Acta Crystallographica Section D
**Biological Crystallography**
ISSN 0907-4449

# Structural and functional characterization of a noncanonical nucleoside triphosphate pyrophosphatase from *Thermotoga maritima*

**Khaldeyah Awwad,[a]‡ Anna Desai,[a]‡ Clyde Smith[b] and Monika Sommerhalter[a]***

[a]Chemistry and Biochemistry, California State University East Bay, 25800 Carlos Bee Boulevard, Hayward, CA 94542, USA, and [b]Stanford Synchrotron Radiation Lightsource, 2575 Sand Hill Road, Menlo Park, CA 94025, USA

‡ These authors made equal contributions.

Correspondence e-mail: monika.sommerhalter@csueastbay.edu

Received 26 July 2012
Accepted 29 October 2012

**PDB Reference:** TM0159, 3s86

The hyperthermophilic bacterium *Thermotoga maritima* has a noncanonical nucleoside triphosphatase that catalyzes the conversion of inosine triphosphate (ITP), deoxyinosine triphosphate (dITP) and xanthosine triphosphate (XTP) into inosine monophosphate (IMP), deoxyinosine monophosphate (IMP) and xanthosine monophosphate (XMP), respectively. The $k_{cat}/K_m$ values determined at 323 and 353 K fall between $1.31 \times 10^4$ and $7.80 \times 10^4$ $M^{-1}$ $s^{-1}$. ITP and dITP are slightly preferred over XTP. Activity towards canonical nucleoside triphosphates (ATP and GTP) was not detected. The enzyme has an absolute requirement for $Mg^{2+}$ as a cofactor and has a preference for alkaline conditions. A protein X-ray structure of the enzyme with bound IMP was obtained at 2.15 Å resolution. The active site houses a well conserved network of residues that are critical for substrate recognition and catalysis. The crystal structure shows a tetramer with two possible dimer interfaces. One of these interfaces strongly resembles the dimer interface that is found in the structures of other noncanonical nucleoside pyrophosphatases from human (human ITPase) and archaea (Mj0226 and PhNTPase).

## 1. Introduction

Noncanonical nucleoside triphosphate pyrophosphatases (noncanonical NTPases; EC 3.6.1.19) hydrolyze a phosphoanhydride bond of inosine triphosphate (ITP), deoxyinosine triphosphate (dITP) or xanthosine triphosphate (XTP) to release pyrophosphate and the corresponding nucleoside monophosphate IMP, dIMP or XMP, respectively (Galperin *et al.*, 2006; Stenmark *et al.*, 2007; Chung *et al.*, 2001). Accumulation of noncanonical nucleoside triphosphates (NTPs) is undesirable because they might become incorporated into RNA or DNA (Auer *et al.*, 1996; Spee *et al.*, 1993). Also, ITP and XTP might interfere with ATP- and GTP-requiring processes (Klinker & Seifert, 1997; Weber & Senior, 2001). ITP can be generated *via* phosphorylation from IMP, which is a direct metabolic precursor for the biosynthesis of AMP and GMP (Vanderheiden, 1979). Another source of noncanonical nucleotides is oxidative damage. Oxidative deamination results in the conversion of adenosine to inosine and of guanosine to xanthosine (Shapiro & Pohl, 1968). Since noncanonical NTPases can readily discriminate between the erroneous nucleotides (d)ITP and XTP and their canonical counterparts (d)ATP and GTP, they help to maintain an intact pool of DNA and RNA precursor molecules and thus have been dubbed 'house-cleaning enzymes' (Galperin *et al.*, 2006).

Noncanonical NTPases have been found in diverse species that belong to all three domains of life: bacteria (Savchenko *et al.*, 2007; Burgis & Cunningham, 2007), archaea (Hwang *et al.*, 1999; Chung *et al.*, 2001; Lokanath *et al.*, 2008) and eukarya

(Stenmark *et al.*, 2007; Lin *et al.*, 2001). Noncanonical NTPases have also been found in hyperthermophilic bacteria, including *Thermotoga maritima* (Galperin *et al.*, 2006). *T. maritima* was selected as a target organism by the Joint Center for Structural Genomics (JCSG), with the goal of covering all possible protein folds within the proteome of this bacterium (Lesley *et al.*, 2002). The structure of the *T. maritima* protein TM0159 was determined in 2004 (PDB entry 1vp2; Joint Center for Structural Genomics, unpublished work). TM0159 has a close structural similarity to other previously characterized non-canonical NTPases, including RdgB from *Escherichia coli* (Savchenko *et al.*, 2007; Burgis & Cunningham, 2007), Mj0226 from *Methanococcus jannaschii* (Hwang *et al.*, 1999; Chung *et al.*, 2001), PhNTPase from *Pyrococcus horikoshii* (Lokanath *et al.*, 2008) and human ITPase (Stenmark *et al.*, 2007; Porta *et al.*, 2006).

Here, we provide experimental evidence for the proposed functional role of TM0159, including enzyme-kinetic data on the hydrolysis of XTP, ITP and dITP. We also present a crystal structure of TM0159 with bound IMP at 2.15 Å resolution and compare it with structures of unliganded TM0159 and other noncanonical NTPases.

## 2. Materials and methods

### 2.1. Expression and purification of TM0159

A glycerol stock of *E. coli* DL41 cells with a pMH1 plasmid harboring the TM0159 gene with an N-terminal polyhistidine tag (MGSDKIHHHHHH) was a generous gift from the Joint Center for Structural Genomics (JCSG). Protein expression was performed in 1 l cultures containing Miller's Luria Broth with 100 µg ml$^{-1}$ carbenicillin. When the culture reached an $OD_{600}$ value of 0.8 ± 0.1, protein expression was induced by adding 5 ml of 20%(*w*/*v*) L-arabinose. After 3 h, the bacteria were harvested by centrifugation at 3400*g* and 277 K for 30 min. The cell pellet was resuspended in lysis buffer [50 m*M* Tris–HCl pH 7.9, 50 m*M* NaCl, 1 m*M* $MgCl_2$, 0.25 m*M* tris(2-carboxyethyl)phosphine (TCEP)] and stored at 253 K. After thawing the suspension, the bacteria were lysed by sonication on ice and the cell debris was removed by centrifugation at 32 000*g* and 277 K for 60 min. The soluble fraction was applied onto a histidine-tag affinity column loaded with nickel (HisLink Protein Purification Resin, Promega). The column was equilibrated with buffer *A* (50 m*M* Tris–HCl pH 7.9, 300 m*M* NaCl, 5% glycerol, 0.25 m*M* TCEP). The column was washed with several column volumes of buffer *A* before a linear gradient with increasing buffer *B* (500 m*M* imidazole pH 8.0, 300 m*M* NaCl, 5% glycerol, 0.25 m*M* TCEP) was applied. Protein elution was monitored *via* the absorbance at 280 nm and all fractions collected during the buffer gradient were analyzed *via* SDS–PAGE. Fractions that contained TM0159 with high purity were pooled and concentrated using centrifugal ultrafiltration devices (Amicon Ultra with a 10 000 Da cutoff). A buffer exchange was performed using prepacked desalting columns (EconoPac 10 DG, Bio-Rad) equilibrated with buffer *C* (10 m*M* Tris–HCl pH 7.9, 50 m*M* NaCl, 0.25 m*M* TCEP). Protein concentration was determined using a bicinchoninic acid protein-concentration assay (Pierce BCA Protein Assay kit). For enzyme-activity measurements, small aliquots with a final protein concentration of 1 mg ml$^{-1}$ were stored at 193 K after flash-cooling in liquid nitrogen. For crystallization trials, the protein was further concentrated to 8 or 13 mg ml$^{-1}$ using a centrifugal ultrafiltration device (Amicon Ultra with a 10 000 Da cutoff) and stored in small aliquots at 193 K after flash-cooling in liquid nitrogen. A 1 l culture corresponding to approximately 5 g of wet cell pellet typically yielded 13 mg purified TM0159.

