Structural Characterization of the Enzyme−Substrate, Enzyme−Intermediate, and Enzyme−Product Complexes of Thiamin Phosphate Synthase[†,‡]

Diane H. Peapus, Hsiu-Ju Chiu,[§] Nino Campobasso,[||] Jason J. Reddick, Tadhg P. Begley,* and Steven E. Ealick*

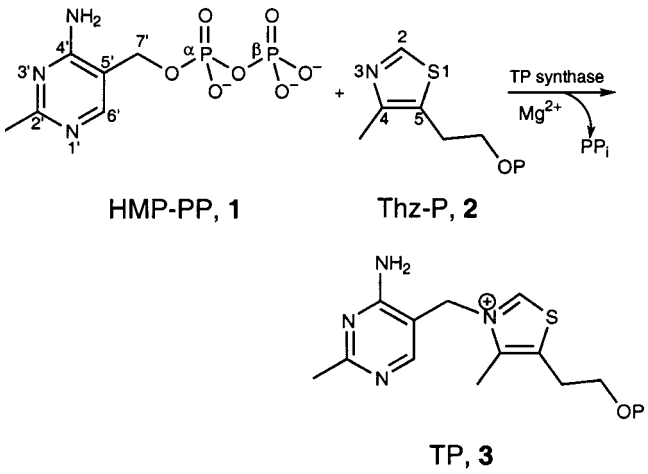
Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853

Received March 8, 2001; Revised Manuscript Received May 31, 2001

ABSTRACT: Thiamin phosphate synthase catalyzes the formation of thiamin phosphate from 4-amino-5-(hydroxymethyl)-2-methylpyrimidine pyrophosphate and 5-(hydroxyethyl)-4-methylthiazole phosphate. Several lines of evidence suggest that the reaction proceeds via a dissociative mechanism. The previously determined crystal structure of thiamin phosphate synthase in complex with the reaction products, thiamin phosphate and magnesium pyrophosphate, provided a view of the active site and suggested a number of additional experiments. We report here seven new crystal structures primarily involving crystals of S130A thiamin phosphate synthase soaked in solutions containing substrates or products. We prepared S130A thiamin phosphate synthase with the intent of characterizing the enzyme−substrate complex. Surprisingly, in three thiamin phosphate synthase structures, the active site density cannot be modeled as either substrates or products. For these structures, the best fit to the electron density is provided by a model that consists of independent pyrimidine, pyrophosphate, and thiazole phosphate fragments, consistent with a carbenium ion intermediate. The resulting carbenium ion is likely to be further stabilized by proton transfer from the pyrimidine amino group to the pyrophosphate to give the pyrimidine iminemethide, which we believe is the species that is observed in the crystal structures.

Thiamin phosphate synthase (TP[1] synthase) catalyzes the coupling of 4-amino-5-(hydroxymethyl)-2-methylpyrimidine pyrophosphate (**1**, HMP-PP) and 5-(hydroxyethyl)-4-methylthiazole phosphate (**2**, Thz-P) to form thiamin phosphate (**3**, TP) (Scheme 1). This reaction is the penultimate step in the biosynthesis of thiamin pyrophosphate. The gene for this enzyme has been cloned and overexpressed from *Escherichia coli* (*1*), *Bacillus subtilis* (*2*), *Saccharomyces cerevisiae* (*3*), and *Arabodopsis thaliana* (*4*), and orthologs can be identified in all thiamin-synthesizing organisms where sequence information is available. The structure of the enzyme from *B. subtilis*, with products bound at the active site, has recently been reported (*5*).

Scheme 1: TP Synthase-Catalyzed Reaction



The TP synthase reaction could occur by either an associative or dissociative mechanism. Several lines of evidence, including substituent effects and positional isotope exchange experiements, support a dissociative mechanism, generating a carbocation intermediate (*6*). We previously determined the crystal structure of TP synthase to identify key active site residues and to provide additional evidence concerning the reaction mechanism (*5*). This TP synthase structure showed that the products of the reaction, TP, pyrophosphate (PP_i), and Mg^{2+}, were bound at the active site. This was surprising because neither reactants nor products were added during the purification or crystallization procedures. We proposed that product release must be triggered by some unknown event such as the availability of substrate.

[†] This work was supported by National Institutes of Health grants to T.P.B. (DK44083) and by Hoffmann-La Roche (T.P.B.). S.E.E. is indebted to the W. M. Keck Foundation and the Lucille P. Markey Charitable Trust.

[‡] The Brookhaven Protein Data Bank codes for thiamin phosphate synthase complexes are 1G4E for S130A, 1G4P for the S130A−CF_3-HMP-PP complex, 1G4S for the S130A−TP−PP_i complex, 1G4T for the WT−CF_3TP−PP_i complex, 1G67 for the S130A−HMP−Thz-P−PP_i complex, 1G69 for the S130A−HMP−Thz-P−PP_i complex (fast soak), and 1G6C for the S130A−CF_3HMP−Thz-P−PP_i complex.

* 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 or tpb2@cornell.edu.

[§] Current address: Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA 91125.

[||] Current address: SmithKline Beecham Pharmaceuticals, King of Prussia, PA 19406.

[1] Abbreviations: HMP-PP, 4-amino-2-methyl-5-(hydroxymethyl)-pyrimidine pyrophosphate; CF_3HMP-PP, 4-amino-2-(trifluoromethyl)-5-(hydroxymethyl)pyrimidine pyrophosphate; Thz-P, 4-methyl-5-(hydroxyethyl)thiazole phosphate; TP, thiamin phosphate; CF_3TP, 2-(trifluoromethyl)thiamin phosphate; PP_i, pyrophosphate.

10.1021/bi0104726 CCC: $20.00
Published on Web 07/21/2001

Table 1: Data Collection Statistics

	S130A	S130A−substrate	S130A−product	WT−product	S130A−intermed 1	S130A−intermed 2	S130A−intermed 3
ligands added (20 mM)	none	CF_3HMP-PP	TP and PP_i	CF_3HMP-PP and Thz-P	HMP-PP and ThiZ-P	CF_3HMP-PP and Thz-P	HMP-PP and ThiZ-P
method	NA	12 h soak	12 h soak	cocrystallized	12 h soak	12 h soak	30 min soak
resolution (Å)	1.6	2.5	1.7	1.55	1.4	1.4	1.5
unit cell *a*, *b* (Å)	76.37, 139.15	76.28, 139.59	76.99, 139.85	76.79, 140.33	76.85, 140.25	76.67, 140.12	76.34, 139.15
source	CHESS F1	Cu Kα	CHESS F1	CHESS F1	CHESS F1	CHESS F1	APS-SBC
detector	ADSC Q4	SAXII CCD	ADSC Q4	ADSC Q4	ADSC Q4	ADSC Q4	SBC-1 CCD
no. of unique reflections	54596	14985	43894	60821	65601	77666	63282
redundancy	6.5	8.1	7.8	5.9	5.4	7.1	6.5
completeness[a] (%)	99.1 (95.0)	100 (99.8)	92.9 (87.8)	98.6 (92.6)	78.7 (64.3)	93.8 (83.1)	94.2 (69.2)
R_{sym}[a]	4.7 (16.2)	11.7 (32.0)	5.7 (15.9)	6.4 (24.8)	5.1 (22.9)	6.1 (12.9)	3.9 (34.4)
I/σ[a]	11.2 (4.5)	7.1 (1.2)	8.6 (3.5)	7.3 (2.1)	7.9 (3.2)	7.4 (4.8)	11.4 (2.0)

[a] Value for the highest-resolution shell in parentheses.

