# Crystal Structure of Thiamin Phosphate Synthase from *Mycobacterium tuberculosis* at 2.35Å Resolution

Keiko Ishida
*Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853*

## ABSTRACT:

Thiamin phosphate synthase (TPS) is a bacterial protein involved in the biosynthesis of thiamin pyrophosphate (TPP), the active form of thiamin (vitamin $B_1$) which is an essential component of the human diet. TPS catalyzes the coupling reaction of pyrimidine pyrophosphate and thiazole phosphate to form thiamin phosphate (TP). *Mt*TPS is a 23 kDa protein and forms a dimer. The crystal structure of thiamin phosphate synthase from *Mycobacterium tuberculosis* (*Mt*TPS) was determined at 2.35 Å resolution. Thiamin phosphate synthase has an α/β structure with a triosephosphate isomerase (TIM barrel) fold. The *Mt*TPS structure clearly shows that it is very similar to the structure of *Bacillus subtilis* TPS. The active site of *Mt*TPS is highly conserved when compared to *Bs*TPS and a phosphate group is bound in approximately the same position as in *Bs*TPS.

## INTRODUCTION:

Thiamin, also known as vitamin $B_1$, is an essential component in the human diet because mammals cannot biosynthesize thiamin. Since it is essential to obtain a daily requirement of 1.4 mg of thiamin in our diet, vitamin $B_1$ is added to many of the commercially produced foods (1). Deficiency in thiamin can lead to the conditions known as beriberi and Wernicke's disease (alcoholics), both of which can cause nervous

system ailments. The active form of thiamin, thiamin pyrophosphate (TPP), is an important cofactor required for carbohydrate metabolism and plays an important role in the stabilization of acyl carbanion intermediates (2). For instance, TPP is required for many of the biological systems ranging from oxidative decarboxylation of pyruvate to form acetyl-coenzyme A in the Krebs cycle to its role as a coenzyme for transketolase in the pentose phosphate pathway (Figure 1), an alternate pathway for glucose oxidation. In addition, aside from its vital role in carbohydrate metabolism, TPP acts as a cofactor in decarboxylation and oxidation reactions of α-keto acids by adding thiamin ylide to the ketone.

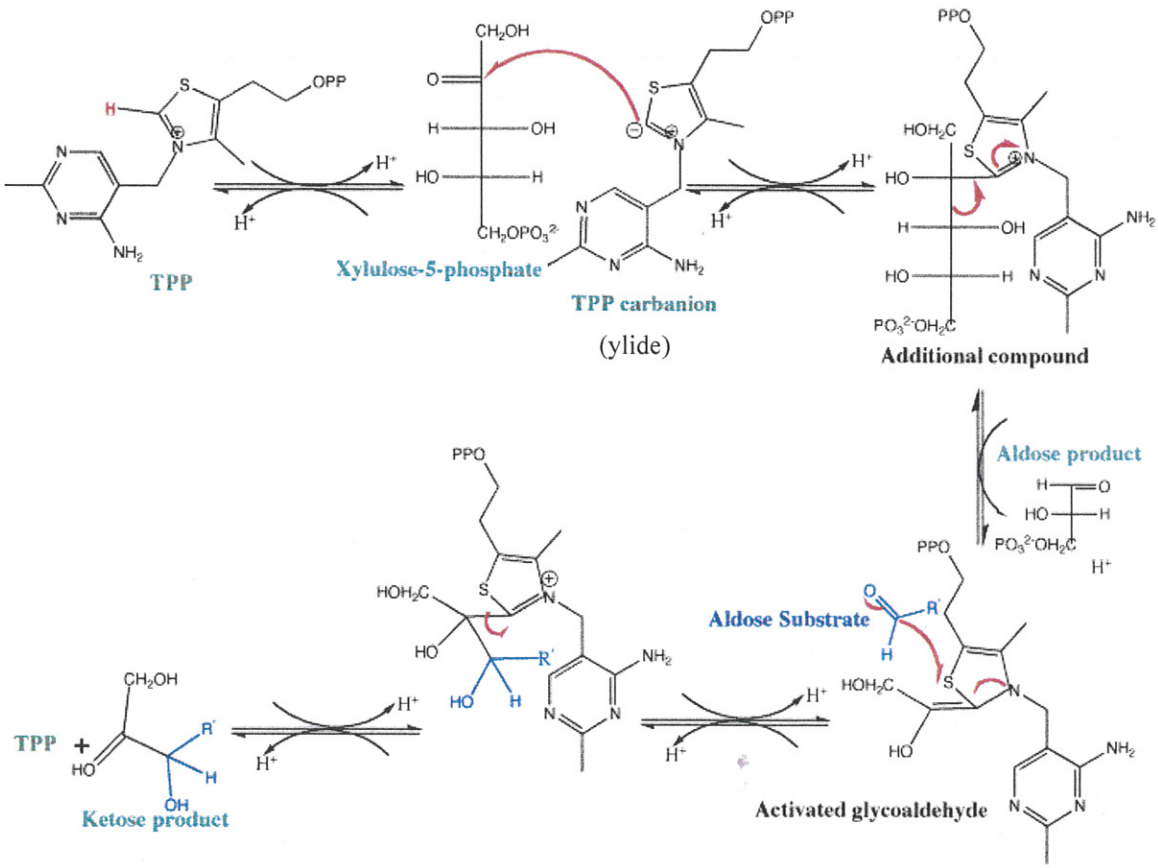


FIGURE 1: The role of TPP in transketolase. Proton dissociation from TPP results in the thiazolium carbanion (ylide) that reacts with the carbonyl group of xylulose-5-phosphate stabilizing acyl carbanion intermediate. Cleavage of C-C bond releases the aldose product while leaving two-carbon fragment tightly bound to TPP. The activated glycoaldehyde intermediate attacks aldose substrate forming a new C-C bond. Once the ketose product is released, TPP becomes available for the next reaction cycle. ChemDraw was used to produce the figure.

In the biosynthesis of TPP, thiamin phosphate synthase (TPS) catalyzes the formation of thiamin phosphate through the coupling of 4-amino-5-hydroxymethyl-2-methylpyrimidine pyrophosphate (HMP-PP) and 5-hydroxyethyl-4-methylthiazole phosphate (Thz-P) (Figure 2b). Both the pyrimidine pyrophosphate and thiazole phosphate moieties are biosynthesized separately as shown in Figure 2a. Six proteins (ThiS, ThiF, ThiG, ThiH, ThiI, and IscS) are involved in the intricate oxidative condensation reactions to form thiazole phosphate. ThiC catalyzes another complex reaction involving the rearrangement and conversion of 5-aminoimidazole ribonuleotide (AIR) to form pyrimidine phosphate. Further phosphorylation by ThiD (HMP-P kinase) gives rise to pyrimidine pyrophosphate (3). The coupling reaction of HMP-P and Thz-P to form thiamin phosphate (TP) occurs by thiazole displacing the pyrophosphate of pyrimidine via a dissociative mechanism catalyzed by TPS. Previous structural and biochemical data and determination of the rate of pyrimidine carbocation formation confirms evidence for a dissociative mechanism for TPS (4, 5). Mechanistic studies on TPS showed that the dissociation of the pyrophosphate (PP) from HMP-PP gives rise to a reactive pyrimidine carbocation intermediate that is then trapped by the thiazole to form TP. In addition, the analysis of the effect of electron-donating and electron-withdrawing substituents at the C2 of the pyrimidine moiety on the reaction rate ($k_{cat}$) allowed the confirmation of a dissociative rather than associative mechanism. In a dissociative mechanism, $k_{cat}$ will be sensitive to C2 substituent group on pyrimidine while in an associative mechanism $k_{cat}$ will not be affected by the C2 substituent group on pyrimidine. Final phosphorylation of TP by ThiL (thiamin monophosphate kinase) generates the active form of thiamin, TPP.

