Structural Studies of Thiamin Monophosphate Kinase in Complex with Substrates and Products[†,‡]

Kathryn M. McCulloch, Cynthia Kinsland, Tadhg P. Begley,* and Steven E. Ealick*

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

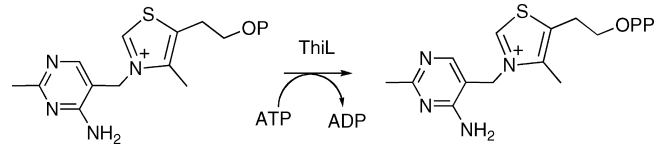
Received January 8, 2008

ABSTRACT: Thiamin monophosphate kinase (ThiL) catalyzes the ATP-dependent phosphorylation of thiamin monophosphate (TMP) to form thiamin pyrophosphate (TPP), the active form of vitamin B_1. ThiL is a member of a small ATP binding superfamily that also includes the purine biosynthetic enzymes, PurM and PurL, NiFe hydrogenase maturation protein, HypE, and selenophosphate synthase, SelD. The latter four enzymes are believed to utilize phosphorylated intermediates during catalysis. To understand the mechanism of ThiL and its relationship to the other superfamily members, we determined the structure of *Aquifex aeolicus* ThiL (AaThiL) with nonhydrolyzable AMP-PCP and TMP, and also with the products of the reaction, ADP and TPP. The results suggest that AaThiL utilizes a direct, inline transfer of the γ-phosphate of ATP to TMP rather than a phosphorylated enzyme intermediate. The structure of ThiL is compared to those of PurM, PurL, and HypE, and the ATP binding site is compared to that of PurL, for which nucleotide complexes are available.

Thiamin, also known as vitamin B_1, is required by all living organisms and is an essential vitamin in the human diet. The active form of thiamin, thiamin pyrophosphate (TPP),[1] stabilizes acyl carbanion intermediates and plays important roles in carbohydrate metabolism and in the pentose phosphate pathway (*1*). Often, the reactions catalyzed by thiamin pyrophosphate involve decarboxylation reactions of α-keto acids. In these reactions, the thiamin ylide adds to the ketone. Loss of carbon dioxide from the resulting adduct gives the thiamin-stabilized acyl carbanion, which can then be protonated by an active site acid. Release of the cofactor generates the aldehyde product (*2*).

The biosynthesis of thiamin pyrophosphate in prokaryotic systems has been studied both structurally and mechanistically (*1, 3*). The thiazole and pyrimidine moieties are biosynthesized separately. Six enzymes act upon deoxy-D-xylulose 5-phosphate, glycine, and cysteine to form the thiazole moiety. The pyrimidine pyrophosphate is formed by a rearrangement of 5-aminoimidazole ribonucleotide catalyzed by ThiC, followed by phosphorylation by HMP-P kinase. The thiazole and pyrimidine moieties are then joined together to form thiamin monophosphate (TMP) in a reaction catalyzed by thiamin phosphate synthase. Thiamin monophosphate kinase (ThiL) is the final enzyme in the thiamin biosynthetic pathway and is responsible for the phosphorylation of TMP to form TPP (Scheme 1), the biologically relevant form of the cofactor (*4*).

Scheme 1



In addition to ThiL's role in thiamin biosynthesis, ThiL has been postulated to be a member of the PurM ATP binding superfamily (*5*). To date, only five proteins have been identified as members of this superfamily: aminoimidazole ribonucleotide synthetase (PurM), formylglycinamide ribonucleotide amidotransferase (PurL), selenophosphate synthetase (SelD), the NiFe hydrogenase maturation protein HypE, and ThiL (*5*). Interestingly, PurL and PurM catalyze sequential reactions in the purine biosynthetic pathway, and the product of the reaction catalyzed by PurM is 5-aminoimidazole ribonucleotide, the small molecule transformed into the pyridine moiety of thiamin pyrophosphate. With the exception of ThiL, these enzymes are believed to proceed via a phosphoimidate intermediate (*6*). It is unclear why ThiL belongs to this superfamily because the TMP phosphorylation reaction is more likely to proceed by an inline phosphate transfer in which a phosphorylated enzyme intermediate would be unnecessary.

[†] This work was supported by National Institutes of Health Grants DK44083 to T.P.B. and DK067081 to S.E.E.

[‡] The coordinates for AaThiL with ATP bound, with AMP-PCP bound, with AMP-PCP and TMP bound, and with ADP and TPP bound have been deposited in the Protein Data Band under the following PDB entries 3C9R, 3C9S, 3C9T, and 3C9U, respectively.

* To whom correspondence should be addressed: Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853. Telephone: (607) 255-7961. Fax: (607) 255-1227. E-mail: see3@cornell.edu.

[1] Abbreviations: TPP, thiamin pyrophosphate; TMP, thiamin monophosphate; ThiL, thiamin monophosphate kinase; PurM, aminoimidazole ribonucleotide synthetase; PurL, formylglycinamide ribonucleotide amidotransferase; SelD, selenophosphate synthetase; AaThiL, *Aquifex aeolicus* ThiL; AMP-PCP, β,γ-methylene adenosine 5′-diphosphate; ADP, adenosine 5′-diphosphate; AaThiL, *A. aeolicus* ThiL; MPD, 2-methyl-2,4-pentanediol; CHESS, Cornell High Energy Synchrotron Source; EcPurM, *Escherichia coli* PurM; StPurL, *Salmonella typhimurium* PurL; TmPurL, *Thermatoga maritima* PurL; TkHypE, *Thermococcus kodakaraensis* KOD1 HypE; FGAR, formylglycinamide ribonucleotide.

10.1021/bi800041h CCC: $40.75
Published on Web 03/01/2008

Table 1: Summary of Data Collection Statistics

	AaThiL−ATP	AaThiL−AMP-PCP	AaThiL−AMP-PCP−TMP	AaThiL−ADP−TPP
beamline	APS 24-ID-C	APS 24-ID-C	APS 24-ID-C	APS 24-ID-C
resolution (Å)	2.30	2.20	2.60	1.48
wavelength (Å)	0.9795	0.9795	0.9795	0.9795
space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
a (Å)	61.3	61.2	60.8	60.6
b (Å)	67.1	67.1	66.8	66.2
c (Å)	200.7	203.5	196.9	197.0
no. of reflections	226383	309648	71312	412005
no. of unique reflections	34823 (1980)	43410 (4269)	23068 (2062)	118539 (11700)
average I/σ	31.5 (1.9)	31.6 (6.0)	23.9 (6.1)	20.3 (3.4)
redundancy	6.5 (3.6)	7.1 (6.6)	3.2 (3.2)	3.6 (2.5)
completeness (%)	92.4 (53.4)	99.8 (99.9)	90.0 (82.6)	89.1 (88.9)
R_{sym}[a] (%)	5.2 (36.5)	5.1 (32.5)	3.8 (14.3)	6.6 (28.6)

[a] $R_{sym} = \sum\sum_i|I_i - \langle I\rangle|/\sum\langle I\rangle$, where $\langle I\rangle$ is the mean intensity of the *N* reflections with intensities I_i and common indices *h*, *k*, *l*.

Table 2: Summary of Data Refinement Statistics

	AaThiL−ATP	AaThiL−AMP-PCP	AaThiL−AMP-PCP−TMP	AaThiL−ADP−TPP
resolution (Å)	50.00−2.30	50.00−2.20	50.00−2.60	50.00−1.48
no. of protein atoms	4778	4750	4881	4906
no. of ligand atoms	62	62	116	116
no. of water atoms	101	90	14	547
no. of reflections in the working set	32928	38005	19849	102350
no. of reflections in the test set	2053	4312	2260	11388
R factor[a] (%)	22.21	21.7	21.3	20.1
R_{free}[b] (%)	24.84	25.7	27.1	21.9
root-mean-square deviation from ideal values				
bonds (Å)	0.006	0.006	0.007	0.004
angles (deg)	1.3	1.3	1.4	1.2
average *B* factor (Å^2)	50.9	45.5	56.1	23.9
Ramachandran plot				
most favored (%)	91.2	92.2	87.8	93.7
additionally allowed (%)	8.4	7.4	11.2	5.6
generously allowed (%)	0.4	0.4	0.4	0.7
disallowed (%)	0.0	0.0	0.6	0.0

[a] *R* factor $= \sum_{hkl}||F_{obs}| - k|F_{cal}||/\sum_{hkl}|F_{obs}|$, where F_{obs} and F_{cal} are observed and calculated structure factors, respectively. [b] For R_{free}, the sum is extended over a subset of reflections (10%) excluded from all stages of refinement.

The structure of *Aquifex aeolicus* ThiL (AaThiL) was previously determined by the New York Structural Genomics Group (Protein Data Bank entry 1VQV), but no substrate or product complexes were available. We initiated additional structural studies to investigate the place of ThiL within the PurM ATP binding superfamily. AaThiL was cocrystallized with β,γ-methylene adenosine 5′-triphosphate (AMP-PCP) as a binary complex, and these crystals were then soaked with TMP to yield the AMP-PCP−TMP ternary complex. Additionally, AaThiL was cocrystallized with the reaction products, TPP and adenosine 5′-diphosphate (ADP), to form the ADP−TPP complex. These structures allow for comparison to the structures of the other members of the PurM ATP binding superfamily, especially at the active sites. Examination of the active site of ThiL, with substrates and products bound, has eliminated the possibility of a phosphoimidate intermediate on ThiL and strongly supports a direct, inline transfer of the γ-phosphate of ATP to TMP.

