Supplementary Material

Structural and functional characterization of a non-canonical nucleoside triphosphate pyrophosphatase from *Thermotoga maritima*

Khaldeyah Awwad, Anna Desai, Clyde Smith, Monika Sommerhalter

Activity assays performed with thin layer chromatography (TLC)

TLC conditions were similar to the procedure described by Marini and Ipata (Marini & Ipata, 2007). Polyethyleneimine impregnated cellulose TLC plates (Flexible TLC Plates, Cellulose PEI F-254 from Selecto Scientific, 20 x 20 cm) were rinsed with 10 % NaCl (w/v), then rinsed twice with deionized water (Milli-Q, Millipore), and finally air dried. Next, the plates were cut into four 10 x10 cm squares. A typical reaction mixture contained 2.3 mM of a nucleoside triphosphate, 20 mM $MgCl_2$, 50 mM CAPS, pH 10.2, and 0.01 mg/mL of TM0159. Each reaction mixture was incubated in a water bath at 40°C, 50°C, and 60°C. The reaction samples (2 µL) were spotted at the following time intervals: 5, 15, 30, 60, and 90 minutes. For reference, the TLC plates were spotted with 2 µL of standard solutions, such as 0.1 mM ITP and 0.1 mM IMP. All spots were eluted with 0.9 M LiCl, dried, and visualized in UV light at 254 nm with the help of a UV lamp (UV-LAMP-Spectroline model ENF-240C). To check whether TM0159 is also active at room temperature, an additional reaction was set up (2.3 mM ITP, 20 mM $MgCl_2$, 50 mM CAPS, pH 10.2, and 0.01 mg/mL of TM0159) and kept at room temperature for 24 hours. After 24 hours, an aliquot was spotted on the TLC plate and visualized. TM0159 readily catalyzes the conversion of ITP to IMP at elevated temperatures. No conversion was detected at room temperature after 24 hours (Fig. S1). We also did not detect any conversion with the nucleoside triphosphate ATP.

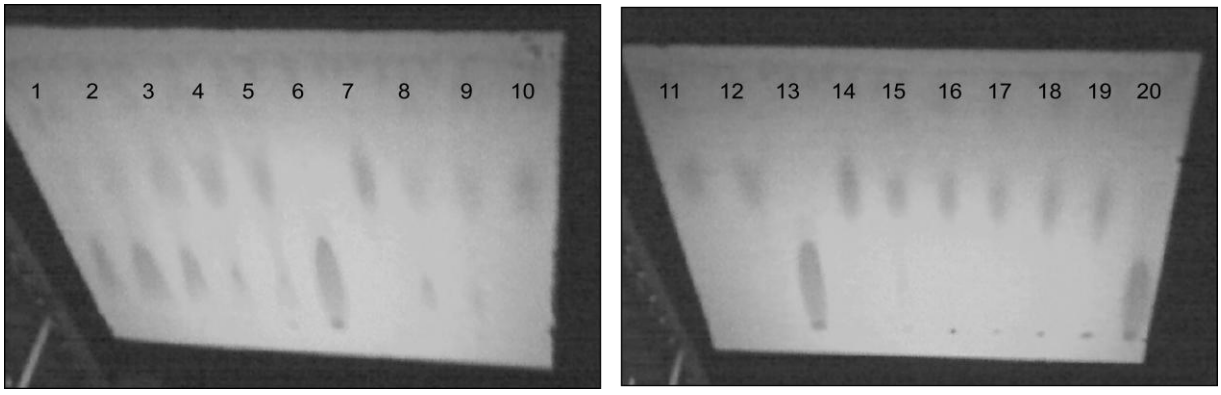


Figure S1: Photographs of two TLC plates that were used to monitor ITP hydrolysis catalyzed by TM0159 at different temperatures. Standard solutions (0.1 mM ITP and 0.1 mM IMP) were spotted on the plates as well: Lanes 6 & 13 ITP standard; lanes 7 & 14 IMP standard. Lanes 1-5: reaction temperature 40°C spotted after 5 min (lane 1), 15 min (lane 2), 30 min (lane 3), 60 min (lane 4), and 90 min (lane 5). Lanes 8-12: reaction temperature 50°C spotted after 5, 15, 30, 60, and 90 min. Lanes 15-19: reaction temperature 60°C spotted after 5, 15, 30, 60, and 90 min. Lane 20 contains a reaction sample incubated at room temperature for 24 hours.

Activity assays performed with HPLC:

A typical reaction mixture consisted of 2 mM nucleotide triphosphate, such as ITP or XTP, 20 mM $MgCl_2$, 50 mM CAPS, pH 10.2, and 0.43 µM TM0159. The enzyme was added last, and the mixture was placed in a water bath set to 50°C and timed. At specific time intervals (0, 2, 4, 6, 8 and 10 minutes), 25 µL aliquots were removed and mixed with 5 µL of 12 % (v/v) trichloroacetic acid (TCA) to denature and precipitate TM0159. After centrifugation for 30 minutes at 15000 rpm, the supernatant was transferred to an HPLC vial and analyzed via HPLC.

The HPLC equipment consisted of an HP 1050 (Agilent 1050) HPLC system. As a mobile phase, we used 15 mM ammonium phosphate buffer pH 6.0 and 1 % to 5 % methanol (Fisher, HPLC grade). The stationary phase was a 1.8 mL reversed-phase C-18 column from Grace, equipped with an All-Guard cartridge system. The injection volume was varied between 1 and 10 µL. The flow rate was 1.5 mL per minute, and 254 nm was chosen as detection wavelength. Every sample was analyzed in triplicate. Before every HPLC run, a mix of standard nucleotides was applied on the HPLC column to establish a retention time specific for each nucleotide and to obtain a peak area corresponding to a specific concentration of each nucleotide. The standard nucleotide samples included 0.1 mM ITP, 0.1 mM IMP, 0.1 mM ITP / 0.1 mM IMP mixture, 0.2 mM ITP / 0.2 mM IMP mixture, 0.1 mM XTP, 0.1 mM XMP and 0.1 mM XTP / 0.1 mM XMP mixture. Fig. S2 (A-E) shows consecutive chromatograms from reaction mixture aliquots (with 2 mM ITP) taken every 2 minutes from the start of reaction until its completion at 8 minutes.