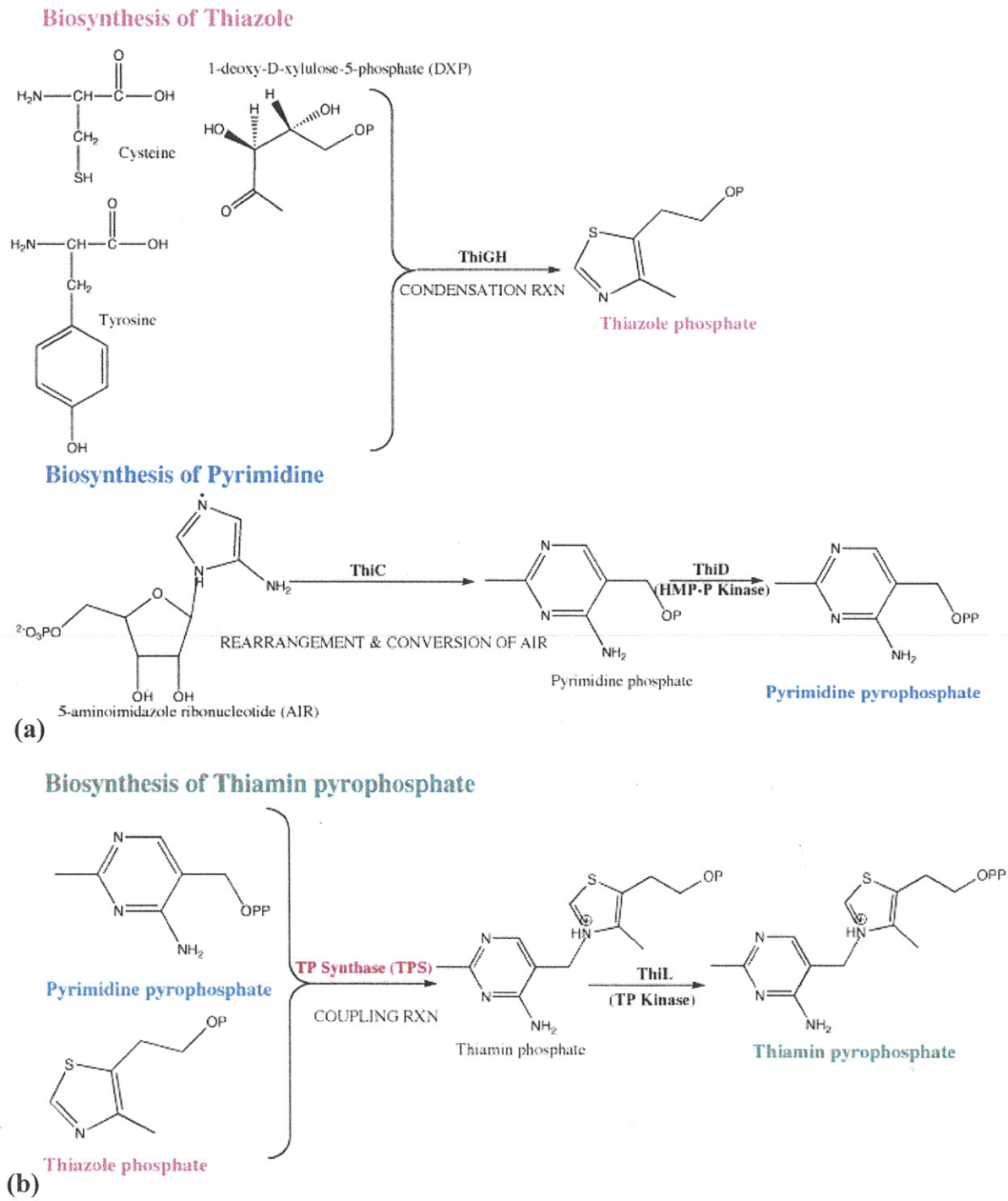


FIGURE 2: (a) Biosynthesis of thiazole phosphate and biosynthesis of pyrimidine pyrophosphate. (b) The biosynthesis of thiamin phosphate in a reaction catalyzed by thiamin phosphate synthase. 4-amino-2-methyl-5-hydroxymethylpyrimidine pyrophosphate (HMP-PP) + 4-methyl-5-hydroxylethylthiazole phosphate (Thz-P) coupled to form Thiamin phosphate (TP). ChemDraw was used to produce the figure.

Living organisms that synthesize TPP must have TPS, which as previously mentioned catalyzes the coupling reaction that is the second to the last step of TPP biosynthesis. However, as humans do not have a biosynthetic pathway for the formation of thiamin, this enzyme could be targeted in pathogens. Therefore, understanding the functional properties of TPS can provide better insight into thiamin biosynthesis and lead to potential benefits in human health. *M. tuberculosis* is a pathogen that is mainly responsible for the cause of tuberculosis (TB), a common and often deadly infectious disease. Humans are the only known reservoir for *M. tuberculosis* and TB continues to be a major global health problem taking lives of 3 million people worldwide each year (6). In this experiment, TPS from *Mycobacterium tuberculosis* was obtained and the crystal structure of *Mt*TPS was determined. Since humans cannot biosynthesize thiamin, the lack of thiamin pathway in humans allows promising use of thiamin biosynthesis as a potential drug target. Potential compound that is identified to target TPS can inhibit thiamin biosynthesis required for the pathogenic function of *M. tuberculosis*, thus inhibiting TB. The ability to effectively target and inhibit their molecular function provides an opportunity to mitigate the deleterious results of TB disease states in humans. Therefore, structural determination and analysis of *Mt*TPS is an imperative project both in the scientific research and human health.

TPS has been overexpressed and purified from *Escherichia coli* (7) and *Bacillus subtilis* (8). Previously, the structure of thiamin phosphate synthase was determined from *B. subtilis* (Figure 3) to 1.25 Å resolution with products (thiamin phosphate, pyrophosphate, and $Mg^{2+}$) bound at the active site (10). The active site of *Bs*TPS was found in a pocket formed by loop regions, residues 59-67 (A loop, joining α3 and β2),

residues 109-114 (B loop, joining α5 and β4), and residues 151-168 (C loop, joining α7 and β6). TPS is an α/β protein with the triosephosphate isomerase fold. In this study TPS was purified from *M. tuberculosis* and the X-ray structure of *Mt*TPS at 2.35 Å resolution was determined. *Mt*TPS is a 23 kDa protein that is structurally similar to the same enzyme from *B. subtilis.* The structure of TPS was obtained in order to provide a detailed picture of the interactions between the enzyme and the substrates that is important for catalysis and could potentially lead to identifying compounds that can inhibit *Mt*TPS.

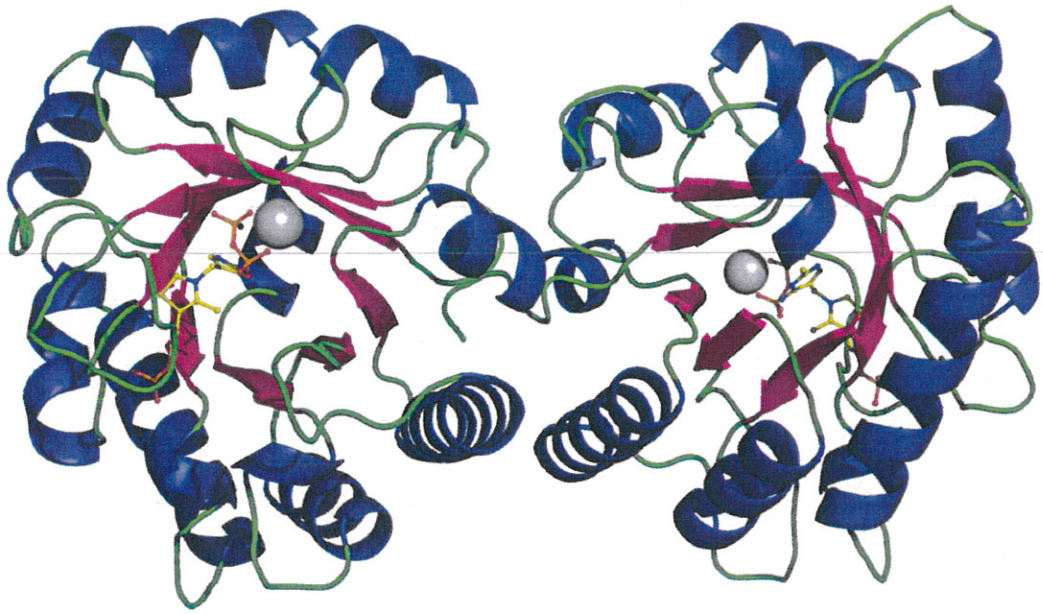

FIGURE 3: Dimeric ribbon representation of the *Bs*TPS that was determined in 1999. The gray sphere represents magnesium ion. The enzymatic products, thiamin phosphate and pyrophosphate are shown in ball-and-stick representations with oxygen atoms in red, carbon in yellow, nitrogen in blue, phosphorous in orange. The diagram is color coded by secondary structures. The figure was created using the PyMol (9).