### 2.2. Enzyme assays

The colorimetric pyrophosphatase assay employed was very similar to the method used by Zheng *et al.* (2005). A typical reaction mixture (500 µl total volume) was set up with 50 m*M* Tris–HCl pH 9.0, 10 m*M* $MgCl_2$ and nucleoside triphosphates in the concentration range from 0.005 to 10 m*M*. Next, 5 µl of 0.1 mg ml$^{-1}$ yeast pyrophosphatase (Roche) was added. A 120 µl negative-control aliquot was transferred to a separate microcentrifuge tube. Finally, 5 µl of 0.1 mg ml$^{-1}$ TM0159 was added to the reaction mixture. The reaction was allowed to progress for 10 min in a 323 K water bath. The reaction was stopped by removing 100 µl aliquots from the reaction mixture into a cuvette containing 900 µl deionized water (Milli-Q, Millipore) and 1000 µl of a colorimetric reagent. The colorimetric reagent was freshly prepared by mixing 10%(*w*/*v*) ascorbic acid, 2.5%(*w*/*v*) ammonium molybdate, 6 *N* sulfuric acid and deionized water in a 1:1:1:2 ratio. Next, the cuvettes were incubated in a 323 K water bath for 20 min. The cuvettes were removed from the water bath and the absorbance at 820 nm was recorded. A phosphate standard curve in a concentration range from 10 to 70 µ*M* was prepared from a 200 µ*M* commercially available phosphate standard solution. The detection limit of our assay system is 0.0002 U or 0.0002 µmol min$^{-1}$. All measurements were performed in triplicate. All enzyme-kinetic data were processed with the *EnzFitter* software (Biosoft).

To test the metal-ion requirement, the concentration of $MgCl_2$ was varied between 0 and 100 m*M*, or other salts ($MnCl_2$, $MnSO_4$, $CoCl_2$, $CaCl_2$, $ZnSO_4$, $CuSO_4$, $Na_2SO_4$, $Li_2SO_4$ or $K_2SO_4$) at concentrations of 10 and 100 m*M* were added instead. The pH was varied by using different 0.05 *M* buffers, including sodium acetate–acetic acid pH 3.14, sodium acetate–acetic acid pH 4.20, MES–NaOH pH 5.86, Tris–HCl pH 6.10, borax–HCl pH 8.02, CHES–NaOH pH 9.00, Tris–HCl pH 9.00, CAPS–NaOH pH 9.65, $NaHCO_3$–NaOH pH 11.04, KCl–NaOH pH 11.95 and KCl–NaOH pH 12.53.

For the activity assays performed at 353 K the method described above had to be adjusted. The yeast pyrophosphatase enzyme was not stable enough, and the second step of the coupled assay became the rate-limiting step. The TM0159-catalyzed reaction was carried out for 10 min in a 353 K water bath without yeast pyrophosphatase. After a rapid cooling step (30–45 s in an ice–water bath) the mixture was placed in a 310 K water bath. Yeast pyrophosphatase was added and,

10 min later, the colorimetric reagent was added. At 310 K TM0159 had a very low activity, and less than 0.2 μ*M* pyrophosphate was released per minute.

### 2.3. Crystallization and X-ray diffraction data collection

At first, we reproduced the original crystallization condition identified by the JCSG for the apo form of TM0159 (PDB entry 1vp2). This condition was composed of 1.6 *M* ammonium sulfate, 0.1 *M* NaCl, 0.1 *M* HEPES pH 7.5. The vapor-diffusion method was used in sitting-drop trays stored at room temperature. Crystals appeared readily after several days. Soaking attempts with nucleotides or nucleotide analogues resulted in visible crystal damage and poor X-ray diffraction quality.

To obtain a structure of TM0159 with a bound nucleotide, TM0159 was incubated with different nucleotides (IMP, ITP, XMP or XTP) at a final concentration of 10 or 20 m*M*. A sparse-matrix screen (The Classics II Suite from Qiagen) was set up at room temperature with the four solutions in sitting-drop clover-leaf trays (Emerald BioSystems). The most promising condition consisted of 200 m*M* L-proline, 100 m*M* Tris–HCl pH 7.4, 16% PEG 3350 with a solution consisting of 10 m*M* IMP and TM0159. This condition was reproduced with home-made solutions in 24-well sitting-drop trays. Crystals typically grew to full size within one week. For cryoprotection, the well solution was mixed with 50%(*v*/*v*) ethylene glycol. Crystals were flash-cooled in liquid nitrogen and loaded into a 96-port cassette for use with the SSRL automated mounting robot (Cohen *et al.*, 2002). X-ray diffraction data were collected on beamline 9-2. Data indexing and integration were performed with *XDS* (Kabsch, 2010). The crystals belonged to space group $P2_12_12_1$, with unit-cell parameters $a$ = 72.4, $b$ = 75.99, $c$ = 157.94 Å.

### 2.4. Structure determination and refinement

The structure of TM0159 with bound IMP was solved by molecular replacement with the program *MOLREP* (Vagin & Teplyakov, 1997, 2010) using the apo TM0159 structure solved by the JCSG (PDB entry 1vp2) as a template. Refinement was carried out with the program *REFMAC* (Murshudov *et al.*, 1997, 1999, 2011). The $R_{free}$ value was calculated using a randomly selected subset of the data (5%). *B* factors for individual atoms were refined isotropically. Structural restraints for the ligand IMP and for sulfate were obtained from the *CCP*4 library (Winn *et al.*, 2011). Map inspection and further model building was performed with *Coot* (Emsley & Cowtan, 2004). Figs. 2, 3, 4 and 6 were prepared with *PyMOL* (DeLano, 2002). The program *PISA* was employed to analyze the dimer interfaces of TM0159 and to compare them with other proteins (Krissinel & Henrick, 2007). Structure-based sequence alignments were performed with the program *VAST* (Gibrat *et al.*, 1996). The structure of IMP-liganded TM0159 has been deposited in the Protein Data Bank as entry 3s86.

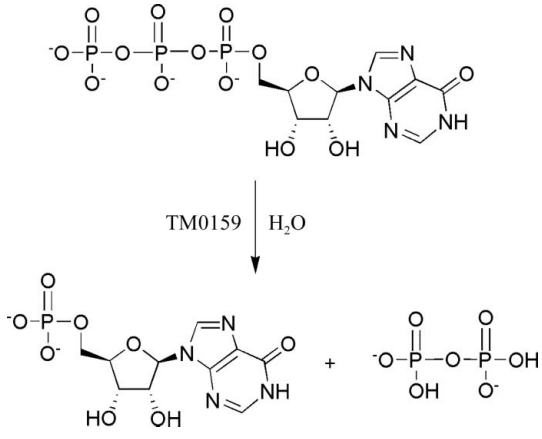


**Figure 1**
Scheme of the reaction catalyzed by TM0159 with the substrate ITP.

**Table 1**
Enzyme-kinetic constants for TM0159 determined at pH 9.0.

| Substrate | $K_m$ (m*M*) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m$ ($M^{-1}$ $s^{-1}$) | Temperature (K) |
|---|---|---|---|---|
| XTP | 0.11 ± 0.03 | 2.70 ± 0.25 | $2.45 \times 10^4$ | 323 |
| ITP | 0.06 ± 0.01 | 3.90 ± 0.35 | $6.50 \times 10^4$ | 323 |
| dITP | 0.07 ± 0.02 | 5.46 ± 0.35 | $7.80 \times 10^4$ | 323 |
| XTP | 1.69 ± 0.67 | 22.08 ± 6.02 | $1.31 \times 10^4$ | 353 |
| ITP | 0.51 ± 0.28 | 13.87 ± 3.63 | $2.72 \times 10^4$ | 353 |
| dITP | 0.45 ± 0.17 | 18.68 ± 2.36 | $4.15 \times 10^4$ | 353 |

## 3. Results

### 3.1. Enzymatic activity of TM0159

TM0159 readily converts noncanonical nucleoside triphosphates [(d)ITP or XTP] into their corresponding nucleoside monophosphates [(d)IMP or XMP] and releases pyrophosphate (see Fig. 1). Canonical nucleoside triphosphates (ATP and GTP) are poor substrates for TM0159, and no significant activity was detected. The release of one pyrophosphate, instead of one or two individual phosphate molecules, was corroborated by the colorimetric assay. In the absence of inorganic pyrophosphatase, no phosphate production was detected. Further assays based on thin-layer chromatography and high-performance liquid chromatography also agree with the proposed function of TM0159 as a noncanonical NTPase (Supplementary Figs. S1 and S2[1]).

