JBC Papers in Press. Published on May 7, 2020 as Manuscript RA119.011811
The latest version is at https://www.jbc.org/cgi/doi/10.1074/jbc.RA119.011811

Crystal structures of the RNA triphosphatase from *Trypanosoma cruzi* provide insights into how it recognizes the 5' end of the RNA substrate.

Yuko Takagi[1*], Naoyuki Kuwabara[2*#], Truong Tat Dang[3],
Koji Furukawa[1] and C. Kiong Ho[3†]

Form the [1]Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki, Japan;
the [2]Structural Biology Research Center, Institute of Materials Structure Science, High Energy Accelerator Research Organization (KEK) Ibaraki, Japan;
the [3]Faculty of Medicine, University of Tsukuba, Tsukuba, Ibaraki, Japan;

Running title: *Structure of T. cruzi RNA Triphosphatase*

[*] Both authors contributed equally to this manuscript.
[#] Present Address: PeptiDream Inc., 3-25-23, Tonomachi, Kawasaki, Kanagawa, 210-0821, Japan.
† To whom correspondence should be addressed: C. Kiong Ho: Faculty of Medicine, University of Tsukuba, Tsukuba, Ibaraki, Japan 305-8575; kiongho@md.tsukuba.ac.jp; Tel. +81 (29) 853-5612

Keywords: RNA triphosphatase, mRNA cap, capping enzyme, *Trypanosoma cruzi*, crystal structure, triphosphate tunnel metalloenzyme (TTM), protozoan parasite, TcCet1

RNA triphosphatase catalyzes the first step in mRNA cap formation, hydrolysis of the terminal phosphate from the nascent mRNA transcript. The RNA triphosphatase from the protozoan parasite *Trypanosoma cruzi*, TcCet1, belongs to the family of triphosphate tunnel metalloenzymes (TTMs). TcCet1 is a promising antiprotozoal drug target because the mechanism and structure of the protozoan RNA triphosphatases are completely different from those of the RNA triphosphatases found in mammalian and arthropod hosts. Here, we report several crystal structures of the catalytically active form of TcCet1 complexed with a divalent cation and an inorganic tripolyphosphate in the active-site tunnel at 2.20–2.51 Å resolutions. The structures revealed that the overall structure, architecture of the tunnel, and arrangement of the metal-binding site in TcCet1 are similar to those in other TTM proteins. On the basis of the position of three sulfate ions that co-crystallized on the positively charged surface of the protein, and results obtained from mutational analysis, we identified an RNA-binding site in TcCet1. We conclude that the 5' end of the triphosphate RNA substrate enters the active-site tunnel directionally. The structural information reported here provides valuable insight into designing inhibitors that could specifically block the entry of the triphosphate RNA substrate into the TTM-type RNA triphosphatases of *T. cruzi* and related pathogens.

Trypanosomatid parasites belong to the order kinetoplastida that causes several neglected diseases of humans and animals affecting close to 100 million people worldwide, including African sleeping sickness caused by *Trypanosoma brucei*, Chagas disease caused by *Trypanosoma cruzi*, and a spectrum of diseases caused by various Leishmania species (1). New therapeutics against trypanosomatids are needed as current medications are generally ineffective at late-stage of infection or have severe side effects (2).

The 5'-cap (m7Gp) is an essential feature in eukaryotic cellular and viral mRNAs that functions to protect mRNA from degradation and to

promote translation initiation. RNA triphosphatase, which catalyzes the first step in mRNA capping, is a promising target for anti-protozoal drug development because the mechanism of protozoan RNA triphosphatase is completely different from mammalian or arthropod host (3, 4). Metazoan and plant RNA triphosphatases belong to the cysteine-phosphatase enzyme superfamily that catalyzes a two-step phosphoryl transfer reaction in which the active site cysteine in the phosphate-binding loop attacks the γ-phosphate of triphosphate-terminated RNA (pppRNA) to form a covalent protein-cysteinyl-S-phosphate intermediate and release the diphosphate RNA product in the absence of metal cofactor (5, 6). The RNA triphosphatases of fungi, protozoa, and several DNA viruses belong to a triphosphate tunnel metalloenzyme (TTM) family that hydrolyzes γ-phosphorus of pppRNA in the presence of magnesium, and NTP in the presence of either manganese or cobalt (7-17). The crystal structure of *Saccharomyces cerevisiae* RNA triphosphatase (Cet1) revealed that the catalytic core is located in a topologically closed hydrophilic tunnel composed of 8 antiparallel β-strands (18). The Cet1 active site constitutes of 15 essential amino acid side chains that either stabilize the topology of the tunnel, or make contacts to the divalent cation or sulfate ion (7, 19-21). The position of the sulfate ion was proposed to reflect the position of γ-phosphate of pppRNA and NTP, or the leaving group phosphate (18). Two glutamate-containing motifs (designated as motif A and C) comprise the metal binding site and are conserved among all TTM family members (22-24).

RNA triphosphatases from kinetoplastids have been characterized in *T. brucei* (TbCet1) and *T. cruzi* (TcCet1) (10, 11). Based on the primary structures of TbCet1 and TcCet1, the kinetoplastid enzymes contain all the putative counterparts of β-strands that comprise the yeast Cet1 triphosphate tunnel but lack extra domains appended to the N-terminal region that are essential for homodimerization and interaction with Ceg1 guanylyltransferase (25-27). TbCet1 is essential for procyclic cell growth (28), and can also complement the function of yeast Cet1 in *S. cerevisiae* as a cap forming enzyme (11). Several classes of small molecular weight compounds have been reported to inhibit the triphosphatase activity of TbCet1 and mutagenesis studies have illuminated functional groups that are important for catalytic activity (28-30). What remains obscure is the structural basis of inhibition and the selectivity for the RNA substrate.

In the present study, we produced a catalytically active form of TcCet1 and crystallized the enzyme in complex with divalent cation and tripolyphosphate (PPPi) bound at the active site. We propose that 5'-end of triphosphate terminus of RNA substrate directionally enters the active site tunnel through a positively charged surface based on the three sulfate ions that co-crystallized with TcCet1. The structural information could provide valuable insight into designing inhibitors that could specifically block the entry of triphosphate RNA substrate in TTM-type RNA triphosphatases.

Results

Defining the minimal catalytic unit of Trypanosoma RNA triphosphatase - Deletion and limited proteolysis analyses of TbCet1 suggest that the hydrophilic N-terminal ~20 amino acid segment is connected by protease-sensitive site and is dispensable for triphosphatase activity (30). To define the minimal catalytic core for structural analysis, we deleted several amino acids from N- and C-termini in TcCet1 [TcCet1(18-243)], and further made an internal deletion to remove a putative protease-sensitive loop [TcCet1(18-243 Δ55-75); Figure 1 and Supporting Information 1]. Similar deletions were made for TbCet1 [TbCet1(29-253) and TbCet1(29-253 Δ62-90)]. The truncated TcCet1 and TbCet1 proteins were expressed in bacteria and purified from a soluble extract by Ni-agarose chromatography, and proteins were assayed for a release of ^{32}Pi from [γ-^{32}P]ATP in the presence of 2 mM manganese. Removal of N-terminal region apparently increased the ATPase activity of TcCet1 and TbCet1, 1.3 and 2.6-fold, respectively, in comparison to their full-length enzymes (Supporting Information 1B). TcCet1(18-243 Δ55-75) and TbCet1(29- 253 Δ62-90) showed a modest reduction in the activity; 90% and 60% of full-length enzymes, respectively. TcCet1(18-243 Δ55-75) had similar activity to the full-length TcCet1 in hydrolyzing pppRNA (Supporting Information 1C). We conclude that removal of putative protease-sensitive loop region

does not significantly affect the triphosphatase activity.

Crystallization and Structure Determination of TcCet1(18-243 Δ55-75) – PPPi is a potent competitive inhibitor of TbCet1, suggesting that PPPi binds tightly to the active site of the enzyme (31). While TbCet1 is capable of hydrolyzing PPPi, tripolyphosphatase activity is only 0.5% of its NTPase activity. We therefore co-crystallized TcCet1(18-243 Δ55-75) in the presence of PPPi and manganese. (Experimental Procedures). The crystal was further soaked with potassium iodine to improve the mosaicity to obtain experimental phase information. The structure was then refined at 2.20 Å resolution with R/Rfree of 22.6% / 26.1% (Supporting Information 3). The electron density corresponding to PPPi was evident and a density for manganese was modeled adjacent to PPPi. We subsequently solved the structure of TcCet1(18-243 Δ55-75) in complex with Mn alone, by means of molecular replacement phasing method using the TcCet1(18-243 Δ55-75) Mn•PPPi structure as a search model. The root mean square deviations between pairs of structures was 0.9 Å for the main chain atoms of residues 37–242, suggests that binding of PPPi may not induce conformational change. Efforts to crystallize full-length TcCet1 and TbCet1 and other derivatives of the proteins were unsuccessful.

Two unique interfaces are formed in the structure of TcCet1(18-243 Δ55-75) (Supporting Information 3). The dimer-forming interface in the Mn-bound and Mn•PPPi-bound forms differ in the position of N-terminal α-helix (from residues Asp18 to Leu36). Analysis with the PDBePISA server (32) reveals the Mn-bound dimer interface buries 663.9 $Å^2$ with a Complex Formation Significance Score (CSS) of 0.000, suggests that this interaction occurs only in the crystal. In the Mn•PPPi-bound form, the N-terminal α-helix is extended and protrudes from one protomer to the other, with an interface area of 654.1 $Å^2$ and the calculated CSS of 0.440, suggests that the interface plays an auxiliary role in complex formation. However, identical contact is formed by the monomeric enzyme, where the N-terminal α-helix folds back by a connecting loop (Supporting Information 3C). Indeed, TcCet1 behaves as a monomer in solution at micromolar concentration, as judged by gel filtration analysis, and as previously noted (10). Thus, the two independent molecules of TcCet1 found in the asymmetric unit in Mn•PPPi-bound form can be described as a dimer formed by a crystal packing.

Structure of TcCet1 with tripolyphosphate in the active site - The overall structure of TcCet1(18-243 Δ55-75) is similar to other TTM proteins, composed of eight-stranded anti-parallel β-barrel with three α-helices surrounding the active site tunnel (Figure 2A). Structural comparison to yeast RNA triphosphatase Cet1 reveals that diameter of the tunnel and the metal binding site is nearly identical between the two enzymes (Figure 2C - 2E). Positions of 15 side-chains in Cet1 active site tunnel are conserved in TcCet1 (Figure 2B and 2E), of which 14 side-chains were shown to be essential for Cet1 (7, 19, 21) and TbCet1 (30) triphosphatase activity. Three glutamates (Glu42, Glu44, and Glu210) from motif A and C from the tunnel floor coordinate the Mn^{2+}, and allows Mn^{2+} to interact with the PPPi at P1 and P2 positions (Figure 2F). The basic side chains (Arg118 and Lys138) extended from the other side of the tunnel likely stabilize the interaction with PPPi. Lys182 and Arg184 interact with central phosphate (P2 position), and Lys138 interacts with phosphate at P3 position of PPPi (Figure 2F). A similar mode of PPPi binding and the metal arrangement was observed in the structures of Arabidopsis protein AtTTM3· Mn•PPPi complex (24, 33). In TcCet1, Arg46 also forms water-mediated contact to Glu44 and Glu210.

Based on the configuration between Mn^{2+} and PPPi, Mn^{2+} could facilitate the hydrolysis of phosphate at P1 position. Hence, we predict that γ-phosphate of pppRNA and NTP will be situated at the position of P1 phosphate. In support of this view, superimposition of TcCet1 and yeast Cet1 structures reveal that P1 phosphate of PPPi is in close proximity to the sulfate ion present in Cet1 structure (Figure 2E), which has been proposed to mimic the position of γ-phosphate of pppRNA and NTP (18).

Screen for TcCet1 ligand by fluorescent-based protein thermal stability shift assay - We performed a fluorescent-based protein thermal stability shift assay to identify potential ligand that

can interact to stabilize the TcCet1 protein. In this procedure, TcCet1(18-243 Δ55-75) was incubated with CYPRO Orange dye, which preferentially binds to the hydrophobic core of the protein and emits fluorescence when the protein is unfolded by temperature shift. We determined that Tm of TcCet1(18-243 Δ55-75) was 44.6°C, and the addition of 2 mM ATP increased the Tm by 1.4°C to 46.0°C (Figure 3A).