The observation of bound products clearly defined the active site and suggested roles for key amino acid side chains, leading us to focus on Ser130. This residue is hydrogen bonded to the PP_i oxygen atom that was originally bonded to the pyrimidine. Subsequent experiments demonstrated that the S130A mutant of TP synthase is 8000 times less active than the native enzyme (*6*). Additional experiments showed that CF_3HMP-PP is not a substrate for wild-type TP synthase (*6*). We used these observations as a basis for designing experiments to obtain structures of unliganded TP synthase and of an enzyme−substrate complex. We reasoned that inactive S130A TP synthase would lack bound product, thus providing a view of the unliganded enzyme. We further reasoned that using either the S130A mutant or the substrate analogue CF_3HMP-PP we would be able to prevent conversion of added substrates to products and thus provide a view of the enzyme−substrate complex. Surprisingly, the crystal structures of several complexes showed active site density that was consistent neither with reactants nor with products. The current studies focus on characterization of structures along the reaction coordinate, including unliganded TP synthase, the TP synthase−substrate complex, and a putative iminemethide intermediate, and further characterization of the TP synthase−product complex.

MATERIALS AND METHODS

Preparation of Native and S130A TP Synthase. Overexpression and purification of wild-type TP synthase and preparation of the inactive S130A TP synthase mutant were performed as previously described (*5*) with the following exception. After expression had been induced with 1 mM isopropyl thio-β-D-galactoside, the cells were incubated overnight and then purified as previously described. Syntheses of the substrates have been described by Reddick et al. (*6*).

Crystallization of TP Synthase Complexes. Crystals of wild-type TP synthase and the S130A mutant were grown in 50 mM Tris-HCl (pH 7.5), 75 mM $MgCl_2$, 21−22% (w/v) PEG4000, and 1 mM DTT, as previously described (*5*). Seeds of wild-type TP synthase were initially used for S130A. The resulting S130A crystals were then used to seed into fresh S130A protein to minimize contamination of wild-type TP synthase in the crystals. All substrate concentrations during soaking were 20 mM with an excess of Mg^{2+} ion already in the mother liquor. In the cocrystallized wild-type TP synthase−product complex, the protein was incubated with 20 mM substrates for 6 h prior to crystallization (Table 1). All complexes crystallize in space group $P4_32_12$ and are isomorphous to the previously reported structure ($a = 76.83$ Å; $c = 140.10$ Å) (*5*).

Data Collection and Processing. Crystals diffracted to very high resolution (1.4−1.7 Å) using synchrotron radiation. Synchrotron data were measured either at the Cornell High Energy Synchrotron Source (CHESS) using a Quantum 4 CCD detector (Area Detector Systems Corp.) or at the Structural Biology Center (SBC) at the Advanced Photon Source using the SBC-1 CCD detector. Data were typically collected in 1° oscillations with 45−60 s exposures at CHESS and 8−10 s exposures at SBC. CHESS data were processed using MOSFLM (*7*) and SCALA in CCP4 (*8*), while SBC data were processed using DENZO and SCALEPACK (*9*). An additional 2.5 Å data set was collected using a SAXII CCD detector (Brucker) and Cu Kα radiation from a rotating anode with 10 min/1° exposures. These data were processed using the accompanying program, SAINT. R_{sym} values for the synchrotron data were 3.9−6.1% (rotating anode data, 11.7%) with 78.7−99.1% completeness (rotating anode data, 100%). Data collection and data processing statistics are summarized in Table 1.

Structure Determination, Model Building, and Refinement. The structure of the S130A mutant TP synthase was determined as follows. The previously reported TP synthase−product complex structure (Protein Data Bank entry 2TPS) with substrate, water, and Ser130 Oγ atoms removed was used for rigid body fitting followed by simulated annealing refinement using CNS (*10*). A $2F_o - F_c$ composite omit map was constructed using the refined model. Side chains and regions showing poorly defined peptide backbone electron density were then manually adjusted using the program O (*11*) running on a COMPAQ Alpha workstation. The rebuilt model was refined, and water molecules were added using an $F_o - F_c$ map. A second composite omit map was then used to examine the active site contents and to identify alternate conformations of side chains. Occupancies of side chain atoms with alternate conformations were set to 50% for each conformer, and another round of least-squares refinement was performed during which individual *B*-factors were allowed to vary. The final model was examined using CNS and PROCHECK (*12*). The root-mean-square deviations between S130A and the starting model were evaluated

Table 2: Refinement Statistics

	S130A	S130A−substrate	S130A−product	WT−product	S130A−intermed 1	S130A−intermed 2	S130A−intermed 3
resolution (Å)	1.6	2.5	1.7	1.55	1.4	1.4	1.5
R-factor	20.9	20.7	20.4	20.9	21.2	21.2	21.0
R-free	23.9	27.4	23.4	23.3	23.6	23.1	22.8
Luzzati error	0.19	0.27	0.20	0.19	0.17	0.17	0.18
residues							
molecule A	10−100, 113−156, 162−235	10−155, 62−234	10−235	10−235	11−235	11−235	10−235
molecule B	9−111, 113−155, 162−235	10−155, 163−234	10−235	9−235	11−235	9−234	8−235
no. of non-H atoms							
no. of proteins	3257	3255	3382	3394	3362	3381	3402
no. of ligands	none	44	64	70	64	70	64
no. of waters	321	176	298	351	263	293	279
B-factors (Å^2)							
protein	14.19	8.65	17.42	14.06	16.01	13.23	13.07
ligand	NA[a]	5.21	14.80	11.31	12.52	9.87	21.53
water	26.14	16.82	27.16	25.91	25.95	23.78	23.03
rms deviations							
bonds (Å)	0.0043	0.0062	0.0051	0.0052	0.0054	0.0049	0.0046
angles (deg)	1.1	1.2	1.1	1.1	1.1	1.1	1.1
dihedrals (deg)	21	22	21	21	21	22	21

[a] Not applicable.

using ProFit (*13*).