The activity of TM0159 with the substrates ITP and XTP was determined at different pH values. TM0159 works best at alkaline pH values of pH 9 for ITP and pH 9.65 for XTP (Supplementary Fig. S3). Since *T. maritima* is a hyperthermophile, most assays were conducted at 313 K or above. No significant activity was detected at room temperature. TM0159 has an absolute requirement for $Mg^{2+}$. At 10 m*M* $Mg^{2+}$ concentration, TM0159 reaches a plateau of optimum activity (Supplementary Fig. S4). Other divalent and monovalent cations barely support the catalytic activity of TM0159. Samples containing up to 100 m*M* $Mn^{2+}$, $Zn^{2+}$, $Co^{2+}$, $Cu^{2+}$, $Na^+$, $K^+$ or $Li^+$ showed less than 10% of the activity measured at 10 m*M* $Mg^{2+}$ (Supplementary Fig. S5).

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: CB5016). Services for accessing this material are described at the back of the journal.

**Table 2**
X-ray data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

| | |
|---|---|
| Data-collection statistics | |
| Resolution limits (Å) | 50.0–2.15 |
| Wavelength (Å) | 0.97946 |
| No. of unique reflections | 48090 (3160) |
| Completeness (%) | 99.8 (100) |
| Multiplicity | 7.17 (7.39) |
| Wilson *B* factor (Å$^2$) | 50.6 (47.9) |
| $\langle I/\sigma(I)\rangle$ | 18.87 (3.05) |
| $R_{meas}$ (%) | 6.2 (73.2) |
| Refinement statistics | |
| Resolution limits (Å) | 28.95–2.15 (2.20–2.15) |
| $R_{free}$ (test set) (%) | 0.278 (0.413) |
| *R* factor (working set) (%) | 0.220 (0.305) |
| No. of protein atoms | 6040 |
| No. of water molecules | 196 |
| No. of heteroatoms | 112 |
| Average isotropic *B* values (Å$^2$) | |
| Overall | 46.9 |
| Protein atoms (Å$^2$) | 46.6 |
| Ligands (IMP and $SO_4^{2-}$) (Å$^2$) | 56.4 |
| Water (Å$^2$) | 55.2 |
| Root-mean-square deviations from ideal geometry | |
| Bond lengths (Å) | 0.021 |
| Bond angles (°) | 1.877 |
| General planes (Å) | 0.009 |

TM0159 was found to exhibit Michaelis–Menten kinetics. The Michaelis–Menten parameters determined at pH 9.0 are given in Table 1. For each substrate, the $k_{cat}/K_m$ values were slightly lower at 353 K compared with 323 K. The threefold to eightfold increases in $k_{cat}$ were accompanied by even larger (sixfold to 15-fold) increases in $K_m$. Notably, ITP and dITP have very similar $K_m$ values. Based on the $K_m$ and $k_{cat}/K_m$ values, ITP and dITP are slightly preferred over XTP.

### 3.2. Crystal structure of TM0159 with bound IMP

The structure of TM0159 with bound IMP was solved by molecular replacement at 2.15 Å resolution (Table 2, Fig. 2). The asymmetric unit contains four polypeptide chains. The root-mean-square deviation (r.m.s.d.) between the C$^\alpha$ atoms of the four chains in the asymmetric unit varies between 0.28 and 0.43 Å. TM0159 has 196 residues and an N-terminal tag. Electron density was visible for residues 3–191 but not for the N-terminal tag. The stereochemical quality of the model is excellent, with 91.9% and 8.1% of all residues (except for glycine and proline) in the core and allowed regions of the Ramachandran plot, respectively (Laskowski *et al.*, 1993). The overall $\alpha/\beta$-fold of each TM0159 monomer houses a central long twisted $\beta$-sheet with three antiparallel $\beta$-strands ($\beta$3, $\beta$4 and $\beta$5). One face of this $\beta$-sheet is covered by two lobes, whereas the other face is more exposed and participates in the formation of an interface with another monomer. The first lobe contains two short $\beta$-strands ($\beta$1 and $\beta$2) arranged in parallel, three $\alpha$-helices ($\alpha$1, $\alpha$2 and $\alpha$9) and several long loops. The short N-terminal $\beta$-strand ($\beta$1) is directly adjacent to the long central $\beta$-strand $\beta$3, thus creating a twisted five-stranded $\beta$-sheet. The second lobe is composed of several short $\alpha$-helices ($\alpha$3–$\alpha$8), one short $\beta$-strand ($\beta$6) and several loops. The IMP binding site is located between the two lobes.

The four polypeptide chains in the asymmetric unit are organized in two dimers that stack upon each other in a larger tetrameric assembly. However, gel-filtration experiments (Supplementary Fig. S6) suggest that TM0159 is a dimer in solution. The first interface (interface 1; Fig. 3*a*) has an extended $\beta$-sheet across both monomers formed by the close contact between the central $\beta$5 strands. Further contacts result from the packing of residues located on helices $\alpha$2 and $\alpha$9, $\beta$-strands $\beta$4 and $\beta$5 and the short loop between $\beta$4 and $\beta$5. Nine hydrophobic and 12 polar side-chain residues are at the interface. A total of eight hydrogen-bonding interactions and seven salt bridges involving the pairs Glu43/Lys58 and Glu47/Lys58 are present. The second dimer copy in the asymmetric unit also displays one salt bridge between Glu124 and Lys182. The total buried surface area is approximately 823 Å$^2$ per monomer. The solvation free-energy gain $\Delta^iG$ upon formation of this interface is −26.0 kJ mol$^{-1}$. A second interface can also be identified (interface 2; Fig. 3*b*). This interface covers only 655 Å$^2$ per monomer, with a $\Delta^iG$ value of −30.6 kJ mol$^{-1}$. Residues that are mostly located on helices $\alpha$3–$\alpha$5 and their connecting loops are involved in this interface. The other lobe of the structure only contributes two residues (Val39 and Asp41) on the loop between $\beta$2 and $\alpha$2. Although more polar residues (11) than hydrophobic residues (nine) are positioned at the interface, a large central hydrophobic contact area is formed by Leu76, Phe79, Met83, Phe87, Met88 and Met103. Up to six hydrogen bonds and one salt bridge between Glu95

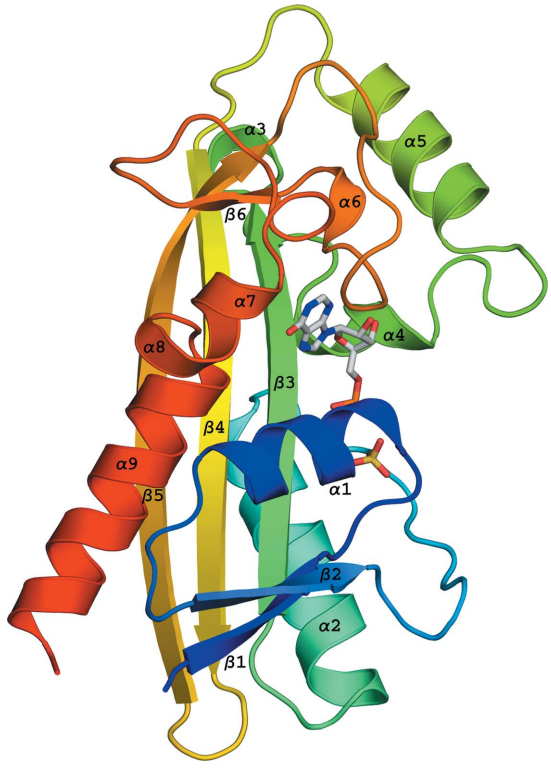


**Figure 2**
Structure of a TM0159 monomer (chain *B*). The protein is shown as a rainbow-colored ribbon (N-terminus blue and C-terminus red) with labels for the secondary-structure elements. The ligands IMP and $SO_4^{2-}$ are shown in stick representation.

and Lys107 are also present. Fig. 3(*c*) displays the full tetrameric assembly with both interfaces.