**MATERIALS AND METHODS**:

*Preparation of Thiamin Phosphate Synthase.* Wild-type *Mt*TPS was overexpressed in *E. coli* and purified. The Cornell Protein Production and Purification Facility provided us with the kanamycin resistant plasmid pMtThiE.28nTEV, which is a

modified pET28a vector (commercially available from Novagen). The modification incorporates a 6x-poly His tag at the N terminus of the protein which can be is cleaved using TEV protease. *Mt*TPS was prepared by transforming pMtThiE.28nTEV into the *E. coli* B834 (DE3) cell line. B834 (DE3) is auxotrophic for methionine, making this cell line ideal for the production of selenomethionyl protein. When grown in rich media, such as Luria-Bertani (LB) broth, B834 (DE3) exhibits normal phenotype. DE3 indicates a copy of the T7 RNA polymerase gene under the control of the lacUV5 promoter. T7 lac promoter allows induction of the desired protein overexpression by addition of isopropyl-β-D-thiogalactopyranoside (IPTG). A 7 mL overnight culture was grown to saturation in LB broth+ kanamycin. This culture was then added to 1L of LB with 50 mg/L kanamycin and grown at 37 ˚C with shaking to an $OD_{600}$ of 0.5. The temperature was then lowered to 15.5 ˚C and overexpression of *Mt*TPS was induced at a final $OD_{600}$ of 0.8 with 1mM IPTG. Cells were further grown by shaking overnight. Cells were then harvested the next morning by centrifugation at 4 °C, 6000 x *g* for 15 minutes. Cell pellets were then frozen and stored at -20 °C until use.

A cell pellet from 4-L of culture was resuspended in approximately 60 mL of lysis buffer (50 mM $NaH_2PO_4$, 300 mM NaCl, 10mM imidazole, 10% glycerol, pH 8.0) followed by sonication. After sonication, the crude lysate was centrifuged for 1 hour at 39, 191 x *g* at 4 °C. The clarified lysate was then passed over a pre-equilibrated 2 mL Ni-NTA column twice, a purification method by immobilized metal ion chromatography where 6x-poly His tag at the N terminus of the protein binds to the Ni column. The Ni-NTA column was then washed with 25 column volumes (c.v.) of lysis buffer, then washed with 25 c.v. wash buffer (50 mM $NaH_2PO_4$, 300 mM NaCl, 20 mM imidazole,

10% glycerol, pH 8.0) and eluted approximately 10 mL using elution buffer (50 mM $NaH_2PO_4$, 300 mM NaCl, 250 mM imidazole, 10% glycerol, pH 8.0). The concentration of imidazole was increased in each wash to improve purity of *Mt*TPS. Imidazole has similar chemical property as histidine residue and therefore it competes with the His tag at the N-terminal of *Mt*TPS to bind to the Ni-NTA column. Thus, increasing the concentration of imidazole used for each wash will allow the removal of other proteins bound non-specifically to the column first and then 250 mM imidazole will then remove the "pure" *Mt*TPS from the column.

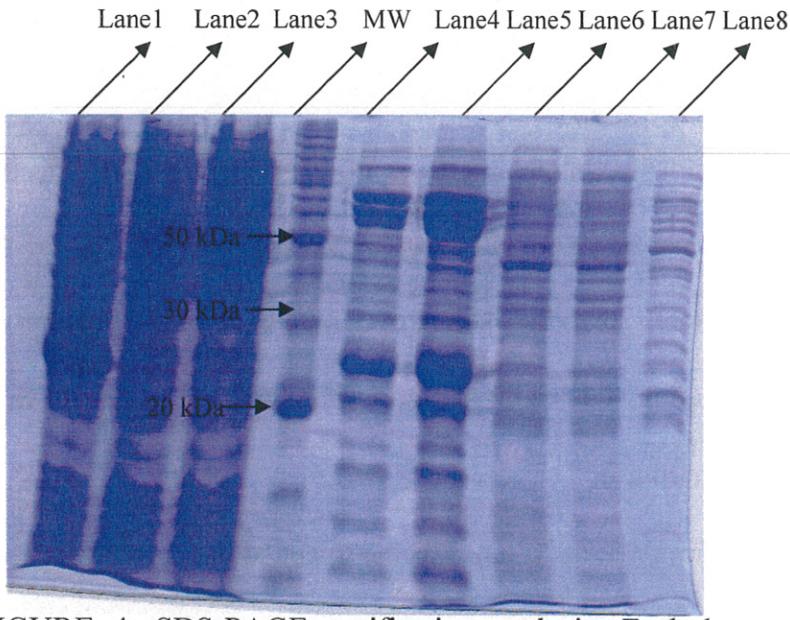


FIGURE 4: SDS-PAGE purification analysis. Each lane represents as follows: Lane 1 = Crude lysate; Lane 2 = clarified lysate; Lane 3 = Flow through; MW= molecular marker (invitrogen) Lanes 4&5 = Elution (250 mM imidazole); Lanes 6&7 = Lysis wash (10 mM imidazole); and Lane 8 = buffer wash (20 mM imidazole).

The purity of protein was analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and judged to be approximately 40 % pure (Figure 4). *Mt*TPS was further purified by size exclusion chromatography using Fast Protein Liquid

Chromatography (FPLC). FPLC was performed using a HiLoad 26/60, Superdex 200 prep grade column with buffer (20 mM Tris pH 8.0, 50 mM NaCl, and 3 % glycerol).The purity of the protein after running through FPLC by analyzing FPLC peaks using SDS-PAGE was judged to be approximately 85 % pure (Figure 5). Finally, fractions 40-48 containing *Mt*TPS were combined and concentrated to approximately 9 mg/mL (as determined by Bradford assay) in 20 mM Tris (pH 8.0) 50 mM NaCl, 3% glycerol. Aliquots of *Mt*TPS were frozen and stored at -80 °C until use.

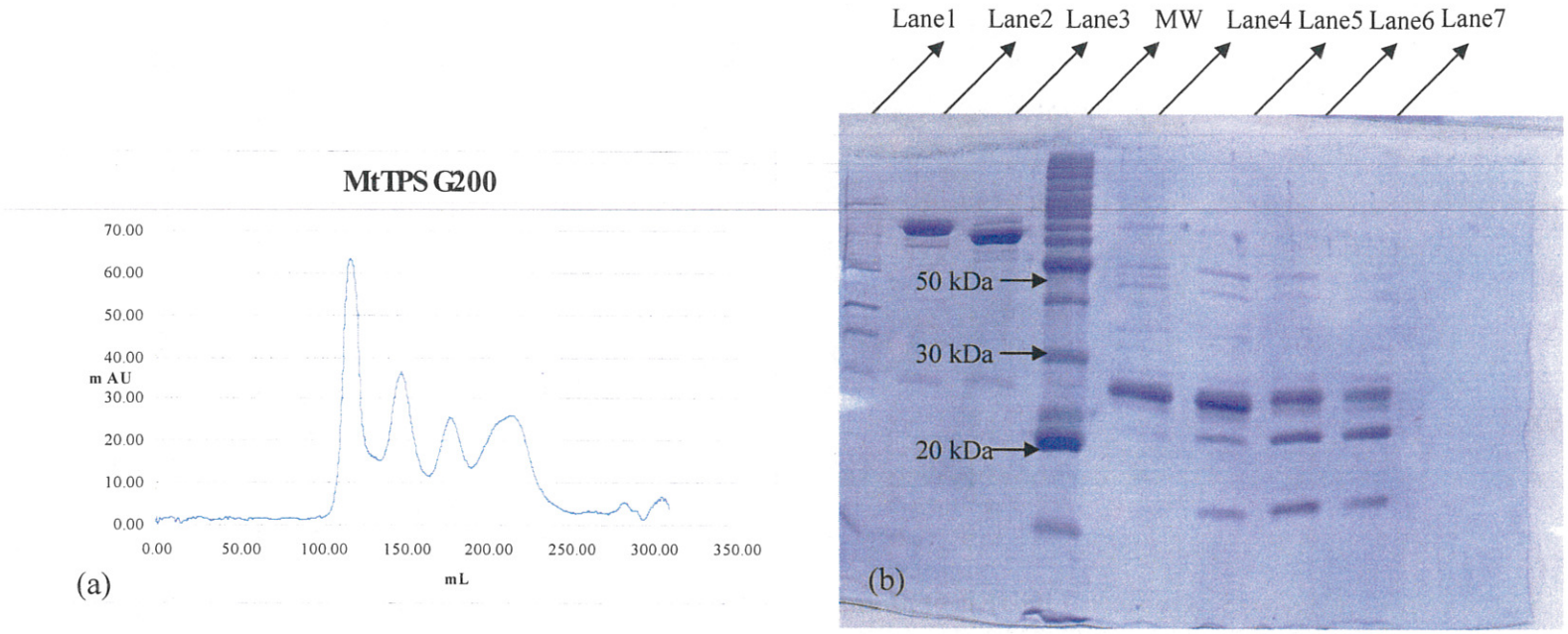


FIGURE 5: *Mt*TPS purification analysis after running through the Fast Protein Liquid Chromatography (FPLC). (a) Results from size exclusion chromatography using FPLC. (b) Analysis of FPLC peaks by SDS-PAGE. Each lane represents as follows: Lane 1= fraction 7 from the 1st peak; Lane 2 = fraction 19 from the 2nd peak; Lane 3= fraction 32 from the 3rd peak; and lanes 4 through 7 are from the last peak. Lane 4= fraction 42; Lane 5 = fraction 45; Lane 6 = fraction 48; and Lane 7 = fraction 50.