Structure determination for the complexes was carried out using the unliganded S130A model as a starting point. After rigid body refinement and simulated annealing refinement, $F_o - F_c$ maps were constructed along with $2F_o - F_c$ composite omit maps. The active site ligands were clearly observed in the $F_o - F_c$ maps, but ligand models were not included in the refinement until later steps to minimize model bias. Adjustment of the protein model and addition of water molecules proceeded as described above for the S130A structure. After all protein and water atoms were refined, the appropriate ligand model was manually fit into the active site followed by a final cycle of refinement. The final refinement statistics are summarized in Table 2.

Thiamin Phosphate Release from TP Synthase. A 500 μL aliquot of 0.8 μM TP synthase in 50 mM Tris-HCl (pH 7.5) and 6 mM $MgCl_2$ was incubated at room temperature in the presence of 50 μM CF_3HMP-PP, 50 μM TP, or both. After 10 min, the small molecule fraction was separated from the protein by ultrafiltration through a 5000 kDa molecular mass cutoff Vivaspin concentrator according to the manufacturer's instructions. The amount of thiamin phosphate in the filtrate was determined using a modification of the thiochrome assay previously described (*5*). Two control experiments were performed with samples containing enzyme but no substrates. The first was incubated and separated as described above to determine the amount of TP released under the experimental conditions. The second control was incubated with 10% trichloroacetic acid for 15 min to determine the amount of unreleased TP in the initial samples. The analyte was then separated from the protein by ultrafiltration as before. TP concentrations in the filtrate of both controls were measured using the modified thiochrome assay (*5*).

RESULTS

Overall Structure of TP Synthase. A total of seven TP synthase structures were determined, including one wild-type TP synthase complex, one unliganded S130A mutant TP synthase, and five S130A mutant TP synthase complexes. Each model consists of two molecules of TP synthase per asymmetric unit, designated as molecules A and B. Except for the unliganded S130A structure, all of the structures showed that both the A and B active sites are occupied by ligands. Each model contains several hundred water molecules, including a subset that is mostly conserved in the active sites of all seven structures as well as in the previously reported structure of the wild-type TP synthase−product complex (*5*).

TP synthase has an α/β structure with a triosephosphate isomerase (TIM barrel) fold. The TIM barrel has two additional helices, an N-terminal α0 helix (Arg19−Leu25) and a single turn of helix, α8a (Ser209−Ser211), between β8 and α8 (Figure 1). Seven of the eight $\beta\alpha$ loops, 2−8, contact the ligands in the active site. Loops 2−5 are involved primarily with HMP-PP interactions, while loops 6−8 are involved primarily in Thz-P interactions.

The structures refined to final *R* values between 20.7 and 22.5%. The coordinate errors based on Luzzati plots (*14*) were between 0.17 and 0.27 Å. Ramachandran plots (*15*) showed that all non-glycine residues were in allowed regions except Asp93, which is ligated to the Mg^{2+} ion in the active site, and Ile208, which contacts Thz-P. The average *B*-factors for the structures range from 8.6 to 17.2 Å^2 for protein atoms, from 5.2 to 21.5 Å^2 for ligand atoms, and from 16.8 to 27.2 Å^2 for water oxygen atoms (Table 2). Weak backbone density is observed at the Gly62-Gly63 portion in all TP synthase structures. In those complexes that lack a thiazole moiety, no electron density is observed between Glu157 and Thr162, which are located in a loop that covers the active site. The models do not include the polyhistidine tag, which was disordered in molecules A and B for all structures. Root-mean-square deviations (rmsd) calculated using Cα atoms show the largest variability between molecule A and molecule B to be in four regions: (1) near α2 (rmsd = 0.4−

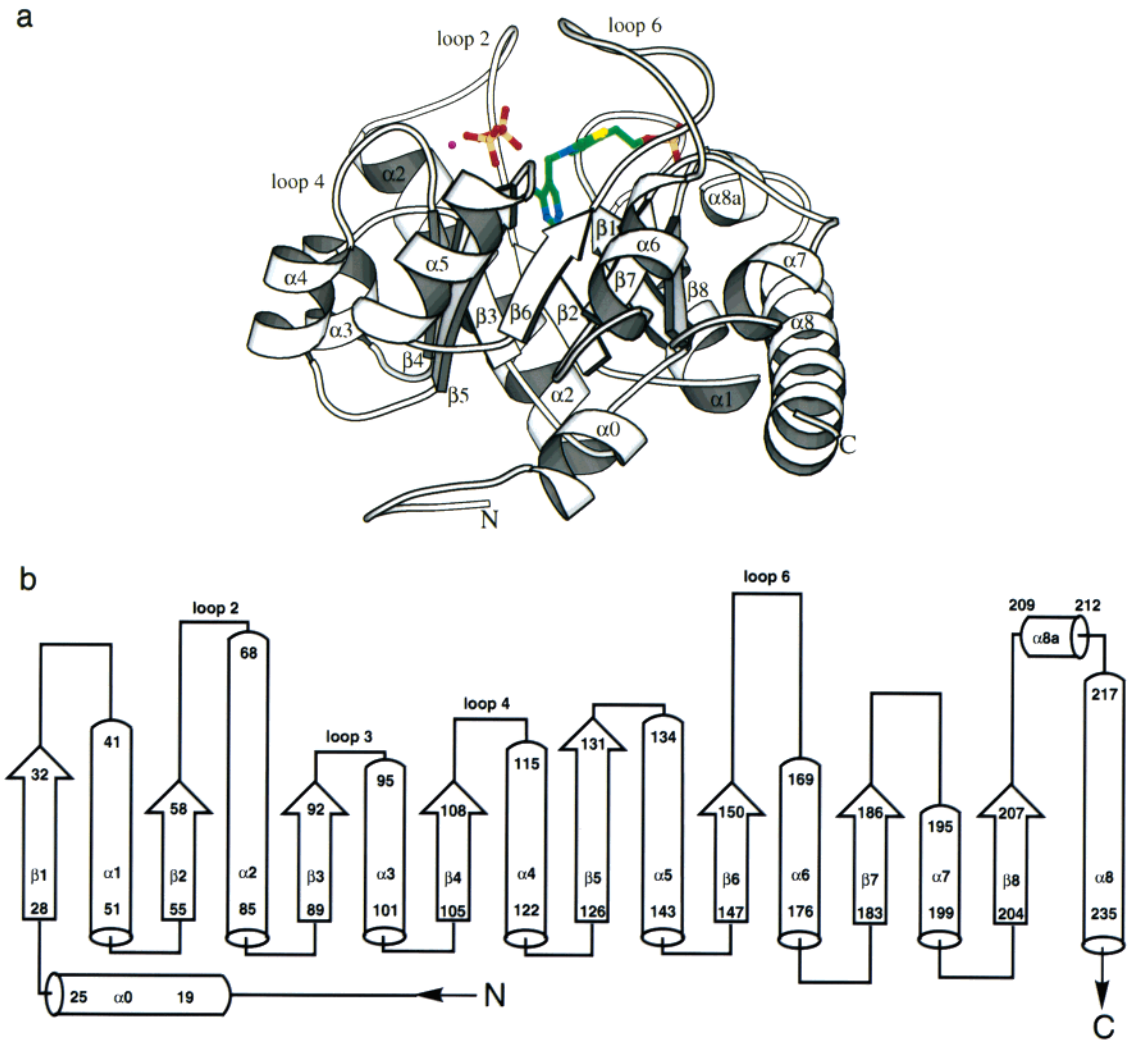


FIGURE 1: Structure of TP synthase. (a) Ribbon drawing of TP synthase with the enzymatic products, thiamin phosphate and PP_i, as well as a Mg^{2+} ion bound. The PP_i, Mg^{2+} ion, and thiazole lie across the C-terminal ends of the β-strands, and the pyrimidine moiety binds in the center of the TIM barrel. Active site loops 2, 4, and 6 are labeled. This figure was created using Rasmol (*20*) and Bobscript (*21*). (b) Topology diagram of TP synthase.

