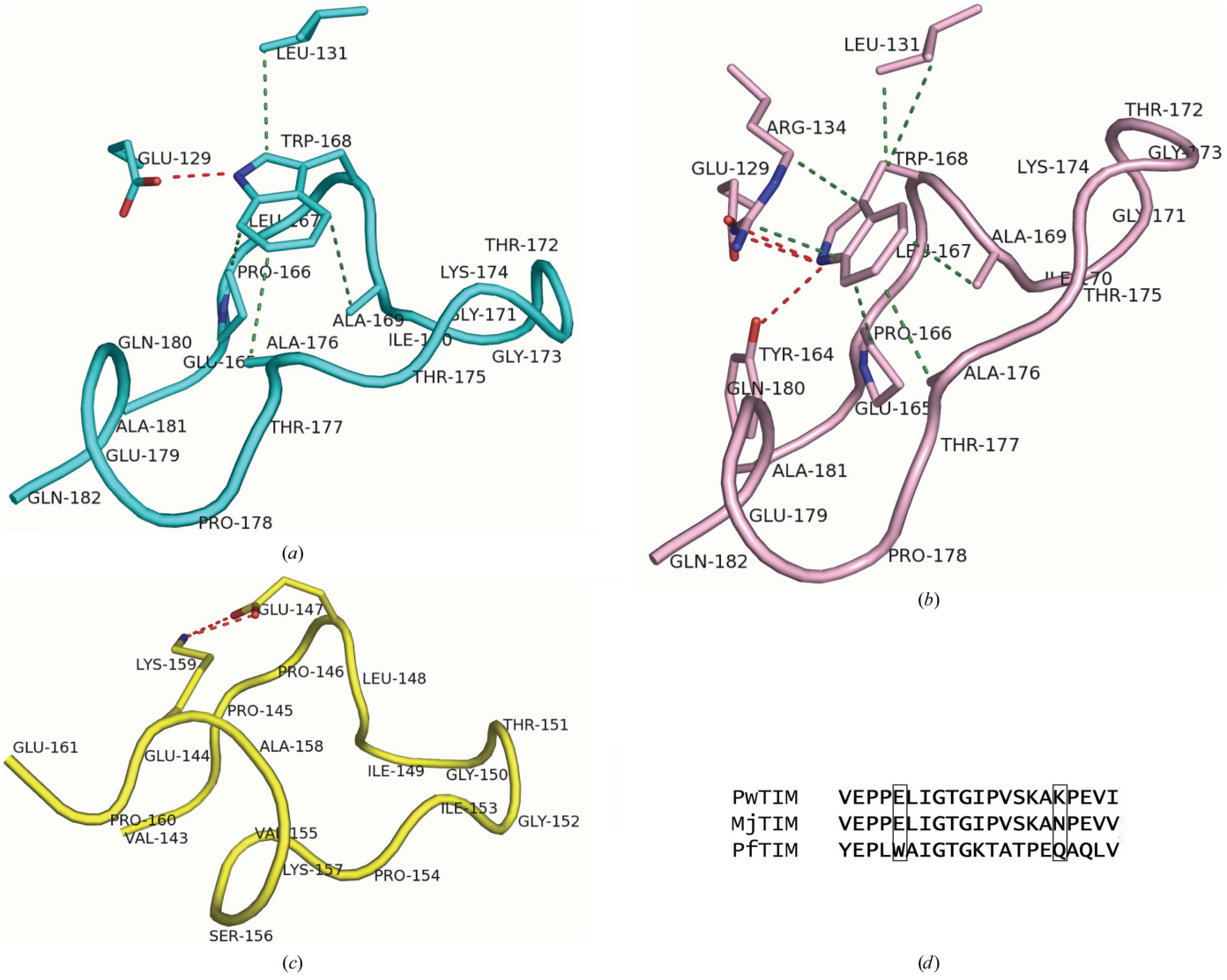


Figure 10
Loop 6 residues in PwTIM and PfTIM. (*a*) Loop 6 in the closed state for PfTIM (PDB code 1m7p). (*b*) Loop 6 in the open state for PfTIM (PDB code 1lyx). The residues interacting with Trp168 are also highlighted. (*c*) Conformation of loop 6 in the PwTIM structure (PDB code 1hg3). The salt bridge formed between Glu147 and Lys159 is highlighted and the residues involved are labelled. (*d*) Alignment of the loop 6 residues of PwTIM, MjTIM and PfTIM. The residues corresponding to Glu147 and Lys159 of PwTIM are highlighted.