*Protein Crystallization.* The hanging drop vapor diffusion method was used to crystallize *Mt*TPS. Each drop consisted of 1μL protein +1μL reservoir solution and was equilibrated against 450μL reservoir solution at 22 °C. CrystalScreen I and CrystalScreen

II from Hapmton Research and Wizard I and Wizard II from Emerald Biosystems were used to screen for initial crystallization conditions. Crystals grew over 3 months in optimized crystallization conditions of 1.6 M $NaH_2PO_4$, 0.4 M $K_2HPO_4$, 100 mM phosphate-citrate buffer (pH 4.2) at 22 °C. The size of the crystal was approximately 300 µm (Figure 6). Crystals were cryoprotected by transferring into a new drop consisting of the crystallization conditions supplemented with 17% glycerol and flash frozen by plunging in liquid nitrogen for data collection.


FIGURE 6: *M. tuberculosis* TPS crystal.

*X-ray Data Collection and Processing.* X-ray diffraction data were collected at the Advanced Photon Source (APS) at the NE-CAT beamline 24-1DC on a single crystal to 2.35 Å using a Quantum315 detector (Area Detector Systems Corporation). Data were collected using a 1° oscillation range over 110°. The data set was indexed, integrated, and scaled using the HKL2000 suite of programs (11). X-ray data collection statistics are summarized in Table 1.

| Table 1: X-ray Data Statistics | *Mt*TPS |
|---|---|
| | APS 24-IDC |
| **Resolution (Å)** | 2.35 |
| **Wavelength (eV)** | 12662 |
| **Space Group** | $C222_1$ |
| **a (Å)** | 84.47 |
| **b (Å)** | 90.91 |
| **c (Å)** | 124.75 |
| **α=δ (°)** | 90.0 |
| **β (°)** | 90.0 |
| **Measured Reflections** | 60007 |
| **Unique Reflections** | 16709 (1119) |
| **Average I/_** | 14.7 (4.4) |
| **Redundancy** | 3.6 (2.8) |
| **Completeness (%)** | 83.1 (56.5) |
| **Rsym (%)** | 7.6 (17.4) |

Values in parentheses refer to the highest resolution shell

*Structure Determination, Model Building, and Refinement.* The structure of *Mt*TPS was determined using molecular replacement (12). NCBI-BLAST (Basic Local Alignment Search Tool) was used to identify potential search models for *Mt*TPS. With the BLAST server, the amino acid sequences of several TPS from different organisms were compared to *Mt*TPS. The structure of *Pyrococcus furiosus* thiamin phosphate pyrophosphorylase (PDBID: 1X13) was used as the search model, which was prepared using the CCP4 program Chainsaw (13). Chainsaw is a model editing program used in molecular replacement that modifies the template structure by deleting nonconserved residues while conserving residues common to both template and target protein (14). The

program MolRep (15) was then used to perform the cross-rotation and translation search functions. Two monomers were located in the asymmetric unit with a corresponding Matthew's coefficient of 2.28 and solvent content of 46%. Refinement of the model was carried out in multiple rounds of refinement using CNS (16, 17) followed by manual model building in COOT (18). Refinement consisted of rigid body refinement, simulated temperature annealing, B-factor refinement, and energy minimization. Final refinement statistics are shown in Table 2.

**Table 2: Structural Refinement Statistics for *Mt*TPS**

| **Source** | **APS 24ID-C** |
|---|---|
| **Resolution (Å)** | 50-2.35 |
| **No. of protein atoms** | 3183 |
| **No. of ligand atoms** | 10 |
| **No. of water atoms** | 118 |
| **No. of refinements in working set** | 15320 |
| **No. of refinements in test set** | 785 |
| **R factor (%)** | 21.9 |
| **R free (%)** | 25.4 |
| **Rms deviation from ideals:** | |
| **Bonds (Å)** | 0.0066 |
| **Angles (°)** | 1.4 |
| **Average B-factor (Å$^2$)** | 23.2 |
| **Ramachandran Plot:** | |
| **Most favored region (%)** | 92.7 |
| **Additional allowed region (%)** | 6.2 |
| **Generously allowed region (%)** | 0.6 |
| **Disallowed region (%)** | 0.6 |

**RESULTS:**

*Purity of Protein.* After metal affinity chromatography, the purity of *Mt*TPS was analyzed by SDS-PAGE gel (Figure 4). From the SDS-PAGE analysis, it is clear that target protein, *Mt*TPS, was not pure but rather contained several impurities. Lanes 4 and 5 of Figure 4 represent elution of *Mt*TPS from the Ni-NTA column and indicate that protein is approximately 40 % pure. Thus, further purification was necessary and performed using size exclusion chromatography using FPLC. The purity of *Mt*TPS was improved up to 85 % purity (Figure 5). The molecular weight of *Mt*TPS dimer is reflected in the size exclusion chromatography by FPLC. Purify of *Mt*TPS was improved by selecting only the fractions that contained the target protein. *Mt*TPS was then concentrated using fractions 40-48 to a final concentration of 9 mg/ml in 20 mM Tris (pH 8.0) 50 mM NaCl, 3% glycerol.

*Model Building.* Amino acid sequence similarity search of *M. tuberculosis* TP synthase using the BLAST server identified sequence similarities in three thiamin phosphate synthase sequences with structures that had been already determined: *P. furiosus* TP pyrophosphorylase (PDB: 1XI3) was 35 % identical and 48 % similar to *M. tuberculosis*, *B. subtilis* TenI (PDB: 1YAD) was 28 % identical and 41% similar, and *B. subtilis* TPS (PDB: 2TPS) was 27 % identical and 40 % similar (Figure 7). The *Mt*TPS structure was determined using *P. furiosus* TP pyrophosphorylase as a search model. The final crystallographic R-factor is 21.9 % and $R_{free}$ is 25.4 %. The refinement statistics are listed in the Table 2.