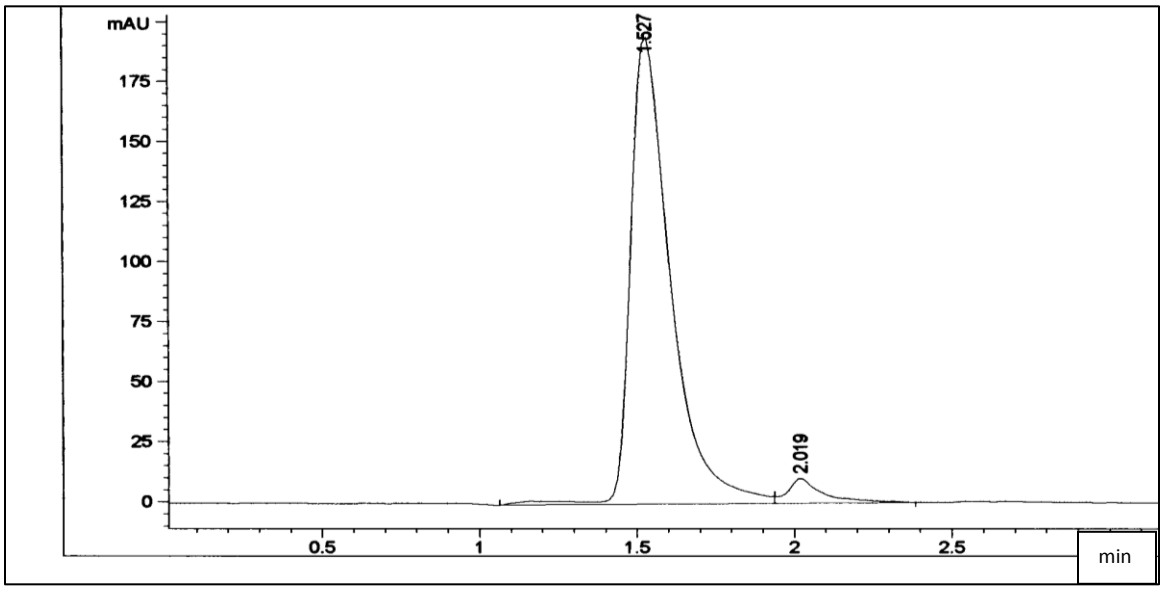


Figure S2(a). HPLC Chromatogram, at 0 min reaction time. The first peak represents ITP (1.527 min), and the second peak represents IMP (2.019 min).

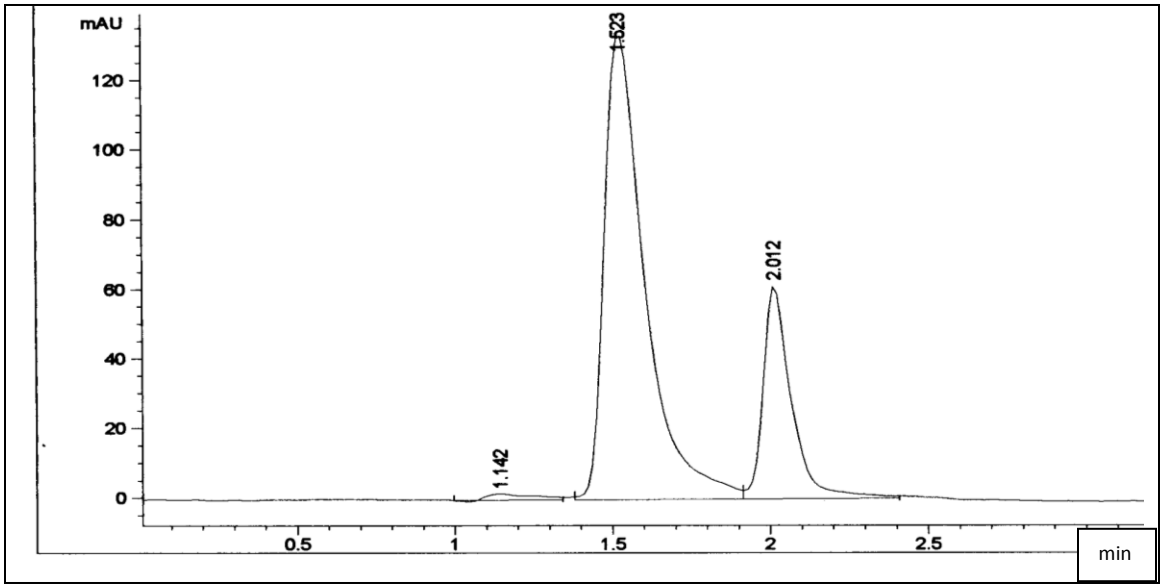


Figure S2(b): HPLC Chromatogram, after 2 min reaction time. The first peak (1.523 min) represents ITP, and the second peak (2.012 min) represents IMP.

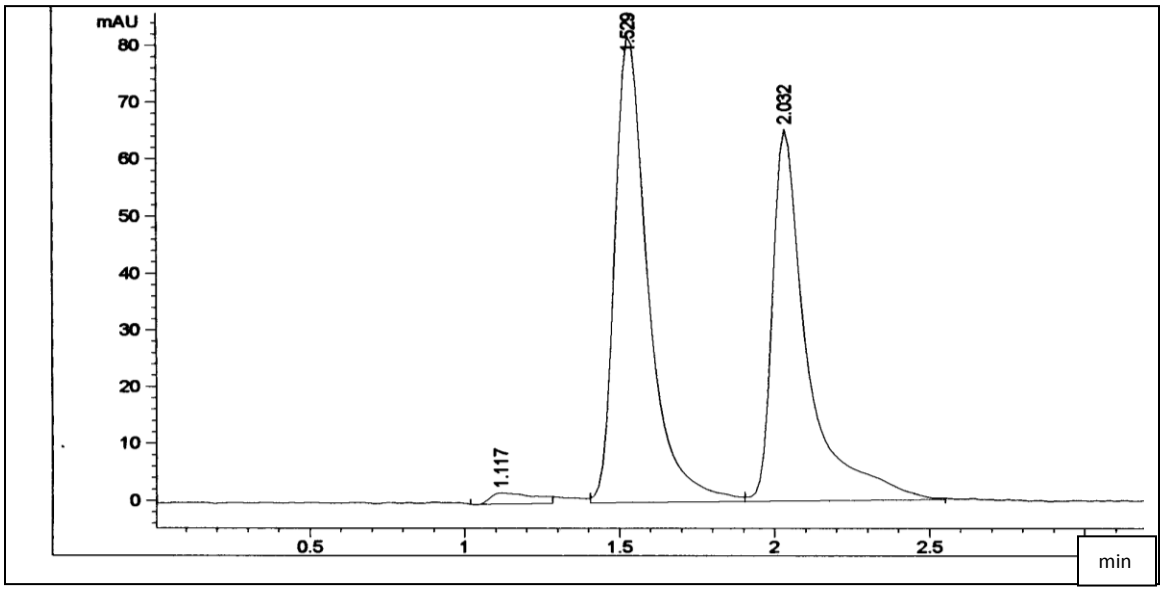


Figure S2(c): HPLC Chromatogram, after 4 min reaction time. The first peak (1.529 min) represents ITP, and the second peak (2.032 min) represents IMP.

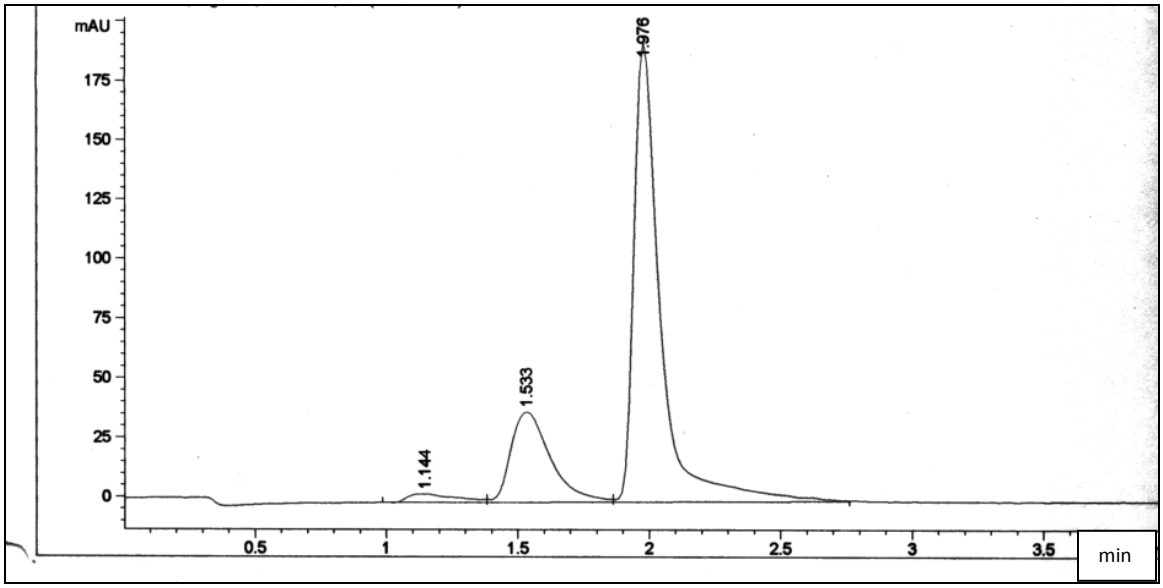


Figure S2(d): HPLC Chromatogram, after 6 min reaction time. The first peak (1.533 min) represents ITP, and the second peak (1.976 min) represents IMP.