We screened for compounds that can bind and increase the Tm of TcCet1(18-243 Δ55-75) using an in-house chemical compound library consists of 965 commercially available low molecular weight compounds (Figure 3A). We identified four compounds that increase the Tm of TcCet1(18-243 Δ55-75) more than 1.0°C when present at 2 mM concentration [#466 ($C_{13}H_{13}NO_2$; Tm = 46.23°C), #695 ($C_{10}H_{11}N_2O_2$; Tm = 45.58°C), #860 ($C_{11}H_{12}N_2O$; Tm = 46.23°C) and #951($C_{10}H_{14}N_4O_2$; Tm = 45.58°C)]. The chemical structures of these compounds are shown in Figure 3B. In addition, we identified ten compounds that lower the Tm of TcCet1(18-243 Δ55-75) to below 37°C. The structures of these compounds are shown in Supporting Information 4.

Structure of TcCet1 in complex with ligands - We attempted to co-crystallize TcCet1(18-243, Δ55-75) with the four compounds identified from the above screen that stabilize the enzyme. Only the complex formed with compound #951 ($C_{10}H_{14}N_4O_2$) crystallized with satisfactory quality for X-ray crystal diffraction analysis. By soaking 2 mM of compound #466 ($C_{13}H_{13}NO_2$) in to $C_{10}H_{14}N_4O_2$ bound crystal, we also obtained the structure of TcCet1(18-243, Δ55-75)· $C_{13}H_{13}NO_2$. The structures of TcCet1(18-243, Δ55-75) with the ligands were solved at 2.39Å (complex with $C_{10}H_{14}N_4O_2$) or 2.51Å (complexed with $C_{13}H_{13}NO_2$) resolution using TcCet1(18-243 Δ55-75) Mn•PPPi bound structure as a search model (Figure 4A and 4B). The bound ligands were well defined in the electron-density maps (Supporting Information 5), and bound on the same surface of the protein that forms a helical loop between amino acid residues 166 -177, a region that was disordered in both Mn and Mn•PPPi-bound structures. Neither Mn nor PPPi is evident in both ligand-bound structures and the tunnel cavities are slightly narrower compared to the Mn- or Mn•PPPi-bound structures (Figure 4).

Two aromatic residues, Phe169 from the loop and Trp117 from β4-strand, interact with $C_{13}H_{13}NO_2$ and $C_{10}H_{14}N_4O_2$ through π-π stacking. The $C_{13}H_{13}NO_2$ oxo-groups are hydrogen bonded by His119 and Asp165 side chains (Figure 4C). Similarly, one of the oxo-groups in $C_{10}H_{14}N_4O_2$ forms a hydrogen bond with His119 side chain, and other oxo-group is stabilized by Arg115 (Figure 4D). Neither compound had a substantial effect on the triphosphatase activity of TcCet1, up to 1 mM concentration (Figure 4E). Point mutations in Phe169 and His119, did not have significant impact on the triphosphatase activity, although Trp117 mutation showed moderate effect with 20 - 50% of WT activity (Figure 4F and 4G). These results indicate that ligand binding site is not essential for TcCet1 activity and ligand bound structures constitute an active form of the enzyme.

Sulfate ions bound on the surface of TcCet1 reveal the entry site for triphosphate RNA - The striking feature of $C_{13}H_{13}NO_2$ and $C_{10}H_{14}N_4O_2$ bound structures is that three sulfate ions, designated as SO_4-a, SO_4-b, and SO_4-c, were bound on the same positively charged surface of the protein (Figures 5A and 5B). The distance between the three consecutive sulfur atoms is ~ 6 Å each (S-S distance), which is similar to the distance between the two phosphorus atoms on a single-stranded RNA (Figure 5C). We predict that three consecutive sulfate ions reflect the position of backbone phosphate on the RNA chain. Superposition of the ligand-bound and the PPPi-bound structures reveals that SO_4-a is situated 12.6 Å (S-P distance) from the nearest phosphate on the PPPi in the active site tunnel, suggesting that SO_4-a likely occupies the position of phosphate between 2nd and 3rd nucleotide of the RNA substrate. SO_4-b and SO_4-c may occupy the positions between 3rd and 4th, and 4th and 5th nucleotides on the RNA, respectively.

We hypothesize if the three sulfate ions found in the structure of TcCet1 reflect the position of phosphodiester bonds of the RNA substrate, mutation in the residues surrounding the bound sulfate ions may affect the RNA triphosphatase activity, but retains its ability to hydrolyze NTP. Alanine substitution was introduced at two

positively charged residues, Arg156 and Lys144, that coordinates SO_4-a and SO_4-c, respectively (Figure 5C). Phe50 and Phe79 were also selected, because these aromatic residues may participate in stacking interaction with the RNA bases. In addition, Arg-58, situated within the internal deletion of TcCet1(18-243 Δ55-75), was substituted to alanine as a positive control.

The full-length wild-type and Ala-substituted TcCet1 proteins were produced in bacteria (Figure 6A). The RNA triphosphatase and ATPase activities were assayed for the release of ^{32}Pi from 1 μM [γ-^{32}P]pppRNA and [γ-^{32}P]ATP, respectively (Figure 6B and 6C). The specific activities were calculated from the average of the slopes of the titration curves in the linear range of enzyme-dependence. Under this condition, TcCet1 can preferentially hydrolyze pppRNA, 5.4-fold higher than that for ATP (Figure 6D). The R156A mutation showed significantly reduced RNA triphosphatase activity compared to the WT enzyme. The ratio of pppRNA hydrolysis / ATP hydrolysis was 0.4, suggests that R156A is selectively impaired for hydrolyzing triphosphate RNA (13.5-fold lower compared to the WT enzyme). F50A and F79A display reduced pppRNA and ATP hydrolysis compared to the WT enzyme. The ratio of pppRNA / ATP hydrolysis by F50A and F79A was 2.4 and 0.6, respectively, implies that F79A was less active in hydrolyzing pppRNA than the ATP. K117A and R58A displayed near WT activity in hydrolyzing pppRNA and ATP. Consistent with our findings, F50A, F79A and R156A mutations show significant reduction in binding to the nucleic acid, while R58A and K144A mutants maintained wild-type binding (Supporting Information 6). Taken together, these results support the notion that three sulfate ions found in the structure of TcCet1 reflect the position of phosphodiester bonds on the RNA, and the 5'- end of the RNA likely enters the active site tunnel directionally through the positively charged nucleic-acid binding surface.

Discussion

The present study provides new insights into the structure and mechanism of TTM-RNA triphosphatase by capturing structures of catalytically active TcCet1 in complex with manganese and PPPi. The catalytic domain of TcCet1 adopts a characteristic TTM-enzyme fold with eight-stranded antiparallel β-barrel, and the arrangement of the metal-binding site is similar to other TTM-enzymes (18, 23). The RNA must bind near one of the tunnel openings to allow the 5'-triphosphate terminus to enter the active site. We further co-crystallized TcCet1 with the two phenolic compounds that were identified to stabilize the TcCet1. The crystals were grown in the presence of ammonium sulfate, and the TcCet1-ligand complexes contained three sulfate ions on an electrostatically positive surface near the tunnel entrance. Each sulfate ions are separated by ~ 6 Å, similar to the distance between consecutive phosphates in single-stranded RNA. The distance between the PPPi found in the active site tunnel and to the closest surfer atom was 12 Å. We predict that sulfate ions reflect positions of 2nd ($pppN_1pN_2\mathbf{p}N_3pN_4p$), 3rd ($pppN_1pN_2pN_3\mathbf{p}N_4p$) and 4th ($pppN_1pN_2pN_3pN_4\mathbf{p}$) backbone phosphate on the pppRNA substrate.

Two metal binding sites were identified in the structures of other TTM members, including AtTTM3 polyphosphatase (24), ygiF polyphosphatase (24), and adenylate cyclase CyaB (34). The second metal ion participates in binding of the substrate, coordinating the triphosphate, or stabilizing the diphosphate leaving group for optimal catalysis. TcCet1 may employs two metal catalytic mechanism, as previous studies suggest that TbCet1 and other RNA triphosphatase exhibit synergistic activation by magnesium and manganese (30, 35). However, only a single metal was found in the structure of TcCet1(18-243 Δ55-75). We speculate that second metal is absent in the structure because PPPi lacks the nucleoside moiety found in NTP and pppRNA. Alternatively, an extra negative charge on the PPPi, which is not present in NTP or pppRNA, may cause interference with the binding of the second metal.

Mutational analysis of selected amino acids on the sulfate binding surface identified Arg156 to be important for the RNA triphosphatase activity. Alanine substitution of Arg156 severely reduced RNA triphosphatase activity, but the mutant protein resulted in a 2-fold increase in NTPase activity. The Arg156 reside is located at the entrance of the tunnel, and is conserved as arginine or lysine in cellular TTM-type RNA triphosphatases

characterized to date, but is not present in other TTM-enzymes (13, 15, 23). TcCet1 Arg156 counterpart for Cet1 is Lys427. Substitution of Cet1 Lys427 with Glu exhibits a cold sensitive growth arrest phenotype in yeast, suggests that positive charge residue at this position could be important for the Cet1 function *in vivo* (36). Two aromatic residues, Phe50 and Phe79, in the vicinity of sulfate ions were also crucial for TcCet1 RNA triphosphatase activity, suggesting that these aromatic residues may participate in stacking interaction with the nucleoside bases on the RNA. The effect of F79A substitution appeared to be much more severe than F50A, as Phe79 is positioned closer to the tunnel entrance than Phe50 and may have greater impact on stabilizing the 5'-end of the RNA.

The TTM-enzyme fold and the active site architecture are evolutionarily conserved but have different substrate specificity to hydrolyze phosphate. We propose that 5'-end of nascent RNA enters the triphosphate tunnel directionally, through the positively charged RNA binding surface to allow the γ-phosphate to be aligned in the active site tunnel. This view is supported by cryo-EM and CX-MS analysis of yeast capping enzyme in complex with RNA polymerase II (37). In their model, the 5'-end of the newly synthesized RNA enters the Cet1 tunnel cavity in the same direction as we proposed for TcCet1. The directionality of substrate binding has been proposed to dictate the cleavage specificity of TTM (22, 24). Thiamine triphosphatase hydrolyzes thiamine triphosphate (TTP) into thiamine diphosphate and Pi, but is unable to hydrolyze pppRNA. The α-helix at the C-terminal end protrudes into the tunnel to interacts with thiamine to orient the terminal phosphate for hydrolysis reaction, in the same direction as we proposed for TcCet1 (24). Similar C-terminal plug-in helix is also present in AtTTM3 tripolyphosphatase (24, 33), but is absent in TcCet1. The C-terminal plug-in helix may also function to block the unwanted substrate from entering the tunnel cavity, which may explain why thiamine triphosphatase and AtTTM3 cannot hydrolyze pppRNA. In the crystal structure of class IV adenylate cyclase, ATP is bound in a reverse orientation in the active site tunnel, compared to the TTP of the Thiamine triphosphatase (24, 34). This conformation of ATP allows the α-phosphate to position in a suitable orientation for an in-line nucleophilic attack by ribose O3' at the cleavage site. Because TcCet1 can also hydrolyze γ-phosphate from NTP, we predict that triphosphate on the NTP will enter the tunnel in the same direction as pppRNA. In contrast, symmetric molecule such as PPPi may enter from either side of the tunnel. This may partly explain why PPPi is a potent inhibitor for the triphosphatase activity (31).

In summary, we solved the structures of TcCet1 and illuminate the mechanism on how RNA substrate is being recognized and how the triphosphate end enters the active site tunnel. The structural information of TcCet1 could be exploited in the development of effective inhibitor that could specifically block the entry of RNA substrate in TTM-type RNA triphosphatase.