Notably, the unliganded structure of TM0159 (PDB entry 1vp2) exhibits the same possible tetrameric assembly composed of two stacked dimers, and the monomer structures of IMP-liganded and unliganded TM0159 align very well with one another. The r.m.s.d. for 189 matching $C^{\alpha}$ atoms is only 0.5 Å. A small change occurs in the position of Thr11, with a shift of 2 Å for the $C^{\alpha}$ atoms. The asymmetric unit of the crystal form containing the unliganded TM0159 houses a dimer that corresponds to the dimer displayed in Fig. 3(*b*) (interface 2), with a buried surface area of 626 Å$^2$ per monomer. Notably, this dimer has one salt bridge between Lys102 and Glu105 instead of between Glu95 and Lys107. The generation of symmetry-related molecules reveals a dimer that corresponds to that displayed in Fig. 3(*a*) (interface 1). The buried surface area of this dimer amounts to approximately 791 Å$^2$ per monomer. The solvation free-energy gain $\Delta^{i}G$ is −28.5 kJ mol$^{-1}$ for the formation of interface 1 and −26.4 kJ mol$^{-1}$ for the formation of interface 2.

### 3.3. IMP binding site

Fig. 4 shows IMP bound to TM0159. Several residues interact with the purine base of IMP: Asp151, Lys170, His175 and Arg176. The carboxylate group of Asp151 interacts with the N1 atom of IMP. The $N^{\varepsilon 2}$ atom of His175 interacts with the N7 atom of IMP. The keto group at position 6 of the IMP purine ring forms hydrogen bonds to Lys170 and Arg176. The positions of these two residues are stabilized by the carboxylate group of Glu18, which forms a bridge between Lys170 and Arg176. This interaction is part of a hydrogen-bonding network that helps to position the side chains of the residues in the IMP binding site and connects Lys15 to Asp67, Arg176, Glu18, Lys170 and finally Asp151. Two phenylalanine residues (Phe113 and Phe148) are positioned above and below the six-membered ring of the purine base and are thereby able to form aromatic stacking interactions.

The two hydroxyl groups of the ribose moiety of IMP are both within hydrogen-bonding distance of Tyr93. The main-chain carbonyl group of Ser84 forms a hydrogen bond to the 3′-hydroxyl group of the ribose. Interactions with the phosphate group of IMP involve the side chains of residues Lys15, Glu40 and Asp68 and the backbone carbonyl group of Ser69. Residual electron density is present close to the IMP nucleotide at coordination distance to the side chains of Lys15 and Lys52 and the amide N atom of Thr11. In the unliganded structure of TM0159 a sulfate group was modeled into this position for both chains of the crystal lattice. Our crystallization condition did not contain sulfate, but a sulfate or another anion with similar electron density would fit well into this residual density. We tentatively modeled a sulfate group that could derive from buffer contamination into this position. However, we cannot rule out that another anionic moiety of similar size and electron density (for example, phosphate) occupies this position.

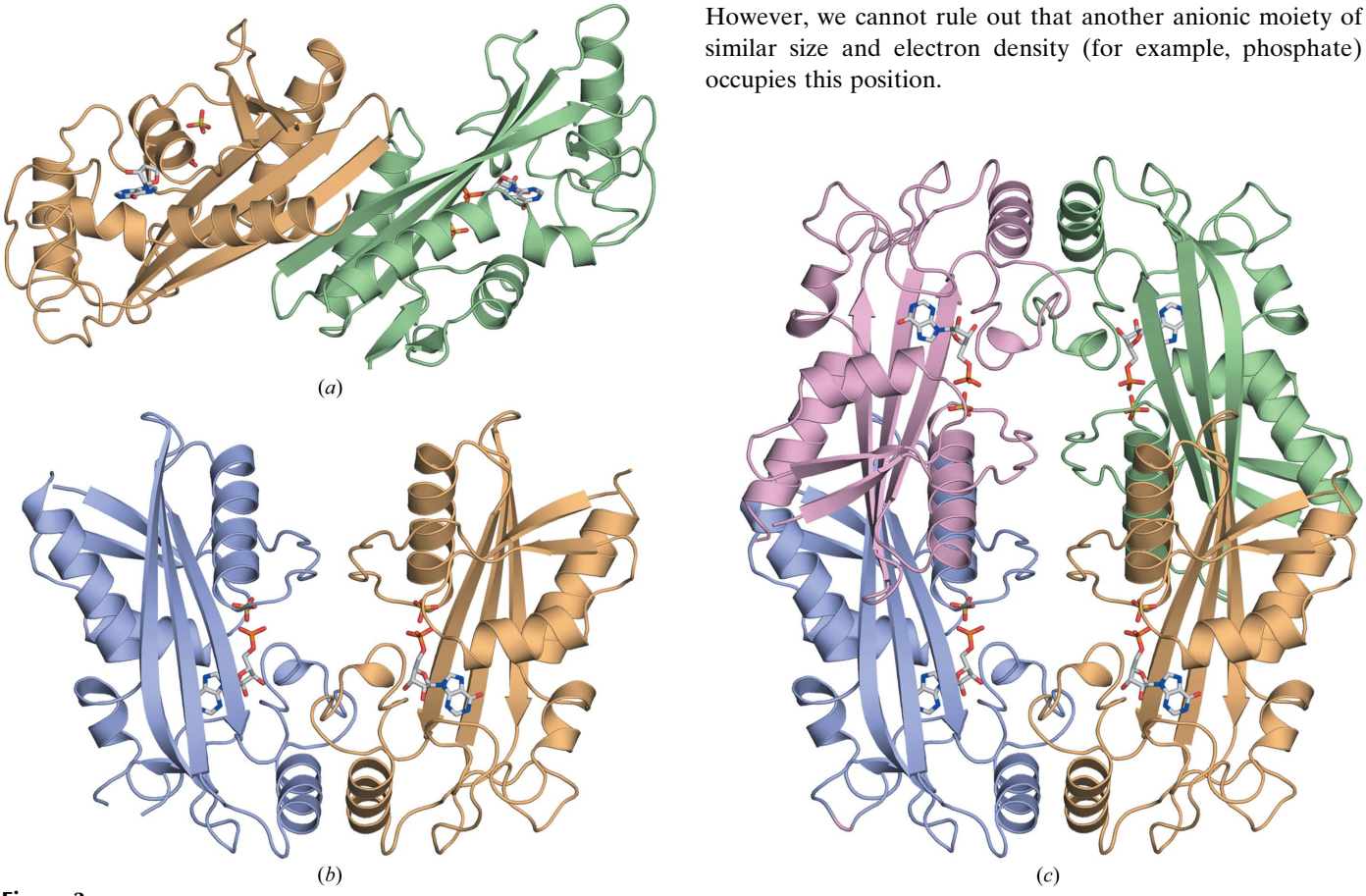


**Figure 3**
(*a*) Dimer with interface 1. (*b*) Dimer with interface 2. (*c*) Tetrameric assembly of TM0159 with bound IMP.

## 4. Discussion

### 4.1. Enzymatic activity of TM0159

TM0159 is a member of a group of enzymes called noncanonical NTPases. These enzymes cleave a phosphoanhydride bond of nucleoside triphosphates, such as ITP, dITP or XTP. The kinetic parameters determined for noncanonical NTPases cover a wide range: RdgB has similar $k_{cat}$ values to TM0159 (5.7–20 $s^{-1}$) but much lower $K_m$ values (5.6–23.3 $\mu M$) (Burgis & Cunningham, 2007; Savchenko *et al.*, 2007). Human ITPase and Mj0226 have similar $K_m$ values to TM0159 (0.03–0.57 m$M$) but much higher $k_{cat}$ values (71–640 $s^{-1}$) (Burgis & Cunningham, 2007; Lin *et al.*, 2001; Chung *et al.*, 2001; Stepchenkova *et al.*, 2009). Overall, this makes TM0159 less efficient than all other characterized noncanonical NTPases. The $k_{cat}/K_m$ values for TM0159 with its substrates (d)ITP and XTP are below $1.0 \times 10^5$ $M^{-1}$ $s^{-1}$, whereas the corresponding values for RdgB, Mj0226 and human ITPase fall between $2.0 \times 10^5$ and $3.4 \times 10^6$ $M^{-1}$ $s^{-1}$.