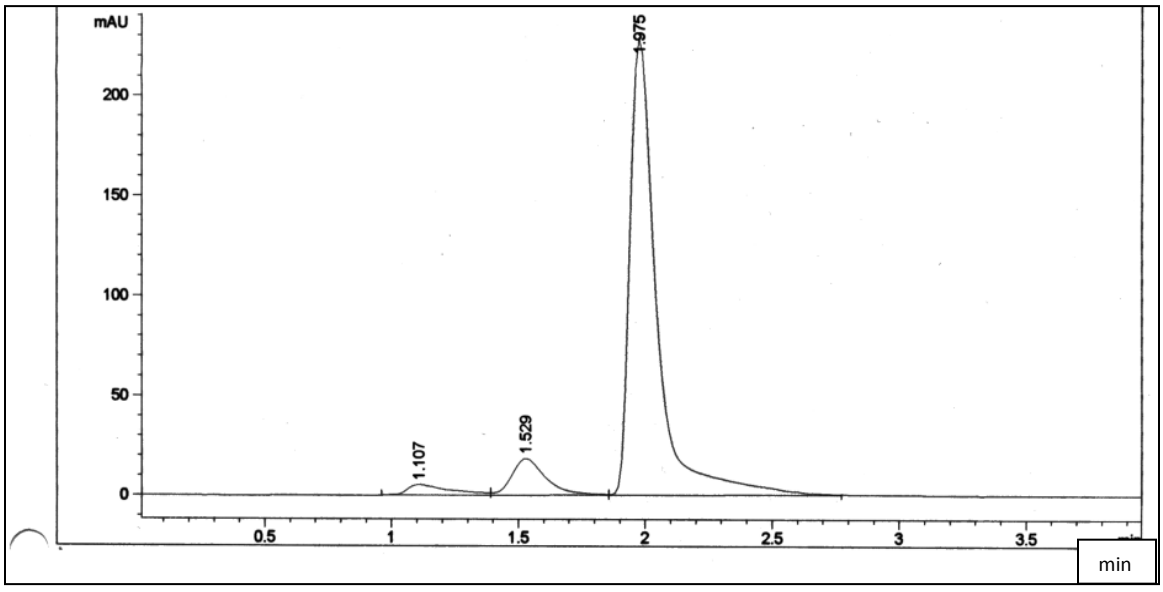


Figure S2(e): HPLC Chromatogram, after 8 min reaction time. The first peak (1.529 min) represents ITP, and the second peak (1.975 min) represents IMP.

Variation of TM0159 activity with pH

The colorimetric assays were performed as described in the experimental section of the main manuscript. Each reaction mixture contained either 0.23 mM ITP or 0.20 mM XTP in addition to 0.01 M MgCl2, 0.001 mg/mL TM0159, 0.001 mg/mL yeast pyrophosphatase, and 0.05 M buffer of varying pH: sodium acetate-acidic acid, pH 3.14; sodium acetate-acidic 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; NaHCO3-NaOH, pH 11.04; KCl-NaOH, pH 11.95; and KCl-NaOH, pH 12.53.All measurements were performed in triplicate.

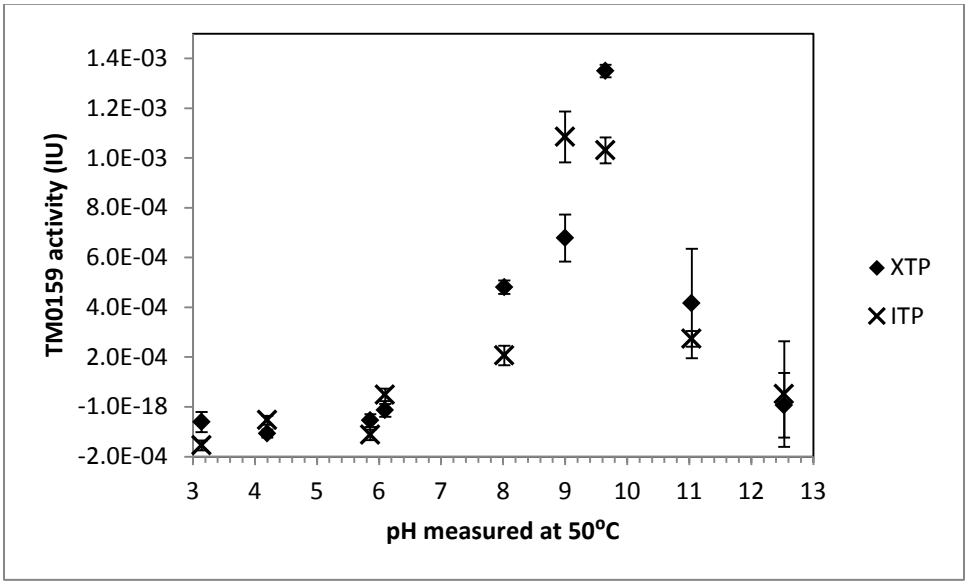


Figure S3: pH profile for TM0159 with XTP and ITP.

Cofactor Requirement

Figure shows the activity of TM0159 in dependence of $MgCl_2$ concentration. Each reaction was performed in the presence of 2.3 mM ITP, 0.001 mg/mL TM0159, 0.05 M Tris, pH 9, and 0.001 mg/mL yeast pyrophosphatase. The following $MgCl_2$ concentrations were used: 0.1 mM, 0.5 mM, 1 mM, 2 mM, 5 mM, 10 mM, 15 mM, and 20 mM. All measurements were performed in triplicate. The temperature was 50°C. It was established that $MgCl_2$ concentrations of 10 mM or higher are needed for optimal catalytic conditions for TM0159. Therefore, all other colorimetric assays were performed in the presence of 10 mM $MgCl_2$.