MATERIALS AND METHODS

Gene Synthesis. Primers for the gene synthesis of *A. aeolicus thiL* were designed using the GeMS Web site (*7*) (http://software.kosan.com/GeMS/). Gene synthesis was carried out using the method described by Kodumal (*8, 9*). The 5′ synthon was cloned into pSTBlue1 (Novagen), and the 3′ synthon was cloned into pCR4-Blunt (Invitrogen). After sequencing had been carried out, the two synthons were merged by single-overlap extension PCR. The resulting PCR fragment was cloned into pENTR-TEV-D-TOPO (Invitrogen). Sequencing revealed a missing base in the overlap region. This was repaired by site-directed mutagenesis using KOD DNA polymerase (New England Biolabs) and a standard PCR protocol. The following primer and its reverse complement were used (5′-GTT AAA CGC GCG TGT GAA TTC TAC AAG TGC GAA GTG GTC GG-3′). Parental DNA was digested with *Dpn*I prior to transformation. Clones were screened for the introduction of an *Eco*RI site and verified by sequencing. A clone with the correct sequence was named pAaThiL.ET and used as the entry vector in an LR reaction using the Gateway system (Invitrogen) with a destination vector based on the Novagen pET system. The destination vector encodes an N-terminal six-His tag, and the pENTR-TEV-D-TOPO vector encodes a TEV protease site between the fusion tag and the AaThiL gene. A correct clone was named pAaThiL.XF1 and used for protein production.

Protein Overexpression. pAaThiL.XF1 was transformed into the B*R2 *Escherichia coli* cell line. Cells were grown in LB medium at 37 °C with shaking (200 rpm) to an OD_{600} of 0.6. Overexpression of ThiL was induced with 1 mM isopropyl β-D-thiogalactopyranoside, and cells were incubated at 37 °C for 3 h. Cells were harvested by centrifugation, resuspended in lysis buffer [300 mM NaCl, 50 mM Na_2HPO_4, and 5 mM imidazole (pH 8.0)], and lysed by

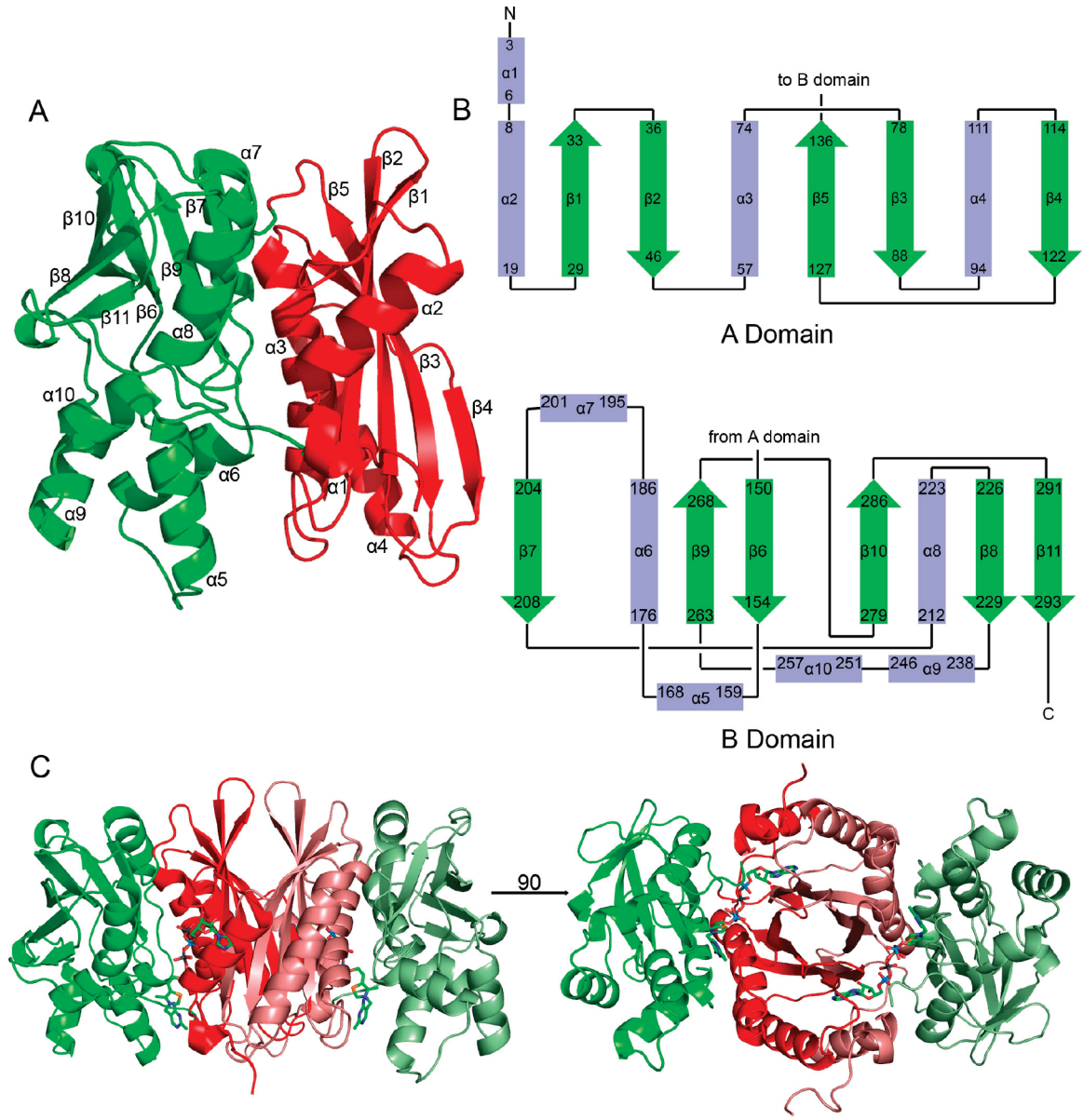


FIGURE 1: Overall structure of AaThiL. (A) Ribbon diagram of the monomer of AaThiL. The A domain is colored red, and the B domain is colored green. All figures were generated using PyMol (*16*). (B) Topology diagram of AaThiL. The A domain and the B domain are shown separately. (C) Ribbon diagram of the dimer of AaThiL. The A domain is colored red, and the B domain is colored green in one monomer; to emphasize domain interactions, the A and B domains of the second monomer are shown in muted tones. The products, TPP and ADP, are shown as stick representations. A 90° rotation provides both a side view of the dimer and a view down the β-barrel.

sonication. The crude lysate was then centrifuged at 40000*g* for 25 min at 4 °C. The cleared lysate was applied to a Ni-NTA column (Qiagen) and washed with the buffer described above in a volume that was 75 times greater than the column volume. The column was then washed with wash buffer [300 mM NaCl, 50 mM Na_2HPO_4, and 30 mM imidazole (pH 8.0)] to remove any nonspecific binding proteins for roughly 20 column volumes. AaThiL was then eluted from the column using elution buffer [300 mM NaCl, 50 mM Na_2HPO_4, and 250 mM imidazole (pH 8.0)] in 5 column volumes. Protein was buffer exchanged into 10 mM Tris and 20 mM NaCl (pH 7.7) by overnight dialysis and was then concentrated using an Amicon Ultra centrifugal filter with a molecular mass cutoff of 10 kDa until a concentration of 13 mg/mL was reached as determined by the Bradford assay (*10*). Protein purity was verified by SDS−PAGE to be 95% (results not shown).

Protein Crystallization, Data Collection, and Structure Determination. Crystallization experiments were conducted using the hanging drop vapor diffusion method at 22 °C by combining equal volumes of protein and reservoir solution. Initial conditions were found using sparse matrix screens Crystal Screen 1 and 2 (Hampton Research) and Wizard Screen 1 and 2 (Emerald Biosystems). To obtain binary complexes, AaThiL was incubated with 4 mM AMP-PCP or ATP and 4 mM $MgCl_2$ on ice for 1 h prior to crystallization experiments. Optimized conditions for the AaThiL binary complexes included 100 mM HEPES (pH 7.4), 6–11% isopropyl alcohol, 200 mM NaCl, and 4–6% 2-methyl-2,4-pentanediol (MPD); 100 mM $CaCl_2$ was used as an additive in a 1:9 ratio with the reservoir solution described above. Crystals reached their maximum size of 200 μm × 100 μm × 100 μm within 3 days. Cryoprotection of these crystals was conducted by increasing the MPD