Experimental Procedures

Expression plasmids for TcCet1 and TbCet1 - The bacterial expression plasmid pET44-TcCet1 encodes *T. cruzi* RNA triphosphatase fused in frame to an N-terminal His-tag with a custom HRV3C protease recognition site. This was accomplished by amplifying the TcCet1 open reading frame from *T. cruzi* (Tulahuen strain) genome by PCR and inserting into modified pET44, in which the plasmid segment encoding for Nus·Tag, S·Tag and thrombin cleavage segments were replaced with His-tag and HRV3C protease recognition site between Nde1-BamH1 sites. Truncated allele TcCet1(18-243) was generated by PCR amplification using mutagenic sense primer that replaces Gly17 with methionine and antisense-strand primer that introduced stop codon in lieu of the codons for Ser244, and the fragment was inserted between Nde1-Sal1 restriction site from pET-TcCet1 template. TcCet1(18-243 Δ55-75) allele spanning from residue 18 to 243 with an internal deletion from 55 to 75 amino acids, and point mutations in TcCet1 gene, were generated by synthetic oligonucleotides using the two-stage PCR-based overlap extension strategy. *T. brucei* RNA triphosphatase was amplified by PCR from pET-His/Smt3-TbCet1 (11) and inserted into pET44a plasmid between NdeI and SalI restriction site. Plasmids for expression of TbCet1(29-253) and TbCet1(29-253 Δ62-90) were generated by

similar strategies described for TcCet1(18-243) and TcCet1(18-243 Δ55-75). The presence of the desired mutation was confirmed in every case by sequencing the entire insert. For deletion and mutational analysis, plasmids were transformed into *E. coli* BL21 (DE3) and proteins were purified from the soluble lysates by Ni-NTA column as described (11).

Large scale purification of TcCet1(18-243 Δ55-75) - *E. coli* BL21 (DE3) transformed with pET-TcCet1(18-243 Δ55-75) was grown in LB medium (6 liters) with 0.1 mg/mL ampicillin at 37°C until the absorbance at 600 nm reached 0.6. The expression of TcCet1(18-243 Δ55-75) was induced by addition of isopropyl β-d-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM. The temperature was reduced to 16°C following induction and the cells were incubated for another 12–16 hours. Bacteria were harvested by centrifugation and resuspended in lysis buffer [50 mM Tris-HCl (pH 8.0), 0.3 M NaCl, 10 mM imidazole, 2 mM $MnCl_2$ and 1 mM DTT], sonicated on ice, and centrifuged. The supernatant was mixed with Ni-NTA sepharose resin that had been equilibrated with lysis buffer, and the suspension was mixed by continuous rotation at 4°C for 1 h. The Ni-sepharose resin was poured into a column. The packed column was washed with lysis buffer containing 20 mM imidazole, and the N-terminal His-tag was cleaved by GST-fused HRV3C protease on the column at 4°C for 16 hrs. Tag-less TcCet1(18-243 Δ55-75) was recovered from Ni-NTA column by elution buffer [50 mM Tris-HCl (pH 8.0), 0.1 M NaCl, 20 mM imidazole, 2 mM $MnCl_2$ and 1 mM DTT]. The eluate was passed through Q-sepharose (3 mL, GE healthcare) column and then to GS4B (3 mL, GE healthcare) column and the flow-through fractions were collected for each step. The GS4B flow-through fraction was concentrated by Amicon-Ultra (M.W.C.O = 10 kDa) between 20-30 mg/ml, and approximately 15 mg of protein was applied to Superdex75 gel filtration column (GE healthcare) equilibrated with 10 mM HEPES (pH 7.5), 2 mM $MnCl_2$ and 100 mM NaCl. Peak fractions containing the recombinant protein were collected, concentrated to ~ 15 mg/ml, and stored at -80°C. Purified TcCet1(18-243 Δ55-75) was used for crystallization and thermal shift assay.

Crystallization, data collection, and structure determination - Crystallization was performed by hanging drop vapor diffusion method at 20°C. Purified TcCet1(18-243 Δ55-75) (7.5 mg/ml) with a drop size of 1 µL was mixed with 1 µL of reservoir solution. Mn-bound form of crystal was obtained in 0.1 M HEPES-NaOH (pH 7.0 - 7.5), 14–20% PEG 20000. Tripolyphosphate (PPPi) was added to a protein sample at 2.5 mM concentration prior to mixing with reservoir solution. The Mn•PPPi bound form was crystalized in 0.1 M sodium acetate (pH 5.0) and 14–18% 1,4-butandiol. Compound #951 ($C_{10}H_{14}N_4O_2$: 1,3-dimethyl-7-propyl-purine-2,6-dione) was purchased from (ChemBridge Corp.) and was added to a protein sample at 2 mM concentration with 2% DMSO, and then crystallized in 0.1 M HEPES (pH 7.5), 1.4–1.6 M ammonium sulfate. Compound #466 ($C_{13}H_{13}NO_2$: 3,4,6,7-tetrahydroacridine-1,8(2H,5H)-dione) was purchased from (Maybridge, Fisher Scientifics). TcCet1(18-243 Δ55-75) •$C_{10}H_{14}N_4O_2$ crystal was soaked with 2 mM $C_{13}H_{13}NO_2$ twice in 0.1 M HEPES (pH 7.5), 1.6 M ammonium sulfate, 2% DMSO and 2 mM $MnCl_2$ for 1 hour to obtain the $C_{13}H_{13}NO_2$ bound form structure. Prior to data collection, crystals were transferred to 35% Xylitol containing reservoir solution for cryo-protection.

Diffraction data was collected at BL15A in NSRRC (Hsinchu, Taiwan) or BL-1A in Photon Factory (Tsukuba, Japan). Data were integrated by XDS (38) and scaled by aimless in CCP4 software package (39). For I-SAD data collection, the crystal was soaked with 0.25 M potassium iodine for 30 min at 293K. Phenix AutoSol was used for the phase determination (40). In the Mn•PPPi complex, seven iodine sites in the asymmetric unit were located and used for phase determination and improvement, which yielded a traceable electron density map. The initial figure of merit was 0.327. Initial model building was performed by Phenix AutoBuild. Iterative rounds of manual adjustment in Coot and refinement using Phenix.refine were performed for structure refinement. B-factor of phosphate atoms (α, β and γ) were 89.21, 88.30, and 99.52, respectively.

Other crystal structures were solved by molecular replacement using Phaser in Phenix. Iodine and Mn•PPPi -bound form structures were

used as a search model for the molecular replacement. The structures were refined to the indicated statistics using iterated rounds of manual adjustments in Coot (41, 42), followed by refinement using Phenix (40). Final coordinates and structure factors were submitted to the PDB under accession code 6L7W (Mn complex), 6L7V (Mn•PPPi complex), 6L7Y ($C_{13}H_{13}NO_2$ complex) and 6L7X ($C_{10}H_{14}N_4O_2$ complex).

Thermal Shift Assay - Library of 965 chemical compounds were assembled in-house from commercially available low molecular weight compounds based on 'Rule of Three' (molecular weight <300, the number of hydrogen bond donors ≤3, the number of hydrogen bond acceptors ≤3 and ClogP is ≤3), as described (43). Compounds were stored desiccated and were resuspended with DMSO to 100 mM prior to the assay. Thermal Shift Assay was carried out with 1 μg of TcCet1-(18-243 Δ55-75) and CYPRO Orange fluorescent dye (Thermo Fisher Scientific) (1:1000 dilution) in 96 well plates. Reaction mixture contained 10 mM HEPES (pH 7.5), 2 mM $MnCl_2$, and 1 mM DTT, and 2 mM of each compound. The temperature of the plate was raised from 25°C to 70°C by an increment of 0.5°C / 5 sec using StepOne Plus Real-time PCR system (Thermo Fisher Scientific). A resulting melting curve was analyzed by Protein Thermal Shift Software v1.1 (Thermo Fisher Scientific) to determine ΔTm.

Preparation of triphosphate terminated RNA - Substrate for RNA triphosphatase assay was prepared by *in vitro* transcription with T7 RNA polymerase from a partially duplexed oligo DNA (44). Oligonucleotides corresponding to T7 promoter sequence (5'TAATACGACTCACTATA3') and complementary sequence with poly(dT) stretch (5'TTTTTTTTTTTTTTTTTTTTATAGTGAGT CGTATTA3') were annealed to produce a template for T7 RNA polymerase. The reaction mixture contained 1 mM of [γ-^{32}P] ATP, 20 μM of duplex oligo DNA, 1 unit/μL of RNase inhibitor (Toyobo Co., Ltd., Japan) and 2.5 units/μL of T7 RNA polymerase (New England Biolabs Inc., USA) in provided reaction buffer. The reaction was carried out for 3 h at 37°C. DNase I (Nippon Gene Co., Ltd., Japan) was then added to digest the template oligo DNAs, according to manufacturer's instruction. The product was extracted with phenol/chloroform/isoamyl alcohol, precipitated with ethanol, and was separated on 20 % denaturing polyacrylamide gel. The radioactive product that migrates between 7 – 9 nts was excised from the gel and eluted by TE. The eluate was passed through a spin column to remove gel debris. The RNA was precipitated with ethanol and was resuspended in water.

Triphosphatase Assay - Reaction mixture containing 50 mM Tris-HCl (pH 7.5), 2 mM DTT, 2 mM of divalent cation ($MgCl_2$ or $MnCl_2$) with either [γ-^{32}P]ATP or [γ-^{32}P]pppRNA, and indicated amount of protein, were incubated for 15 min at 30°C. The reaction was quenched by the addition of EDTA or formic acid. The reaction products were separated by PEI-cellulose thin layer chromatography (TLC) with 0.45 M ammonium sulfate. The TLC plate was exposed to a PhosphorImager plate, scanned by BAS-2000 (FujiFilm, Japan), and quantitated by Image Gauge software.

Data Availability - The atomic coordinates and structure factors have been deposited into the Protein Data Bank (http://www.rcsb.org) under PDB entries 6L7V, 6L7W, 6L7X and 6L7Y. All other information and data are available from the authors upon request.

Acknowledgments

We thank Osamu Tani from AIST for performing thermal shift assay, Yanchang Liu and Moe Goto from University of Tsukuba for their help with experiments, beamline staffs of Photon factory (Tsukuba, Japan) and NSRRC (Shinchu, Taiwan) for X-ray crystallographic analysis, Dr. Toshiya Senda from KEK and Dr. Hitoshi Sakashita from AIST for valuable discussion.

Funding

This material is based upon work supported by the JSPS Grants-in-Aid for Scientific Research KAKENHI Grant Number 16H05180 (C.K.H.) and 18K15141 (Y.T).

Conflict of interest statement. None declared

References

1. McCall, L.-I., and McKerrow, J. H. (2014) Determinants of disease phenotype in trypanosomatid parasites. *Trends Parasitol*. **30**, 342–349
2. Field, M. C., Horn, D., Fairlamb, A. H., Ferguson, M. A. J., Gray, D. W., Read, K. D., De Rycker, M., Torrie, L. S., Wyatt, P. G., Wyllie, S., and Gilbert, I. H. (2017) Anti-trypanosomatid drug discovery: an ongoing challenge and a continuing need. *Nat Rev Microbiol*. **15**, 217–231
3. Shuman, S. (2001) The mRNA capping apparatus as drug target and guide to eukaryotic phylogeny. *Cold Spring Harb Symp Quant Biol*. **66**, 301–312
4. Shuman, S. (2002) What messenger RNA capping tells us about eukaryotic evolution. *Nat Rev Mol Cell Biol*. **3**, 619–625
5. Takagi, T., Moore, C. R., Diehn, F., and Buratowski, S. (1997) An RNA 5'-triphosphatase related to the protein tyrosine phosphatases. **89**, 867–873
6. Changela, A., Ho, C. K., Martins, A., Shuman, S., and Mondragón, A. (2001) Structure and mechanism of the RNA triphosphatase component of mammalian mRNA capping enzyme. *EMBO J*. **20**, 2575–2586
7. Ho, C. K., Pei, Y., and Shuman, S. (1998) Yeast and viral RNA 5' triphosphatases comprise a new nucleoside triphosphatase family. *J Biol Chem*. **273**, 34151–34156
8. Gross, C. H., and Shuman, S. (1998) RNA 5'-triphosphatase, nucleoside triphosphatase, and guanylyltransferase activities of baculovirus LEF-4 protein. *J Virol*. **72**, 10020–10028
9. Pei, Y., Lehman, K., Tian, L., and Shuman, S. (2000) Characterization of Candida albicans RNA triphosphatase and mutational analysis of its active site. *Nucleic Acids Res*. **28**, 1885–1892
10. Massayuki Kikuti, C., Tersariol, I. L. S., and Schenkman, S. (2006) Divalent metal requirements for catalysis and stability of the RNA triphosphatase from Trypanosoma cruzi. *Mol Biochem Parasitol*. **150**, 83–95
11. Ho, C. K., and Shuman, S. (2001) Trypanosoma brucei RNA triphosphatase. Antiprotozoal drug target and guide to eukaryotic phylogeny. *J Biol Chem*. **276**, 46182–46186
12. Ho, C. K., Gong, C., and Shuman, S. (2001) RNA triphosphatase component of the mRNA capping apparatus of Paramecium bursaria Chlorella virus 1. *J Virol*. **75**, 1744–1750
13. Hausmann, S., Vivarès, C. P., and Shuman, S. (2002) Characterization of the mRNA capping apparatus of the microsporidian parasite Encephalitozoon cuniculi. *J Biol Chem*. **277**, 96–103
14. Ho, C. K., and Shuman, S. (2001) A yeast-like mRNA capping apparatus in Plasmodium falciparum. *Proc Natl Acad Sci USA*. **98**, 3050–3055
15. Kyrieleis, O. J. P., Chang, J., la Peña, De, M., Shuman, S., and Cusack, S. (2014) Crystal structure of vaccinia virus mRNA capping enzyme provides insights into the mechanism and evolution of the capping apparatus. *Structure.* **22**, 452–465
16. Benarroch, D., Smith, P., and Shuman, S. (2008) Characterization of a trifunctional mimivirus mRNA capping enzyme and crystal structure of the RNA triphosphatase domain. *Structure*. **16**, 501–512
17. Hausmann, S., Altura, M. A., Witmer, M., Singer, S. M., Elmendorf, H. G., and Shuman, S. (2005) Specificity and mechanism of RNA cap guanine-N2 methyltransferase (Tgs1). *J Biol Chem*. **280**, 12077–12086