0.5 Å), (2) residues near the C-terminus (rmsd = 0.1−1.0 Å), (3) loop 2 (rmsd = 0.2−1.0 Å), and (4) residues near α8 (0.5−0.7 Å). The first two regions are involved in crystal packing, while loop 2 contains Gly62 and Gly63 for which the density is always weak.

TP Synthase Active Site. Approximately 13 highly conserved residues are involved in ligand binding (Figure 2). Arg59 and Lys61 (loop 2), Asn92 (loop 3), Lys159 (loop 6), and Mg^{2+} form interactions with the β-phosphate group of HMP-PP, while the α-phosphate group interacts with Ser130 (loop 5), Lys159 (loop 6), Mg^{2+}, and four water molecules. In addition to the PP_i, Mg^{2+} is ligated to Asp93 (loop 3), Asp112 (loop 4), and two water molecules. Two hydrogen bonds are formed between the edge of the pyrimidine and the side chain of Gln57 (loop 2). The phosphate group of Thz-P forms hydrogen bonds with Thr156 and Thr 158 (loop 6), with Gly188 (loop 7), and with Ile208 and Ser 209 (loop 8). Only van der Waals contacts, primarily from loop 6, are formed with the thiazole ring.

The seven structures reported here, together with the original structure, can be divided into four groups: (1) unliganded TP synthase, (2) complexes with substrates, (3) complexes with products, and (4) complexes with intermediates. The substrates, intermediates, and products are all oriented in a similar fashion at the C-terminal face of the TIM barrel (Figure 1). The PP_i moiety is the most solvent accessible, forming hydrogen bonds with water molecules at the solvent interface and with solvent accessible residues in loops 2−4 and 6 and strand β5. The plane of the Thz-P moiety is roughly perpendicular to the TIM barrel axis with the thiazole group covered by loop 6 and the phosphate of Thz-P exposed to a small solvent channel near α8a. The pyrimidine moiety binds deepest into the TIM barrel and is oriented with its plane nearly parallel to the barrel axis. In addition to the hydrogen bonds with Gln57, the pyrimidine is sandwiched between His107 and Ile186. The β-phosphate of HMP-PP is close to the pyrimidine amino group forming a hydrogen bond. This hydrogen bond completes a 10-membered ring (···H−N4′−C4′−C5′−C7′−O−Pα−O−Pβ−O···). This hydrogen bond may be mechanistically important (see the Discussion).

Two different conformations of the PP_i group are observed among the eight structures (Figure 3). In one conformation, the α-phosphate group points toward the C7′ atom to which it would be bound in the HMP-PP substrate molecule. The

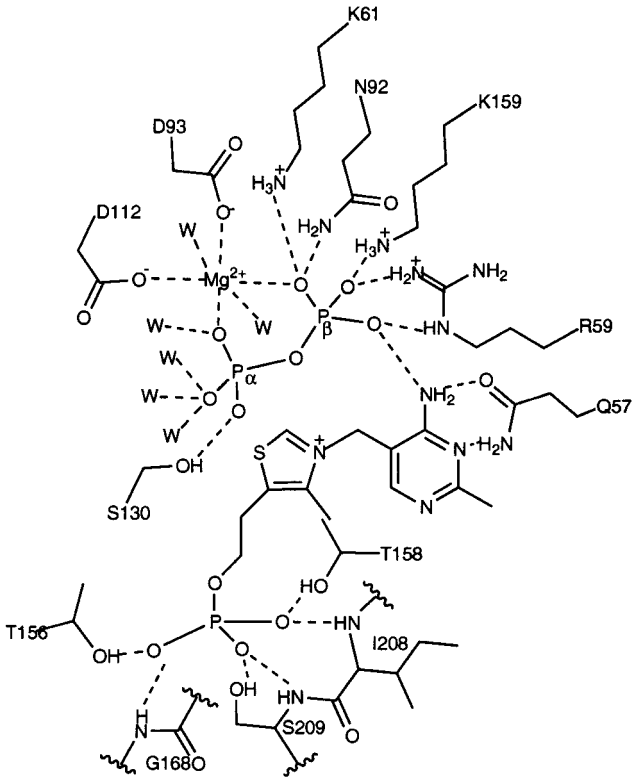


FIGURE 2: Active site of TP synthase. The products TP and PP_i and Mg^{2+} are shown along with 13 highly conserved residues. Hyrdogen bonds are shown as dotted lines. The figure was created using ChemDraw.

O−Pα torsion angle is staggered, and the O−Pβ torsion angle is nearly eclipsed. In the alternate conformation, the PP_i rotates around the O−Pβ bond such that the α-phosphate group points away from C7′ and both the O−Pα torsion angle and the O−Pβ torsion angle are nearly eclipsed. The alternate conformation of the PP_i leads to displacement of a well-ordered water molecule and formation of a hydrogen bond to Lys159 Nζ. In the first PP_i conformation, Lys159 forms a hydrogen bond with a different water molecule. Only the first conformation is observed for wild-type structures, while both conformations were observed in the S130A structures. In S130A crystals where both substrates were added, the PP_i group shows both conformations. When products were added directly to S130A crystals, only the second conformation of PP_i was observed.

Structure of Unliganded S130A TP Synthase. In the original crystal structure of TP synthase, the reaction products, TP and PP_i, along with Mg^{2+} were unexpectedly found in the active sites of both molecules A and B even though neither reactants nor products were added at any time during protein production or crystallization. S130A TP synthase is inactive, and the structure revealed that active sites of both molecules A and B were empty, except in those cases where ligands were added during or after crystallization. The unliganded S130A TP synthase model contains 439 residues and 323 water molecules. Fifteen of the water molecules (seven for molecule A and eight for molecule B) are located in the active site. Six of the water molecules are also seen in the previously reported structure, and nine occupy positions where ligand oxygen or nitrogen atoms appear in the TP synthase complexes.