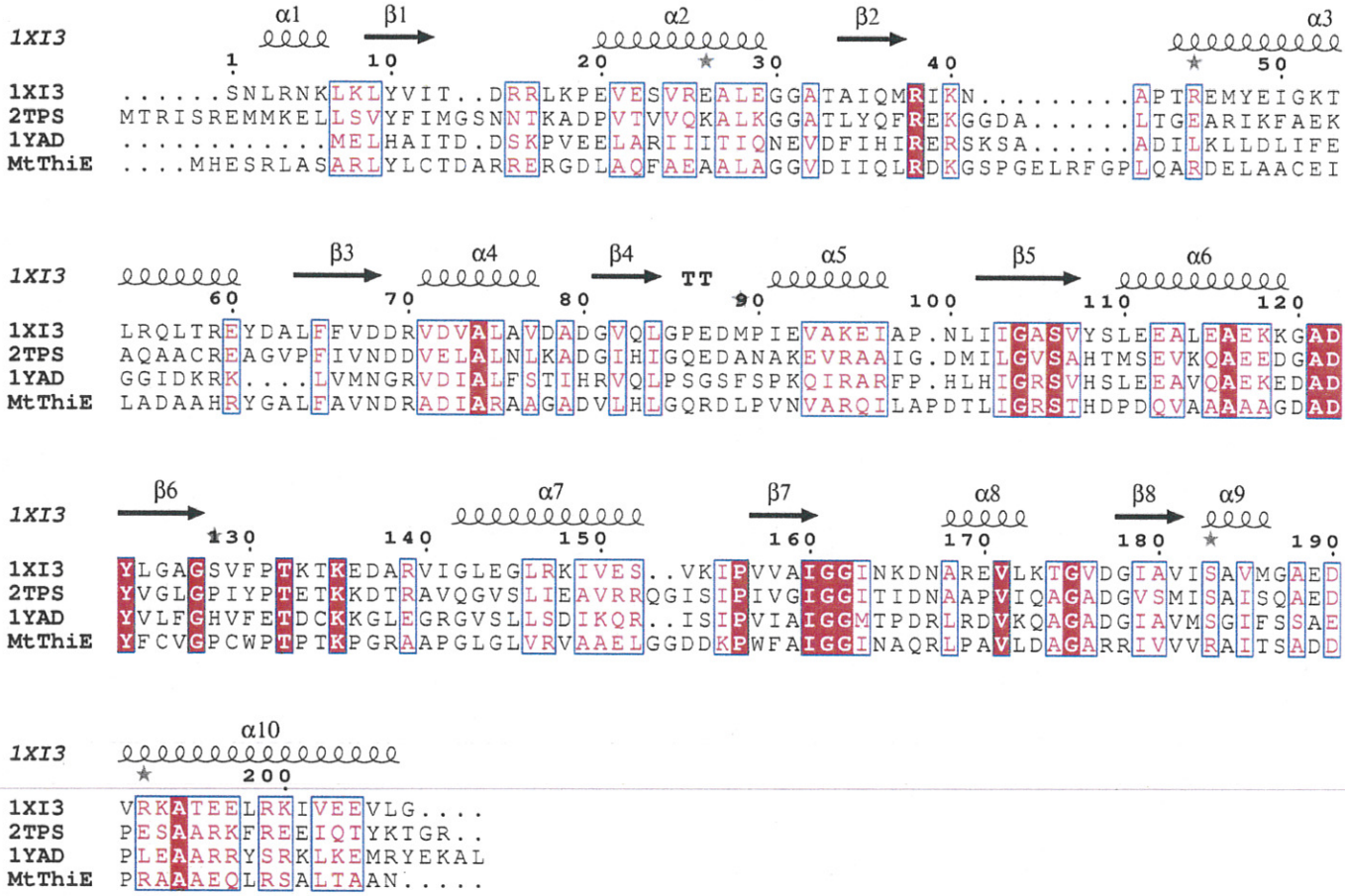


FIGURE 7: Alignment of the amino acid sequences of thiamin phosphate synthase. *P. furiosus* TP pyrophosphorylase (PDB: 1XI3) was 35 % identical and 48 % similar to *M. tuberculosis*, *B. subtilis* TenI (PDB: 2TPS) was 28 % identical and 41% similar, and *B. subtilis* TP synthase (PDB: 1YAD) was 27 % identical and 40 % similar. Residues identical in all sequences are highlighted in red and similar residues are highlighted in pink.

The final model included 222 residues and 2 phosphate groups. Electron density was missing for residues 146-153 and 167-169 of monomeric structure thus these residues were not included in the final model. The breaks in the backbone of *Mt*TPS monomer seen in the ribbon diagram (Figure 8) reflect these excluded residues. Only those residues with electron density were included in the final model (Figure 9).

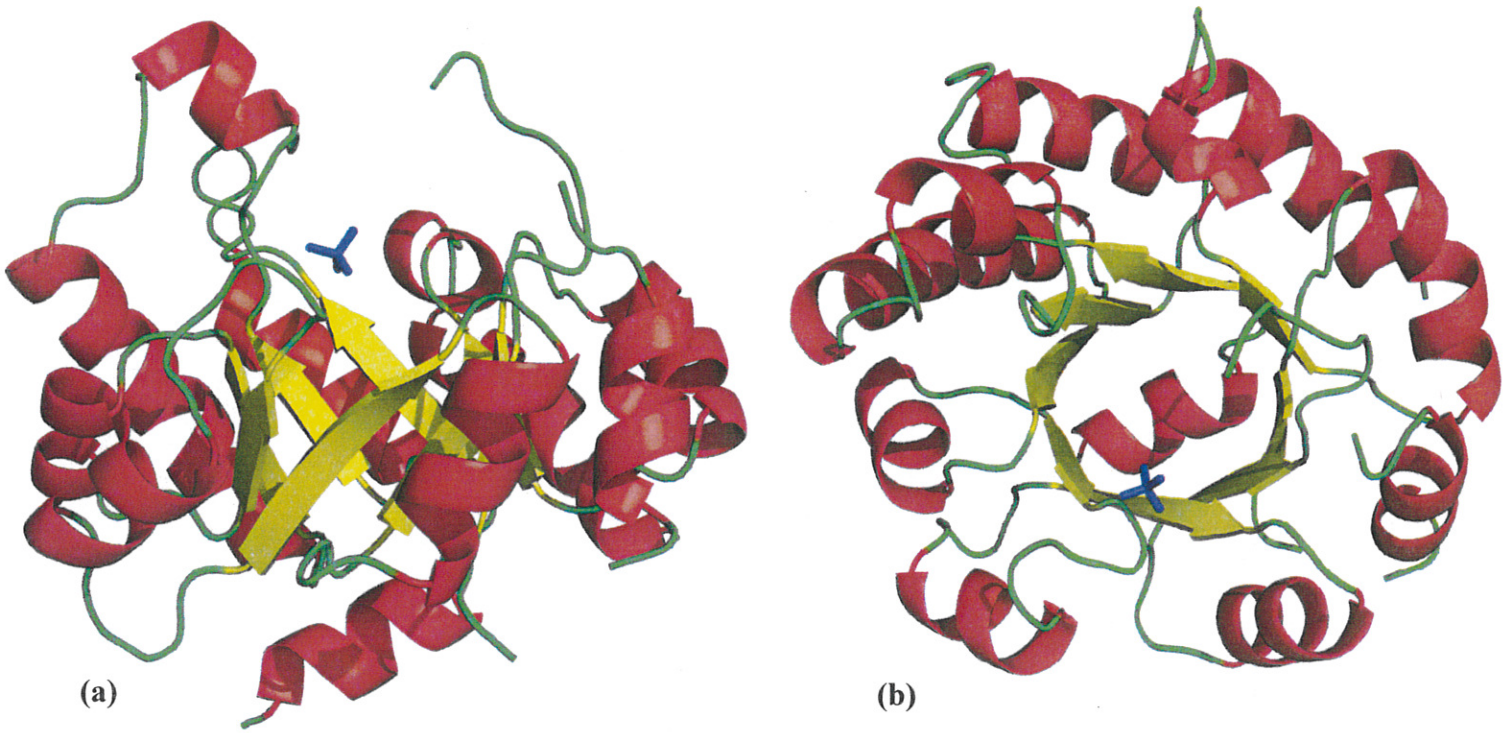


FIGURE 8: Monemeric ribbon representation of the *Mt*TPS. (a) Shows side view of the monomeric structure rotated 90° along the vertical with respect to a top view. (b) Shows top view looking down the barrel. These ribbon structures were created using the PyMOL program (9). Ribbon structures are colored by secondary structures and the blue stick molecule represents phosphate group.

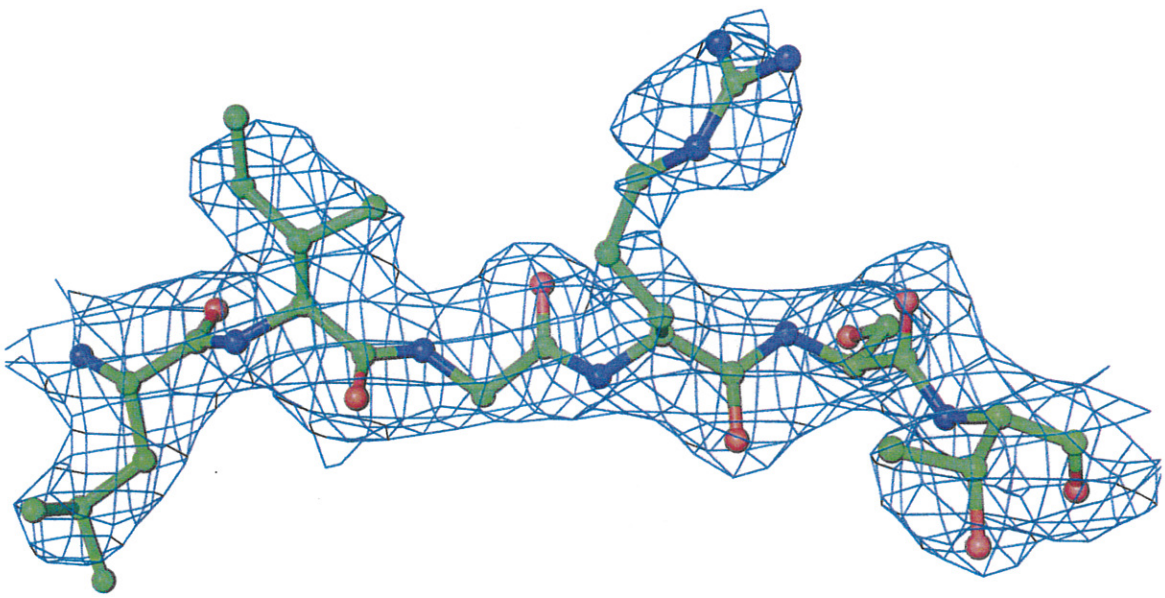

FIGURE 9: Electron density map of *Mt*TPS. Represent sample of electron density map of *Mt*TPS structure. Blue density indicates fitness of the crystal structure. Map was created in the COOT program (18).

*Overall Structure.* *Mt*TPS has an α/β structure which adopts the TIM barrel fold. It has a core of eight parallel β-strands tilted 45° from the axis of the barrel flanked by eight helices (α2-α8, α10) with topology N-terminus α1β1↑α2β2↑α3β3↑α4β4↑α5β5↑α6β6↑α7β7↑α8β8↑α9α10. At the C-terminal entrance of the β–barrel, α/β2-α/β8 loops interact with ligands in the active site. Residues of β-strands form the hydrophobic core of the barrel. *Mt*TPS adopts dimeric conformation but is not an obligate dimer; the active site is wholly contained within one monomer. For each dimer composing of monomers of chain A and chain B of *Mt*TPS, there are two active sites present in the dimer conformation (Figure 10). In each monomeric conformation, the negatively charged phosphate group is stabilized by surrounding water-mediated hydrogen bonds. Several salt bridges and hydrogen bonds are observed at the interactions of dimer interphase of chain A and chain B. Hydrogen bond formation between Ser4-Arg73, both Arg208 and Tyr74 to Gly75 and salt bridges between Glu212-Arg215 and Gly34-Tyr74 stabilize the dimeric interface structure of the *Mt*TPS.