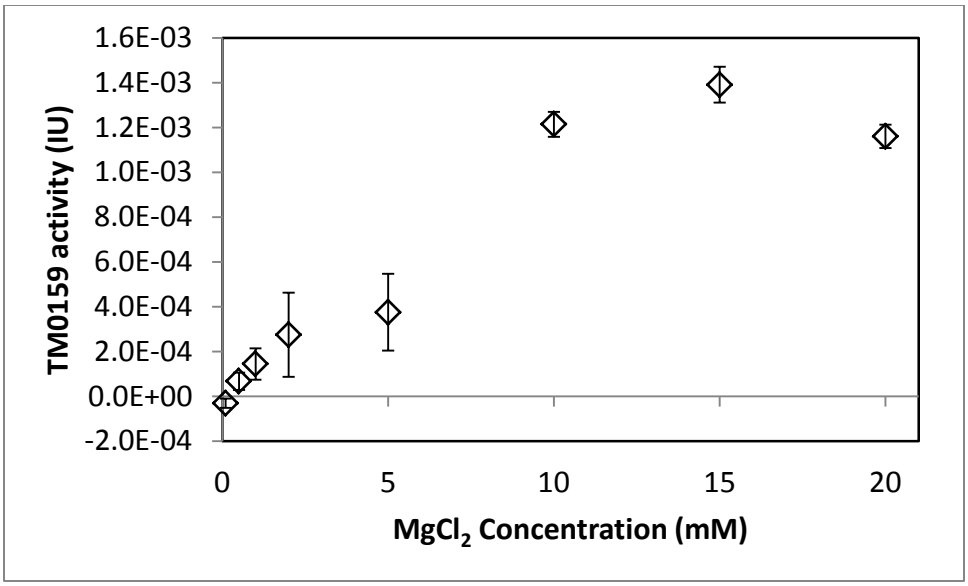


Figure S4: Kinetic activity of TM0159 with varying concentrations of $MgCl_2$.

Figure S5 shows TM0159's activity with various cofactors. Each reaction mixture contained 0.23 mM ITP, 0.05 M Tris, pH 9, 0.001 mg/mL TM0159 and 0.001 mg/mL yeast pyrophosphatase and 0.1 M salt. The following salts were used: manganese (II) sulfate ($MnSO_4$), manganese chloride ($MnCl_2$), zinc sulfate ($ZnSO_4$), potassium sulfate (K_2SO_4), lithium sulfate (Li_2SO_4), copper (II) sulfate ($CuSO_4$), sodium sulfate (Na_2SO_4), cobalt (II) chloride ($CoCl_2$), magnesium chloride ($MgCl_2$), and calcium chloride ($CaCl_2$). All measurements were performed at 50°C and in triplicate. These colorimetric assays confirmed that TM0159 requires Mg^{2+} for its catalytic activity.

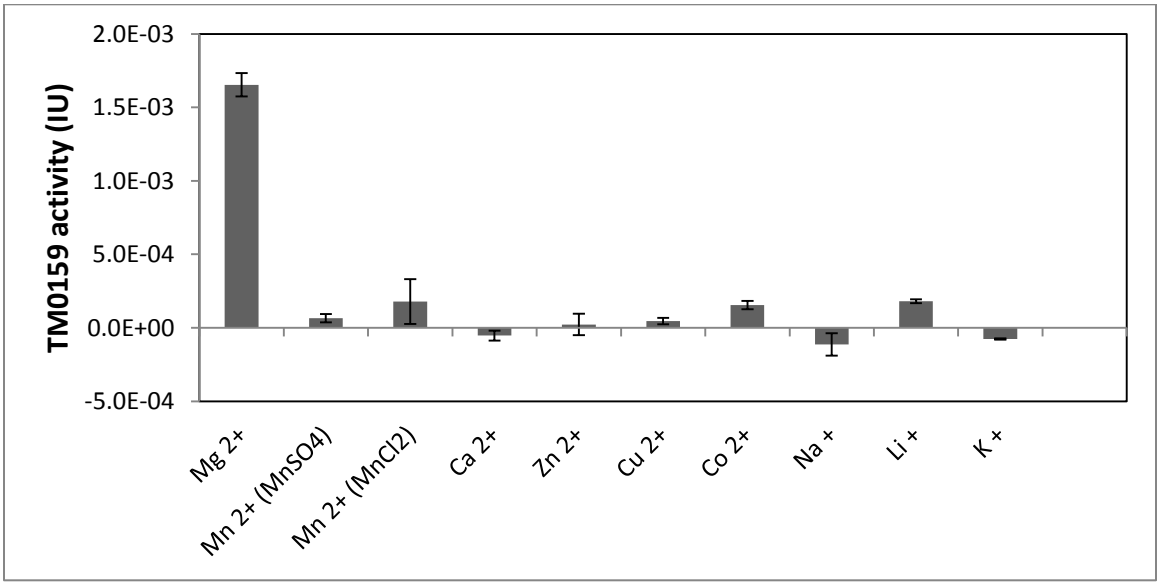


Figure S5: TM0159 activity in the presence of various cations.

Gel filtration

A column with a diameter of 1 cm was packed with Sephadex-G75 to a height of 27 cm. The column was attached to a BioRad FPLC system (BioLogic LP system). The elution buffer composed of 20 mM TRIS, 5 mM $MgCl_2$, pH 7.2 was applied with a flow rate of 0.25 mL/min. The elution of protein samples was monitored by recording the absorbance at 280 nm. The following samples were used to calibrate the column: 2 mg/mL Dextran Blue (2000 kDa; to determine the void volume V_0 = 3.8 mL corresponding to 15.2 minutes), 2 mg/mL bovine serum albumin (BSA, 66 kDa; V_e = 3.85 mL or 15.4 min), 1 mg/mL carbonic anhydrase (CA, 29 kDa; V_e = 5.5 mL or 22.0 min), and 5 mg/mL equine cytochrome c (Cyt c, 12.4 kDa; V_e = 7.2 mL or 28.8 min). The calibration curve is shown in Fig. S6. TM0159 was injected in concentrations of 4 mg/mL and 6 mg/mL and eluted at V_e = 4.2 mL or 16.7 min. This elution time corresponds to an apparent molecular weight of 56 kDa. Since a single TM0159 polypeptide chain has a molecular weight of 23.8 kDa, the best match for the oligomerization state of TM0159 is a dimer.

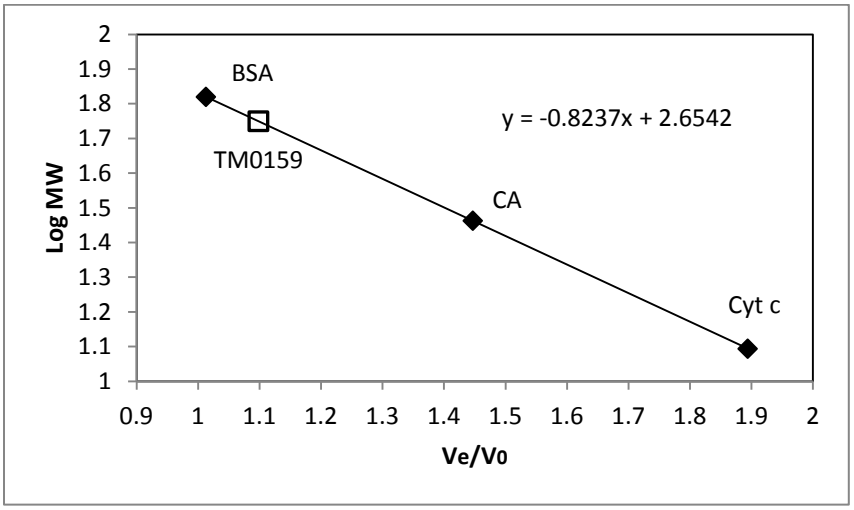


Figure S6: Calibration curve for the determination of the apparent molecular weight of TM0159 using a Sephadex-G75 gel filtration column.