The rmsd from the previously reported TP synthase product complex is 0.3 Å (using Cα atoms) with the largest differences seen in the active site. For the active site loop 6, electron density is not seen for Glu157−Asp161 in molecule A and Thr156−Thr162 in molecule B. These side chains are involved primarily in hydrogen bonds to the phosphate groups in both Thz-P and HMP-PP. For the loop 6 residues that are visible, the rmsd is 0.6 Å for molecule A and 0.4 Å for molecule B. Loop 4, which contains the Mg^{2+} ligating residue Asp112, also exhibits missing backbone density in the unliganded S130A mutant TP synthase and higher rmsd values (1.1 Å for molecule A and 0.8 Å for molecule B). The β5 strand, where the Ser130 mutation is located, does not appear to be significantly perturbed by deletion of the hydroxyl group.

Structure of the TP Synthase−Substrate Complex. One of the structures reported here provides a direct view of an enzyme−substrate binary complex. S130A crystals soaked overnight in 20 mM CF_3HMP-PP, a slowly reacting substrate, showed the position of the pyrimidine substrate and its interactions with the protein in the absence of the thiazole moiety (Figure 4a). The model contains 437 protein residues, one molecule of CF_3HMP-PP and one Mg^{2+} ion in each active site, and 176 water molecules. Thirteen of the water molecules (seven for molecule A and six for molecule B) are in active site positions equivalent to the previously reported structure. Four additional active site water molecules occupy positions that are filled by phosphate oxygen atoms of Thz-P in other structures.

The overall structure of the S130A−CF_3HMP-PP complex is similar to both the previously reported TP synthase−product complex (rmsd = 0.3 Å) and the unliganded S130A mutant (rmsd = 0.3 Å). The largest deviations are seen between the S130A−substrate complex and unliganded S130A in loop 3 (1.0 Å for molecule A and 0.5 Å for molecule B) and loop 4 (0.4 Å for molecule A and 0.8 Å for molecule B). Loop 3 and loop 4 contain the Mg^{2+} ligating residues Asp93 and Asp112, respectively. Unlike the unliganded S130A structure, the S130A−substrate complex contains Mg^{2+} and shows an ordered Asp112. This region (loop 4) is almost identical to the TP synthase−product complex. However, the active site loop 6, which interacts primarily with the thiazole moiety, is disordered here, as it is in the unliganded S130A mutant with residues Thr156−Asp161 missing from molecule A and residues Thr156−Thr162 missing from molecule B.

A shift in both the PP_i and the pyrimidine moieties occurs in the S130A−substrate complex with respect to the previous TP synthase−product complex. The pyrimidine ring tilts out of plane, resulting in a 1.4 Å difference in the C7′ positions. Ile186 Cδ and His107 Nϵ, which are in van der Waals contact with the pyrimidine ring, shift by 1.0 and 0.4 Å, respectively, but there are no backbone shifts at these residues. The α-phosphate group is rotated toward the pyrimidine moiety of the substrate, resulting in a less staggered conformation in the substrate than in the product, and a displacement of the oxygen closest to the C7′ of 0.9 Å. A consequent shift in the Mg^{2+} ion (0.6 Å) and the Asp93 and Asp112 ligands (Oδ atoms shift 0.4 Å for each) is also observed.

Structures of the TP Synthase−Product Complex. Two of the structures reported here show additional views of the TP

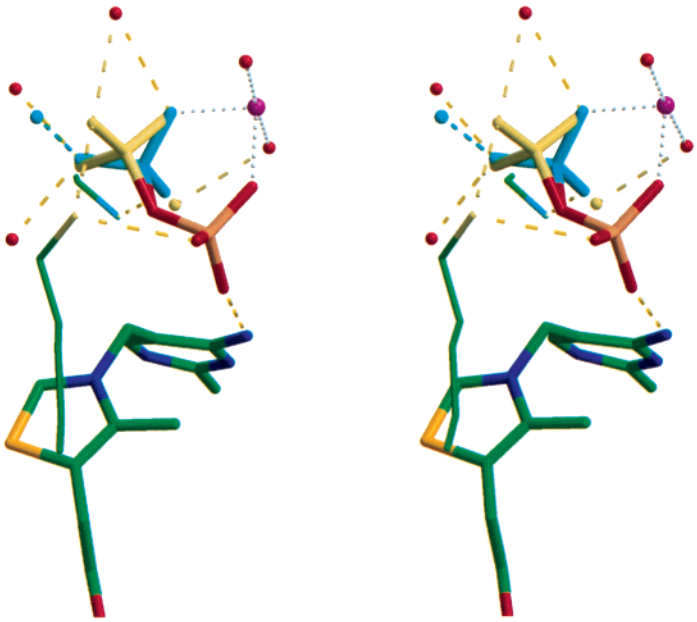


FIGURE 3: Two conformations for PP_i. The main conformation of PP_i, a single water molecule, and the hydrogen bond between that water and the PP_i are shown in cyan. The alternate conformation, seen only in the S130A mutant, and all unique contacts to that conformation are colored yellow. All atoms that are the same in both conformations are colored by atom type. This figure was created using Rasmol (*20*), Bobscript (*21*), and Raster3d (*22*).