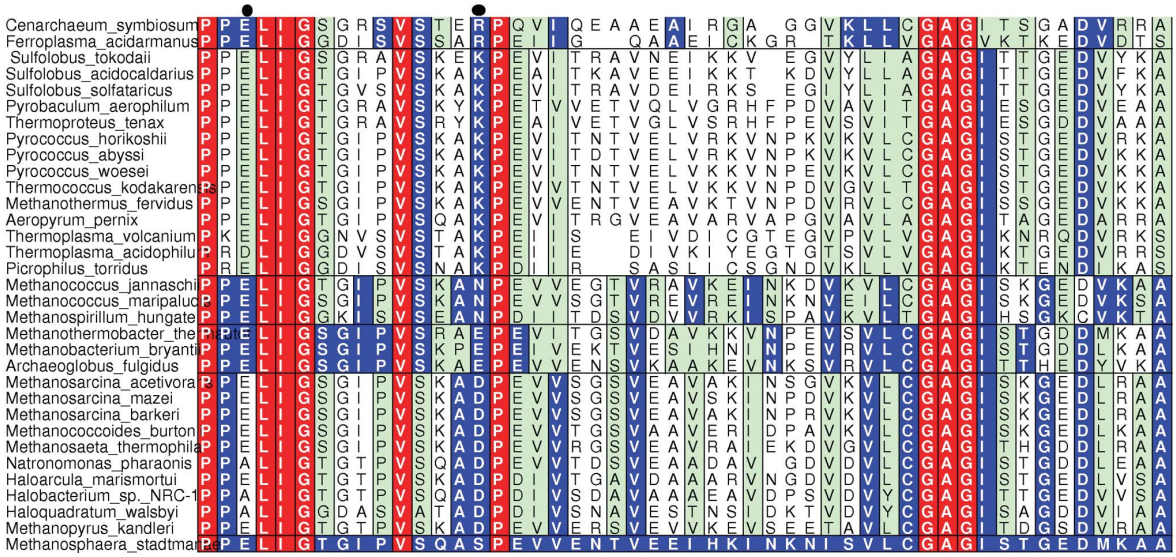


Figure 11
Alignment of the loop 6 residues of archaeal TIMs. A representation of the sequence alignment of all archaeal TIMs from the loop 6 to loop 7 residues is shown. The figure was prepared using the *AMAS* server (Livingstone & Barton, 1993). The sequences are grouped according to the residue present (R, K, N, E, D or S) at the position corresponding to Lys159 (PwTIM numbering). The positions of the pair of the residues involved in the salt bridge are highlighted by a black circle. The residues are shaded according to their conservation: red for identity along the whole set of sequences, blue for identity within the group and green for similarity within the group.

means of multiple interactions with neighbouring residues. Interestingly, the only instances of loop 6 disorder in PfTIM are reported for the three crystal structures of the W168F mutant (Eaazhisai *et al.*, 2004). Fluorescence studies in solution also confirm the absence of local motion involving Trp168 in the W11F single-tryptophan PfTIM mutant (Pattanaik *et al.*, 2003). A second type of loop 6 structure is observed in the archaebacterial enzymes PwTIM and TtTIM. The archaeal sequences are characterized by the N-terminal hinge segment PPE. Glu147 (MjTIM) is structurally equivalent to Trp168 of dimeric TIMs. In the PwTIM structure, Glu147 forms a strong salt bridge with Lys159. Importantly, this lysine is mutated to asparagine in the case of MjTIM, resulting in loss of the salt bridge. This presumably leads to enhanced loop flexibility and consequent disorder of the loop in the crystal structure of MjTIM. The observed enzymatic activity of MjTIM suggests that the absence of internal structuring within loop 6 does not affect enzyme function. In a substrate/ligand-bound form, loop 6 will presumably adopt an ordered structure facilitated by interactions with the ligand. A phylogenetic analysis of TIM sequences from archaebacteria (Fig. 11) shows that a basic lysine or arginine residue is conserved at this position in most cases, except for a cluster of sequences including some methanogenic and halophilic archaea, which have an Asn, Asp or Glu. *Methanosphaera stadtmanae* TIM is an exception, as it has a serine at this position.