TM0159 shares a preference for alkaline pH values with other noncanonical NTPases (Savchenko *et al.*, 2007), but differs in its metal-cofactor requirement. In contrast to TM0159, other noncanonical NTPases perform very well with $Mn^{2+}$ (Savchenko *et al.*, 2007; Chung *et al.*, 2001). The two divalent cations $Mg^{2+}$ and $Mn^{2+}$ can often replace one another since they have similar ionic radii (0.65 Å for $Mg^{2+}$, 0.75 Å for $Mn^{2+}$) and both prefer sixfold coordination (Bock *et al.*, 1999). Another *T. maritima* enzyme, 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXR), showed activity with $Mg^{2+}$ but not with $Mn^{2+}$ or $Co^{2+}$ when the assay was conducted at 358 K (Takenoya *et al.*, 2010). At lower temperatures both $Mg^{2+}$ and $Mn^{2+}$ supported DXR activity. In agreement with this study, our results also support the claim that *T. maritima* mainly utilizes $Mg^{2+}$ *in vivo* (Takenoya *et al.*, 2010).

### 4.2. Oligomerization state of TM0159

Gel-filtration experiments (Supplementary Fig. S6) suggest that TM0159 is a dimer in solution. All other characterized noncanonical NTPases are also described to be dimers (Lokanath *et al.*, 2008; Savchenko *et al.*, 2007; Stenmark *et al.*, 2007; Hwang *et al.*, 1999). Thus, crystal structures of TM0159 might show two possible dimer interfaces, but only one of them would be physiologically relevant. The size of an interface is not a suitable measure to distinguish specific (biologically relevant) contacts from crystal contacts, as long as the total buried surface area exceeds 1000 $Å^2$ (corresponding to 500 $Å^2$ per monomer; Bahadur & Zacharias, 2008). Other criteria, including the presence of conserved residue clusters, also have to be considered (Guharoy & Chakrabarti, 2010). Interface 2 of TM0159 exhibits a strong similarity to the interfaces described for human ITPase, Mj0226 and PhNTPase (Supplementary Fig. S9 and Table S2). This supports the assignment of the smaller interface (interface 2) as the physiologically relevant interface. Notably, the crystal forms of these other noncanonical NTPases house only one possible dimeric assembly. In both crystal forms of TM0159 two dimers are stacked on top of each other, which distorts the twofold symmetry found in other noncanonical NTPase dimers. In these more symmetric dimers both lobes contribute to the interface, with several contacts resulting in large buried surface areas of roughly 1000 $Å^2$ per monomer. The solvation free-energy gain $\Delta^i G$ ranges from −44.8 to −72.4 kJ $mol^{-1}$ upon dimer formation. In the TM0159 dimer the C-terminal lobe still contributes several residues to form contacts, but the N-terminal lobe has only one main contact point (Asp41). The consequence is a smaller buried surface area with only 655 $Å^2$ per monomer for IMP-liganded TM0159 and 626 $Å^2$ per monomer for unliganded TM0159 and less exergonic $\Delta^i G$ values of −30.5 and −26.4 kJ $mol^{-1}$, respectively.

However, it is also conceivable that TM0159 adopts a tetrameric oligomerization state in solution. Notably, both crystal structures of TM0159 yield the same tetrameric assembly, which is predicted to be the most stable quaternary structure based on the free-energy values of assembly dissociation ($\Delta G^{diss}$) calculated using the program *PISA*. The $\Delta G^{diss}$ values for a tetramer were +28.5 kJ $mol^{-1}$ for IMP-liganded TM0159 and +13.4 kJ $mol^{-1}$ for unliganded TM0159. In contrast, the most stable dimer (IMP-liganded TM0159 with interface 1) yielded a calculated $\Delta G^{diss}$ value of only 4.6 kJ $mol^{-1}$. The gel-filtration experiments might be hampered by interactions between the protein and the gel-filtration matrix or an inability to detect weak tetrameric interactions. An example of an enzyme with various oligomerization states is 2-deoxyribose-5-phosphate aldolase (DERA; Lokanath *et al.*, 2004). The living temperature of the source organism was found to be correlated with the quaternary structure of DERA (Lokanath *et al.*, 2004).

RdgB has a unique dimer interface which was not found in any other characterized noncanonical NTPase. For unliganded and IMP-liganded RdgB, however, electron

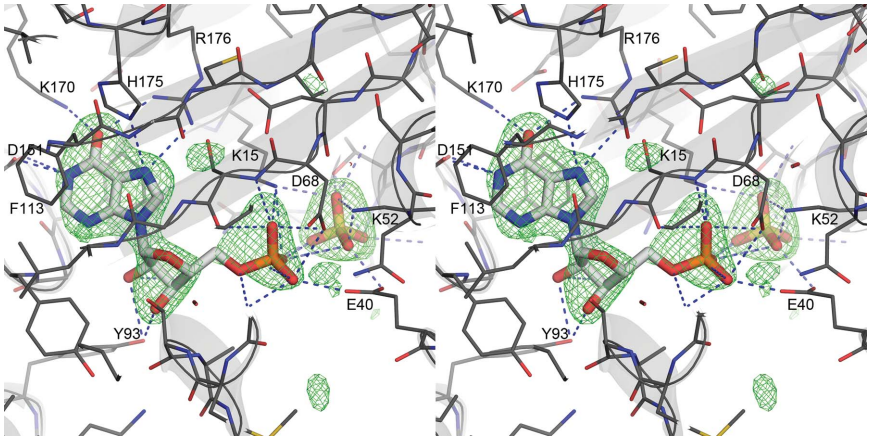


**Figure 4**
Stereoview of the IMP-binding site (chain *B*). The $F_o - F_c$ electron-density map colored in green and contoured at 3$\sigma$ was calculated without the IMP and sulfate anion. For clarity, a 16 Å clipping plane was chosen. Hydrogen-bonding interactions of IMP and the sulfate anion with atoms of the protein are displayed as blue dashed lines.

density for the N-terminal tag was visible, which contributed four residues to the interface in each case. It is conceivable that a construct without an N-terminal tag would form a different interface. Notably, a deviation from Michaelis–Menten kinetics caused by catalytic cooperativity was only described for RdgB (Savchenko *et al.*, 2007). The enzyme-kinetic analysis of RdgB revealed Hill coefficients of 1.7–2.3, indicating that the two dimer halves strongly support each other (Savchenko *et al.*, 2007).

### 4.3. Nucleotide-binding site

The nucleotide-binding site is remarkably similar in all characterized noncanonical NTPases (Stenmark *et al.*, 2007; Savchenko *et al.*, 2007; Lokanath *et al.*, 2008). Two aromatic side chains are stacked on top and below the purine base. A network of hydrogen bonds aids in the recognition and positioning of the noncanonical nucleotide substrate. These hydrogen bonds involve very well conserved residues: Glu18, Asp151, Lys170, His175 and Arg176 (TM0159 sequence numbering). The ability of TM0159 to discriminate between (d)ITP and (d)ATP is readily explained, since ATP has an amino group instead of a keto group at position 6 and therefore cannot participate in the hydrogen-bonding interactions with Lys170 and Arg176. GTP, on the other hand, has a keto group at position 6. GTP and XTP differ at position 2, at which XTP has another keto group but GTP has an amino group. A docking model of XTP and TM0159 showed that the 2-keto group of XTP can form a hydrogen bond to the backbone amide of residue Asp151 (Galperin *et al.*, 2006). An amino group at position 2 cannot be accommodated in the same manner.