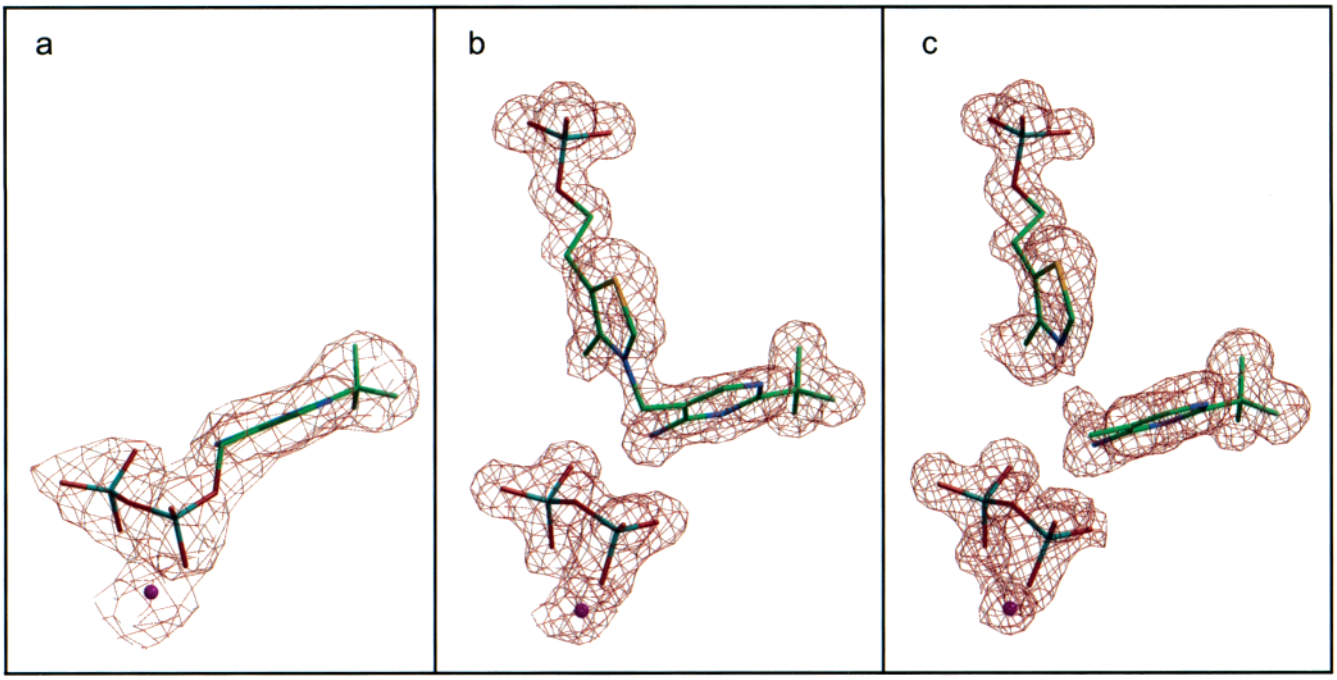


FIGURE 4: Electron density for TP synthase compelxes. (a) Substrate taken from molecule A of the S130A−CF_3HMP-PP complex, (b) products taken from molecule A of TP, and (c) intermediates taken from molecule B of the S130A−CF_3HMP−Thz-P−PP_i complex. Figures were created using Rasmol (*20*), Bobscript (*21*), and Raster3d (*22*).

synthase−product complex. The structure of the S130A mutant soaked with products was determined to compare it to the original TP synthase−product complex. S130A crystals soaked overnight in 20 mM TP and PP_i reveal both products bound to the inactive mutant (Figure 4b). The final model contains 452 protein residues, one molecule each of TP and PP_i and one Mg^{2+} ion in each active site, and 298 water molecules. Thirty-four (17 each for molecules A and B) of the 38 active site water molecules are in positions equivalent to those in the previously reported TP synthase−product complex. A comparison of S130A−product complex to the previously reported TP synthase−product complex gives an overall rmsd of 0.2 Å using Cα atoms. The largest differences between these two structures are seen in the N-terminal region, which is poorly ordered in all structures. The TP, PP_i, and Mg^{2+} ions overlap with those in the previously reported structure, except that the PP_i is only found in the second conformation for both molecules A and B.

An attempt to prepare a TP synthase−substrate ternary complex using wild-type TP synthase and the slowly reacting substrate CF_3HMP-PP resulted in an additional view of the TP synthase−product complex. The wild-type TP synthase

was incubated in 20 mM CF_3HMP-PP and 20 mM Thz-P and cocrystallized. Surprisingly, the resulting structure showed CF_3TP, PP_i, and Mg^{2+} bound at the active site rather than the substrates that were added. The structure could be distinguished from the TP synthase−product complex by the clear density for the trifluoromethyl group in both molecules A and B. This result demonstrates that the (trifluoromethyl)-pyrimidine is reactive enough that given sufficient time it can be converted to product. Furthermore, the experiment demonstrated that one of the substrates or the combination of both substrates was able to displace the bound product, TP, that copurified with the enzyme. The final model contains 453 protein residues, one molecule each of CF_3TP and PP_i and one Mg^{2+} ion in each active site, and 351 water molecules. All of the 38 (19 each for molecules A and B) active site waters are in positions equivalent to those in the previously reported structure.

This TP synthase−product complex differs very little from the previously reported structure. The largest difference with respect to the previously reported TP synthase−product complex is seen in active site loop 6 (rmsd = 0.5 Å for molecule A and rmsd = 0.3 Å for molecule B), and the largest difference with respect to the S130A−product complex is in the N-terminal coil (rmsd = 0.5 Å for molecule A and rmsd = 0.4 Å for molecule B). The three fluorine atoms of CF_3TP are within 3.5 Å of seven protein residues, Tyr29, Gln57, His107, Tyr147, Gly149, Val184, and Ser206; however, the rmsd deviation for these residues was very small when it was compared to the rmsd of either of the other TP synthase−product complexes. The only notable difference in these residues is seen for Ser206 Oγ, which is found in two alternate conformations in the previously reported structure but is sterically restricted to the conformation that is furthest from the fluorine atoms in this complex.

Structure of the TP Synthase−Intermediate Complex. Three of the structures reported here appear to be distinctly different from either the TP synthase−substrate complex or the TP synthase−product complexes and have been modeled as TP synthase−intermediate complexes. All three of the structures were determined in an attempt to obtain the TP synthase−substrate complex but clearly showed that the reaction had progressed to an intermediate stage.

Crystals of the S130A mutant TP synthase were treated overnight in 20 mM HMP-PP and 20 mM Thz-P, the preferred TP synthase substrates. After preliminary refinement, electron density maps revealed strong density for the pyrimidine, the thiazole phosphate, and the PP_i (Figure 4c). However, only weak electron density, less than one standard deviation, appears either between HMP C7′ and PP_i as would be expected for products or between HMP C7′ and Thz-P N3 as would be expected for substrates. Furthermore, neither substrates nor products fit the electron density with reasonable geometry, and the fitted ligand models showed significant differences compared to either the TP synthase−substrate complex or the TP synthase−product complex described above. The final model contains 450 protein residues and 263 water molecules. Each active site contains three separate fragments, one each of the chemical moieties HMP, PP_i, and Thz-P, and one Mg^{2+} ion. The α-phosphate of PP_i is disordered and is a mixture of the two conformations described above. Thirty-three (16 for molecule A and 17 for molecule B) active site waters are in conserved positions relative to the other models. The largest deviation for protein atoms between this TP synthase−intermediate complex and the TP synthase−product complex is in active site loop 6.

In another attempt to trap the TP synthase−substrate ternary complex, crystals of the S130A mutant TP synthase were soaked in 20 mM HMP-PP and 20 mM Thz-P, but the soaking time was limited to 30 min. The crystals were removed from the soaking solution frozen, and X-ray intensity data were recorded. As was the case for the two other TP synthase−intermediate complexes, only density at the level of noise was observed between the HMP and PP groups or between the HMP and Thz-P groups. Interestingly, in this structure the PP_i group showed no indication of the alternate conformation seen in the other intermediate structures. The final model contains 454 protein residues and 279 water molecules. Each active site contains three separate fragments corresponding to HMP, PP_i, and Thz-P, and one Mg^{2+} ion.