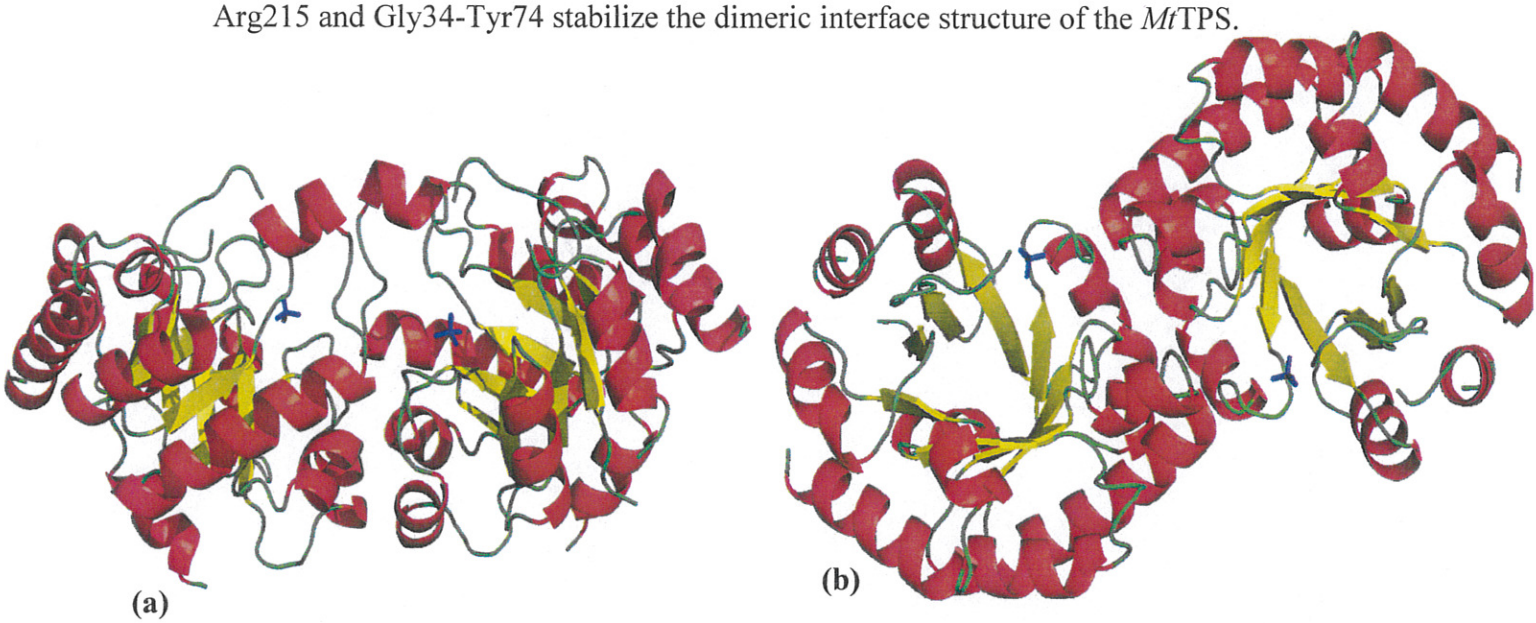


FIGURE 10: Dimeric ribbon representation of the *Mt*TPS. (a) A side view of the dimeric structure rotated 90° along the vertical with respect to a top view. (b) A top view looking down the barrel of the dimeric structure. These ribbon structures were created using the program PyMOL (9). Ribbon structures are colored by secondary structures and blue stick molecule represents phosphate group.

*Active Site of MtTPS.* The active site of *Mt*TPS is located at the C-terminal end of the $(\alpha/\beta)_8$ barrel (Figure 10). Phosphate group from the crystallization condition is bound in the active site of *Mt*TPS. In the active site of *Mt*TPS consist of following residues surrounding phosphate group: Tyr137, Leu97, Tyr12, Asn81, His96, Asp101, Gln99, Pro142, Ile179, Arg42, Cyc14, Ile176, Ser120, Gly98, Asp82, Gly178, Gly40,Gly177,and His122 (Figure 11). The C loop joining α7 and β6 surrounds the active site and is seen as 'open' conformation as seen from the side view of the ribbon diagram presentation in Figure 8a.

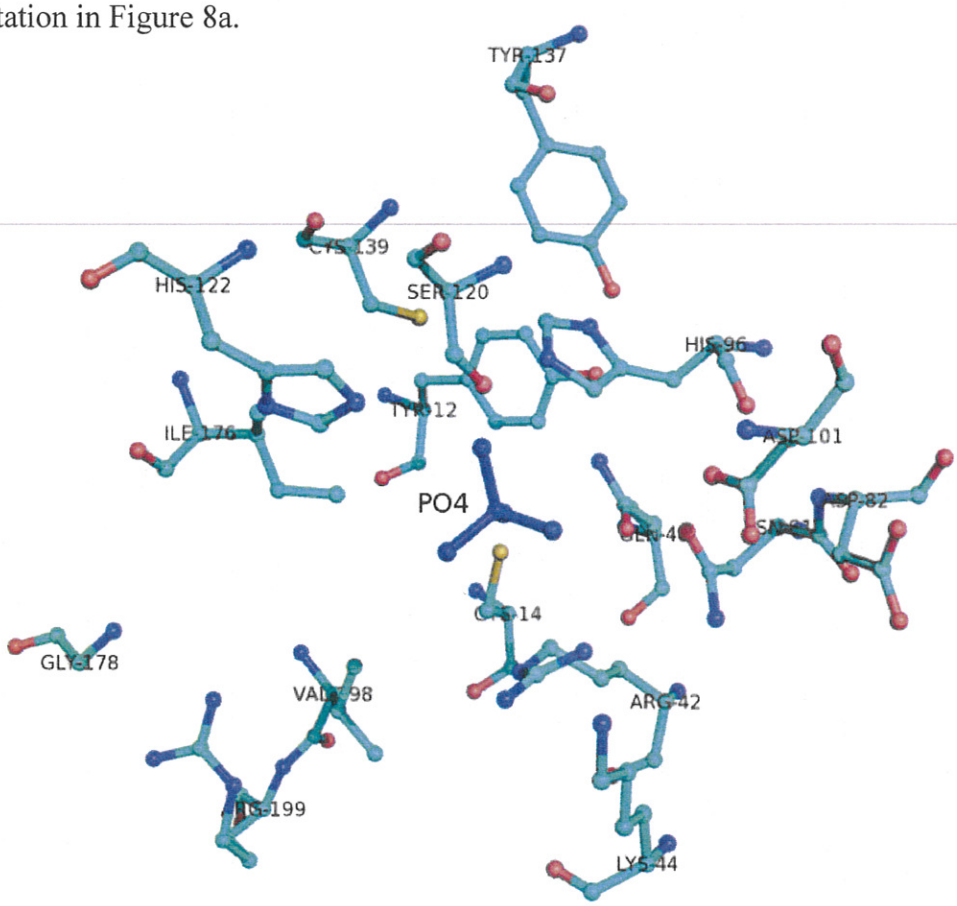


FIGURE 11: Active sites of *Ms*TPS. Active site residues are shown in ball-and-stick representations with oxygen atoms in red, carbon in sky-blue, nitrogen in blue, sulfur in yellow, and $PO_4$ group in blue ball-and-stick. The figure was created using the PyMOL (9) program.

**DISCUSSION:**

*Active Site of MtTPS and BsTPS.* Both *Mt*TPS and *Bs*TPS have a very similar structure. They have an α/β structure with a TIM barrel fold and has a core of eight parallel β-strands flanked by eight α-helices. They both adopt a dimeric conformation but they have different dimeric assemblies. *Mt*TPS dimer adopts a *cis* conformation (Figure 12a) with the ends of the β-barrels facing the same direction as seen in *Bs*TenI (Figure 12b). On the other hand, *Bs*TPS dimer adopts a *trans* conformation with the ends of the β-barrels oriented in antiparallel direction (Figure 12c). Also, from the structural studies of *Mt*TPS and *Bs*TPS, many of the active site residues are superimposed with one another. As it was determined from the structural analysis of *Mt*TPS, the active site of *Mt*TPS is located at the C-terminus end on top of the $(\alpha/\beta)_8$ barrel as well as in *Bs*TPS (8). In *Bs*TPS a pocket is formed by loop A, B, and C in the active site, while these loops are shifted or disordered in *Mt*TPS since neither thiamin phosphate nor pyrophosphate were found bound to the active site. In comparison to *Bs*TPS, the active site of *Mt*TPS is highly conserved. Some of the residues that were missing in *Mt*TPS but present in *Bs*TPS is mostly due to fact that *Bs*TPS structure was complexed with reaction products whereas *Mt*TPS was not complexed with any of the products. If products were present to be complexed, then those disordered loops of *Mt*TPS would hopefully become ordered and then see the missing residues that are present in *Bs*TPS.