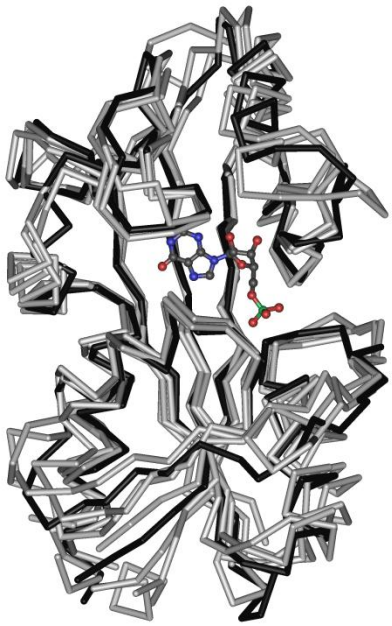
Structure-based sequence alignments:

The program VAST (Gibrat *et al.*, 1996) was used to generate structure-based sequence alignments. We observed that our TM0159 structure with bound IMP has a close resemblance to structures of non-canonical NTPase that are present in the closed conformation (mostly due to ITP binding). The overlay of the backbone traces in Figures S7 and S8 was generated with Cn3D (Wang *et al.*, 2000). Figure S7 focuses on structures present in a closed conformation. In contrast, Figure S8 compares TM0159 to non-canonical NTPases present in an open conformation. Table S1 summarizes the numerical results of the structure-based sequence alignment for TM0159 with bound IMP. Since this structure has 4 chains, each entry has four values for chain A, B, C, and D.

In our assessment of the statistics summarized in table S1 we focused on alignment length, VAST score, P-value, and RMSD values. All non-canonical NTPases showed similar alignment lengths without major interruptions (or in other words the alignment encompasses the whole structure). Structures of other proteins listed by VAST reached only shorter alignment lengths with multiple interruptions. The alignment length corresponds to the number of equivalent C-alpha atoms. The RMSD value is the root mean square superposition of these equivalent C-alpha atoms. Some structures of non-canonical NTPases with an open conformation show lower RMSD values than their counterparts with a closed conformation. However, this lower RMSD value is associated with a shorter alignment length. Several of the non-aligned residues are highlighted with a red box in Figure S8. These residues are located in a region that moves upon closure of the nucleotide binding site. The VAST score is a similarity score that takes into account the superposition of secondary structure elements. Together with the P-value, a probability value for alignment by pure chance, the VAST score provides a very general measure of correlation between structural alignments. Notably, all non-canonical NTPases reach VAST scores well above 10.0 and very low P-values. The highest P-value of 10e-5.2 represents a likelihood of roughly 1 to 100,000 for finding a structurally matching protein by pure chance.

We observed a deviation from the alignment proposed by VAST in Figure S7 in a manual inspection of superimposed structures that were generated using Pymol. Introducing a small gap (one bar corresponding to one amino acid) shortly after residue Glu47 (TM0159 numbering, residue Lys3 (K3) corresponds to location 1 for 3S86) in all structures except for 2dvn from PhNTPase and increasing the alignment gap for 2dvn by one amino acid position in front of His32 is more consistent with the Pymol alignment and brings about two more conserved amino acids emphasized in red below:

```
                  10        20        30        40        50        60
             ....*....|....*....|....*....|....*....|....*....|....*....|
3S86_A  1 KLTVYLATTNPHKVEEIKMIA-P-E--WMEILPS--PEKIEVVEDGETFLE-NSVKKAVVY 54
2Q16_A 21 xqKVVLATGNVGKVRELASLLsD-F--GLDIVAQtdLGVDSAEETGLTFIE-NAILKARHA 77
2DVN_A  1 -MKIFFITSNPGKVREVANFL-GtF--GIEIVQLk----HEYPEIQAEKLEDVVDFGISWL 53
2J4E_   9 GKKIVFVTGNAKKLEEVVQIL-G-DkfPCTLVAQk----IDLPEYQGEPDE-ISIQKCQEA 62
```



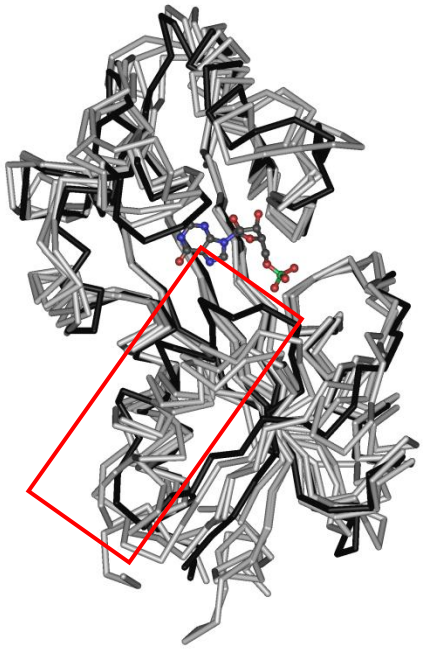
```
                10        20        30        40        50        60
           ....*....|....*....|....*....|....*....|....*....|....*....|
3S86_A   1 KLTVYLATTNPHKVEEIKMIA-P-E--WMEILPS--PEKIEVVEDGETFLENSVKKAVVY 54
2Q16_A  21 xqKVVLATGNVGKVRELASLLsD-F--GLDIVAQtdLGVDSAEETGLTFIENAILKARHA 77
2DVN_A   1 -MKIFFITSNPGKVREVANFL-GtF--GIEIVQLk---HEYPEIQAEKLEDVVDFGISWL 53
2J4E_G   9 GKKIVFVTGNAKKLEEVVQIL-G-DkfPCTLVAQk----IDLPEYQGEPDEISIQKCQEA 62

                70        80        90       100       110       120
           ....*....|....*....|....*....|....*....|....*....|....*....|
3S86_A  55 GKKLKHPVMADDSGLVIYSLGGFPGVMSARFMEEHs-YKEKMRTILKMLEG-K---DRRA 109
2Q16_A  78 AKVTALPAIADASGLAVDVLGGAPGIYSARYSGEDatDQKNLQKLLETXKD-VpddQRQA 136
2DVN_A  54 KGKVPEPFMIEDSGLFIESLKGFPGVYSSYVYRTI-----GLEGILKLMEG-Ae--DRRA 105
2J4E_G  63 VRQVQGPVLVEDTCLCFNALGGLPGPYIKWFLEK-----LKPEGLHQLLAGfE---DKSA 114

               130       140       150       160       170       180
           ....*....|....*....|....*....|....*....|....*....|....*....|
3S86_A 110 AFVCSATFFDPv-ENTLISVEDRVEGRIANEIRGTGGFGYDPFFIPDGYDKTFGEI-P-H 166
2Q16_A 137 RFHCVLVYLRHaedPTPLVCHGSWPGVITREPAGTGGFGYDPIFFVPSEGKTAAEL-T-R 194
2DVN_A 106 YFKSVIGFYI---DGKAYKFSGVTWGRISNEKRGTHGFGYDPIFIPEGSEKTFAEM-TiE 161
2J4E_G 115 YALCTFALSTgdpsqPVRLFRGRTSGRIVA-PRGCQDFGWDPCFQPDGYEQTYAEMpK-A 172

               190       200
           ....*....|....*....|...
3S86_A 167 -LKEKISHRSKAFRKLFSVLEKI 188
2Q16_A 195 eEKSAISHRGQALKLLLDALRNG 217
2DVN_A 162 -EKNALSHRGKALKAFFEWLKVN 183
2J4E_G 173 -EKNAVSHRFRALLELQEYFGsl 194
```