18. Lima, C. D., Wang, L. K., and Shuman, S. (1999) Structure and mechanism of yeast RNA triphosphatase: an essential component of the mRNA capping apparatus. *Cell*. **99**, 533–543
19. Pei, Y., Ho, C. K., Schwer, B., and Shuman, S. (1999) Mutational analyses of yeast RNA triphosphatases highlight a common mechanism of metal-dependent NTP hydrolysis and a means of targeting enzymes to pre-mRNAs in vivo by fusion to the guanylyltransferase component of the capping apparatus. *J Biol Chem*. **274**, 28865–28874
20. Bisaillon, M., and Shuman, S. (2001) Functional groups required for the stability of yeast RNA triphosphatase in vitro and in vivo. *J Biol Chem*. **276**, 30514–30520
21. Bisaillon, M., and Shuman, S. (2001) Structure-function analysis of the active site tunnel of yeast RNA triphosphatase. *J Biol Chem*. **276**, 17261–17266
22. Jain, R., and Shuman, S. (2008) Polyphosphatase activity of CthTTM, a bacterial triphosphate tunnel metalloenzyme. *J Biol Chem*. **283**, 31047–31057
23. Bettendorff, L., and Wins, P. (2013) Thiamine triphosphatase and the CYTH superfamily of proteins. *FEBS J.* **280**, 6443–6455
24. Martinez, J., Truffault, V., and Hothorn, M. (2015) Structural Determinants for Substrate Binding and Catalysis in Triphosphate Tunnel Metalloenzymes. *J Biol Chem*. **290**, 23348–23360
25. Gu, M., Rajashankar, K. R., and Lima, C. D. (2010) Structure of the Saccharomyces cerevisiae Cet1-Ceg1 mRNA capping apparatus. *Structure*. **18**, 216–227
26. Ho, C. K., Lehman, K., and Shuman, S. (1999) An essential surface motif (WAQKW) of yeast RNA triphosphatase mediates formation of the mRNA capping enzyme complex with RNA guanylyltransferase. *Nucleic Acids Res*. **27**, 4671–4678
27. Lehman, K., Schwer, B., Ho, C. K., Rouzankina, I., and Shuman, S. (1999) A conserved domain of yeast RNA triphosphatase flanking the catalytic core regulates self-association and interaction with the guanylyltransferase component of the mRNA capping apparatus. *J Biol Chem*. **274**, 22668–22678
28. Smith, P., Ho, C. K., Takagi, Y., Djaballah, H., and Shuman, S. (2016) Nanomolar Inhibitors of Trypanosoma brucei RNA Triphosphatase. *MBio*. **7**, e00058–16
29. Antczak, C., Shum, D., Radu, C., Seshan, V. E., and Djaballah, H. (2009) Development and validation of a high-density fluorescence polarization-based assay for the trypanosoma RNA triphosphatase TbCet1. *Comb. Chem. High Throughput Screen.* **12**, 258–268
30. Gong, C., Martins, A., and Shuman, S. (2003) Structure-function analysis of Trypanosoma brucei RNA triphosphatase and evidence for a two-metal mechanism. *J Biol Chem*. **278**, 50843–50852
31. Gong, C., and Shuman, S. (2002) Chlorella virus RNA triphosphatase. Mutational analysis and mechanism of inhibition by tripolyphosphate. *J Biol Chem*. **277**, 15317–15324
32. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol*. **372**, 774–797
33. Moeder, W., Garcia-Petit, C., Ung, H., Fucile, G., Samuel, M. A., Christendat, D., and Yoshioka, K. (2013) Crystal structure and biochemical analyses reveal that the Arabidopsis triphosphate tunnel metalloenzyme AtTTM3 is a tripolyphosphatase involved in root development. *Plant J.* **76**, 615–626
34. Gallagher, D. T., Kim, S.-K., Robinson, H., and Reddy, P. T. (2011) Active-site structure of class IV adenylyl cyclase and transphyletic mechanism. *J Mol Biol*. **405**, 787–803
35. Martins, A., and Shuman, S. (2003) Mapping the triphosphatase active site of baculovirus mRNA capping enzyme LEF4 and evidence for a two-metal mechanism. *Nucleic Acids Res*. **31**, 1455–1463
36. Takase, Y., Takagi, T., Komarnitsky, P. B., and Buratowski, S. (2000) The essential interaction between yeast mRNA capping enzyme subunits is not required for triphosphatase function in vivo. *Mol Cell Biol*. **20**, 9307–9316

37. Martinez-Rucobo, F. W., Kohler, R., van de Waterbeemd, M., Heck, A. J. R., Hemann, M., Herzog, F., Stark, H., and Cramer, P. (2015) Molecular Basis of Transcription-Coupled Pre-mRNA Capping. *Mol Cell*. **58**, 1079–1089
38. Kabsch, W. (2010) XDS. *Acta Crystallogr D*. **66**, 125–132
39. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D*. **50**, 760–763
40. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J., and Adams, P. D. (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol*. **75**, 861–877
41. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr D*. **66**, 486–501
42. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D*. **60**, 2126–2132
43. Congreve, M., Carr, R., Murray, C., and Jhoti, H. (2003) A “rule of three” for fragment-based lead discovery? *Drug Discov Today*. **8**, 876–877
44. Ogino, T., and Banerjee, A. K. (2007) Unconventional mechanism of mRNA capping by the RNA-dependent RNA polymerase of vesicular stomatitis virus. *Mol Cell*. **25**, 85–97

Figure Legends

Figure 1. Structure-based alignment of *T. cruzi* RNA triphosphatase. Secondary structure of *T. cruzi* RNA Triphosphatase (TcCet1) is shown above the amino acid sequence. TcCet1 amino acids shaded in blue correspond to the segments that are removed in TcCet1(18-243 Δ55-75) protein. TcCet1 sequence shaded in green (amino acid residues 166–177) correspond to the segment that interacts with $C_{13}H_{13}NO_2$ and $C_{10}H_{14}N_4O_2$, but was disordered in TcCet1(18-243 Δ55-75) Mn and Mn•PPPi-bound structures. The amino acid sequences of TcCet1 is aligned with *L. major* (LmCet1), *T. brucei* (TbCet1), *S. pombe* (SpPct1) and *S. cerevisiae* (ScCet1) RNA triphosphatases. The secondary structure of ScCet1 is indicated below the aligned sequences. Identical side-chains found in all polypeptides are highlighted in red. Amino acids with similar side chains are highlighted in gray. Position of conserved motifs A and C, located within the catalytic domain of the metal-dependent RNA triphosphatases are indicated. The alignment was prepared by ESPript. The secondary structure assignment was based on DSSP program

Figure 2: Structure of TcCet1(18-243 Δ55-75) in complex with tripolyphosphate in the active site. (A) Crystal structure of TcCet1(18-243 Δ55-75) is depicted as cartoon models. Tripolyphosphate (PPPi) is shown as stick, and manganese (Mn) is shown as magenta sphere. (B) Cross section of the triphosphate tunnel of TcCet1(18-243 Δ55-75). (C) The structure of TcCet1(18-243 Δ55-75) colored in green was superimposed on to yeast Cet1(210-549) colored in yellow [(18), PDB 1D8H)]. The sulfate and Mn in Cet1(210-549) are depicted as yellow stick and yellow sphere, respectively. (D) The amino acids of TcCet1 that correspond to active site residues of Cet1, as judged from the structural comparison. Arg180 side chain was not visible in Mn•PPPi-bound structure, and therefore was modeled from the Mn-bound structure. (E) Superimposition of triphosphate tunnel of TcCet1(18-243 Δ55-75) and Cet1(210-549). The image was prepared with PyMOL. (F) The figure highlights the network of bonding interactions that coordinate PPPi, Mn and two water molecules (red sphere) in TcCet1(18-243 Δ55-75). The images were prepared with PyMOL and LigPlus.

Figure 3. Screening of small molecule that stabilizes TcCet1. (A) Melting temperature (Tm) of TcCet1(18-243 Δ55-75) in the presence of 2 mM compounds were determined by thermal shift assay. Tm of TcCet1(18-243 Δ55-75) with DMSO (control) is indicated by the red line (44.6 °C), and with ATP is indicated by blue line (46.0°C). Four compounds (#466, #695, #860 and #951) that significantly increase the Tm are indicated. (B) Structures of compound #466 ($C_{13}H_{13}NO_2$: 3,4,6,7-tetrahydroacridine-1,8(2H,5H)-dione], compound #695 ($C_{10}H_{11}N_2O_2$: 2-(1H-benzimidazol-1-yl)propanoic acid), compound #860 ($C_{11}H_{12}N_2O$: 3-ethyl-2-methyl-4H-pyrido[1,2-a]pyrimidin-4-one), and compound #951 ($C_{10}H_{14}N_4O_2$: 1,3-dimethyl-7-propyl-purine-2,6-dione).

Figure 4. Structures of TcCet1 in complex with $C_{10}H_{14}N_4O_2$ and $C_{13}H_{13}NO_2$. (A) Crystal structures of TcCet1(18-243 Δ55-75) with $C_{13}H_{13}NO_2$ (colored in violet) and (B) $C_{10}H_{14}N_4O_2$ (colored in wheat) was superimposed with the TcCet1(18-243 Δ55-75) Mn•PPPi bound complex (colored in green). Two segments, amino acids 35-39 and 166 -177, were ordered in both $C_{13}H_{13}NO_2$ and $C_{10}H_{14}N_4O_2$ bound structures but was disordered in Mn•PPPi-bound structure. Close up view of (C) $C_{13}H_{13}NO_2$ and (D) $C_{10}H_{14}N_4O_2$ binding site on TcCet1(18-243 Δ55-75). The images were prepared with PyMOL and LigPlus. (E) Effect of $C_{13}H_{13}NO_2$ or $C_{10}H_{14}N_4O_2$ on triphosphatase activity. Reaction mixtures (10 µL) containing 50 mM Tris-HCl, pH 7.5, 5 mM DTT, 1 mM $MnCl_2$, 0.2 mM [γ-^{32}P]ATP, 4 ng of TbCet1, and 0, 0.002, 005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5 or 1.0 mM of either $C_{13}H_{13}NO_2$ (blue circle) or $C_{10}H_{14}N_4O_2$ (orange circle), were incubated for 15 min at 30°C. The extents of ATP hydrolysis were plotted as a function of ligand concentration. The data shown represent the average of three separate experiments with SE bars. (F) Aliquots (5 µg) of native wild-type (WT) TcCet1 and mutant proteins were analyzed by SDS-PAGE. Polypeptides were visualized by staining with Coomassie Blue dye. Position of WT and mutant TcCet1 polypeptide is denoted by an arrow. (G) ATPase activity. Reaction mixtures (10 µL) containing 50 mM

Tris-HCl (pH 7.5), 5 mM DTT, 1 mM $MnCl_2$, 0.2 mM [γ-^{32}P]ATP, and indicated amount of WT and mutant TcCet1 proteins were incubated for 15 min at 30 °C. The extents of ATP hydrolysis were plotted as a function of input protein. The data shown represent the average of three separate experiments with SE bars.