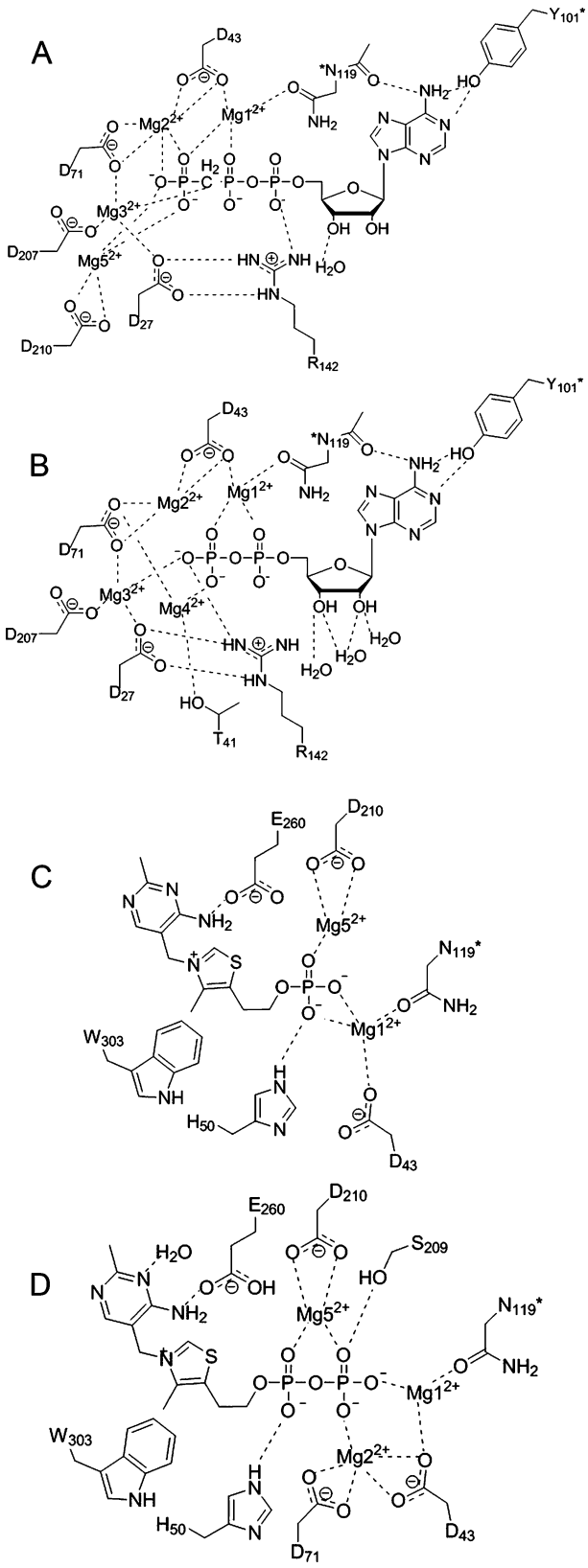


FIGURE 2: Schematic diagrams for the active sites of the substrate and product complexes. Interactions between ligand and residues, ions, and water molecules are shown with dashed lines. (A) Binding of AMP-PCP within the active site of AaThiL. (B) Binding of ADP within the active site of AaThiL. (C) Binding of TMP within the active site of AaThiL. (D) Binding of TPP within the active site of AaThiL.

concentration to 15% and adding 2% ethylene glycol to the crystallization conditions. To form a ternary substrate complex, TMP was soaked into the cocrystallized AaThiL and AMP-PCP crystals by transferring the crystals to a solution containing the mother liquor and 20 mM TMP and allowing the crystals to soak for 5 min before being frozen. To obtain AaThiL ADP−TPP crystals, AaThiL was incubated on ice with 4 mM ADP, 4 mM TPP, and 4 mM $MgCl_2$. Optimized conditions for the product complex included 100 mM imidazole (pH 8.0), 7–10% PEG 8000, and 100−250 mM $Ca(C_2H_3O_2)_2$. These crystals reached their maximum size of 200 μm × 80 μm × 80 μm within 1 week and were cryoprotected using 15% ethylene glycol in addition to the crystallization conditions.

Data sets were collected for AaThiL with AMP-PCP, AMP-PCP−TMP, and ADP−TPP complexes at the Advanced Photon Source 24-ID-C beamline using a Quantum315 detector (Area Detector Systems Corp.) at the wavelength 0.9795 Å. The data set for AaThiL complexed with ATP was collected at the A1 station at the Cornell High Energy Synchrotron Source (CHESS) using a Quantum210 detector (Area Detector Systems Corp.) and a wavelength of 0.9771 Å. All data sets were collected using an oscillation range of 0.5° to resolve the diffraction patterns. Crystals were indexed in the $P2_12_12_1$ space group with the following cell dimensions: $a = 61.2$ Å, $b = 67.1$ Å, and $c = 203.5$ Å for the binary complexes (giving a Matthews coefficient of 2.82 and a solvent content of 56%). The binary structure with AMP-PCP bound was determined to 2.20 Å resolution, and the structure with ATP bound was determined to 1.98 Å resolution. Cell dimensions were as follows: $a = 60.8$ Å, $b = 66.8$ Å, and $c = 197.0$ Å for the ternary complexes containing either TMP or TPP with a smaller Matthews coefficient of 2.67 and 54% solvent content (*11*). The substrate analogue ternary complex, AMP-PCP−TMP, was determined to 2.60 Å resolution, and the ADP−TPP product ternary complex was determined to 1.48 Å resolution. The data collection and refinement statistics are given in Tables 1 and 2.

Data were indexed, integrated, and scaled using the HKL2000 suite of programs (*12*). Structures were determined using molecular replacement with the thiamin monophosphate kinase structure (Protein Data Bank entry 1VQV) as the search model (*13*). Ligands and missing residues were manually added using COOT (*14*). Structures were then refined using rigid body, simulated annealing, *B* factor refinement, and minimization with CNS (*15*). Figures were generated using ChemDraw and PyMOL (*16*).

RESULTS

Structure of the AaThiL Protomer. The ternary complex models of AaThiL contain all 306 possible residues, as well as two to six residues of the His tag. However, the binary complexes lacking TMP or TPP have no density for the final 6−11 C-terminal residues, and they are not included in the model. The AaThiL protomer is composed of two domains, A and B, as seen in Figure 1A. The A domain consists of the first 140 residues and adopts an α/β-fold. The mixed β-sheet has a strand order of $\beta1\uparrow\beta2\downarrow\beta5\uparrow\beta3\downarrow\beta4\downarrow$, shown in Figure 1B. The β-strands are unusually long, with $\beta2$ and $\beta3$ each having 11 residues, $\beta5$ 10 residues, and $\beta9$ nine residues. The fifth strand, $\beta1$, is shorter with only five residues. The β-sheet is flanked on one side by four α-helices.

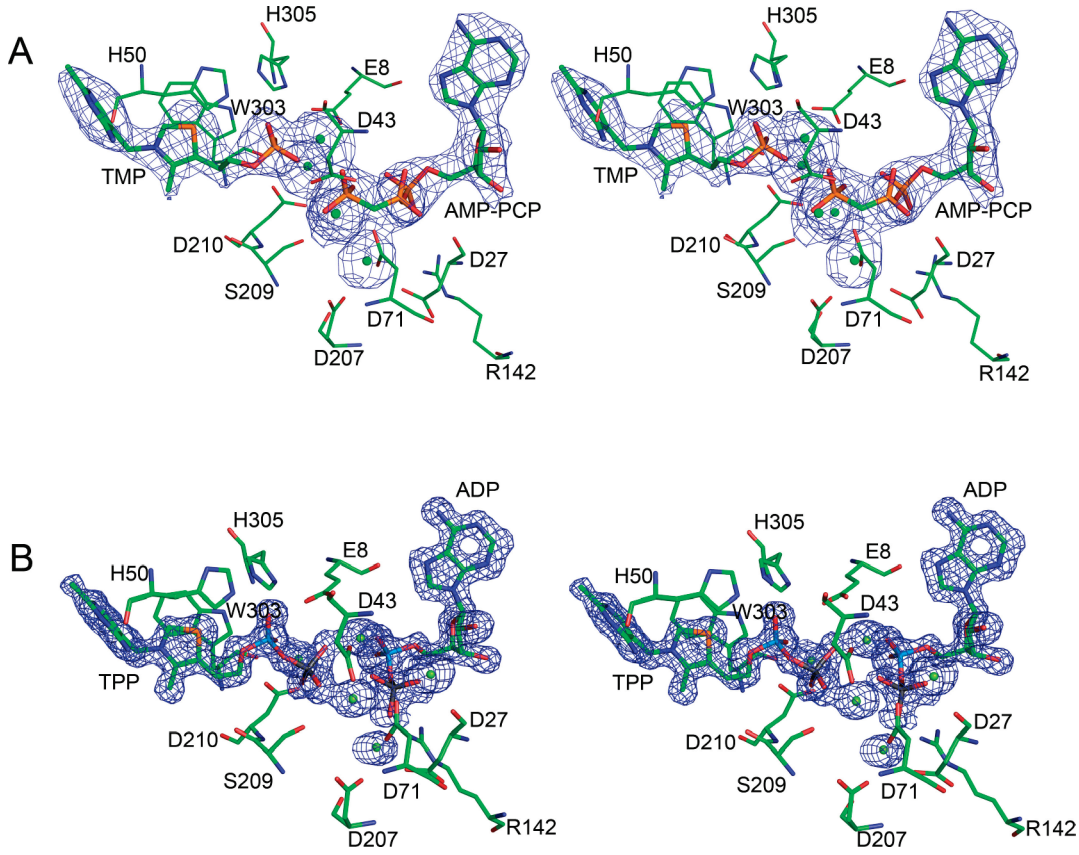


FIGURE 3: Stereoviews of the active site of the AaThiL ternary complexes. (A) Stereoview of the active site in the AMP-PCP and TMP structure. The $1F_O - F_C$ map is contoured at 2.5σ and is colored blue. (B) Stereoview of the active site in the ADP and TPP structure. The magnesium ions, shown as green spheres, have moved to adjust to the shift of the transferred phosphate group. The $1F_O - F_C$ map is again contoured at 2.5σ and is colored blue.