Figure S7: Structure-based sequence alignment for TM0159 with bound IMP (3S86 chain A) with other non-canoncial NTPases in closed conformations: RdgB from *E. coli* with bound ITP (2Q16 chain A), PhNTPase from *Pyrococcus horikoshii* with bound IMP (2DVN chain A), human ITPase with bound ITP (2J4E chain G). The backbone trace of TM0159 is shown in black, all other non-canonical NTPases are colored in grey.



```
                 10        20        30        40        50        60
            ....*....|....*....|....*....|....*....|....*....|....*....|
3S86_A    1 KLTVYLATTNphkveEIKMIAP-----EWM--EILPS--PEKIEVVEDGETFLENSVKKA 51
2PYU_A   21 MQKVVLATgn----vgKVRELasllsdFGL--DIVAQtdLGVDSAEETGLTFIENAILKA 74
2CAR_A    9 GKKIVFVTGnakkleeVVQILG-----DKFpdTLVAQ----KIDLPEYQGEPDEISIQKC 59
2MJP_A    8 imKIYFATGNpnkikEANIILK-----DLkdvEIEQI----KISYPEIQGTLEEVAEFGA 58

                  70        80        90       100       110       120
            ....*....|....*....|....*....|....*....|....*....|....*....|
3S86_A   52 VVYGKKLKHPVMADDSGLVIYSLGGFPGVMSARFME-EHSYKEKMRTILKMLEG-K---D 106
2PYU_A   75 RHAAKVTALPAIADASGLAVDVLGGAPGIYSARYSGeDATDQKNLQKLLETMKD-Vpdd Q 133
2CAR_A   60 QEAVRQVQGPVLVEDTCLCFNALGGLPGPYIKWFLE-K----LKPEGLHQLLAGfE---D 111
2MJP_A   59 KWVYNILKKPVIVEDSGFFVEALNGFPGTYSKFVQE-T----IGNEGILKLLEG-Kd--N 110

                 130       140       150       160       170       180
            ....*....|....*....|....*....|....*....|....*....|....*....|
3S86_A  107 RRAAFVCSATFFDPVE-NTLISVEDRVEGRIANEIRG-TGGFGYDPFFIPDGYDKTFGEI 164
2PYU_A  134 RQARFHCVLVYLRHaedptPLVCHGSWPGVITREPAG-TGGFGYDPIFFVPSEGKTAAEL 192
2CAR_A  112 KSAYALCTFALSTGDPsQPVRLFRGRTSGRIVAPRGC-QDF-GWDPCFQPDGYEQTYAEM 169
2MJP_A  111 RNAYFKTVIGYCD--E-NGVRLFKGIVKGRVSEEIRSkGYGFAYDSIFIPEEEERTFAEM 167

                 190       200
            ....*....|....*....|....*..
3S86_A  165 -PHL-KEKISHRSKAFRKLFSVLEKIL 189
2PYU_A  193 tREE-KSAISHRGQALKLLLDALRNgg 218
2CAR_A  170 pKAE-KNAVSHRFRALLELQEYFGSla 195
2MJP_A  168 -TTEeKSQISHRKKAFEEFKKFLLDRI 193
```

Figure S8: Structure-based sequence alignment for TM0159 with bound IMP (3S86 chain A) with other non-canonical NTPases in open conformations: RdgB from *E. coli* with bound IMP (2PYU chain A), human ITPase without nucleotide (2CAR chain A), Mj0226 from *Methanoccus jannashii* with AMPNP (2MJP chain A). The backbone trace of TM0159 is shown in black, all other non-canonical NTPases are colored in grey. The region of poor alignment is boxed in red.

Table S1: Structure-based sequence alignment for TM0159 with bound IMP (3S86 chains A,B,C,D) using the program VAST.

PDB-ID chain #	Description	Alignment length	Score	P-Value	RMSD	% ID
1VP2 chain A	TM0159 unliganded	188	15.2	10e-13.8	0.4	100
		188	15.4	10e-14.3	0.4	100
		189	15.2	10e-13.9	0.4	100
		189	15.3	10e-14.1	0.4	100
1VP2 chain B		188	17.7	10e-19.6	0.5	100
		189	17.7	10e-19.6	0.4	100
		189	17.0	10e-16.8	0.5	100
		189	17.7	10e-19.2	0.4	100
2Q16 chain A	RdgB from *E. coli* with bound ITP (closed conformation)	183	10.8	10e-5.2	1.8	34.4
		186	11.5	10e-6.3	1.9	33.3
		186	13.1	10e-8.9	1.8	33.3
		184	11.9	10e-7.0	1.7	34.2
2Q16 chain B		182	11.7	10e-6.7	1.8	34.1
		182	12.3	10e-7.6	1.6	33.3
		182	13.7	10e-10.1	1.7	34.6
		182	12.7	10e-8.2	1.7	34.6
2PYU chain A	RdgB from *E. coli* with IMP (open conformation)	174	15.6	10e-12.9	2.0	32.8
		175	16.1	10e-13.9	2.1	33.7
		174	15.2	10e-12.1	2.1	33.3
		174	16.2	10e-14.2	2.0	33.3
1K7K chain A	RdgB from *E. coli*, unliganded (open conformation)	177	15.4	10e-12.7	2.1	34.5
		173	15.9	10e-13.5	2.0	34.1
		175	14.9	10e-11.6	2.1	33.7
		177	16.0	10e-13.8	2.2	33.3
2MJP chain A	Mj0226 from *Methanoccus jannashii* with AMPNP (open conformation)	173	14.5	10e-12.1	1.7	38.2
		178	14.4	10e-11.8	1.8	38.2
		168	11.9	10e-7.3	2.3	32.7
		171	13.8	10e-10.7	1.8	37.4
2MJP chain B		176	16.8	10e-16.0	1.7	36.9
		177	16.8	10e-15.8	1.9	37.3
		175	15.3	10e-12.6	1.9	37.7
		174	16.2	10e-14.6	1.8	38.5
1B78 chain A	Mj0226 from *Methanoccus jannashii* without ligand (open conformation)	172	14.6	10e-12.1	1.7	37.8
		172	14.3	10e-11.7	1.8	37.8
		157	11.9	10e-7.3	1.9	36.3
		173	13.8	10e-10.6	1.8	34.7
1B78 chain B		175	14.6	10e-12.2	1.8	37.7
		177	14.4	10e-11.9	1.9	37.9
		169	12.1	10e-7.7	2.3	34.3
		175	13.9	10e-10.9	1.9	36.6