Figure 5. Sulfate ions bound on the positive surface of TcCet1 (A) The space-filling surface images of TcCet1(18-243 Δ55-75) bound to $C_{13}H_{13}NO_2$ was prepared with the program GRASP. Positive surface charge potential is shown in blue, and negative potential in red. $C_{13}H_{13}NO_2$ and sulfate ions are shown as stick models. (B) Cartoon model of panel (A). PPPi and Mn were modeled on to TcCet1(18-243 Δ55-75) •$C_{13}H_{13}NO_2$ structure from the structure of PPPi •Mn bound complex. (C) Close-up view of SO_4 binding site. Distance between the three sulfur atoms and the distance between sulfur on SO_4-(a) to the nearest phosphorous of PPPi are indicated by dotted lines. Amino acid residues 54 to 76 are colored in red to indicate the region that was deleted in TcCet1, and the asterisk denotes a probable position of Arg58.

Figure 6. Mutational Analysis of putative RNA binding surface of TcCet1. (A) Aliquots of the nickel-agarose preparations of the wild-type and mutant TcCet1 proteins (5 μg) indicated were analyzed by SDS-PAGE, and visualized by staining with Coomassie Blue dye. (B) RNA triphosphatase activity. Reaction mixtures contained 1 μM of [γ-^{32}P]triphosphate-terminated RNA, 2 mM $MgCl_2$ with WT or mutant proteins as specified. The data shown represent the average of three separate experiments with SE bars. (C) ATPase activity. Reaction mixtures contained 1 μM of [γ-^{32}P]ATP, 2 mM $MnCl_2$ with WT or mutant proteins as specified. The extent of Pi release is plotted as a function of input enzyme. The data shown represent the average of three separate experiments with SE bars. (D) Efficiency of RNA hydrolysis vs ATP hydrolysis. The specific activity of WT and mutant TcCet1 RNA triphosphatase and ATPase activates were calculated from the data presented in (B) and (C), and expressed as a ratio of RTPase / ATPase activates and plotted for each protein indicated.

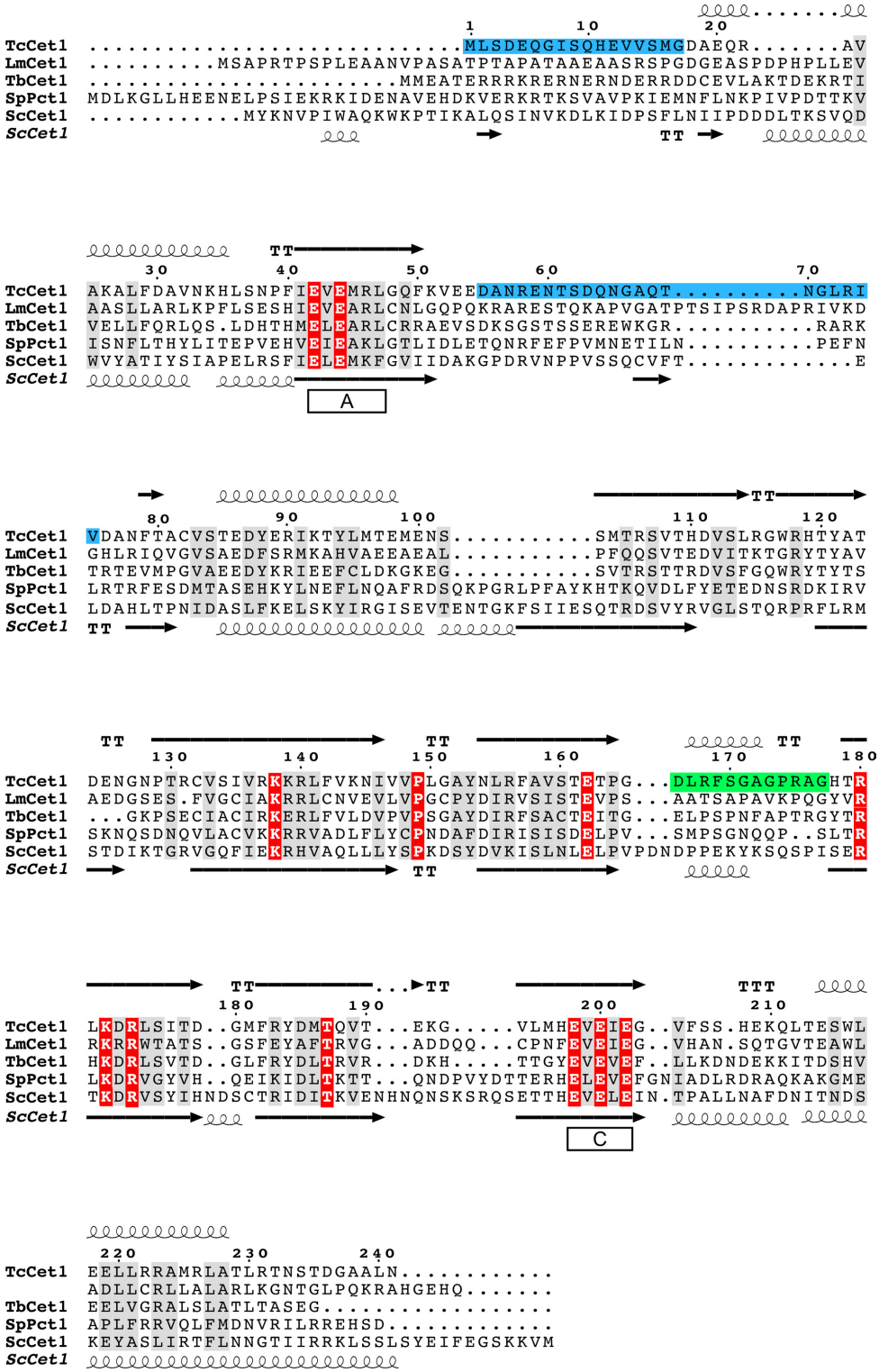
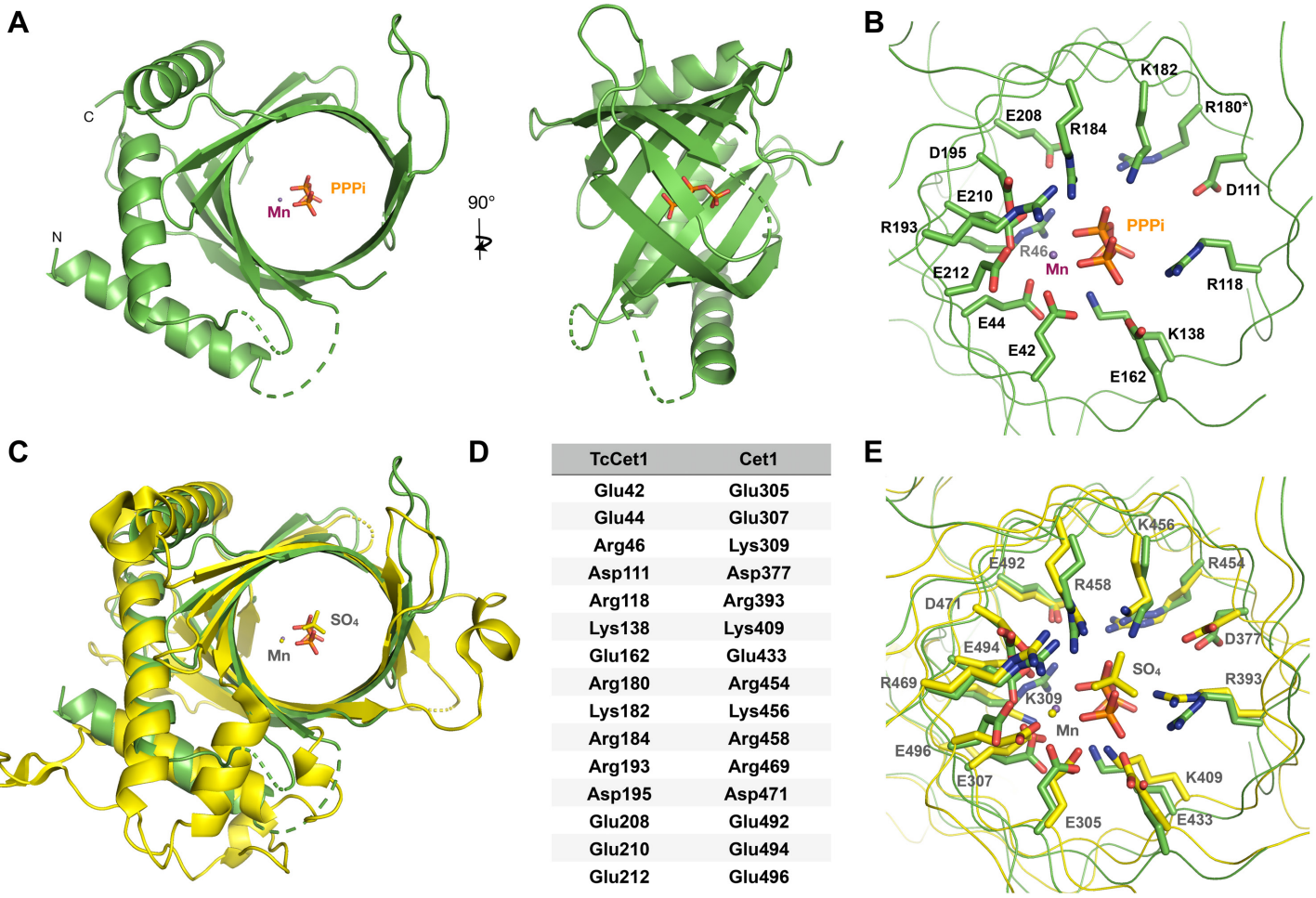


Figure 1



TcCet1	Cet1
Glu42	Glu305
Glu44	Glu307
Arg46	Lys309
Asp111	Asp377
Arg118	Arg393
Lys138	Lys409
Glu162	Glu433
Arg180	Arg454
Lys182	Lys456
Arg184	Arg458
Arg193	Arg469
Asp195	Asp471
Glu208	Glu492
Glu210	Glu494
Glu212	Glu496