Two of the α-helices, α3 and α4, are also longer than average, each with 18 residues. The B domain, consisting of the 160 C-terminal residues, has an α/β-fold as well. The mixed β-sheet has a strand order of $\beta7\downarrow\beta9\uparrow\beta6\downarrow\beta10\uparrow\beta8\downarrow\beta11\downarrow$. The β-strands are much shorter than the β-strands found in the A domain. The longest β-strand, $\beta10$, has eight residues, and the two shortest strands, $\beta8$ and $\beta11$, each have four residues. $\beta10$ has a sharp twist, seen in Figure 1A, that results in a bend in the β-sheet. The β-sheet is flanked on both sides by α-helices. Five α-helices, α5, α6, and α8−α10, are found on one side of the β-sheet, while α7 is the only α-helix on the other side of the β-sheet. The A domain and B domain are linked by a short connective loop between $\beta5$ and $\beta6$.

Structure of the AaThiL Dimer. The unit cell contains one dimer per asymmetric unit (Figure 1C). The surface area of the dimer is approximately 25500 Å^2, with 6300 Å^2 buried at the interface between protomers (*17*). The primary interaction between protomers occurs in the A domain in which the 2-fold related β-sheets come together to form an eight stranded β-barrel. The short $\beta1$ strand stacks on top of $\beta4$ to form one long strand that is of a length comparable to that of the other β-strands in the β-barrel. This β-barrel is flanked by α-helices from the A domains. α2 of one protomer and α4 of the second protomer run parallel to each other, and hydrophobic residues along each α-helix interact with the opposite α-helix. A noncrystallographic 2-fold axis runs through the β-barrel. The interface between protomers is stabilized primarily through hydrophobic interactions and the packing of β-strands against each other. A disulfide bond forms between Cys34 and Cys34*, where the asterisk indicates the 2-fold-related protomer. Ten hydrogen bonds form between protomers. The carboxylate group of Glu20 forms a hydrogen bond to the amine of Lys105*; the amide group of Asn46 is hydrogen bonded to both the hydroxyl group of Ser121 and the oxygen atom of the amide group of the side chain of Asn87, and the carboxylate of Asp91 is hydrogen bonded to the amine of Lys4. The hydroxyl group of Tyr33 is hydrogen bonded to the carboxylate group of Glu135, and the amine of Lys125 forms a hydrogen bond to the carboxylate of Glu124. In addition to these hydrogen bonds, four hydrogen bonds are formed between the backbone atoms of the two protomers. The carbonyl oxygen atoms of Met1, Asp91, Leu92, and Glu90 form hydrogen bonds to the nitrogen atoms of the amide bonds in Val94, Arg2, Leu3, and Lys4, respectively. The ends of the β-strands are composed of mostly hydrophilic residues that are exposed to solvent. Additionally, several highly conserved aspartate residues that contribute to the formation of the active site are found within the β-barrel facing away from the β-barrel core.

ATP Binding Site. The AaThiL dimer has two equivalent active sites, located mostly within a single protomer but with some interactions from the opposite protomer. The active site is located in a cleft between the A and B domains and also involves β-strands $\beta3$ and $\beta4$ of the A domain of the 2-fold-related protomer. The interactions between the residues of the active site and ATP are shown in Figures 2A and 3. ATP binds deeply within the active site cleft and has no exposure to the solvent. The adenine base of ATP binds in a highly hydrophobic pocket, rich with highly conserved

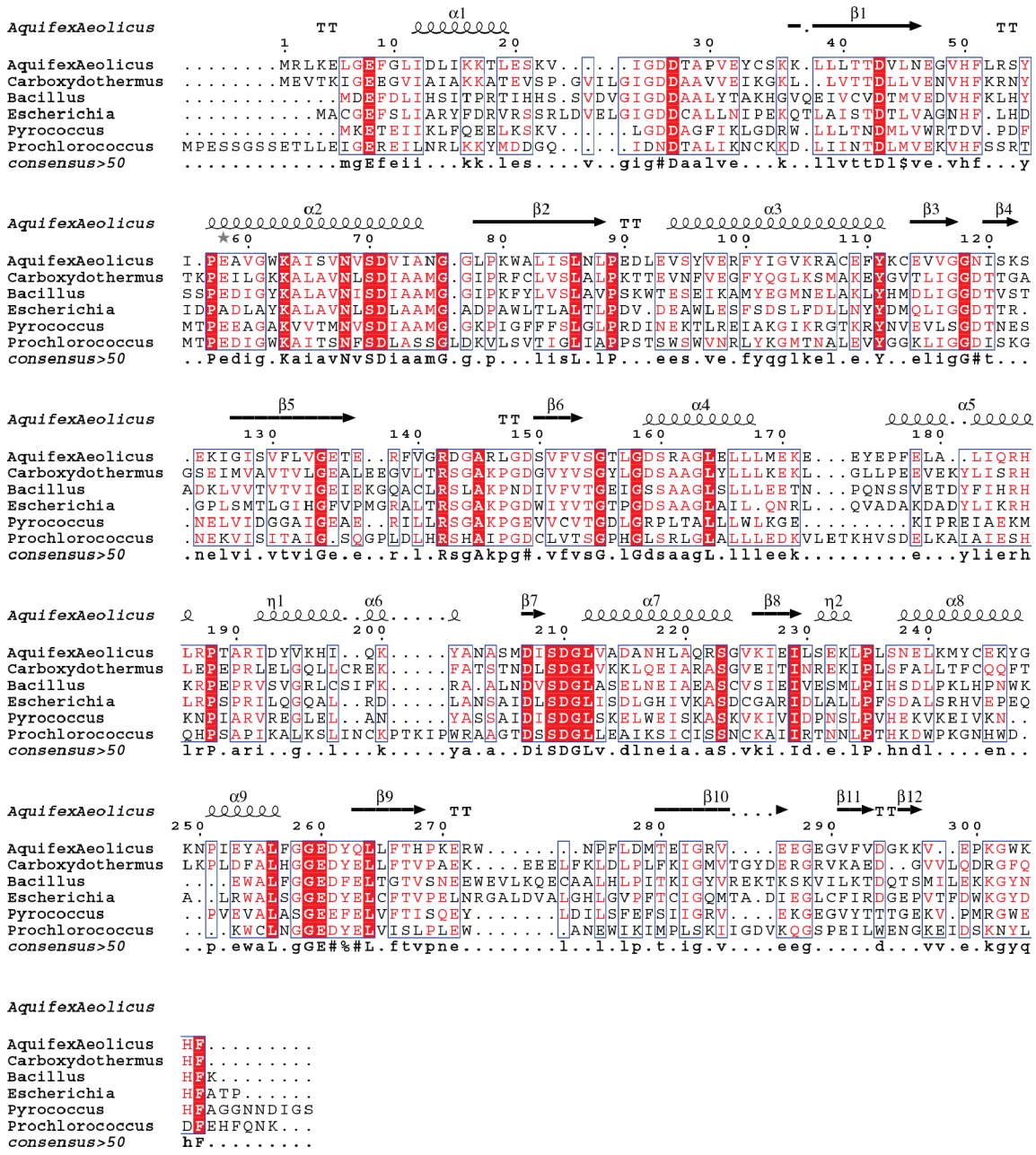


FIGURE 4: Sequence alignment of randomly selected ThiL proteins. Strictly conserved residues are colored white and highlighted with a red background, while less conserved residues are colored red. The secondary structure from AaThiL is shown above the alignment. Alignments were performed using ClustalW (*34*) and ESPript (*35*).