Continued:						
PDB-ID chain #	**Description**	**Alignment length**	**Score**	**P-Value**	**RMSD**	**% ID**
1V7R chain A	PhNTPase from *Pyrococcus horikoshii* Ot3, unliganded	177 179 179 177	16.7 17.1 16.2 17.2	10e-15.6 10e-16.8 10e-14.6 10e-17.0	1.6 1.6 1.7 1.5	37.9 37.4 37.4 37.9
2DVP chain A		179 180 180 177	16.8 17.3 16.3 17.4	10e-16.0 10e-17.5 10e-14.8 10e-17.7	1.7 1.7 1.7 1.5	37.4 37.2 37.2 37.9
2DVN chain A	PhNTPase from *Pyrococcus horikoshii* Ot3, IMP and sulfate	179 178 180 180	16.9 17.4 16.5 17.4	10e-16.0 10e-17.6 10e-15.2 10e-17.9	1.7 1.6 1.7 1.6	36.9 37.1 37.2 37.2
2DVN chain B	PhNTPase from *Pyrococcus horikoshii* Ot3, IMP and propentriol	180 181 181 181	14.3 14.6 14.6 14.8	10e-11.6 10e-12.3 10e-12.1 10e-12.6	1.7 1.8 1.8 1.8	36.7 37.0 36.5 37.0
2ZTI chain A	PhNTPase from *Pyrococcus horikoshii* Ot3 with Mn(2+)	179 180 180 180	16.9 17.4 16.5 17.4	10e-16.2 10e-17.7 10e-15.3 10e-17.8	1.7 1.7 1.7 1.7	36.9 37.2 37.2 37.2
2E5X chain A	PhNTPase from *Pyrococcus horikoshii* Ot3 with ITP	178 178 180 179	17.0 17.5 16.5 17.5	10e-16.3 10e-18.0 10e-15.3 10e-17.9	1.8 1.7 1.8 1.7	37.1 37.1 37.2 36.9
2DVO chain A		180 179 180 177	16.7 17.2 16.5 17.3	10e-15.6 10e-16.9 10e-15.3 10e-17.9	1.7 1.7 1.7 1.6	36.7 36.9 37.2 37.9
2CAR chain A	Human ITPase, unliganded	173 173 172 167	11.4 11.8 13.2 11.7	10e-6.5 10e-7.1 10e-9.5 10e-6.9	2.3 2.5 2.4 2.2	26.0 26.6 26.7 25.7
2CAR chain B		169 171 169 168	11.5 11.9 13.3 11.8	10e-6.6 10e-7.3 10e-9.8 10e-7.1	2.2 2.3 2.4 2.3	27.2 27.5 26.0 26.8
2I5D chain A		169 169 170 167	15.2 15.1 13.8 14.7	10e-12.1 10e-12.0 10e-9.5 10e-11.3	2.2 2.3 2.4 2.3	26.0 26.6 26.5 26.9

Continued:						
PDB-ID chain #	**Description**	**Alignment length**	**Score**	**P-Value**	**RMSD**	**% ID**
2J4E chain A	Human ITPase with bound ITP	174	13.5	10e-10.0	1.9	29.3
		174	13.9	10e-10.8	1.9	29.3
		174	14.6	10e-12.2	1.9	29.9
		171	13.8	10e-10.7	1.8	30.4
2J4E chain B		179	16.4	10e-15.0	1.9	29.1
		178	16.9	10e-16.2	1.9	29.2
		179	15.3	10e-12.6	2.0	29.1
		179	16.9	10e-16.1	2.0	27.9
2J4E chain C		178	13.0	10e-9.1	1.9	29.2
		175	13.5	10e-10.2	1.9	29.1
		178	14.5	10e-12.0	2.1	28.1
		177	13.5	10e-10.1	2.0	28.8
2J4E chain D		179	12.9	10e-9.0	2.0	28.5
		177	13.5	10e-10.1	1.9	28.2
		180	14.4	10e-11.7	2.1	27.2
		181	13.4	10e-10.0	2.1	27.1
2J4E chain E		178	12.9	10e-8.9	1.9	29.8
		177	13.4	10e-9.9	1.8	29.4
		178	14.3	10e-11.6	2.0	29.2
		178	13.3	10e-9.7	1.9	29.2
2J4E chain F		178	12.8	10e-8.9	1.9	29.8
		178	13.5	10e-10.1	1.9	28.7
		179	14.3	10e-11.7	1.9	29.1
		179	13.6	10e-10.3	1.9	29.6
2J4E chain G		174	17.0	10e-16.3	1.8	29.3
		175	17.2	10e-17.0	1.9	29.7
		174	15.6	10e-13.3	1.8	30.5
		175	17.0	10e-16.3	1.9	29.7
2J4E chain H		172	13.1	10e-9.0	1.8	29.7
		170	13.5	10e-9.8	1.8	29.4
		171	14.5	10e-11.6	1.8	29.8
		172	13.6	10e-9.9	1.8	29.1

Comparison to dimer interfaces of other non-canonical NTPases

We used the program PISA (Krissinel & Henrick, 2007) to look for dimer interfaces that are similar to the dimer interfaces found for our TM0159 structure with bound IMP. Figure S9 illustrates the overlay of the following non-canonical NTPases: TM0159, human ITPase, and PhNTPase. Table S2 summarizes the statistics of all search results for interface-2 (see Figure 3B in the manuscript). Table S2 contains interfaces with a Q-score higher than 0.4. The Q-score is a measure for interface similarity and ranges from zero to one. Identical interfaces yield a score of one. The comparison between the interfaces of unliganded and liganded TM0159, for example, resulted in a Q-score of 0.982.

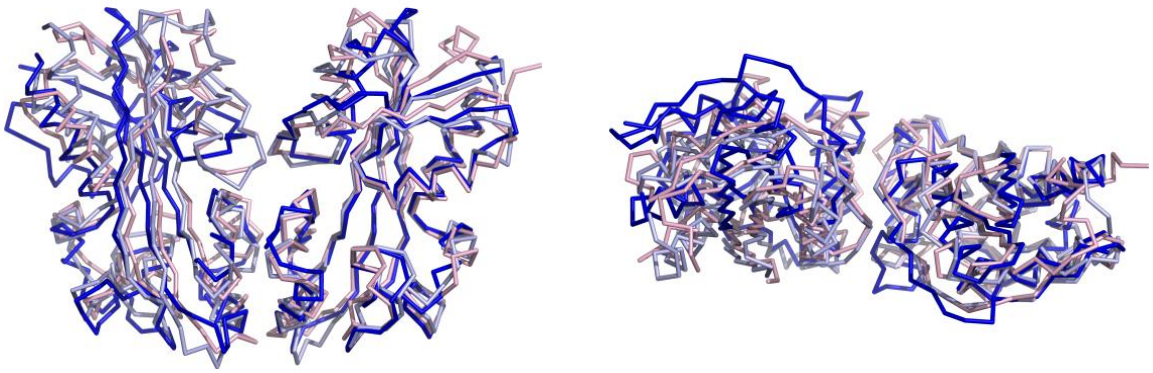


Figure S9: The dimer interface-2 of TM0159 with bound IMP (3S86, chains B &D, symmetry operation (x, y-1, z)) is colored in dark blue. The superimposed dimers of human ITPase (2J4E) and PhNTPase from *Pyrococcus horikoshii* (2DVN) are colored in light pink and light blue, respectively. The view is rotated by 90 degree.

We also searched for dimer interfaces that resemble interface-1 (see Figure 3A in the manuscript). One dimer interface of dUTP complexed Maf (PDB-ID 1exc) scored a Pisa Q-Score of 0.312 (Minasov *et al.*, 2000). The buried surface area of this interface is only 512 $Å^2$ per monomer in comparison to another potential Maf dimer interface with a buried surface area of 1100 $Å^2$ per monomer. Maf has some structural similarity to non-canonical NTPases, and it was suggested that Maf might function as a non-canonical NTPase with O6-methyl-dGTP as likely substrate (Galperin *et al.*, 2006). The structures of Maf and TM0159 might share a similar crystal packing interface.

Table S2: Search for similar interfaces using PISA (Krissinel & Henrick, 2007). Interface-2 between monomer B and D of IMP liganded TM0159 was used for this search.