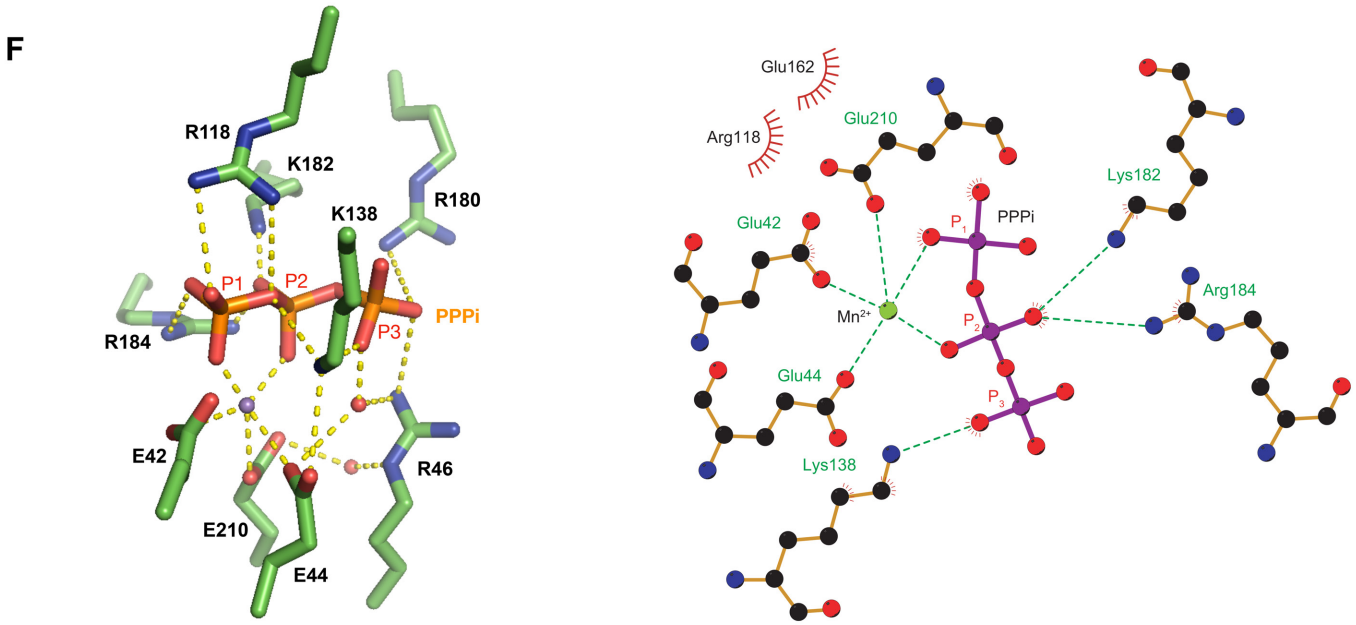
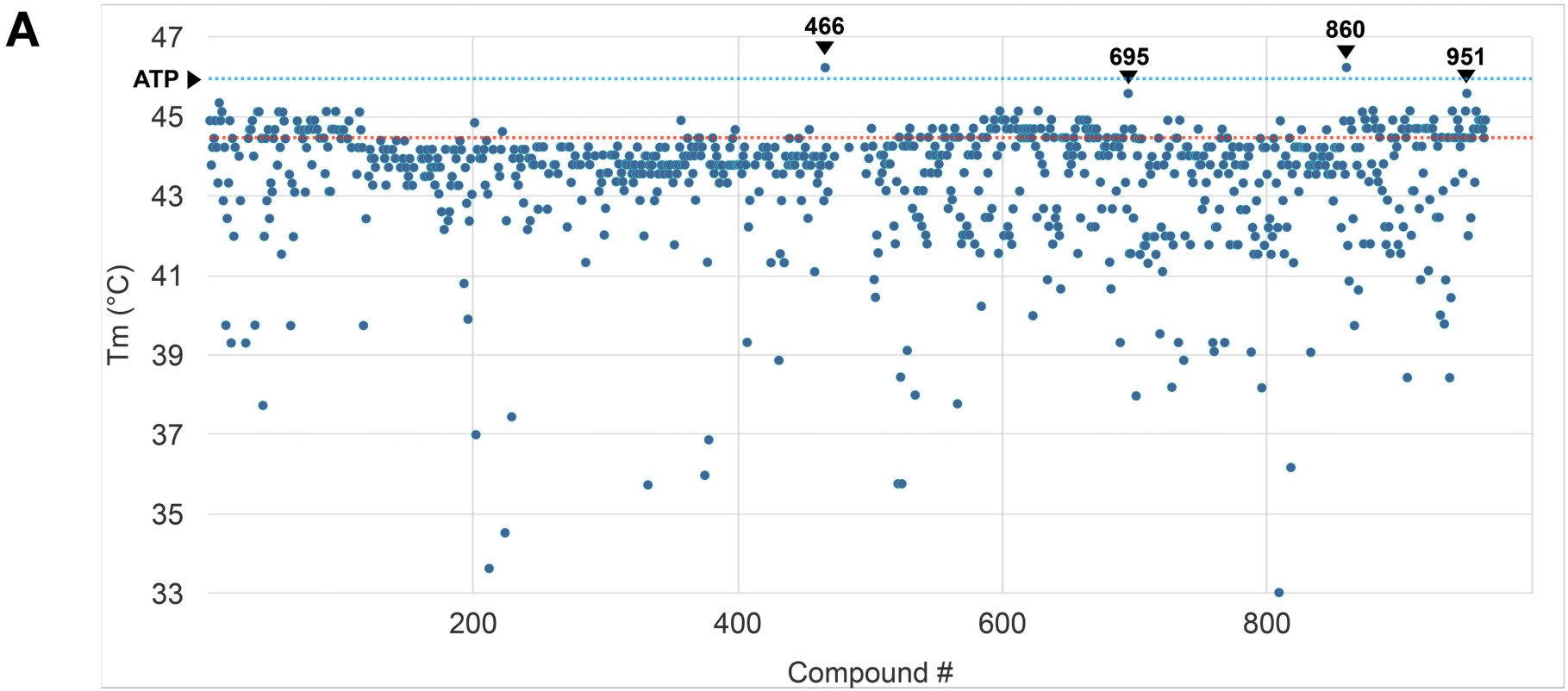


Figure 2



Compound # 466
($C_{13}H_{13}NO_2$)

Compound # 695
($C_{10}H_{11}N_2O_2$)

Compound # 860
($C_{11}H_{12}N_2O$)

Compound # 951
($C_{10}H_{14}N_4O_2$)

Figure 3

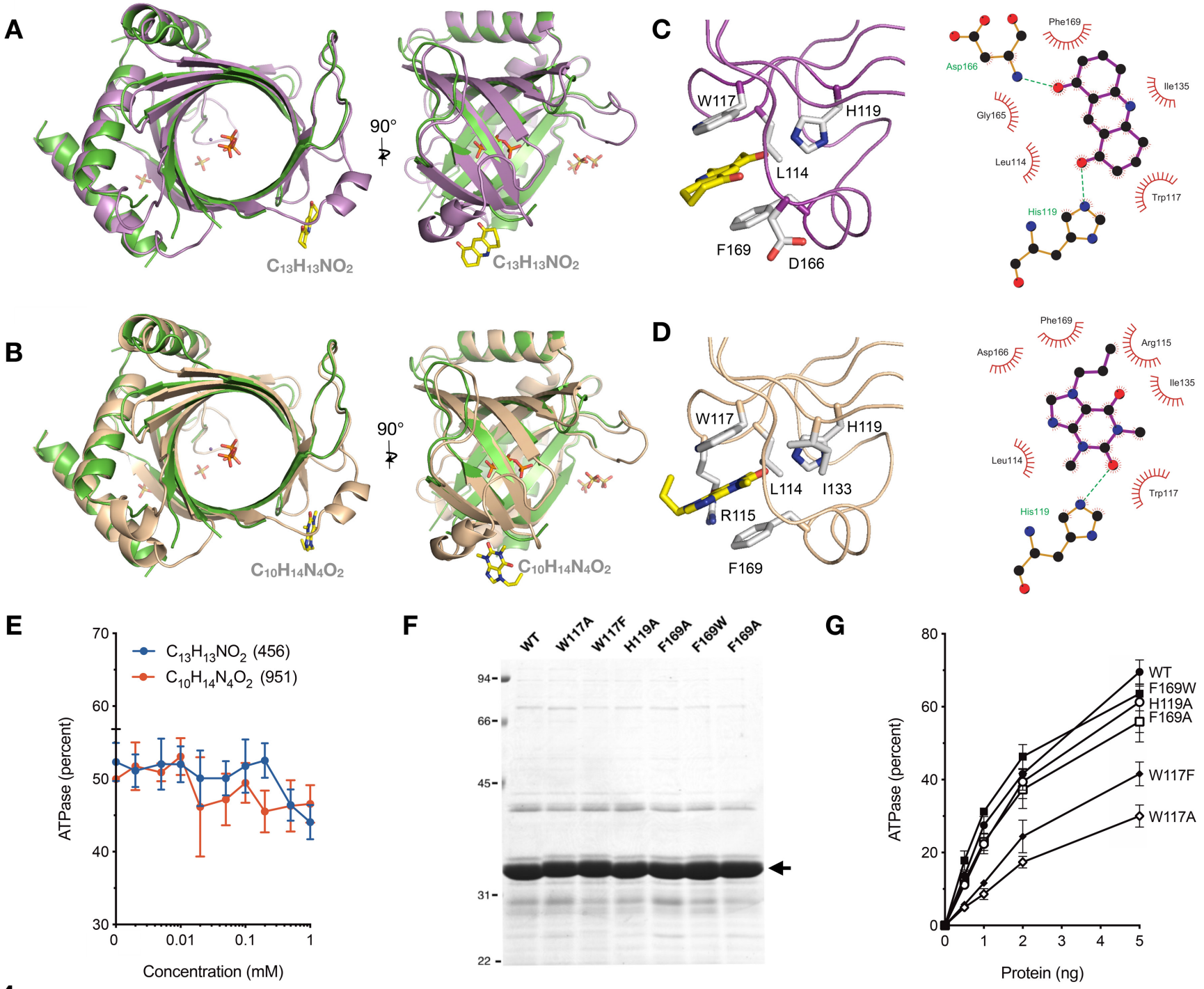

A
90°
$C_{13}H_{13}NO_2$
$C_{13}H_{13}NO_2$
B
90°
$C_{10}H_{14}N_4O_2$
$C_{10}H_{14}N_4O_2$
C
W117
H119
L114
F169
D166
Phe169
Asp166
Ile135
Gly165
Leu114
Trp117
His119
D
W117
H119
L114
I133
R115
F169
Phe169
Asp166
Arg115
Ile135
Leu114
Trp117
His119
E
$C_{13}H_{13}NO_2$ (456)
$C_{10}H_{14}N_4O_2$ (951)
ATPase (percent)
Concentration (mM)
F
WT
W117A
W117F
H119A
F169A
F169W
F169A
94
66
45
31
22
G
WT
F169W
H119A
F169A
W117F
W117A
ATPase (percent)
Protein (ng)

Figure 4

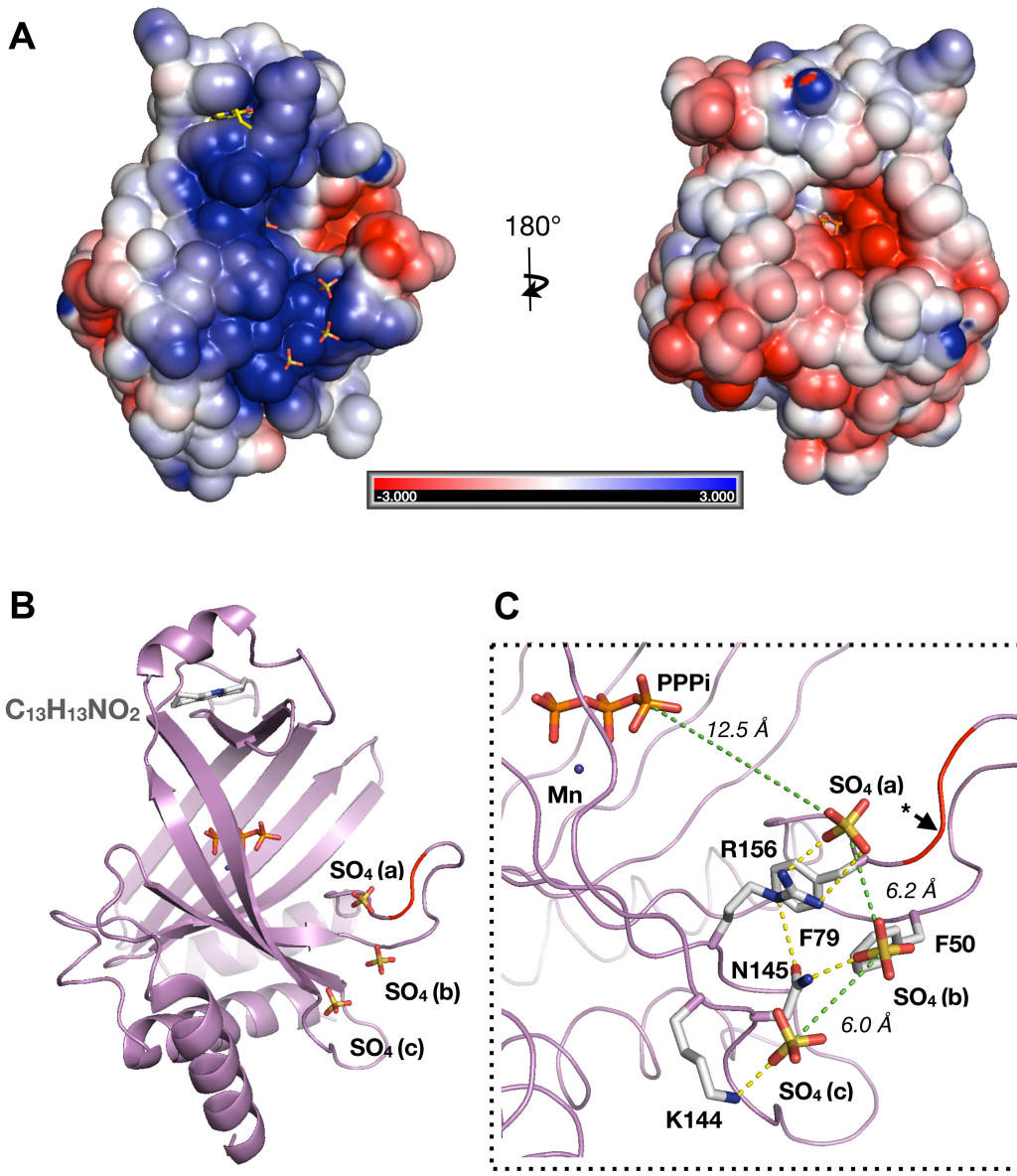

A
180°
-3.000
3.000
B
C13H13NO2
SO4 (a)
SO4 (b)
SO4 (c)
C
PPPi
12.5 Å
Mn
SO4 (a)
*
R156
6.2 Å
F79
F50
N145
SO4 (b)
6.0 Å
SO4 (c)
K144