Table 3: Summary of DALI Search Results

protein	PDB entry	*Z* score	rmsd	no. of aligned residues
ThiL	1yaw	42.2	1.1	287
TkHypE	2z1f	28.2	2.6	263
TmPurL (AMP-PCP/FGAR)	2hs4	19.6	3.1	249
EcPurM	1cli	19.6	3.1	242
StPurL	1t3t	14.2	3.0	240
TmPurL	1vk3	13.3	3.5	210
YjgF	1qu9	8.4	2.4	94
YjgH	1pf5	6.6	2.9	91
YjgF	2ewc	5.4	3.2	92

isoleucine and valine residues, formed on one side by the β4 strand from the opposing protomer. The hydroxyl group of Tyr101* forms hydrogen bonds to N6 and N1 of the adenine base. The amino group of the adenine base is also hydrogen bonded to the backbone carbonyl group of Asn119*. The hydroxyl groups of the ribose sugar form hydrogen bonds to three ordered water molecules that are involved in water-mediated hydrogen bonding to the protein backbone at Gly25, Asp26, Asp27, Thr28, and Ala29; these residues form a turn between strands β1 and β2. The phosphate tail in both binary complex structures is precisely positioned by the presence of three magnesium ions. These magnesium ions are coordinated to several strictly conserved aspartate residues, Asp27, Asp43, Asp71, and Asp207. Mg1 is coordinated to both the α-phosphate and β-phosphate groups of AMP-PCP or ATP and is also coordinated by the carboxylate group of Asp43, the carbonyl groups of the backbone at Thr42, Asn119*, and two water molecules. Mg2 coordinates the β-phosphate and γ-phosphate groups and is positioned by Asp43, Asp71, and one water molecule. The third magnesium ion is coordinated to the same phosphate

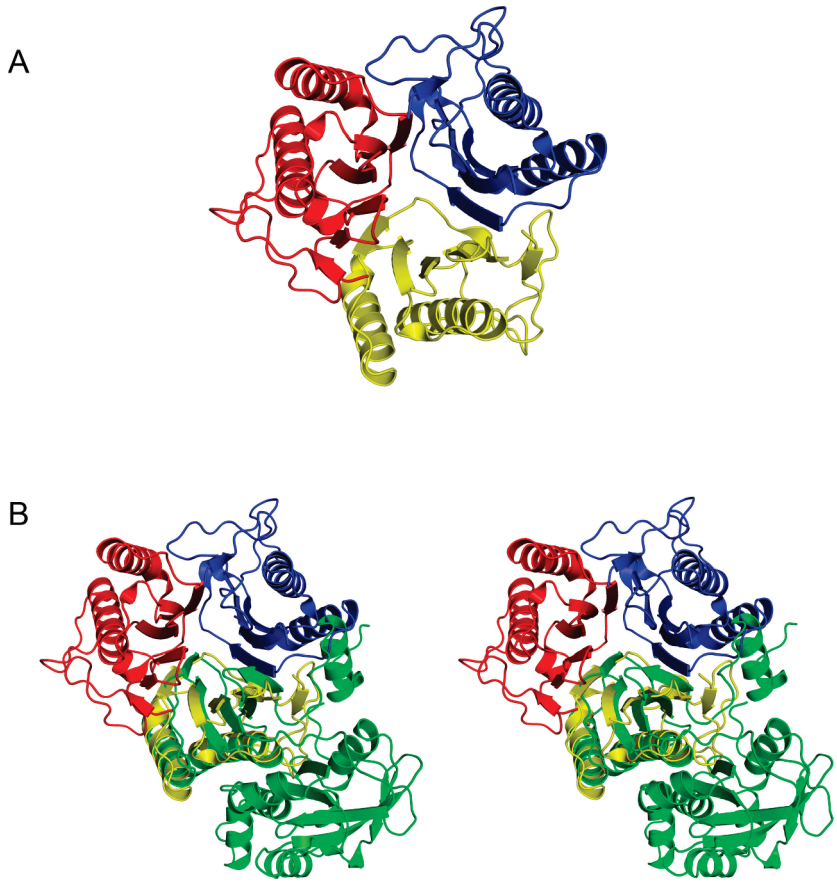


FIGURE 5: Comparison of the YjgF gene product from *E. coli* and AaThiL. (A) Structure of the trimer of YjgF. (B) Stereoview of the superpositioning of the YjgF trimer on the monomer of AaThiL. AaThiL is colored green, while the monomers of YjgF are colored yellow, blue, and red.

groups as Mg2 and is held in position by Asp27, Asp71, Asp207, and one water molecule. The β-phosphate group is also hydrogen bonded to Arg142, the only positively charged residue located within the active site. In the AMP-PCP−TMP complex, Mg1 moves 2.9 Å to coordinate to both ligands while maintaining its coordination to the protein. The other magnesium ions remain coordinated to the same residues and phosphate groups observed in the binary complexes. In addition to the three magnesium ions seen in the binary complexes, a fourth magnesium ion is found coordinated to the oxygen atom of the β-phosphate group, which is not coordinated to any of the other magnesium ions. Mg4 is also coordinated to Thr41, Asp71, the carbonyl group of Asp27, and three water molecules. A fifth magnesium ion, Mg5, is found coordinated to the α-phosphate and β-phosphate groups of TPP, as well as three water molecules and Asp210.

The conversion of ATP to ADP results in movement of the magnesium ions. Mg1 remains coordinated to the α-phosphate and β-phosphate groups but loses its coordination to Thr42 and is instead bound by only the two water molecules, Asp43, and Asn119*. Mg2 coordinates to the same oxygen atom as Mg1 as well as Asp71, Asp43, and two water molecules. Asp27, Asp71, Asp207, a water molecule, and the β-phosphate group of ADP are coordinated to Mg3.

Thiamin Binding Site. ATP binds deeply within the active site, and in both binary complexes, the C-terminal end (6−10 residues) is disordered. Binding of TMP leads to ordering of the C-terminus. In the structure with AMP-PCP, the electron density of the A chain can be seen through Pro300 and for the B chain through Lys297. Ordering of the C-terminus in the ternary complex results in a lid forming over the active site, shielding the substrates from the solvent. Additionally, the ordering of the C-terminus causes the *c* axis of the unit cell to be shortened by approximately 6 Å (from 203 to 197 Å).

Several residues of the lid interact directly with TMP. Trp303 stabilizes the thiazole ring through π-stacking, and this residue, while not strictly conserved, is always an aromatic residue, as seen in the sequence alignment (Figure 4). Strictly conserved His305 is coordinated to an oxygen atom of the α-phosphate group of TMP, one of the few interactions between a positively charged residue and the phosphate tail. The pyrimidine portion of TMP is partially exposed to the solvent with few interactions between the pyrimidine ring and the active site residues. The amino group of the pyrimidine ring forms a hydrogen bond to Glu260, which is conserved among ThiL's. N3 of the pyrimidine ring is hydrogen bonded to an ordered water molecule, which provides water-mediated hydrogen bonding to the backbone carbonyl oxygen atom of His50 and the hydroxyl group of Tyr55. The phosphate group of TMP is positioned by a magnesium ion. Mg1 is coordinated to the phosphate group as well as to phosphate groups of AMP-PCP, Asp43, and Asn119*.

Conversion of TMP to TPP results in some movement of the magnesium ions. These interactions are shown in Figures 2B and 3. The γ-phosphate group of ATP moves approximately 2.3 Å to form the products ADP and TPP. Three magnesium ions, Mg1, Mg2, and Mg5, and Ser209 coordi-

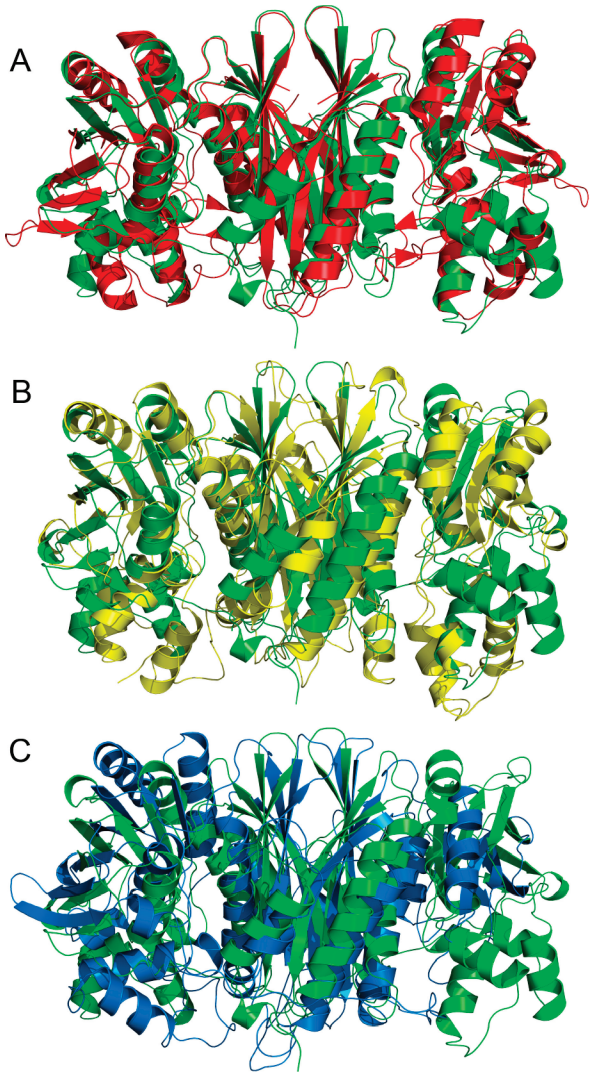


FIGURE 6: Comparison of AaThiL to other members of the PurM ATP binding superfamily. (A) Ribbon diagram of TkHypE superimposed on AaThiL. TkHypE is colored red and AaThiL green. (B) Ribbon diagram of EcPurM superimposed on AaThiL. EcPurM is colored yellow and AaThiL green. (C) Ribbon diagram of TmPurL superimposed on AaThiL. TmPurL is colored blue and AaThiL green.

nate the β-phosphate group of TPP. Mg1 is also coordinated to Asn119* and Asp43; Mg2 is positioned between the β-phosphate group of TPP and the β-phosphate group of ADP, and Mg5 interacts with both the α-phosphate and β-phosphate groups of TPP. Attempts to prepare a binary complex with TMP were unsuccessful, suggesting that ATP binding precedes TMP binding.