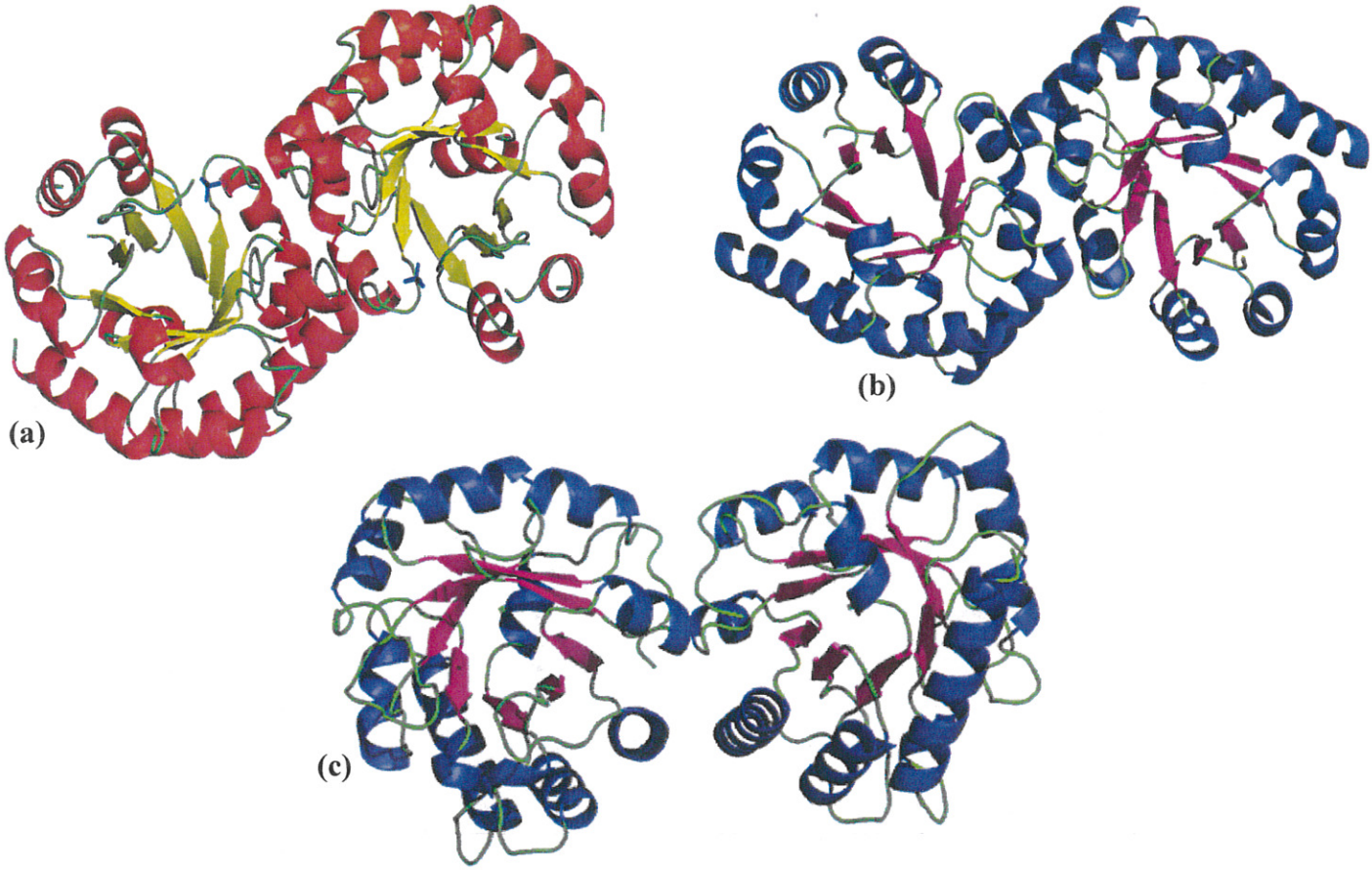


FIGURE 12: Dimeric ribbon representation of the TPS. (a) A top view of the *cis*-dimeric structure of *Mt*TPS. (b) A top view of the *cis*-dimeric structure of *Bs*TenI. (c) A top view of the *trans*-dimeric structure of *Bs*TPS. These ribbon structures were created using the PyMOL program (9). Ribbon structures are colored by secondary structures.

Previous crystallographic experiments on *Bs*TPS showed that this enzyme selectively crystallized with thiamin phosphate and pyrophosphate as well as a $Mg^{2+}$ ion bound at the active site (10). Although no $Mg^{2+}$ ion was bound in *Mt*TPS there are other positively charged and polar residues such as lysine, serine, cysteine, glutamine, and tyrosine residues present in the structure. In *Mt*TPS, phosphate group originated from the crystallization conditions and was bound in the same active site position as seen in *Bs*TPS. Superposition of active site residues of *Mt*TPS and *Bs*TPS show that phosphate group from the crystallization condition of *Mt*TPS is shifted 2.5 Å from phosphate group of pyrophosphate bound at the active site of *Bs*TPS (Figure 13). In both *Bs*TPS and *Mt*TPS, phosphate groups are stabilized by positively charged residues with extensive

electrostatic and hydrogen bonding interactions. Interactions betweens Arg59 and Lys61 from the A loop, Lys159 from the C loop, and $Mg^{2+}$ ion stabilized phosphates in *Bs*TPS and hydrogen bonding interactions between Arg108 and Asp136 in *Mt*TPS.

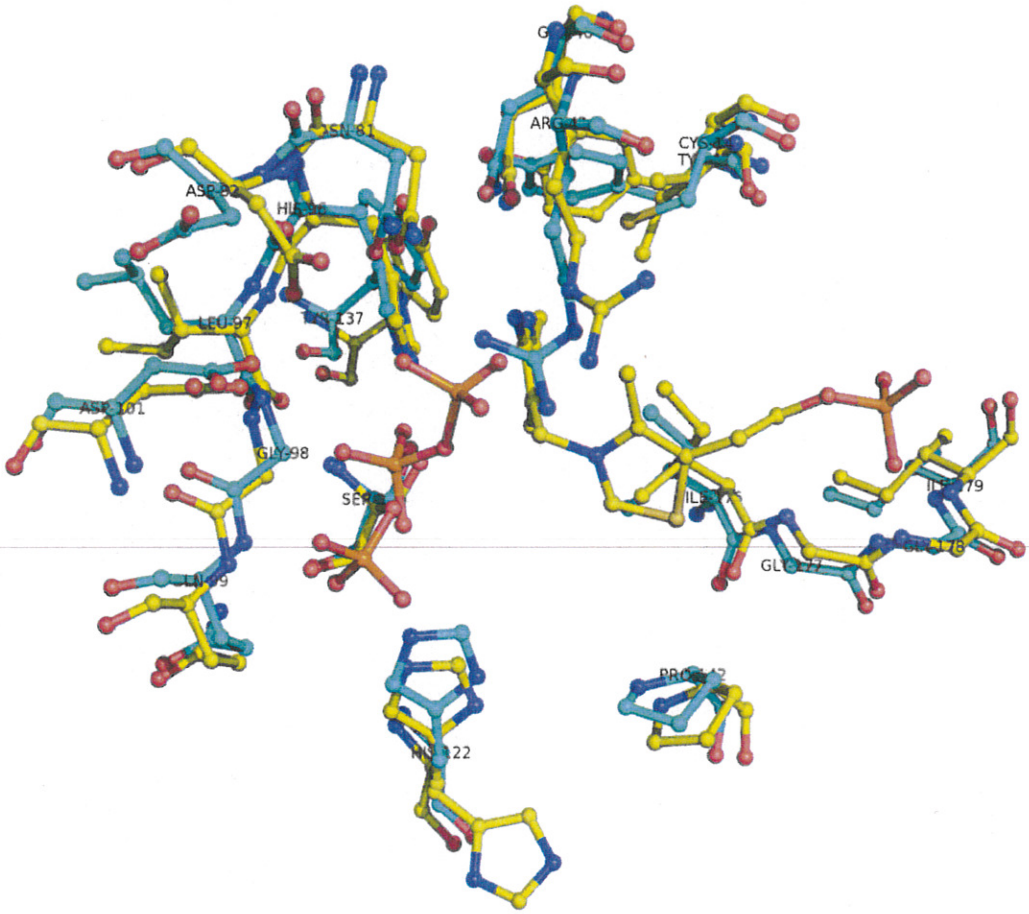


FIGURE 13: Stick-and-ball figure showing a superposition of active sites of *Ms*TPS and *Bs*TPS. *Ms*TPS active site residues are shown in ball-and-stick representations with oxygen atoms in red, carbon in sky-blue, nitrogen in blue, sulfur in orange. *Bs*TPS active site residues are shown in ball-and-stick representations with oxygen atoms in red, carbon in yellow, nitrogen in blue, sulfur in orange. Phosphate groups of pyrophosphate, thiamin phosphate, and phosphate from the crystallization conditions are represented in ball-and-stick with phosphorus in orange and oxygen in red. The figure was created using the PyMOL program (9).