In conclusion, the MjTIM structure reported here provides an interesting example of a tetrameric TIM from an archaebacteria and a valuable addition to tetrameric archaeal TIM structures. Together with the structures of *P. woesei* TIM and *T. tenax* TIM, the present results allow a detailed comparison of the dimeric and tetrameric TIMs. Tetramerization in archaeal TIM structures is facilitated by the shortening of the sequence and also by differences in the surface charge at the vicinity of the tetramer interfaces. Oligomerization is likely to be a major contributor to the thermostability of archaeal TIMs. In addition, more compact packing of the protein core involving both inner barrel and barrel–helix interactions appear to be contributory factors. The active-site loop 6 region of the archaeal TIMs reveals unique features compared with the other TIMs. The absence of a structured loop 6 in the unliganded TIM structure from *Methanocaldococcus jannaschii* can be attributed to the missing salt bridge that holds the hinge regions of the loop in the other tetrameric TIM structures solved to date. The role of the salt bridge in maintaining a structured loop 6 can be confirmed by further mutational analysis of the residues involved.

We thank Soumya S. Ray and S. K. Burley for the plasmid containing the MjTIM gene. We thank Dr Lokanath and the beamline staff at BL44XU for assistance during data collection. PG and MRNM thank Professor T. Tsukihara, Institute of Protein Research, Osaka University and Professor N. Kunishima, RIKEN SPring-8 Center, Harima Institute for the travel grant to Japan. SA thanks the Indian Council for Agricultural Research for a postdoctoral fellowship and leave of absence from Central Plantation Crops Research Institute, Kasargode, Kerala. This research was supported by program grants from Department of Biotechnology (DBT) and Department of Science and Technology (DST), Government of India. The mass-spectroscopic facility was supported under the Proteomics program of the DBT. This work was performed under the International Collaborative Research Program of

the Institute of Protein Research, Osaka University. PG and MB acknowledge the Council for Scientific and Industrial Research, Government of India for Senior Research Fellowships.

References

Barends, T. R., de Jong, R. M., van Straaten, K. E., Thunnissen, A. M. & Dijkstra, B. W. (2005). *Acta Cryst.* D**61**, 613–621.

Beaucamp, N., Hofmann, A., Kellerer, B. & Jaenicke, R. (1997). *Protein Sci.* **6**, 2159–2165.

Cohen, G. E. (1997). *J. Appl. Cryst.* **30**, 1160–1161.

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

DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* http://www.pymol.org.

Delboni, L. F., Mande, S. C., Rentier-Delrue, F., Mainfroid, V., Turley, S., Vellieux, F. M., Martial, J. A. & Hol, W. G. (1995). *Protein Sci.* **4**, 2594–2604.

Eaazhisai, K., Balaram, H., Balaram, P. & Murthy, M. R. (2004). *J. Mol. Biol.* **343**, 671–684.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Gokhale, R. S., Ray, S. S., Balaram, H. & Balaram, P. (1999). *Biochemistry*, **38**, 423–431.

Gonzalez-Mondragon, E., Zubillaga, R. A., Saavedra, E., Chanez-Cardenas, M. E., Perez-Montfort, R. & Hernandez-Arana, A. (2004). *Biochemistry*, **43**, 3255–3263.

Gopal, B., Ray, S. S., Gokhale, R. S., Balaram, H., Murthy, M. R. & Balaram, P. (1999). *Biochemistry*, **38**, 478–486.

Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. (1999). *Bioinformatics*, **15**, 305–308.

Gunasekaran, K., Ramakrishnan, C. & Balaram, P. (1996). *J. Mol. Biol.* **264**, 191–198.

Hubbard, S. J. & Thornton, J. M. (1996). *NACCESS*, version 2.1.1. Department of Biochemistry and Molecular Biology, University College London.

Joseph, D., Petsko, G. A. & Karplus, M. (1990). *Science*, **249**, 1425–1428.

Konagurthu, A. S., Whisstock, J. C., Stuckey, P. J. & Lesk, A. M. (2006). *Proteins*, **15**, 559–574.

Krissinel, E. V. & Henrick, K. (2005). *CompLife 2005*, edited by M. R. Berthold, R. Glen, K. Diederichs, O. Kohlbacher & I. Fischer, pp. 67–78. Berlin: Springer-Verlag.

Laskowski, R. A. (2003). *Methods Biochem. Anal.* **44**, 273–303.

Livingstone, C. D. & Barton, G. J. (1993). *Comput. Appl. Biosci.* **9**, 745–756.

Maes, D., Zeelen, J. P., Thanki, N., Beaucamp, N., Alvarez, M., Thi, M. H., Backmann, J., Martial, J. A., Wyns, L., Jaenicke, R. & Wierenga, R. K. (1999). *Proteins*, **37**, 441–453.

Maithal, K., Ravindra, G., Balaram, H. & Balaram, P. (2002). *J. Biol. Chem.* **277**, 25106–25114.

Maithal, K., Ravindra, G., Nagaraj, G., Singh, S. K., Balaram, H. & Balaram, P. (2002). *Protein Eng.* **15**, 575–584.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

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

Oesper, P. & Meyerhof, O. (1950). *Arch. Biochem.* **27**, 223–233.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Parthasarathy, S., Balaram, H., Balaram, P. & Murthy, M. R. (2002). *Acta Cryst.* D**58**, 1992–2000.

Parthasarathy, S., Ravindra, G., Balaram, H., Balaram, P. & Murthy, M. R. (2002). *Biochemistry*, **41**, 13178–13188.

Pattanaik, P., Ravindra, G., Sengupta, C., Maithal, K., Balaram, P. & Balaram, H. (2003). *Eur. J. Biochem.* **270**, 745–756.

Plaut, B. & Knowles, J. R. (1972). *Biochem. J.* **129**, 311–320.

Ranie, J., Kumar, V. P. & Balaram, H. (1993). *Mol. Biochem. Parasitol.* **61**, 159–169.

Ray, S. S., Balaram, H. & Balaram, P. (1999). *Chem. Biol.* **6**, 625–637.

Robinson-Rechavi, M., Alibes, A. & Godzik, A. (2006). *J. Mol. Biol.* **356**, 547–557.

Rosendal, K. R., Sinning, I. & Wild, K. (2004). *Acta Cryst.* D**60**, 140–143.

Stites, W. E., Meeker, A. K. & Shortle, D. (1994). *J. Mol. Biol.* **235**, 27–32.

Sun, J. & Sampson, N. S. (1998). *Protein Sci.* **7**, 1495–1505.

Sun, J. & Sampson, N. S. (1999). *Biochemistry*, **38**, 11474–11481.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.

Walden, H., Bell, G. S., Russell, R. J., Siebers, B., Hensel, R. & Taylor, G. L. (2001). *J. Mol. Biol.* **306**, 745–757.

Walden, H., Taylor, G. L., Lorentzen, E., Pohl, E., Lilie, H., Schramm, A., Knura, T., Stubbe, K., Tjaden, B. & Hensel, R. (2004). *J. Mol. Biol.* **342**, 861–875.

Williams, J. C., Zeelen, J. P., Neubauer, G., Vriend, G., Backmann, J., Michels, P. A., Lambeir, A. M. & Wierenga, R. K. (1999). *Protein Eng.* **12**, 243–250.

Yeates, T. O. (1997). *Methods Enzymol.* **276**, 344–358.