DISCUSSION

Conformational Changes in ThiL upon Substrate Binding. The lack of substrate or product bound to ThiL in the deposited structure (Protein Data Bank entry 1VQV) presented several questions about the enzyme's mechanism of action. Our results showed that when ATP or AMP-PCP binds in the binary complexes, several disordered regions in the unliganded structure become ordered, although the C-terminus remains disordered. Residues 7−9, 25, 32, 33, 289, and 301–306 from the A chain and residues 25, 33–36, and 301–306 from the B chain are disordered in unliganded AaThiL. Gly7, Glu8, and Phe9 make up a loop between α1 and α2. Binding of ATP appears to order this loop, and the side chain of Glu8 points into the active site. The loop near the ATP molecule between α2 and β1, composed of residues 20–28, lacks density for Gly25, and the conformation of this loop changes upon binding of ATP. The unliganded AaThiL structure includes two phosphate ions bound to each protomer; however, these phosphate groups are located on the surface of the protein and do not correspond to phosphate group locations in the AaThiL active site.

Binding of TMP results in some additional changes in the structure of AaThiL. The rmsd is 0.6 Å between the AMP-PCP and AMP-PCP−TMP structures. The most noticeable and important structural change is the ordering of the C-terminal residues when TMP binds. The last five residues are mostly conserved among ThiL's, as seen in Figure 4. Trp303 is consistently an aromatic residue and stabilizes the bound thiamin group by providing π-stacking with the thiazole ring. His305 is strongly conserved, as is Phe306. His305 directly interacts with the α-phosphate group of TMP, providing stabilization. Phe306 fits neatly into a pocket of hydrophobic residues Leu3, Leu11, Leu88, Val97, and Ile120. This pocket shows no movement or conformational changes of residues resulting from TMP binding. Within the active site, two additional magnesium ions, Mg4 and Mg5, are bound in the ternary substrate complex. These magnesium ions play a key role in positioning the phosphate group of TMP. The highly conserved aspartate residues, Asp27, Asp43, Asp71, and Asp207, retain the side chain conformations seen in the binary complexes when TMP binds.

Little change occurs with the conversion of substrates to products, with an rmsd between these two structures of 0.4 Å. The γ-phosphate group moves 2.4 Å toward TMP to become the β-phosphate group of TPP, and in turn, Mg1 moves 2.5 Å to maintain its coordination to this phosphate group and help stabilize the transition state. The β-phosphate and α-phosphate groups of ADP and ATP do not move, and the α-phosphate of TMP and TPP also remains in place. The side chains of all active site residues, including the residues of the C-terminal tail, are in the same conformation in the substrate complex and the product complex.

Comparison of ThiL to Other Proteins. The structure determination of *E. coli* PurM (EcPurM) revealed that this enzyme adopted a novel fold that binds ATP (*5*). BLAST searching using this enzyme identified four other enzymes, PurL, SelD, HypE, and ThiL, as likely members of this ATP binding superfamily. These five enzymes each contain a short signature amino acid sequence, Dx_4GA/GxP, which characterizes members of the superfamily. Thus far, no other sequence motif has been identified for this superfamily.

Structures have been determined for every member of the superfamily except SelD. There are two forms of PurL, designated large and small. The structure of *Salmonella typhimurium* PurL (StPurL) (*6*) represents the large PurL's, which are found in eukaryotes and Gram-negative bacteria. StPurL (140 kDa) is a multifunctional enzyme in which the formylglycinamide ribonucleotide synthetase domain shows 2-fold pseudosymmetry and adopts the fold of a PurM dimer. Small PurL's are found in Gram-positive bacteria and are represented by the structure of *Thermatoga maritima* PurL (TmPurL) (*18*). TmPurL (80 kDa) corresponds to the formylglycinamide ribonucleotide synthetase domain of large

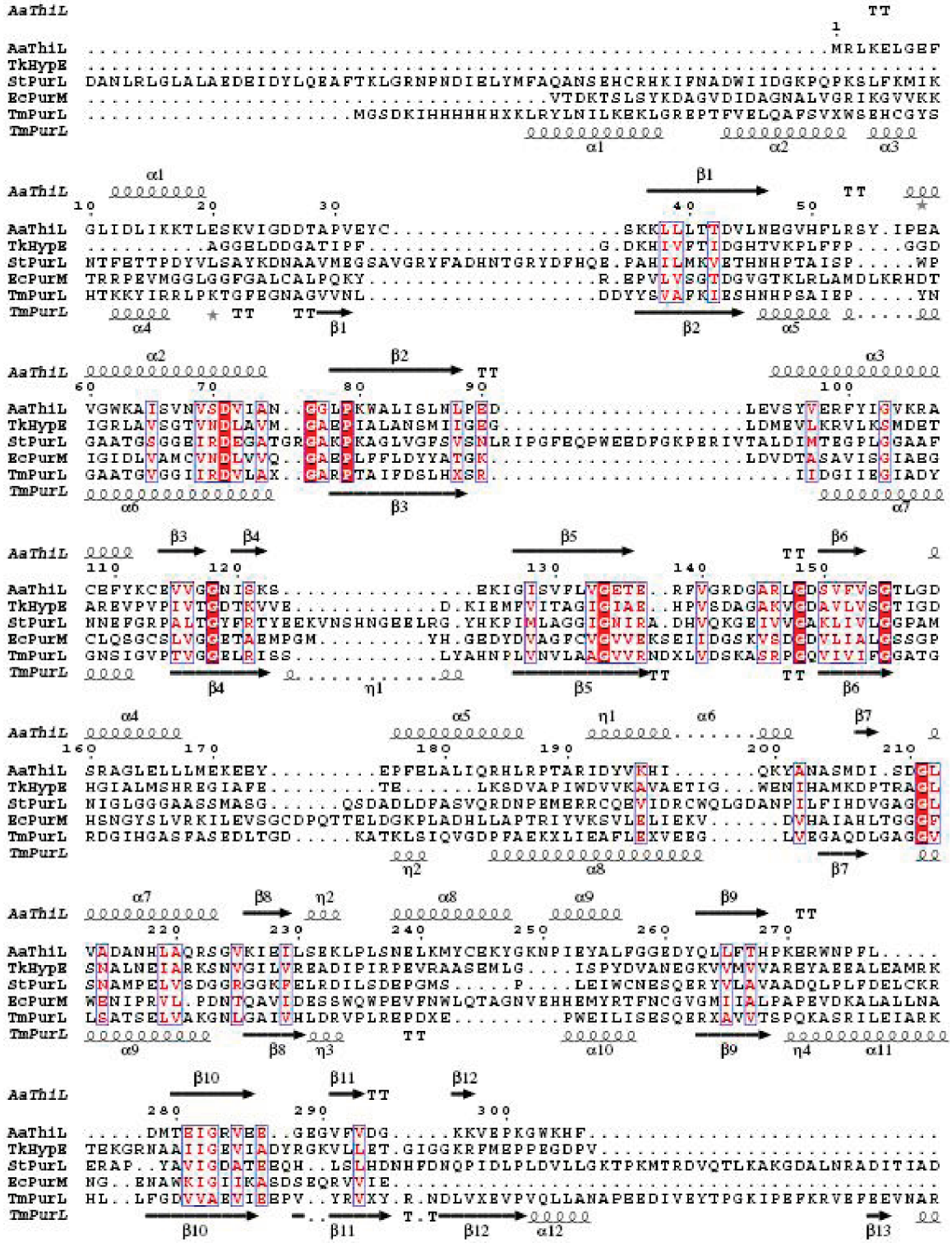


FIGURE 7: Structural sequence alignment of all available structures of the PurM ATP binding superfamily. The conserved residues are colored white and highlighted with a red background, while less conserved residues are colored red. The only conserved residues are the signature sequence for the superfamily and glycine residues important for folding. The structural alignment was done using VAST (*23*).

PurL. TmPurL and other small PurL's require two additional enzymes, PurS and PurQ, to be active. Recently, the structure of *Thermococcus kodakaraensis* KOD1 HypE (TkHypE), involved in the maturation of NiFe hydrogenase, was determined (*19*). Like ThiL and PurM, HypE is a homodimer in the crystal structure.

A DALI search was performed using AaThiL as the search structure, and the results are summarized in Table 3 (*20*). The proteins most structurally similar to ThiL are as expected the other members of the PurM ATP binding superfamily. The top DALI score corresponded to HypE. HypE requires carbamoylation at a conserved cysteine residue at the C-terminus; however, this modification was not present in the crystal structure (*19*). An ATP-dependent dehydration transfers the resulting cyanide ligand to a HypCD complex. Although TkHypE forms a dimer crystallographically, the

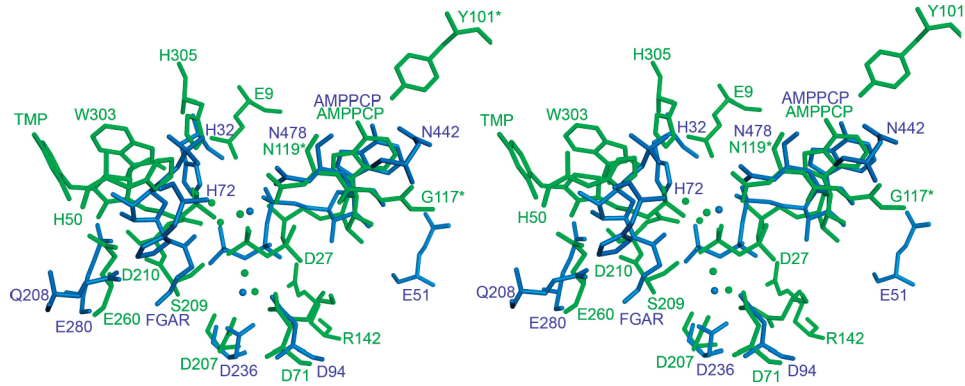


FIGURE 8: Superposition of the active sites of AaThiL and TmPurL. The ThiL structure is colored green and the PurL structure blue.