Figure 5

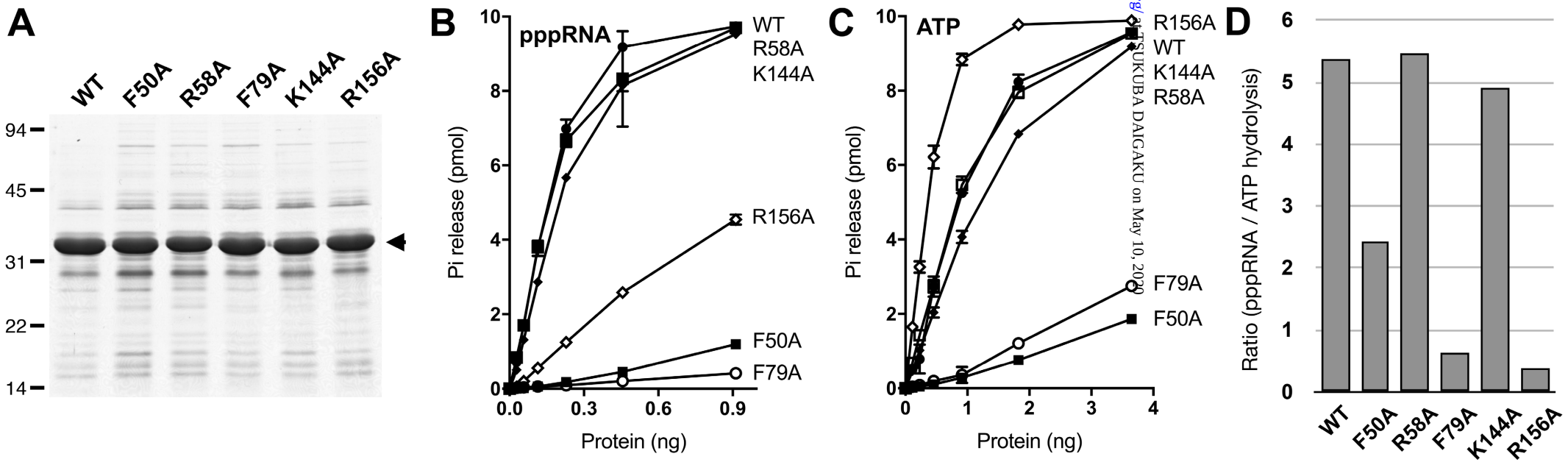
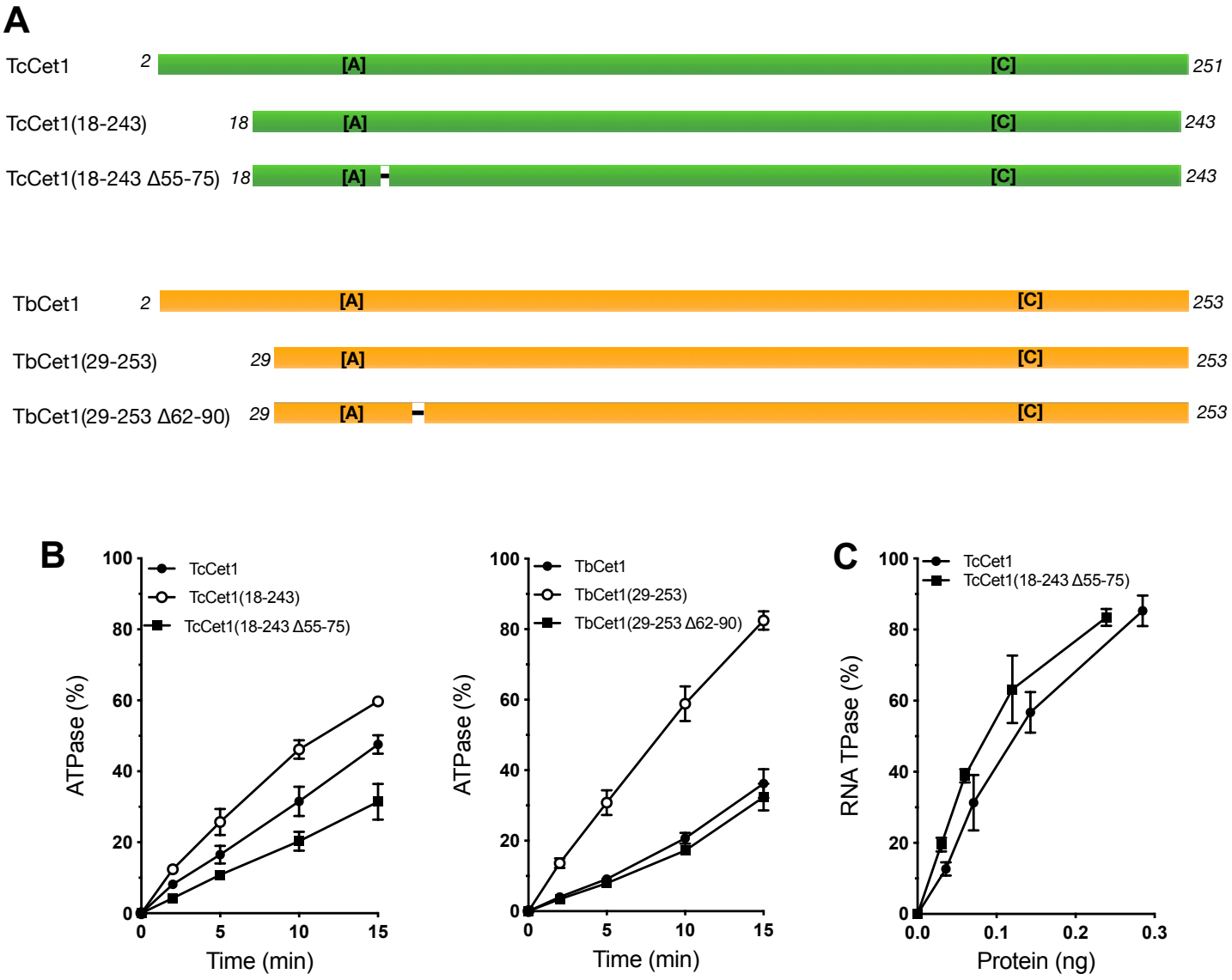


Figure 6

Crystal structures of the RNA triphosphatase from Trypanosoma cruzi provide insights into how it recognizes the 5' end of the RNA substrate

Yuko Takagi, Naoyuki Kuwabara, Truong Tat Dang, Koji Furukawa and C. Kiong Ho

J. Biol. Chem. published online May 7, 2020



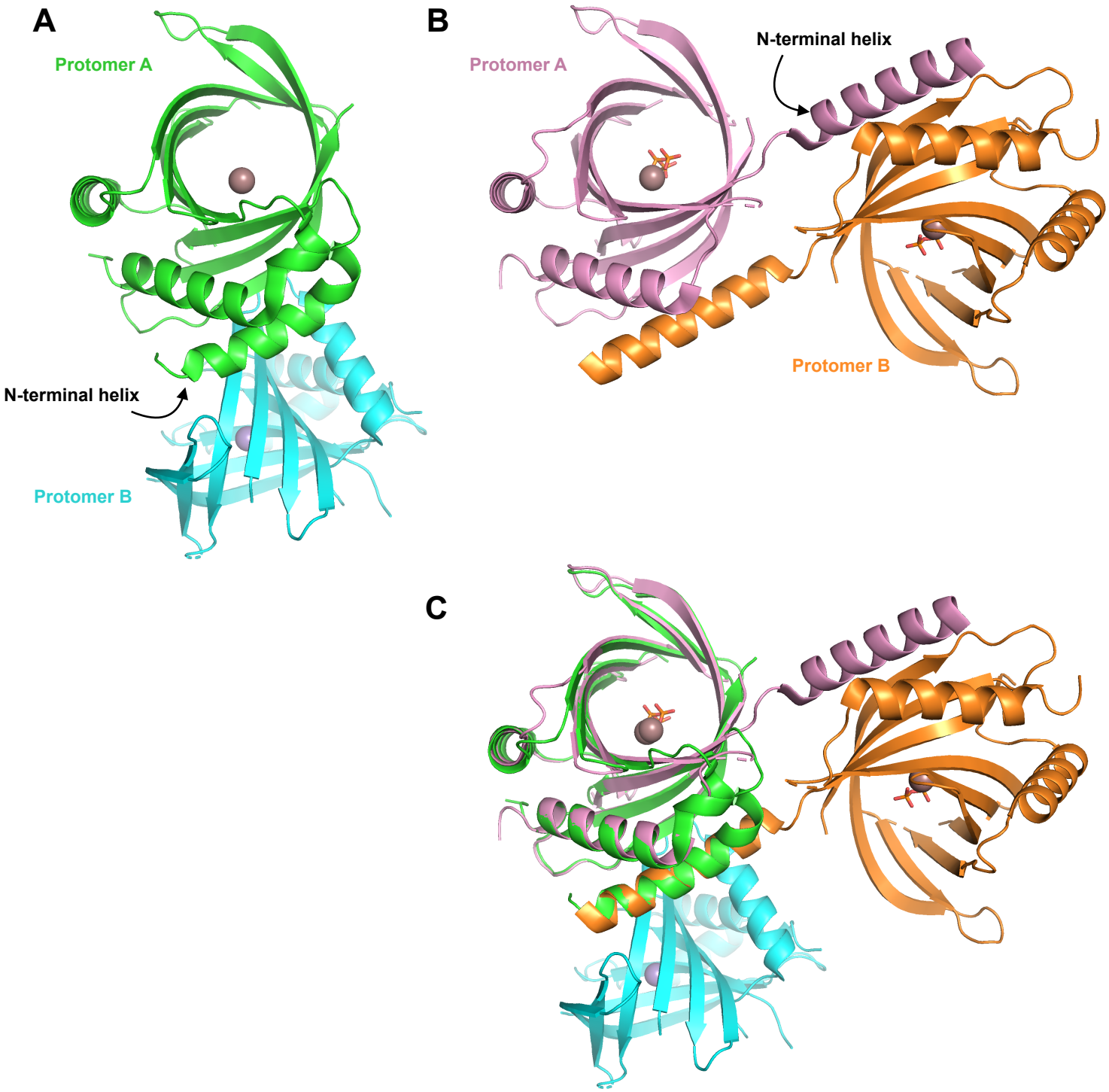
Supporting Information 1. Deletion analysis of Trypanosoma RNA triphosphatase. (A) Schematic diagram of full-length TcCet1, TcCet1(18-243), TcCet1(18-243 Δ55-75), TbCet1, TbCet1(29-253) and TbCet1(29-253 Δ62-90) are shown. Internal deletion is indicated by a dash. Position of conserved motifs, A and C, are indicated. (B) ATPase activity. Reaction mixtures (50 µL) containing 50 mM Tris-HCl (pH 7.5), 2 mM DTT, 2 mM $MnCl_2$, 0.2 mM [γ-^{32}P]ATP, and 5 ng of full-length and truncated proteins were incubated at 30°C. Aliquots were withdrawn at the indicated time and Pi release was plotted as a function of time. (C) RNA triphosphatase activity of TcCet1 and TcCet1(18-243 Δ55-75). Reaction mixtures (10 µL) containing 50 mM Tris-HCl, pH 7.5, 2 mM DTT, 2 mM $MgCl_2$, 1 µM [γ-^{32}P]pppRNA, and indicated amount of full-length TcCet1 and TcCet1(18-243 Δ55-75) were incubated at 30°C for 15 min. Pi release was plotted as a function of input protein. The data shown represent the average of three separate experiments with SE bars.