Entry	Description	Intf No	mm Size	Space group	Q score	Seq. Id	Interface area, Å^2	Δ^iG † kcal/mol	CSS§
1vp2	TM0159 unliganded	3	4	P 41 21 2	0.982	1.000	626.1	-6.3	0.101
2dvp	PhNTPase from *Pyrococcus horikoshii* Ot3, unliganded	1	2	P 31 1 2	0.721	0.394	978.7	-14.0	1.000
2dvn	PhNTPase from *Pyrococcus horikoshii* Ot3, with IMP	1	4	P 41 21 2	0.711	0.394	1045.0	-13.5	0.074
2dvo	PhNTPase from *Pyrococcus horikoshii* Ot3, with ITP	1	2	C 2 2 21	0.709	0.397	1062.4	-10.8	0.302
1v7r	PhNTPase from *Pyrococcus horikoshii* Ot3, unliganded	1	2	C 2 2 21	0.708	0.385	1057.2	-11.3	0.783
2zti	PhNTPase from *Pyrococcus horikoshii* Ot3, with Mn(2+)	1	2	P 32 2 1	0.707	0.397	1092.3	-10.7	0.590
2e5x	PhNTPase from *Pyrococcus horikoshii* Ot3, with ITP	1	2	P 32 2 1	0.696	0.391	990.2	-12.8	0.158
2mjp	Mj0226 from *Methanoccus jannashii* with AMPNP	1	2	P 21 21 21	0.654	0.363	950.0	-12.9	0.689
1b78	Mj0226 from *Methanoccus jannashii*	1	2	P 21 21 21	0.650	0.358	931.2	-15.6	1.000
2j4e	Human ITPase with bound ITP	3	2	P 1	0.621	0.286	1091.2	-17.3	0.401
2j4e		2	2	P 1	0.605	0.286	1093.9	-17.0	0.401
2car	Human ITPase, unliganded	1	2	P 1 21 1	0.490	0.267	1100.5	-16.0	1.000

† Solvation free energy gain upon formation of the interface. § Complexation Significance Score.

Table S3: Comparison of amino acid composition among structurally characterized non-canonical NTPases

	Mj0026	**PhNTPase**	**TM0159**	**RdgB**	**Human ITPase**
Source organism	*Methanococcus jannaschii*	*Pyrococcus horikoshii*	*Thermotoga maritima*	*Escherichia coli*	*Homo sapiens*
Superkingdom of source organism	Archaea		*Bacteria*		*Eukaryota*
Optimal growth temperature of source organism	358 K	368 K	353 K	310 K	310 K
Number of Amino Acids	193	186	196	199	194
Amino acid composition					
Charged amino acids					
Asp	7 (3.6%)	6 (3.2%)	9 (4.6%)	14 (7.0%)	9 (4.6%)
Glu	**25 (13.0 %)**	**19 (10.2 %)**	**22 (11.2%)**	12 (6.0%)	15 (7.7 %)
Lys	**21 (10.9 %)**	**18 (9.7 %)**	**20 (10.2%)**	10 (5.0%)	13 (6.7%)
Arg	8 (4.1%)	7 (3.8%)	10 (5.1%)	10 (5.0%)	8 (4.1%)
Sum	**61 (31.6 %)**	**50 (26.9 %)**	**61 (31.1 %)**	46 (23.1 %)	45 (23.2 %)
Branched chain amino acids					
Leu	12 (6.2%)	13 (7.0%)	14 (7.1%)	**23 (11.6%)**	**21 (10.8 %)**
Ile	20 (10.4%)	16 (8.6%)	13 (6.6 %)	8 (4.0%)	7 (3.6 %)
Val	11 (5.7%)	11 (5.9%)	15 (7.7%)	15 (7.5%)	13 (6.7%)
Sum	43 (22.3 %)	40 (21.5 %)	42 (21.4 %)	**46 (23.1 %)**	41 (21.1 %)
Aromatic amino acids					
Tyr	8 (4.1%)	9 (4.8 %)	6 (3.1 %)	4 (2.0%)	6 (3.1 %)
Trp	1 (0.5%)	3 (1.6 %)	1 (0.5%)	1 (0.5%)	2 (1.0 %)
Phe	**14 (7.3%)**	**16 (8.6%)**	**12 (6.1 %)**	6 (3.0 %)	**11 (5.7%)**
His	1 (0.5%)	3 (1.6%)	5 (2.6%)	5 (2.5%)	2 (1.0%)
Sum	**24 (12.4 %)**	**31 (16.7 %)**	**24 (12.2 %)**	16 (8.0 %)	**22 (11.3 %)**
Other amino acids					
Thr	9 (4.7%)	7 (3.8%)	9 (4.6%)	12 (6.0%)	7 (3.6%)
Ser	8 (4.1%)	11 (5.9%)	12 (6.1 %)	11 (5.5%)	8 (4.1 %)
Cys	1 (0.5%)	0 (0%)	1 (0.5 %)	2 (1%)	**7 (3.6 %)**
Pro	5 (2.6%)	7 (3.8 %)	10 (5.1 %)	9 (4.5%)	13 (6.7%)
Ala	9 (4.7 %)	9 (4.8 %)	10 (5.1 %)	**24 (12.1 %)**	**17 (8.8%)**
Asn	9 (4.7 %)	5 (2.7 %)	5 (2.6 %)	4 (2.0 %)	3 (1.5%)
Gln	5 (2.6 %)	2 (1.1 %)	0 (0.0 %)	7 (3.5%)	**13 (6.7%)**
Gly	16 (8.3 %)	20 (10.8 %)	14 (7.1 %)	20 (10.1 %)	17 (8.8 %)
Met	3 (1.6 %)	4 (2.2 %)	**8 (4.1 %)**	2 (1.0 %)	2 (1.0 %)

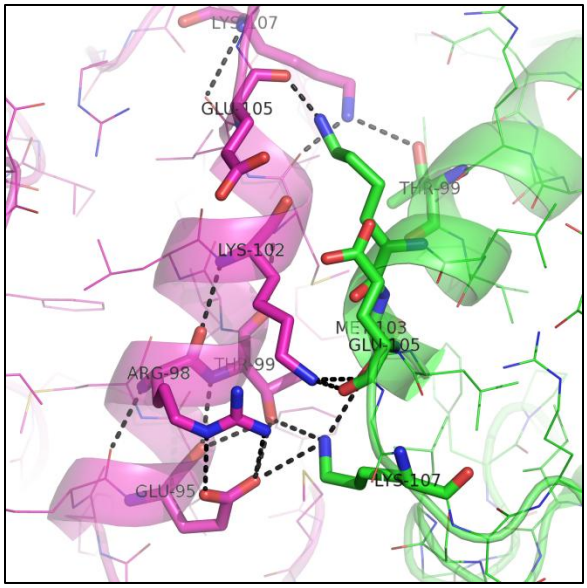


Figure S10: A network of hydrogen bonding interactions and one salt bridge between Lys107 (from monomer D) and Glu95 (from monomer B) stabilizes the dimer interface-2 of TM0159 with bound IMP (3S86, chains B (pink) &D (green), symmetry operation (x, y-1, z)).

References

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.
Krissinel, E. & Henrick, K. (2007). J Mol Biol 372, 774-797.
Marini, I. & Ipata, P. L. (2007). BAMBED 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.
Wang, Y., Geer, L. Y., Chappey, C., Kans, J. A. & Bryant, S. H. (2000). TiBS 25, 300-302.