A unique feature of the IMP-binding pocket of TM0159 is the contact between the ribose ring of IMP and the residues located on this loop. Tyr93 forms a hydrogen bond to the 2-hydroxyl group of the ribose moiety. Only RdgB has a loop of similar length, but none of the side chains in this loop are within hydrogen-bonding distance of the hydroxyl groups of the ribose moiety. Since TM0159 has virtually identical $K_m$ values for dITP and ITP, we conclude that this single interaction does not confer any selectivity for the discrimination of dNTP and NTP nucleotides.

### 4.4. Nucleotide hydrolysis

All crystal structures of noncanonical NTPases, with the exception of ITP-liganded PhNTPase (Lokanath *et al.*, 2008), were obtained at neutral or acidic pH values. However, TM0159 and all other characterized noncanonical NTPases work best at alkaline pH values (Savchenko *et al.*, 2007; Chung *et al.*, 2001; Lin *et al.*, 2001). This preference for alkaline pH values might be linked to improved substrate binding and/or an enhancement of the hydrolysis of the phosphoanhydride bond. At pH 9 and above, both (d)ITP and XTP are mostly present in their anionic forms (Fig. 5). XTP has a remarkably low p$K_a$ of 5.4 for the deprotonation of its purine moiety (Sigel *et al.*, 2009). The corresponding p$K_a$ value for ITP (p$K_a$ = 9.0) is much higher (Naumann *et al.*, 1974). Even at neutral pH values, XTP (but not ITP) should be present in the anionic form depicted on the right in Fig. 5. More negative charge on the 6-keto group would promote the interaction with Lys170 and Arg176 and enhance substrate binding. However, a hydrogen bond between Asp151 and the purine ring is only possible if a proton is still attached to N1. At alkaline pH values this interaction should turn into an electrostatic repulsion for ITP. XTP would already mostly be present in its anionic form at neutral pH values. The alkaline pH optimum of noncanonical NTPases might therefore be better explained by an enhancement of the hydrolysis reaction. The mechanism proposed by Savchenko and coworkers involves the generation of a nucleophilic hydroxide ion and several proton-transfer steps among well conserved lysine and aspartate residues in RdgB (Savchenko *et al.*, 2007). The corresponding residues in TM0159 are Lys15, Lys52 and Asp68. Savchenko and coworkers noted that dUTPase, which also removes pyrophosphate from a noncanonical nucleotide (in this case dUTP), has a very similar reaction mechanism (Savchenko *et al.*, 2007). Notably, the rate-limiting step for dUTPase is also the hydrolysis reaction and not the substrate-binding step (Tóth *et al.*, 2007).

### 4.5. Product release

Some noncanonical NTPases, including RdgB from *E. coli* and human ITPase, exhibit structures with an open conformation in the unliganded state and a closed conformation in the ITP-liganded state (Savchenko *et al.*, 2007; Stenmark *et al.*, 2007). Curiously, both TM0159 structures (unliganded and with bound IMP) exhibit the typical features of a closed conformation. This is illustrated in Fig. 6 *via* structural superpositions of TM0159 with RdgB from *E. coli* and human ITPase. The most striking difference between the open and

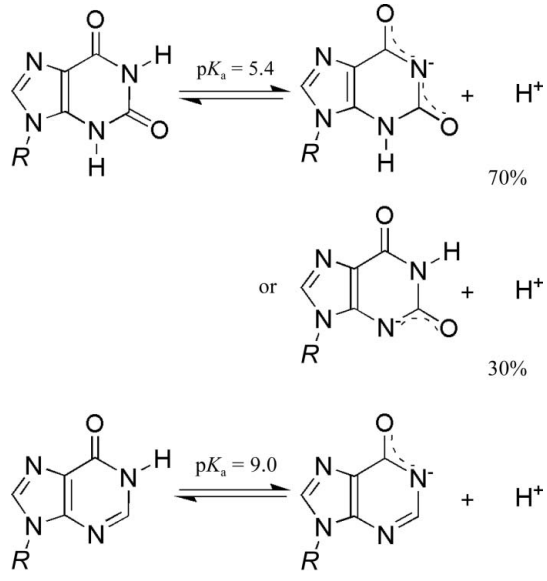


**Figure 5**
Protonation/deprotonation equilibria for XTP and ITP. *R* represents the ribose sugar and three phosphate groups. The p$K_a$ values were obtained from Sigel *et al.* (2009) and Naumann *et al.* (1974).

the closed conformations is the position of helix $\alpha$1 and its preceding loop (residues 11–24 in TM0159). This region of the N-terminal lobe only aligns if TM0159 is compared with the other noncanonical NTPases in a closed conformation in the ITP-liganded state. Additional differences occur in small loop regions of the C-terminal lobe (residues 145–149 and 89–95 in TM0159). Despite low sequence identity ranging from 26% for human ITPase to 34% for RdgB, the overall alignments with TM0159 are excellent. Between 167 and 186 (out of 189) $C^{\alpha}$ atoms align with r.m.s.d.s in the range 1.6–2.5 Å. Additional figures and information on the structure-based sequence alignments performed with the program *VAST* (Gibrat *et al.*, 1996) are provided in the Supplementary Material (Table S1, Fig. S7 and Fig. S8).

The IMP-liganded structure of RdgB from *E. coli* strongly suggests that product release is accompanied by a change in the local conformation at the nucleotide-binding site. The average isotropic *B* factor for IMP (41.72 Å$^2$) is clearly elevated in comparison to the average isotropic *B* factor for all protein atoms (31.6 Å$^2$) in this 2.02 Å resolution structure (Savchenko *et al.*, 2007). In contrast, the average isotropic *B* factors for IMP in the 2.15 Å resolution structure of TM0159 range from 50.41 Å$^2$ in monomer *B* to 57.43 Å$^2$ in monomer *D* and are only somewhat higher compared with the average isotropic *B* factors for protein atoms, which range from 42.13 Å$^2$ in monomer *B* to 50.63 Å$^2$ in monomer *C*. This indicates some disorder, but not necessarily mobility, of IMP in the TM0159 structure.

The observation of closed conformations in both TM0159 structures is most likely owing to the presence of an anion in a position that accommodates the last phosphate group of ITP in the structures of RdgB and human ITPase. The backbone amide group of Thr11 and the side chains of Lys15 and Lys52 coordinate the modeled sulfate anion in both TM0159 structures. The first two residues (Thr11 and Lys15) are part of the region that moves upon ITP coordination in human ITPase and RdgB. The presence of a sulfate or another anion is most likely to trigger the closure of the two lobes by moving helix $\alpha$1 and its preceding loop toward the other lobe. The structures of TM0159 therefore suggest a critical role for the $\gamma$-phosphate group of the substrates in closing the nucleotide-binding pocket. The structure of TM0159 with bound IMP, however, cannot illustrate product release *via* a local conformation change, since the open conformation was not observed.

An alternative or additional mechanism for product release could involve inter-monomer rotations. The structures of the noncanonical NTPase PhNTpase from *P. horikoshii* support this mechanism (Lokanath *et al.*, 2008). Most structures of PhNTPase, including unliganded PhNTPase, are present in a closed conformation. Only one structure, the IMP complex of PhNTPase (PDB entry 2dvn; Lokanath *et al.*, 2008), differs: one subunit (chain *A*) is in the closed conformation and the other subunit (chain *B*) exhibits domain opening. Chain *A* houses a sulfate group in the same position as in TM0159, whereas chain *B* contains a glycerol molecule in the corresponding location. Lokanath *et al.* (2008) propose that PhNTPase has a constitutively closed active site and that inter-monomer rotations play a role in the reaction mechanism, in particular the recruitment of the metal-ion cofactor and the release of the IMP product. One interface residue (Tyr80 in PhNTPase) is always located near the rotation axis. The interface residue in the corresponding position of TM0159 is Met83. Since both TM0159 structures align best with the structures of other noncanonical NTPases in a closed conformation, product release from TM059 might thus also require inter-monomer rotations. The tilted interface of the TM0159 dimer suggests that inter-monomer rotations are readily possible.