Supporting Information 2

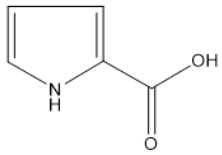
Data collections and refinement statistics of the *Trypanosoma cruzi* RNA Triphosphatase, TcCet1(18-243 Δ55-75) , crystals

Ligand(s)	Mn^{2+}, PPPi and I^-	Mn^{2+}	Compound #951	Compound #466
PDB ID	6L7V	6L7W	6L7X	6L7Y
Data Collection				
Beamline	Photon Factory, BL1A	NSRRC, BL15A	Photon Factory, BL1A	Photon Factory, BL1A
Wavelength (Å)	1.9000	1.0000	1.1000	1.1000
Resolution range	33.63 - 2.20 (2.28 - 2.2)	47.9 - 2.60 (2.72 - 2.60)	49.65 - 2.39 (2.48 - 2.39)	49.40 - 2.51 (2.61 - 2.51)
Space group	$P3_221$	$P2_1$	$P3_221$	$P3_221$
Unit cell a,b,c	67.3, 67.3, 76.8	52.4, 73.2, 63.7 β = 114.1	114.7, 114.7, 56.7	115.9,115.9, 56.7
Total reflections	409654	51598	176432	147972
Unique reflections	10613	13646	17288	15320
Multiplicity	38.6	3.8	10.2	9.7
Completeness (%)	100.0 (99.7)	99.6 (98.7)	100.0 (100.0)	99.7 (98.2)
Mean I/sigma(σ)	33.5 (0.7)	9.9 (1.7)	14.9 (1.7)	16.6 (1.8)
Wilson B-factor	39.4	40.1	45.1	57.2
R-merge	0.095 (5.854)	0.121 (0.860)	0.101 (1.843)	0.083 (1.208)
R-meas	0.096 (5.939)	0.142 (1.054)	0.106 (1.934)	0.087 (1.273)
CC1/2	1.000 (0.955)	0.996 (0.709)	0.999 (0.771)	0.999 (0.927)
Phasing				
Number of sites	7	-	-	-
Figure of merit	0.327	-	-	-
Refinement				
R-work	22.6	23.9	20.4	20.1
R-free	26.1	29.4	23.8	24.0
Number of non-hydrogen atoms	1449	2893	1675	1657
macromolecules	1402	2850	1612	1613
ligands	21	2	36	36
solvent	26	41	27	8
Protein residues	180	362	203	203
RMS(bonds)	0.004	0.005	0.004	0.004
RMS(angles)	0.97	0.98	0.95	0.99
Ramachandran favored (%)	95.9	97.7	98.0	97.5
Ramachandran allowed (%)	4.1	2.3	2.0	2.5
Ramachandran outliers (%)	0.0	0.0	0.0	0.0
Rotamer outliers (%)	3.4	1.3	1.2	0.6
Average B-factor	61.2	56.0	67.0	67.6
macromolecules	60.3	56.2	66.8	66.8
ligands	133.5	53.5	98.8	101.7
Mn annd PPPi	97.6	-	-	-
iodine	205.2	-	-	-
solvent	49.4	43.6	56.5	74.8

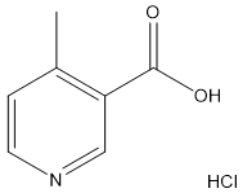
*Highest resolution shells are shown in parenthesis



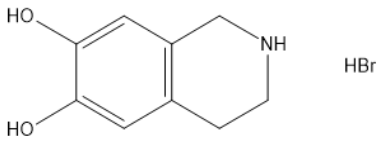
Supporting Information 3. Crystal structures of TcCet1(18-243 Δ55-75). (A) Mn-bound structure. Ordered amino acids in protomer A/B include 18–50, 79-165, 178-218, 220-242/ 20-51, 79-114, 116-164, 180-215, 219-242. (B) Mn•PPPi-bound structure. The ordered amino acids include 18–51, 77-112, 117-164, 179–217, 220–242. (C) Superimposition of Mn-bound and Mn•PPPi-bound structure of TcCet1(18-243 Δ55-75). Note that position of N-terminal helix (amino acid residues 19-36) in the protomer A of the Mn-bound form (green) and protomer B of the Mn•PPPi bound form (orange) overlap each other.



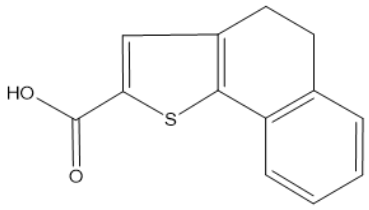
Compound # 202
Tm = 37.0°C



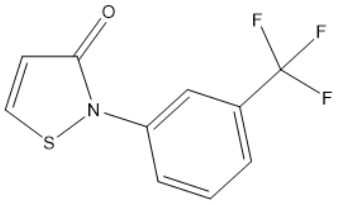
Compound # 212
Tm = 33.6°C



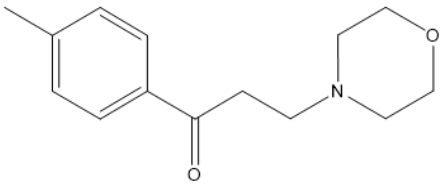
Compound # 224
Tm = 34.5°C



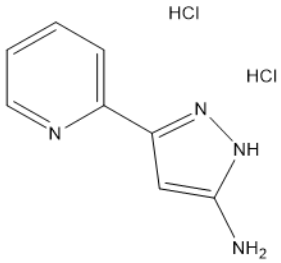
Compound # 332
Tm = 35.7°C



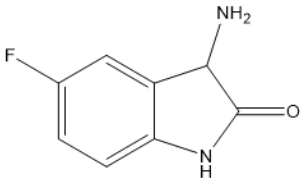
Compound # 375
Tm = 36.0°C



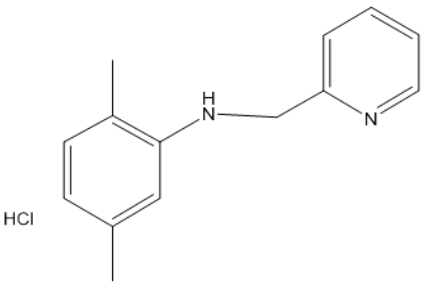
Compound # 378
Tm = 36.9°C



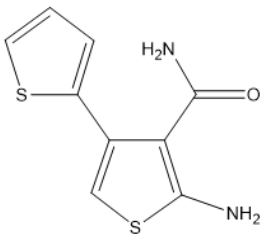
Compound # 521
Tm = 35.8°C



Compound # 524
Tm = 35.8°C

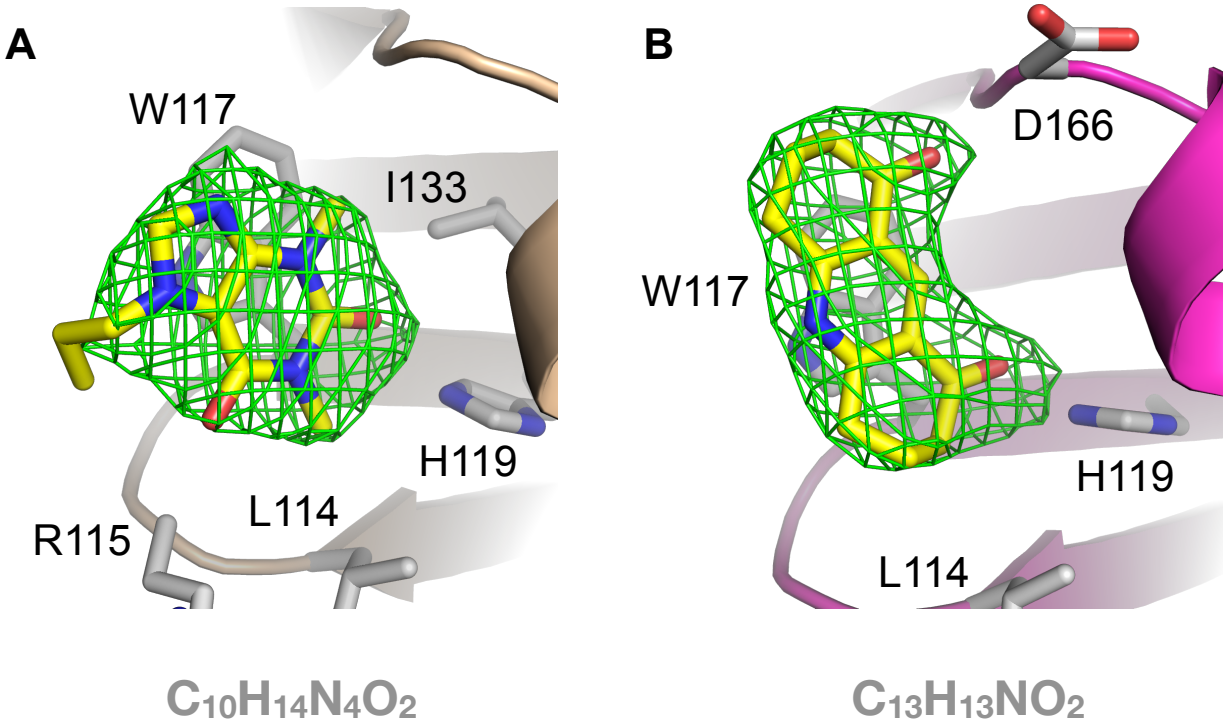


Compound # 809
Tm = 33.0°C

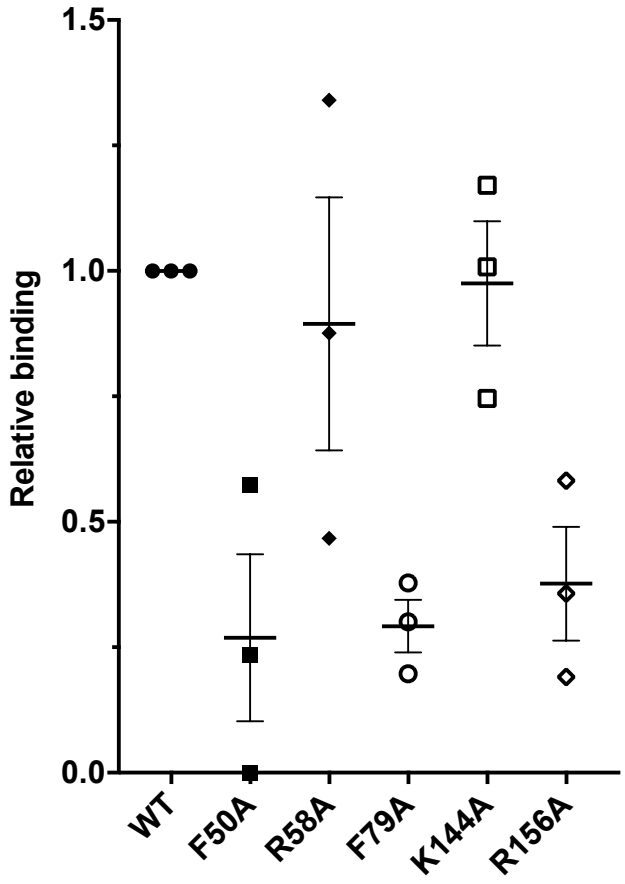


Compound # 818
Tm = 36.2°C

Supporting Information 4. Small molecules that destabilize TcCet1. Melting temperature (Tm) of TcCet1(18-243 Δ55-75) in the presence of 2 mM compounds were determined by thermal shift assay (Figure 3A). Structures of ten compounds that reduce the Tm to below 37°C are shown.



Supporting Information 5. Electron-density of ligands bound to TcCet1(18-243 Δ55-75). Omit (Fo-Fc) map of (A) $C_{10}H_{14}N_4O_2$; compound #951 and (B) $C_{13}H_{13}NO_2$; compound #466. Omit maps are shown as green mesh, contoured at 3 σ. Maps are carved at 2.0 Å around the ligands. Proteins are shown in cartoon representations and ligands as sticks colored according to the atom type (nitrogen in blue, oxygen in red and carbon in yellow).



Supporting Information 6. Filter binding Assay. Nucleic acid binding of wild type and mutant TcCet1 proteins were analyzed by filter binding assay. One pmol of 5' [^{32}P]-labeled 17-mer oligonucleotide was incubated with 10 pmol of protein in a 20 μL buffer containing 50 mM Tris-HCl (pH 7.5), 2 mM DTT and 50 mM NaCl for 15 min at 23°C. The mixture was then spotted onto MF-Millipore membrane filter (Merck), and washed twice with the buffer. The dried membrane was exposed to a PhosphorImager plate, scanned by BAS-2000, and quantitated by Image Gauge software. The amount of labeled oligonucleotide bound by the WT was defined as an arbitrary unit of 1.0, against which all other values were normalized. The average of three independent experiments and standard error of the mean are plotted.