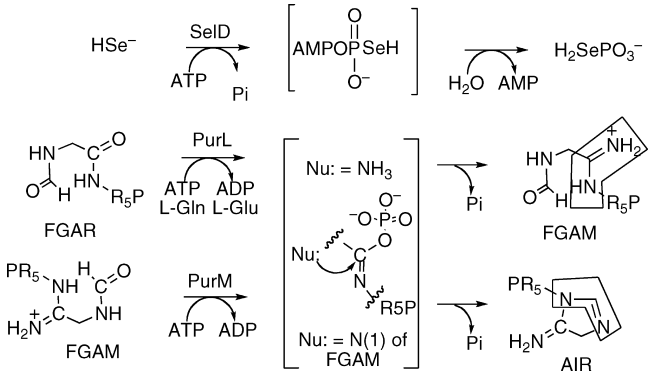


FIGURE 9: Proposed phosphoimidate intermediates for SelD, PurL, and PurM.

biologically active form is uncertain because *E. coli* HypE has also been reported to be a monomer (*21*).

AaThiL and TkHypE were structurally aligned using the DaliLite Pairwise comparison (*22*). The *Z* score for these two structures was 28.2, and the rmsd was 2.6 Å, reflecting the similarity of these two structures (Figure 6A). No density for ATP is observed in the HypE structure; however, the C-terminal tail shows different conformations depending on whether ATP was added during crystallization. Inclusion of ATP in the crystallization conditions resulted in an inward conformation in which the C-terminal tail folded back into the active site, which suggests that after carbamoylation at Cys338 the C-terminal end could enter the active site for dehydration. The position of Cys338 corresponds to the thiazole ring in the AaThiL structures. For both AaThiL and TkHypE, the C-terminal tail appears to form a lid over the active site when both substrates are present.

The superpositions of EcPurM and TmPurL upon AaThiL are shown in panels B and C of Figure 6, respectively. In both superpositions, the A domains are more structurally conserved than the B domains. The rmsd for the entire EcPurM structure when compared to AaThiL is 3.2 Å and drops to 2.1 Å when only the A domains are compared. The TmPurL structure has an overall rmsd of 3.2 Å, and the rmsd of the A domain is 2.0 Å. Comparisons of the B domains yield rmsd values similar to that of the overall structure, 3.1 Å for EcPurM and 3.0 Å for TmPurL. The higher degree of similarity in the A domains for the members of the PurM ATP binding superfamily is not surprising as the A domains are responsible for forming the dimer interface and the binding of ATP occurs almost entirely within the A domain. The B domain interacts with the second substrate and is expected to show greater variation because each enzyme has evolved to catalyze a different chemical reaction.

Despite this strong structural similarity, the level of sequence identity is low between AaThiL and other PurM superfamily members, ranging from 20% identity with TkHypE to 13% identity with StPurL. The only conserved sequence based on multiple-sequence alignments is the signature sequence motif of Dx_4GA/GxP. To identify any other conserved sequence motifs, a structural alignment was carried out using VAST (*23*). As seen in Figure 7, the signature sequence motif is structurally conserved, as well as five glycine residues, but no other residue is conserved within the superfamily. The aspartate residue from the signature motif is responsible for coordinating a magnesium ion in the active sites of TmPurL and AaThiL and, by inference, the active sites of StPurL, EcPurM, and HypE for which no complex structures are available. This magnesium ion coordinates the β-phosphate and γ-phosphate groups of ATP. The conserved glycine residues are found at the ends of β-strands or α-helices and most likely play structural roles within the PurM ATP binding superfamily.

Also identified through the DALI search with *Z* scores ranging from 8.4 to 5.4 and rmsd values between 2.4 and 3.1 Å for 90−95 residues aligned were a series of proteins belonging to the YjgF family. The function of the YjgF family of proteins is largely unassigned, although one family member, YabJ, has been implicated in the regulation of purines (*24*). Members of the YjgF family of proteins (Protein Data Bank entry 1QU9) adopt a fold that is similar to that of chorismate mutase with a six-stranded mixed β-sheet, $\beta1\uparrow\beta2\downarrow\beta3\uparrow\beta6\downarrow\beta4\uparrow\beta5\uparrow$, flanked on one side by two α-helices (*25*). Like the β-strands of the A domain of the PurM superfamily, these β-strands are unusually long and contain 8−10 residues. ThiL has a strand order of $\beta1\uparrow\beta2\downarrow\beta5\uparrow\beta3\downarrow\beta4\downarrow$, and in the DALI alignment, strands $\beta1$, $\beta2$, $\beta5$, $\beta3$, and $\beta4$ of AaThiL superimpose on strands $\beta2$, $\beta3$, $\beta6$, $\beta4$, and $\beta5$, respectively, of YjgF. YjgF is a trimer and forms a 12-stranded central β-barrel that is reminiscent of the eight-stranded central β-barrel of the AaThiL dimer. Clefts that are postulated to be in the active site are formed between the subunits of the YjgF trimer. After the DALI alignment, the cleft of YjgF superimposes on the active site of ThiL (Figure 5B). Examination of the YjgF cleft and the active site of ThiL failed to reveal the conserved aspartate residue from the signature motif. Therefore, the shared structural motif is unlikely to bind ATP in the YjgF family.

Active Site Comparison. While structures of EcPurM, StPurL, and TkHypE have nothing bound, the structures of both AaThiL and TmPurL have been determined with substrates or products bound to the active site (*18*). The superposition of the AaThiL and TmPurL active sites is shown in Figure 8. The two AMP-PCP molecules bind similarly, although the ribose moiety is slightly rotated in TmPurL. This results in a slight shift in the phosphate tail placing the γ-phosphate groups 2.3 Å apart. Overall, the secondary structural elements surrounding AMP-PCP superimpose well. Two magnesium ions, Mg1 and Mg2, are present in the active sites of both structures and precisely position ATP for phosphate transfer. A third magnesium ion in the AaThiL structure, Mg3, has no structural equivalent in the TmPurL structure and coordinates the α-phosphate and β-phosphate groups of AMP-PCP. In other ATP-dependent enzymes, this role of positioning the phosphate tail is often played by positively charged residues, which are lacking in the PurM ATP binding superfamily.

The TMP and formylglycinamide ribonucleotide (FGAR) binding sites are in similar locations but show fewer similarities to the ATP binding sites. Both substrates have a phosphate group; however, the phosphate group plays a different role in the two enzymes. The phosphate group of TMP is directly involved in the reaction as the acceptor of the γ-phosphate group of ATP. The TMP phosphate group is pointed into the active site and is aligned with the γ-phosphate group of AMP-PCP. The transfer is facilitated by Mg1 and Mg2 in AaThiL. In contrast, the phosphate group of FGAR is not directly involved in the reaction catalyzed by PurL; rather, it is the carbonyl group of FGAR that is phosphorylated. The oxygen atom of the carbonyl group of FGAR is 2.1 Å from the oxygen atom of the phosphate group of TMP, and no magnesium ions are located near FGAR in the TmPurL structure. Each structure has two conserved histidine residues that form hydrogen bonds to TMP or FGAR; no other positively charged residues hydrogen bond to TMP, and FGAR has only one additional hydrogen bond to a positively charged residue.

Comparison to Proteins Binding Thiamin. TPP in cells usually acts as a cofactor where it participates in a variety of enzymatic reactions. TPP-dependent enzymes require either a magnesium ion or a calcium ion, and the TPP cofactors exist in similar conformations when bound to TPP-dependent enzymes (*26*). For example, in yeast pyruvate decarboxylase (Protein Data Bank entry 1PYD) (*27*) and yeast transketolase (Protein Data Bank entry 1AY0) (*28*), N4 is positioned to deprotonate C2 of the thiazole ring. In contrast, TPP when not serving as a cofactor, as well as thiamin and TMP, is usually found in different conformations when bound to proteins. For example, in thiamin phosphate synthase (Protein Data Bank entry 1G4S), TMP is bound such that N4 is pointed away from C2 (*29*). In this conformation, N4 is unable to deprotonate C2, preventing the carbanion from forming (*26*). In AaThiL, the pyrimidine ring of TPP is rotated such that N4 is pointed toward C2 of the thiazole ring; however, the conformation TPP adopts in ThiL is different from both described above. While N4 is pointed toward C2, the pyrimidine ring is rotated 100° compared to TPP as a cofactor, which may prevent TPP activation in the ThiL binding site.