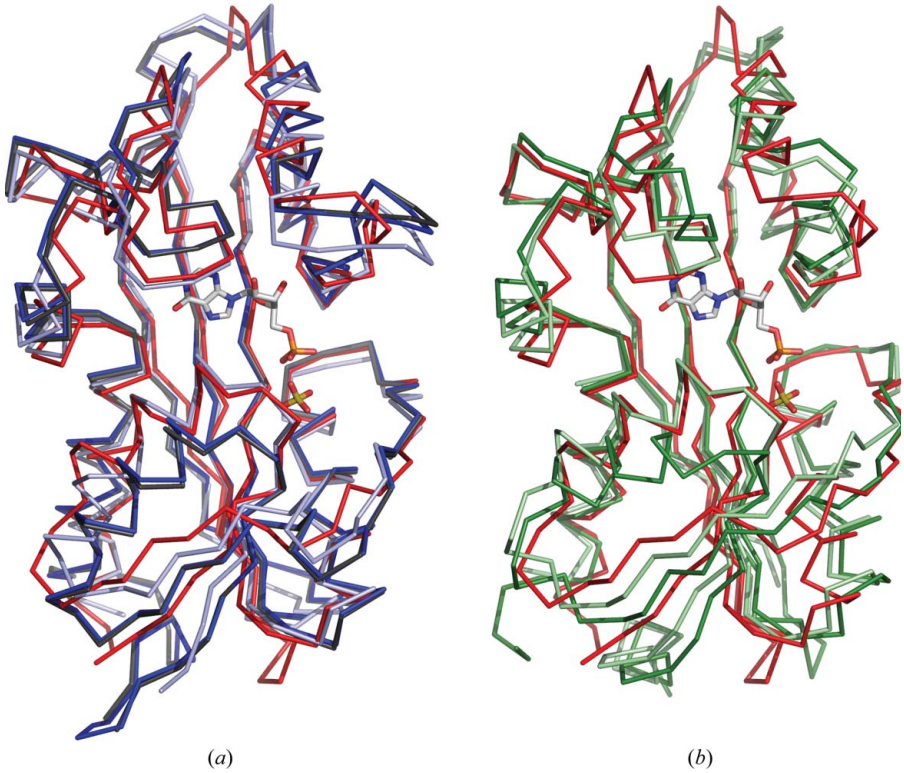


**Figure 6**
Structural alignment of TM0159 with the structures of RdgB and human ITPase. The ribbon trace of TM0159 (chain *C*) is colored red; IMP and $SO_4^{2-}$ are shown as small sticks. (*a*) RdgB with bound ITP (PDB entry 2q16, chain *A*) is colored light blue, RdgB with bound IMP (PDB entry 2pyu) dark blue and RdgB without ligand (PDB entry 1k7k) dark gray (Savchenko *et al.*, 2007). (*b*) Human ITPase with bound ITP (PDB entry 2car, chain *A*) is colored light green and human ITPase without ligand (PDB entry 2j4e, chain *G*) dark green (Stenmark *et al.*, 2007).

### 4.6. Thermophilic features of TM0159

Extended salt-bridge networks and compact protein cores with branched

hydrophobic side chains have been proposed to be particularly important for thermostability (Kumar & Nussinov, 2001). Other features, such as shortened loops at the protein surface or the abundance of specific amino acids (*e.g.* proline), do not make consistent contributions to thermostability when large structural data sets are considered (Kumar & Nussinov, 2001). Robinson-Rechavi and coworkers compared 94 protein structures from *T. maritima* with their homologs from mesophilic bacteria (Robinson-Rechavi *et al.*, 2006). The unliganded structures of RdgB and TM0159 served as an example to demonstrate that thermophilic proteins show increased compactness and residue connectivity compared with mesophilic proteins (Robinson-Rechavi *et al.*, 2006). Unfortunately, the authors did not consider that unliganded RdgB and unliganded TM0159 adopt different conformations. The closed conformation of TM0159 already contributes to an increase in compactness. Nevertheless, a striking difference between the amino-acid compositions of RdgB and TM0159 still prevails. TM0159 contains a total of 61 charged amino acids. In contrast, RdgB only contains 46 charged amino acids, similar to human ITPase, with 45 charged amino acids. The two noncanonical NTPases from thermophilic archaea, Mj0226 and PhNTPase, contains 61 and 50 charged amino acids, respectively. PhNTPase from *P. horikoshii* is somewhat shorter than the other noncanonical NTPases (186 residues compared with 193–197 residues). The number of branched hydrophobic amino-acid side chains is similar in the five structurally characterized noncanonical NTPases. These numbers range from 46 branched amino acids for RdgB to 40–43 for the other four noncanonical NTPases. More notable differences exist in the numbers of aromatic amino acids, which are 16, 22, 24, 24 and 31 for RdgB, human ITPase, TM0159, PhNTPase and Mj0226, respectively. Supplementary Table S3 summarizes the amino-acid compositions of all five noncanonical NTPases. Here, we describe some noteworthy examples of salt bridges that aid in enhancing the thermostability of TM0159 from the thermophilic bacterium *T. maritima* in comparison to PhNTPase and Mj0226 from thermophilic archaea.

One salt bridge that is unique to TM0159 involves Asp41 and Lys51. Lys51 has an additional hydrogen-bonding connection to the main-chain carbonyl group of Val39. These interactions aid in the stabilization of the position of Glu40 within the long loop encompassing residues 32–44. The side chain of the well conserved Glu40 plays an important functional role as it coordinates the modeled sulfate and the phosphate group of IMP in the TM0159 structure. A second salt bridge between Glu40 and Lys52 is also found in RdgB from *E. coli* and human ITPase, but only TM0159 has another charged residue in the adjacent position, Lys51, which creates a unique closed network of stabilizing interactions. The noncanonical NTPases from thermophilic archaea demonstrate a different approach to stabilizing the position of the long loop. Lys52 (TM0159 sequence numbering) is replaced by a glycine residue to make space for a larger tyrosine side chain. This tyrosine side chain (Tyr34 in PhNTPase and Tyr43 in Mj0226) is part of a YPE motif which includes the well conserved Glu residue (Glu40 in TM0159). Both Glu and Tyr are in a position that is well suited to form hydrogen bonds to the $\gamma$-phosphate group of a nucleoside triphosphate substrate. The proline residue of the YPE motif is part of a hydrophobic pocket which includes a phenylalanine and a tryptophan residue (Phe48 and Trp52 in PhNTPase, Phe56 and Trp60 in Mj0226) and accommodates a phenylalanine (Phe48 in PhNTpase, Phe56 in Mj0226) from the second monomer. In contrast to the two proteins from thermophilic archaea, TM0159 does not employ such hydrophobic interactions across this part of the dimer interface. In fact, the position of phenylalanine Phe48 in PhNTPase (or Phe56 in Mj0226) is occupied by Lys51 in TM0159. The dimer interface of TM0159 is smaller and Lys51 interacts with Asp41 and Val39 of the same monomer, as mentioned above. In contrast to all other dimeric noncanonical NTPases, TM0159 might be able to form a tetramer. This different oligomerization state would significantly enhance the thermostability of TM0159.

Salt bridges across interfaces can make particularly important contributions to the thermostability of proteins (Takenoya *et al.*, 2010). Thus, we investigated interface 2, which is found in most noncanonical NTPases. Human ITPase does not contain a salt bridge at the dimer interface regardless of the crystal structure. The thermophilic proteins show salt bridges across the dimer interface, but not all crystal structures or all dimeric arrangements within a crystal structure show the same salt bridge. For example, only some structures of PhNTPase show a salt bridge between Arg98 and Glu40 across the dimer interface. As an alternative or in addition to this salt bridge, other PhNTPase structures display a salt bridge between Lys95 and Glu98. Notably, this second salt bridge is observed in all three thermophilic proteins and involves Lys102 and Glu105 in unliganded TM0159, and Lys103 and Glu106 in Mj0226. In IMP-liganded TM0159, however, the Glu105 side chain points toward the solvent channel and the main-chain carbonyl group of Glu105 forms a hydrogen bond to the Lys102 side chain. The IMP-liganded structure of TM0159 exhibits another salt bridge between Glu95 and Lys107. This salt bridge is unique to TM0159 and is part of a larger network of interactions (Supplementary Fig. S10). Glu95 forms an internal salt bridge with Arg98 and a hydrogen bond to the side chain of Thr99. The side chain of Thr99 in turn forms a hydrogen bond to the main-chain carbonyl group of Glu95. The side-chain amino group of Lys107 forms a hydrogen bond to the side chain of Thr99 across the dimer interface and an internal hydrogen bond to the main-chain carbonyl group of Met103. This main-chain carbonyl group forms a hydrogen bond to Lys102 across the dimer interface. As mentioned above, the Lys102 side chain then engages in a hydrogen bond to the main-chain carbonyl group of Glu105. The dimer interface of TM0159 therefore always has one active salt bridge and a network of connections between the two possible salt-bridge pairs Lys102/Glu105 and Glu95/Lys107. This feature is noteworthy because intermonomer rotations might assist in the release of product and adjustable dimer-interface interactions are therefore advantageous.