The active sites of *Bs*TPS, and *Mt*TPS were superimposed in the COOT program to compare their structural similarities. *BsTPS* has thirteen conserved active site residues: Ser130, Arg59, Lys61, Lys159, Asp112, Asp93, Asn92, Thr158, Thr156, Ser209,

Gly188, Ile208, and Gln57 (19). These *Bs*TPS active sites residues are highly conserved because when a sequence similarity search of *Bs*TPS with the BLAST server was run, eleven additional TPS sequences were determined that had significant amino acid sequences identities. Of the thirteen conserved active site residues, ten of thirteen *Bs*TPS active site conserved residues were superimposed in *Mt*TPS: Ser120, Arg42, Lys44, Asp101, Asp82, Asn81, Arg199, Gly178, Val198, and Cys139. The root mean square deviation (RMSD) between *Bs*TPS and *Mt*TPS was calculated to be 2.1 Å. Despite the fact that no substrates and products were bound at the active site of *Mt*TPS, based on these findings the binding sites of pyrophosphate and thiamin phosphate can be predicted. Similar residues from *Bs*TPS will probably participate in binding substrates and products in *Mt*TPS. Aside from the binding characteristics of *Mt*TPS, an additional structural trait of *Mt*TPS can be interpreted from structure of *Bs*TPS. In *Bs*TPS the pyrimidine ring is buried in the hydrophobic pocket formed by adjacent domains conserving hydrophobic character of binding site. Moreover, in *Bs*TPS, TP is bound in the "V" conformation and both the pyrimidine amino group and thiazol group are in a *trans* position to one another. The bulky hydrophobic residue interacts with between the pyrimidine and the thiazole rings of TP. These findings from *Bs*TPS can offer deep insight into how *Mt*TPS would behave in the presence of its substrates and products.

Comparison of buried surface area and both initiation and dissociation ΔGs of dimeric *Mt*TPS and dimeric *Bs*TPS offers insight into how each protein assemblies (Table 3). Buried surface area of *Mt*TPS is 2240Å$^2$ and that of *Bs*TPS is 4202.2Å$^2$. *Bs*TPS has almost the double the amount of buried surface area compared to *Mt*TPS and this can indicate that *Bs*TPS has a wider interface while *Mt*TPS has a narrower interface.

ΔG of initiation measures the solvation free energy upon protein assembly in kcal/mol while ΔG of dissociation measures the free energy barrier of assembly dissociation in kcal/mol (19). Negative ΔG *Mf*TPS assembly initiation value of -30.8 kcal/mol and positive ΔG *Mf*TPS assembly dissociation value of 16.3 kcal/mol indicate that obtained *Mf*TPS structure is thermodynamically stable.

**TABLE 3: Comparison of *Mf*TPS and *Bs*TPS Assemblies**

| | **Buried Area (Å$^2$)** | **Initiation ΔG (kcal/mol)** | **Dissociation ΔG (kcal/mol)** |
|---|---|---|---|
| ***Mf*TPS** | 2440 | -30.8 | 16.3 |
| ***Bs*TPS** | 4202.2 | -16.4 | -5.3 |

Note: Buried area and ΔG values were calculated using the Protein interfaces, surfaces and assemblies (PISA) program.

A closer look at the structures of *Mf*TPS and *Bs*TPS reveal some differences in their structural conformation. One of the striking differences between active site residues between *Bs*TPS and *Mf*TPS is that Gly149 in *Bs*TPS is not conserved and Cys139 is conserved in *Mf*TPS. Gly139 is a small amino acid residue and usually Gly or Ala residues are present in most TPS probably since it is small enough to accommodate the large pyrimidine ring. TenI is evolutionarily related to many of TPS structures and it has rather bulky residues such as Leu than Gly or Ala residues. This presence of bulky residues is the key element in differentiating most TPS from TenI. Therefore, TenI only aromatizes the thiazole moiety and the bulky residues prevent HMP-PP from binding to its active site. However, unlike most TPS, *Mf*TPS has Cys residue, which is not a small residue and this indicates the potential for steric clashes in *Mf*TPS with HMP-PP.

*Future Directions.* Future plan include crystallization of *Mt*TPS crystals in alternate crystallization conditions lacking phosphate solution and then soak these crystals in TP product itself and sodium pyrophosphate ($Na_2HP_2O_7$) to facilitate interactions between TPS and its substrates to better understand the biosynthesis pathway of TP. In addition to soaking experiments, virtual drug screening will be performed on *Mt*TPS by our collaborators. Virtual drug screening is an efficient approach to visualize the relationship between the potential inhibitor drug, in our case TB inhibitor drug, and protein in target (TPS) using software programs. From these analyses, we can screen for potential drug candidates before performing experiments on living organisms.

**ACKNOWLEDGEMENTS:**

I deeply thank Dr. Steve E. Ealick and Ms. Kathryn McCulloch of Cornell University for providing me with never ending support and assistance during the collection and processing of my undergraduate research at Cornell. I would also like to thank members of Dr. Ealick's lab especially Ms. Leslie Kinsland for their patience and assistance through out my undergraduate study. I thank staff at the NE-CAT at the Advanced Photon Source for providing and allowing the collection of crystallization data. I would like to thank members of the Honors Committee, thesis reviewers, Dr. George P. Hess, Ms. Susan Coombs, Ms. Pamela Davis, and Dr. Laurel Southard for providing opportunity and guidance in honors thesis.

**REFERENCES:**


1. Begley, T.P., Downs, D., Ealick, S.E., McLafferty, F., Van Loon, D., Taylor, S., Campobasso, N., Chiu, J., Kinsland, C., Reddick, J., *et al.*(1999) *Arch Microbiol.* 171: 293-300.
2. Schowen, R.L. (1998) *Thiamin-dependent enzymes. Comprehensive Biological Catalysis* (Sninott, M., Ed.) Vol. II, pp 217-266, Academic Press, San Diego.
3. Settembre, E., Begley, T.P., and Ealick S.E. (2003) *Current Opinion in Structural Bio.* 13: 739-747.
4. Peapus, D.H., Chiu, H.J., Campobasso, N., Reddick, J.J., Begley, T.P., and Ealick S.E. (2001) *Biochemistry* 40, 10103-10114.
5. Hanes, J.W., Ealick S.E., and Begley, T.P. (2006) *J. Am. Chem. Soc.* 129, 4860-4861.
6. World Health Organization. Global tuberculosis control report – surveillance, planning, finance. WHO report 2008. World Health Organization. Available at <http://www.who.int/tb/publications/global_report/en/index.html>
7. Backstrom, A.D., McMordie, A., and Begley, T.P. (1995) *J. Am. Chem. Soc.* 117, 2351-2351.
8. Zhang, Y., Taylor, S.V., Chiu, H.J., and Begley, T.P. (1997) *J. Bacteriol.* 179, 3030-3035.
9. DeLano, W.L. (2002) *DeLano Scientific*, San Carlos, CA.
10. Chiu, H., Reddick, J., Begley, T., and EAlick, S. (1999) *Biochemistry* 38, 6460-6470.
11. Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode, *Methods Enzymol.* 276, 307-326.
12. Schwarzenbacher, R., and *et al.* (2004) *Acta Cryst.* D60, 1229 - 1236.
13. Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003) *Acta Cryst.* D59, 1131-1137.
14. Stein, N. (2008) *J. Appl. Cryst.* 41, 641 - 643.
15. Vagin, A., and Teplyakov, A. (1997) *J. Appl. Cryst.* 30, 1022-1025.
16. Brunger, A.T. (2007) Nature Protocols 2, 2728-2733.
17. Brunger, A.T., Adams, P.D., Clore, G.M., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, N., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T., and Warren, G.L. (1998) Acta Cryst.D54, 905-921.
18. Emsley, P., and Cowtan, K. (2004) *Acta Cryst.* D60, 2126-2132.
19. Reddick, J.J., Nicewonger, R., and Begley, T.P. (2001) *Biochemistry* 40, 10095-10102.
20. E. Krissinel and K. Henrick (2007). *Inference of macromolecular assemblies from crystalline state*. J. Mol. Biol. **372**, 774--797.