Mechanistic Implications. The presence of a phosphoimidate intermediate (Figure 9) was proposed as a common theme among the members of the PurM superfamily (*6*). Positional isotope exchange studies performed on SelD suggest a phosphorylated enzyme intermediate, although the phosphorylated residue has not been definitively identified (*30, 31*). Studies on PurL in which the amide carbonyl oxygen atom from FGAR is labeled with ^{18}O show that this oxygen atom is found in the inorganic phosphate product, also supporting a phosphoimidate intermediate (*32*). These observations suggest that ThiL might transfer the phosphate to TMP indirectly through a phosphoenzyme intermediate rather than directly through an inline mechanism (*33*). Inspection of the AaThiL complexes revealed no candidate (e.g., backbone amide or amino acid side chain) with suitable geometry for formation of a phosphoenzyme intermediate. The binary complexes showed that Ser209 is possibly positioned for phosphorylation; however, the ternary complexes with AMP-PCP and TMP or with ADP and TPP showed that Ser209 is better positioned to provide a hydrogen bond to the γ-phosphate of ATP or the β-phosphate of TPP and is poorly positioned to take part in a double-displacement mechanism. Additionally, the γ-phosphate of ATP is positioned only 2.7 Å from the β-phosphate of TPP with an ideal geometry for inline transfer. The active site of ThiL is relatively simple; only two of the protein side chains participate directly in catalysis (Arg142 and Ser209). The five magnesium ions and several water molecules present in the ternary substrate and product complexes orient the substrates, stabilize the trigonal bipyramidal intermediate, and activate ADP as a leaving group (Figure 2). These observations support a direct inline transfer rather than an indirect double transfer via a phosphoenzyme intermediate.

ACKNOWLEDGMENT

This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines of the Advanced Photon Source, supported by Grant RR-15301 from the National Center for Research Resources at the National Institutes of Health. Use of the Advanced Photon Source is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, under Contract W-31-109-ENG-38. We also thank the staff at the CHESS A1 beamline as well as the staff at the Advanced Photon Source's 24-ID-C beamline for their assistance during data collection. Leslie Kinsland's help in preparing the manuscript is greatly appreciated.

REFERENCES

1. Settembre, E., Begley, T. P., and Ealick, S. E. (2003) Structural biology of enzymes of the thiamin biosynthesis pathway. *Curr. Opin. Struct. Biol. 13*, 739–747.
2. Jordan, F. (2003) Current mechanistic understanding of thiamin diphosphate-dependent enzymatic reactions. *Nat. Prod. Rep. 20*, 184–201.
3. Begley, T. P., Downs, D. M., Ealick, S. E., McLafferty, F. W., Van Loon, A. P., Taylor, S., Campobasso, N., Chiu, H. J., Kinsland, C., Reddick, J. J., and Xi, J. (1999) Thiamin biosynthesis in prokaryotes. *Arch. Microbiol. 171*, 293–300.
4. Webb, E., and Downs, D. (1997) Characterization of thiL, encoding thiamin-monophosphate kinase, in *Salmonella typhimurium*. *J. Biol. Chem. 272*, 15702–15707.
5. Li, C., Kappock, T. J., Stubbe, J., Weaver, T. M., and Ealick, S. E. (1999) X-ray crystal structure of aminoimidazole ribonucleotide

synthetase (PurM), from the *Escherichia coli* purine biosynthetic pathway at 2.5 Å resolution. *Structure 7*, 1155–1166.
6. Anand, R., Hoskins, A. A., Stubbe, J., and Ealick, S. E. (2004) Domain Organization of *Salmonella typhimurium* Formylglycinamide Ribonucleotide Amidotransferase Revealed by X-ray Crystallography. *Biochemistry 43*, 10328–10342.
7. Jayaraj, S., Reid, R., and Santi, D. V. (2005) GeMS: An advanced software package for designing synthetic genes. *Nucleic Acids Res. 33*, 3011–3016.
8. Kodumal, S. J., Patel, K. G., Reid, R., Menzella, H. G., Welch, M., and Santi, D. V. (2004) Total synthesis of long DNA sequences: Synthesis of a contiguous 32-kb polyketide synthase gene cluster. *Proc. Natl. Acad. Sci. U.S.A. 101*, 15573–15578.
9. Kodumal, S. J., and Santi, D. V. (2004) DNA ligation by selection. *BioTechniques 37*, 34.
10. Sambrook, J., Fritsch, G. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Guide*, Cold Spring Harbor Laboratory Press, Plainview, NY.
11. Matthews, B. W. (1968) Solvent content of protein crystals. *J. Mol. Biol. 33*, 491–497.
12. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307–326.
13. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res. 28*, 235–242.
14. Emsley, P., and Cowtan, K. (2004) Coot: Model-Building Tools fo Molecular Graphics. *Acta Crystallogr. D60*, 2126–2132.
15. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D54*, 905–921.
16. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA.
17. Krissinel, E., and Henrick, K. (2005) Detection of Protein Assemblies in Crystals, *CompLife 2005*, pp 163–174, Springer-Verlag, Berlin.
18. Morar, M., Anand, R., Hoskins, A. A., Stubbe, J., and Ealick, S. E. (2006) Complexed structures of formylglycinamide ribonucleotide amidotransferase from *Thermotoga maritima* describe a novel ATP binding protein superfamily. *Biochemistry 45*, 14880–14895.
19. Watanabe, S., Matsumi, R., Arai, T., Atomi, H., Imanaka, T., and Miki, K. (2007) Crystal structures of [NiFe] hydrogenase maturation proteins HypC, HypD, and HypE: Insights into cyanation reaction by thiol redox signaling. *Mol. Cell 27*, 29–40.
20. Holm, L., and Sander, C. (1998) Touring protein fold space with Dali/FSSP. *Nucleic Acids Res. 26*, 316–319.
21. Blokesch, M., Paschos, A., Bauer, A., Reissmann, S., Drapal, N., and Bock, A. (2004) Analysis of the transcarbamoylation-dehydration reaction catalyzed by the hydrogenase maturation proteins HypF and HypE. *Eur. J. Biochem. 271*, 3428–3436.
22. Holm, L., and Park, J. (2000) DaliLite workbench for protein structure comparison. *Bioinformatics 16*, 566–567.
23. Madej, T., Gibrat, J. F., and Bryant, S. H. (1995) Threading a database of protein cores. *Proteins 23*, 356–369.
24. Sinha, S., Rappu, P., Lange, S. C., Mantsala, P., Zalkin, H., and Smith, J. L. (1999) Crystal structure of *Bacillus subtilis* YabJ, a purine regulatory protein and member of the highly conserved YjgF family. *Proc. Natl. Acad. Sci. U.S.A. 96*, 13074–13079.
25. Volz, K. (1999) A test case for structure-based functional assignment: The 1.2 Å crystal structure of the yjgF gene product from *Escherichia coli*. *Protein Sci. 8*, 2428–2437.
26. Malandrinos, G., Louloudi, M., and Hadjiliadis, N. (2006) Thiamine models and perspectives on the mechanism of action of thiamine-dependent enzymes. *Chem. Soc. Rev. 35*, 684–692.
27. Dyda, F., Furey, W., Swaminathan, S., Sax, M., Farrenkopf, B., and Jordan, F. (1993) Catalytic centers in the thiamin diphosphate dependent enzyme pyruvate decarboxylase at 2.4-Å resolution. *Biochemistry 32*, 6165–6170.
28. Wikner, C., Nilsson, U., Meshalkina, L., Udekwu, C., Lindqvist, Y., and Schneider, G. (1997) Identification of catalytically important residues in yeast transketolase. *Biochemistry 36*, 15643–15649.
29. Peapus, D. H., Chiu, H. J., Campobasso, N., Reddick, J. J., Begley, T. P., and Ealick, S. E. (2001) Structural characterization of the enzyme-substrate, enzyme-intermediate, and enzyme-product complexes of thiamin phosphate synthase. *Biochemistry 40*, 10103–10114.
30. Mullins, L. S., Hong, S.-B., Gibson, G. E., Walker, H., Stadtman, T. C., and Raushel, F. M. (1997) Identification of a Phosphorylated Enzyme Intermediate in the Catalytic Mechanism for Selenophosphate Synthetase. *J. Am. Chem. Soc. 119*, 6684–6685.
31. Walker, H., Ferretti, J. A., and Stadtman, T. C. (1998) Isotope exchange studies on the *Escherichia coli* selenophosphate synthetase mechanism. *Proc. Natl. Acad. Sci. U.S.A. 95*, 2180–2185.
32. Schendel, F. J., Mueller, E., Stubbe, J., Shiau, A., and Smith, J. M. (1989) Formylglycinamide ribonucleotide synthetase from *Escherichia coli*: Cloning, sequencing, overproduction, isolation, and characterization. *Biochemistry 28*, 2459–2471.
33. Prasher, D. C., Carr, M. C., Ives, D. H., Tsai, T. C., and Frey, P. A. (1982) Nucleoside phosphotransferase from barley. Characterization and evidence for ping pong kinetics involving phosphoryl enzyme. *J. Biol. Chem. 257*, 4931–4939.
34. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res. 22*, 4673–4680.
35. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) ESPript: Analysis of multiple sequence alignments in PostScript. *Bioinformatics 15*, 305–308.

BI800041H