## 5. Conclusion

In conclusion, we provide experimental evidence for the proposed function of TM0159 as a noncanonical NTPase and present the protein structure of TM0159 with bound IMP. The structure of TM0159 shares key features with other noncanonical NTPases, such as the fold of the individual monomers and well conserved hydrogen-bonding networks among residues critical for substrate recognition and catalysis. Intermonomer rotations as well as more local conformational changes can contribute to product release. Most noncanonical NTPases, with the exception of RdgB, share the same type of dimer interface. In contrast to human ITPase, however, the thermophilic noncanonical NTPases have at least one salt bridge at the dimer interface. In addition, TM0159 might employ a tetrameric oligomerization state for further stabilization. The fact that noncanonical NTPases have evolved in all domains of life indicates that these enzymes accomplish an important task, which is to safeguard the pool of correct RNA and DNA building blocks.

This work was supported by a grant from the CSU Faculty–Student Collaborative Research Seed Grant Program (CSUPERB; to MS), a Faculty Support Grant (CSU East Bay; to MS) and two Graduate Student Research Grants (CSU East Bay; to AD). Portions of this research were carried out at the Stanford Synchrotron Radiation Lightsource (SSRL), a Directorate of SLAC National Accelerator Laboratory and an Office of Science User Facility operated for the US Department of Energy Office of Science by Stanford University. The W. M. Keck Foundation Center for Molecular Structure coordinated the CSU research partnership with SSRL. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program (P41RR001209) and the National Institute of General Medical Sciences. We would like to thank Professor Dr K. Kantardjieff for facilitating beam time at SSRL. Maggie Kit-Man Ho, Chieu Phan and Vu Tran assisted during the early stages of protein purification. Suhasini Kulkarni and Shuang Huang assisted with colorimetric assays. Wendy Roberts-Waitkus helped with the preparation of crystallization solutions.

## References

Auer, T., Sninsky, J. J., Gelfand, D. H. & Myers, T. W. (1996). *Nucleic Acids Res.* **24**, 5021–5025.
Bahadur, R. P. & Zacharias, M. (2008). *Cell. Mol. Life Sci.* **65**, 1059–1072.
Bock, C. W., Katz, A. K., Markham, G. D. & Glusker, J. P. (1999). *J. Am. Chem. Soc.* **121**, 7360–7372.
Burgis, N. E. & Cunningham, R. P. (2007). *J. Biol. Chem.* **282**, 3531–3538.
Chung, J. H., Back, J. H., Park, Y. I. & Han, Y. S. (2001). *Nucleic Acids Res.* **29**, 3099–3107.
Cohen, A. E., Ellis, P. J., Miller, M. D., Deacon, A. M. & Phizackerley, R. P. (2002). *J. Appl. Cryst.* **35**, 720–726.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System*, v.1.2r3pre. http://www.pymol.org.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Galperin, M. Y., Moroz, O. V., Wilson, K. S. & Murzin, A. G. (2006). *Mol. Microbiol.* **59**, 5–19.
Gibrat, J. F., Madej, T. & Bryant, S. H. (1996). *Curr. Opin. Struct. Biol.* **6**, 377–385.
Guharoy, M. & Chakrabarti, P. (2010). *BMC Bioinformatics*, **11**, 1–17.
Hwang, K. Y., Chung, J. H., Kim, S.-H., Han, Y. S. & Cho, Y. (1999). *Nature Struct. Biol.* **6**, 691–696.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Klinker, J. F. & Seifert, R. (1997). *Biochem. Pharmacol.* **54**, 551–562.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Kumar, S. & Nussinov, R. (2001). *Cell. Mol. Life Sci.* **58**, 1216–1233.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lesley, S. A. *et al.* (2002). *Proc. Natl Acad. Sci. USA*, **99**, 11664–11669.
Lin, S., McLennan, A. G., Ying, K., Wang, Z., Gu, S., Jin, H., Wu, C., Liu, W., Yuan, Y., Tang, R., Xie, Y. & Mao, Y. (2001). *J. Biol. Chem.* **276**, 18695–18701.
Lokanath, N. K., Pampa, K. J., Takio, K. & Kunishima, N. (2008). *J. Mol. Biol.* **375**, 1013–1025.
Lokanath, N. K., Shiromizu, I., Ohshima, N., Nodake, Y., Sugahara, M., Yokoyama, S., Kuramitsu, S., Miyano, M. & Kunishima, N. (2004). *Acta Cryst.* D**60**, 1816–1823.
Marini, I. & Ipata, P. L. (2007). *Biochem. Mol. Biol. Educ.* **35**, 293–297.
Minasov, G., Teplova, M., Stewart, G. C., Koonin, E. V., Anderson, W. F. & Egli, M. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 6328–6333.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. (1999). *Acta Cryst.* D**55**, 247–255.
Naumann, C. F., Prijs, B. & Sigel, H. (1974). *Eur. J. Biochem.* **41**, 209–216.
Porta, J., Kolar, C., Kozmin, S. G., Pavlov, Y. I. & Borgstahl, G. E. O. (2006). *Acta Cryst.* F**62**, 1076–1081.
Robinson-Rechavi, M., Alibés, A. & Godzik, A. (2006). *J. Mol. Biol.* **356**, 547–557.
Savchenko, A., Proudfoot, M., Skarina, T., Singer, A., Litvinova, O., Sanishvili, R., Brown, G., Chirgadze, N. & Yakunin, A. F. (2007). *J. Mol. Biol.* **374**, 1091–1103.
Shapiro, R. & Pohl, S. H. (1968). *Biochemistry*, **7**, 448–455.
Sigel, H., Operschall, B. P. & Griesser, R. (2009). *Chem. Soc. Rev.* **38**, 2465–2494.
Spee, J. H., de Vos, W. M. & Kuipers, O. P. (1993). *Nucleic Acids Res.* **21**, 777–778.
Stenmark, P., Kursula, P., Flodin, S., Gräslund, S., Landry, R., Nordlund, P. & Schüler, H. (2007). *J. Biol. Chem.* **282**, 3182–3187.
Stepchenkova, E. I., Tarakhovskaya, E. R., Spitler, K., Frahm, C., Menezes, M. R., Simone, P. D., Kolar, C., Marky, L. A., Borgstahl, G. E. & Pavlov, Y. I. (2009). *J. Mol. Biol.* **392**, 602–613.
Takenoya, M., Ohtaki, A., Noguchi, K., Endo, K., Sasaki, Y., Ohsawa, K., Yajima, S. & Yohda, M. (2010). *J. Struct. Biol.* **170**, 532–539.
Tóth, J., Varga, B., Kovács, M., Málnási-Csizmadia, A. & Vértessy, B. G. (2007). *J. Biol. Chem.* **282**, 33572–33582.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Vanderheiden, B. S. (1979). *J. Cell. Physiol.* **99**, 287–301.
Wang, Y., Geer, L. Y., Chappey, C., Kans, J. A. & Bryant, S. H. (2000). *Trends Biochem. Sci.* **25**, 300–302.
Weber, J. & Senior, A. E. (2001). *J. Biol. Chem.* **276**, 35422–35428.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Zheng, J., Singh, V. K. & Jia, Z. (2005). *Structure*, **13**, 1511–1520.