Because the preceding attempts to prepare a TP synthase−substrate complex led to an unexpected result, we then tried incubating the S130A mutant with the trifluoromethyl substrate to further retard the rate of intermediate formation. Crystals of S130A TP synthase were soaked overnight in 20 mM CF_3HMP-PP and 20 mM Thz-P. As with the S130A−HMP-PP−Thz-P complex, only weak electron density near the noise level appeared between the intermediate fragments and could not be modeled by either substrates or products. The final model contains 451 protein residues and 293 water molecules. Each active site contains three separate fragments, one each of the chemical moieties CF_3HMP, PP_i, and Thz-P, and one Mg^{2+} ion. As was the case for the first TP synthase−intermediate complex, the PP_i is disordered over two possible conformations. Thirty-two (16 each for molecules A and B) active site water molecules are in conserved positions. The largest variation between the S130A−CF_3HMP-PP−Thz-P and S130A−HMP-PP−Thz-P complexes is seen in loop 2 which interacts with both the PP_i and HMP fragments.

Substrate-Induced Product Release. The previously reported TP synthase structure unexpectedly contained tightly bound reaction products even though neither substrates nor products had been added at any time. Therefore, we speculated that product release might be triggered by some unknown mechanism such as availability of substrates. To determine whether substrates for this enzyme could accelerate product release, it was necessary to treat the enzyme with a substrate analogue that would not be converted to thiamin as this would interfere with the assay for released TP. Since CF_3HMP-PP is not a substrate for the enzyme, it is an ideal substrate analogue for these studies. When the enzyme is treated with a mixture of this analogue and with Thz-P, we observed the release of 68% of the bound thiamin phosphate (Figure 5). We then examined the role of the individual substrates in stimulating thiamin phosphate release and found that CF_3HMP-PP is more effective in stimulating TP release (74%) than Thz-P release (29%). It is unclear why we do not observe complete release in any of these experiments.

DISCUSSION

Previous studies on TP synthase demonstrate that the reaction rate is retarded >7800-fold when Ser130, which is

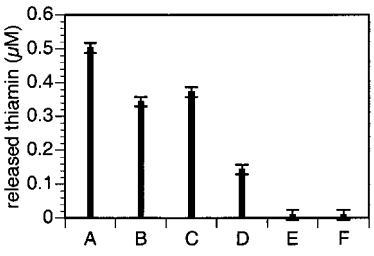


FIGURE 5: Release of thiamin phosphate from TP synthase under various conditions. The columns are as follows: (A) TCA-treated sample, (B) CF_3HMP-PP- and Thz-P-treated sample, (C) CF_3HMP-PP-treated sample, (D) Thz-P-treated sample, (E) control, with no substrates added, and (F) control, with no enzyme added.

hydrogen bonded to the oxygen leaving group in the PP_i, is replaced with alanine (*6*) and >7800-fold when the methyl group of HMP-PP is replaced with the electron-withdrawing trifluoromethyl group (*6*). In addition, the enzyme catalyzes positional isotope exchange (PIX) in the pyrimidine pyrophosphate (*6*). The high sensitivity of the reaction rate to perturbations designed to destabilize an ion pair intermediate and the observation of PIX strongly suggest that the reaction proceeds via a dissociative (S_N1-like) rather than an associative (S_N2-like) mechanism (Scheme 2). The crystal structures suggest that the resulting carbenium ion intermediate **4** is likely to be further stabilized by proton transfer from the pyrimidine amino group to the PP_i to give the pyrimidine iminemethide **5**. Addition of the thiazole to C7′ of **4** followed by proton transfer from the PP_i to the pyrimidine amino group would complete the reaction. In this research, the structures of the free enzyme and the enzyme−substrate, enzyme−intermediate, and enzyme−product complexes were obtained. These enable us to structurally define the reaction coordinate for TP synthase.

Structure of the Enzyme−Product Complex. The TP synthase−product complex is directly observed in three complexes: the previously reported wild-type TP synthase complex, a wild-type TP synthase−CF_3TP−PP_i complex, and an S130A−TP−PP_i complex. Two differences are observed in the comparison of the three structures. First, Ser206, which shows a mixture of two conformations in other TP synthase complexes, is sterically restricted to one of these in the CF_3-TP complex. More importantly, the α-phosphate group of PP_i, originally bound to C7′ of the pyrimidine and hydrogen bonded to Ser130 Oγ, is rotated away from the TP product in the S130A mutant, which cannot form the hydrogen bond. Nevertheless, the C7′···Oα distance in all three structures is similar, ranging from 3.42 to 3.56 Å, and the N4′···O hydrogen bond distance is similar, ranging from 2.91 to 3.00 Å. Because the three structures are so similar, the previously reported wild-type complex was chosen to represent the enzyme−product complex modeled in Figure 6c.

The structures of the TP synthase−product complexes suggest that the large rate retardation observed for the S130A mutant is caused by the removal of an important hydrogen bond that must be involved in the catalytic mechanism rather than by a structural change in the mutant enzyme. If carbocation formation were fully rate limiting, this hydrogen bond would need to be 5.5 kcal/mol stronger in the transition state than it is in the enzyme−substrate complex to explain the rate retardation. This could occur due to the increasing basicity of the ester oxygen in the transition state. However, it is also possible that the altered conformation of the PP_i moiety at the active site of the S130A mutant retards the reaction by decreasing the mobility of the thiazole or the pyrimidine (see the discussion of the enzyme−intermediate complex below).

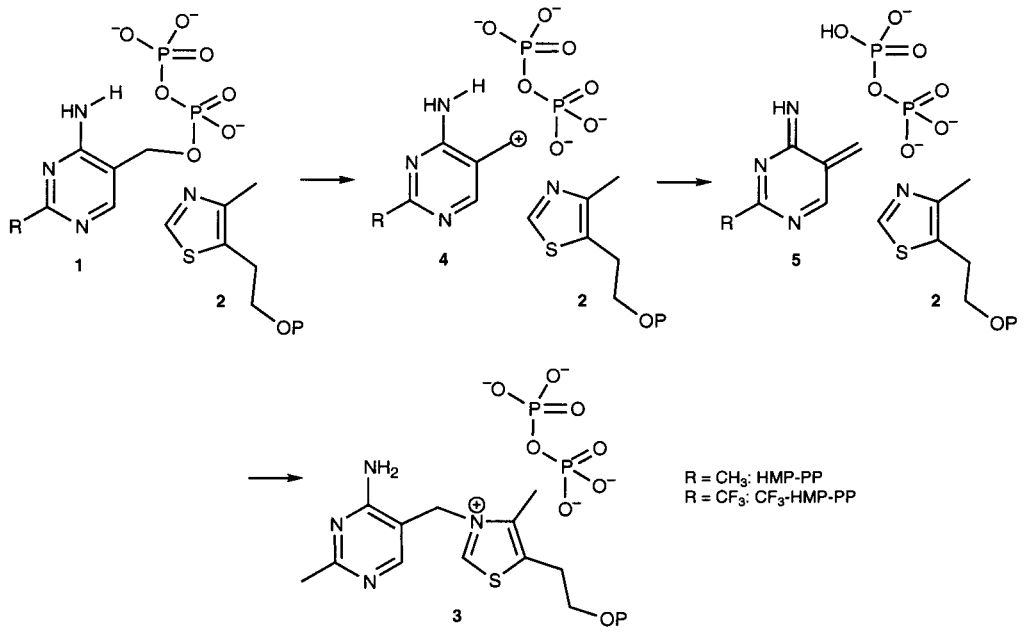
CF_3HMP-PP is not a substrate for the enzyme, suggesting that this pyrimidine analogue might be useful for obtaining a structure of the enzyme−substrate complex. However, cocrystallization of the enzyme in the presence of CF_3HMP-PP and Thz-P established the presence of product at the active site. This unexpected observation suggests that the large rate retardation caused by substituting the methyl group of the pyrimidine with a trifluoromethyl group is due to the electron-withdrawing effect of the trifluoromethyl substituent and not due to structural perturbations caused by binding of this analogue to the active site.

Structure of the Enzyme−Substrate Complex. Only the pyrimidine substrate is directly observed in a crystal structure. Soaking of S130A crystals in a solution of CF_3HMP-PP gave the expected binary S130A−substrate complex. This structure demonstrates that pyrimidine C7′ is tilted away from the thiazole binding site of the enzyme−product complex and that the overall protein structure is similar to the structure of the unliganded enzyme (S130A). The main difference is that loop 6, which interacts with the Thz-P and PP_i moieties in the enzyme−product complex, is disordered. The intramolecular NH_2···Oβ hydrogen bond appears to be slightly shorter (2.82 Å) then in the product complexes.

The position of the Thz-P substrate is not directly observed in any X-ray structure, but can be inferred from three complexes where a thiazole moiety is present. In all such complexes, active site loop 6 is well-ordered and the phosphate ends of all six Thz-P molecules (three each from molecules A and B) superimpose closely. Because of the tight interactions, it is unlikely that the phosphate group of the Thz-P substrate is any different from what is observed in these complexes. The thiazole ring positions show slight variation among the three complexes, but always shift away from HMP C7′ compared to the TP synthase−product complexes. The largest difference is about 1 Å and occurs at the N3 position. It is likely that the thiazole ring position in the substrate complex more closely resembles the nonbonded ring position in the set of intermediate structures, and might shift even farther from the HMP. However, ring movement away from HMP C7′ is bounded by contacts with the extended Thr158-Lys159 aliphatic chain. The Thz-P ring position of the substrate complex was therefore modeled in the observed conformation that is farthest from C7′ (Figure 6a).

We have previously observed (*6*) that the binary enzyme−HMP-PP complex shows a higher rate of positional isotope exchange than the tertiary enzyme−HMP-PP−isomeric Thz-P complex. In the isomeric Thz-P, the nitrogen and sulfur atoms are interchanged to prevent product formation. Comparison of loop 6 in the S130A−substrate complex with that in all other complexes in this study provides an explanation for the observation of PIX. The presence of a thiazole moiety causes loop 6 to become ordered and establishes the interaction of Lys159 Nζ on loop 6 with PP_i O4 and Oβ3. Though the S130A−substrate complex shows PP_i in the same location as in other complexes, and therefore in position to form those hydrogen bonds, these interactions

Scheme 2: Mechanistic Proposal for TP Synthase



cannot form when loop 6 is disordered. The enzyme−HMP-PP complex in the PIX study would therefore have a lower barrier to PA−O4 bond rotation, showing a higher rate of PIX than the ternary complex.

Structure of the Enzyme−Intermediate Complex. Our first attempt to obtain a structure of the enzyme−substrate ternary complex involved soaking S130A crystals overnight in a solution of HMP-PP and Thz-P. The structure of this complex surprisingly contained an intermediate in the active site rather than the expected substrates. While this structure is consistent with a bound carbenium ion, it is more likely that the carbenium ion has been deprotonated by the PP_i to give the more stable iminemethide (Scheme 2). Further evidence for the protonation state of the intermediate is provided by S130A TP synthase crystals that were soaked overnight in a solution of CF_3HMP-PP and Thz-P. As was the case for S130A treated with HMP-PP and Thz-P, an intermediate was observed in the active site. The observation of an intermediate in the presence of a trifluoromethyl-substituted pyrimidine, which would strongly destabilize a carbenium ion, strongly suggests that the observed intermediate is the pyrimidine iminemethide rather than the carbenium ion. For these two intermediate complexes, the average distance from C7′ of the pyrimidine to the closest PP_i oxygen is 2.83 Å and the average distance of C7′ to the thiazole nitrogen is 2.94 Å. A typical C−N single bond length is 1.47 Å, and a typical C−O single bond length is 1.43 Å; thus, in this TP synthase complex, C7′ is clearly not bonded to either nucleophile. The average distance between the amine nitrogen (N4′) and the closest PP_i oxygen is 3.00 Å, which is similar to the values observed for the TP synthase−product complex.

The third attempt to obtain an enzyme−substrate complex also resulted in an intermediate. In this case, the S130A mutant enzyme was treated for less than 30 min with substrates and then flash-frozen. The overall active site structure was similar to the structures of the two intermediate complexes discussed above. Both the C7′−O and C7′−N3 distances indicated the absence of a single bond. Interestingly, the N4′···O hydrogen bond was noticeably shorter (average value of 2.64 Å) then in any other complex. The shorter observed hydrogen bond distance suggests that in addition to the pyrimidine iminemethide, a small fraction of an amino-substituted pyrimidine species may still remain. Presumably, the hydrogen bond between an amino group and a negatively charged pyrophosphate oxygen atom would be stronger then the hydrogen bond between the iminemethide nitrogen atom and a protonated PP_i oxygen atom. Further extension of this reasoning suggests that the N4′···O hydrogen bond of the enzyme−substrate ternary complex would be even shorter than the one observed in this or either of the other two enzyme−intermediate complexes. An intermediate model based on an average of the three complexes (six copies) is presented in Figure 6b.

The detection of the pyrimidine iminemethide intermediate at the active site of the S130A mutant TP synthase is an unanticipated novel discovery arising from our much more rational attempts to prepare crystals of the ternary enzyme−substrate complex. The high stability of this intermediate is intriguing and must be due to the retarded rate of its trapping by PP_i and Thz-P. The rate of the ion recombination reaction may be suppressed due to PP_i (Oα), originally attached to the pyrimidine, freely rotating away from C7′ as it is no longer hydrogen bonded to Ser130 Oγ. It is unclear however why the thiazole does not react with the iminemethide. Apparently, movement of the thiazole toward C7′ of the pyrimidine iminemethide or movement of the pyrimidine iminemethide toward the thiazole is restricted. This putative immobilization of the thiazole−iminemethide intermediate may be an indirect consequence of the S130A mutation or may be a consequence of the crystalline state of the enzyme.

Flexibility and Conformational Changes. The protein secondary structure shows no evidence of concerted conformational change; however, minor movement in active site loops, which is consistent with ligand binding, is observed. In complexes that lack the thiazole moiety, active site loop 6 is disordered, thus implying flexibility. The Asp112 backbone density also suggests flexibility in complexes

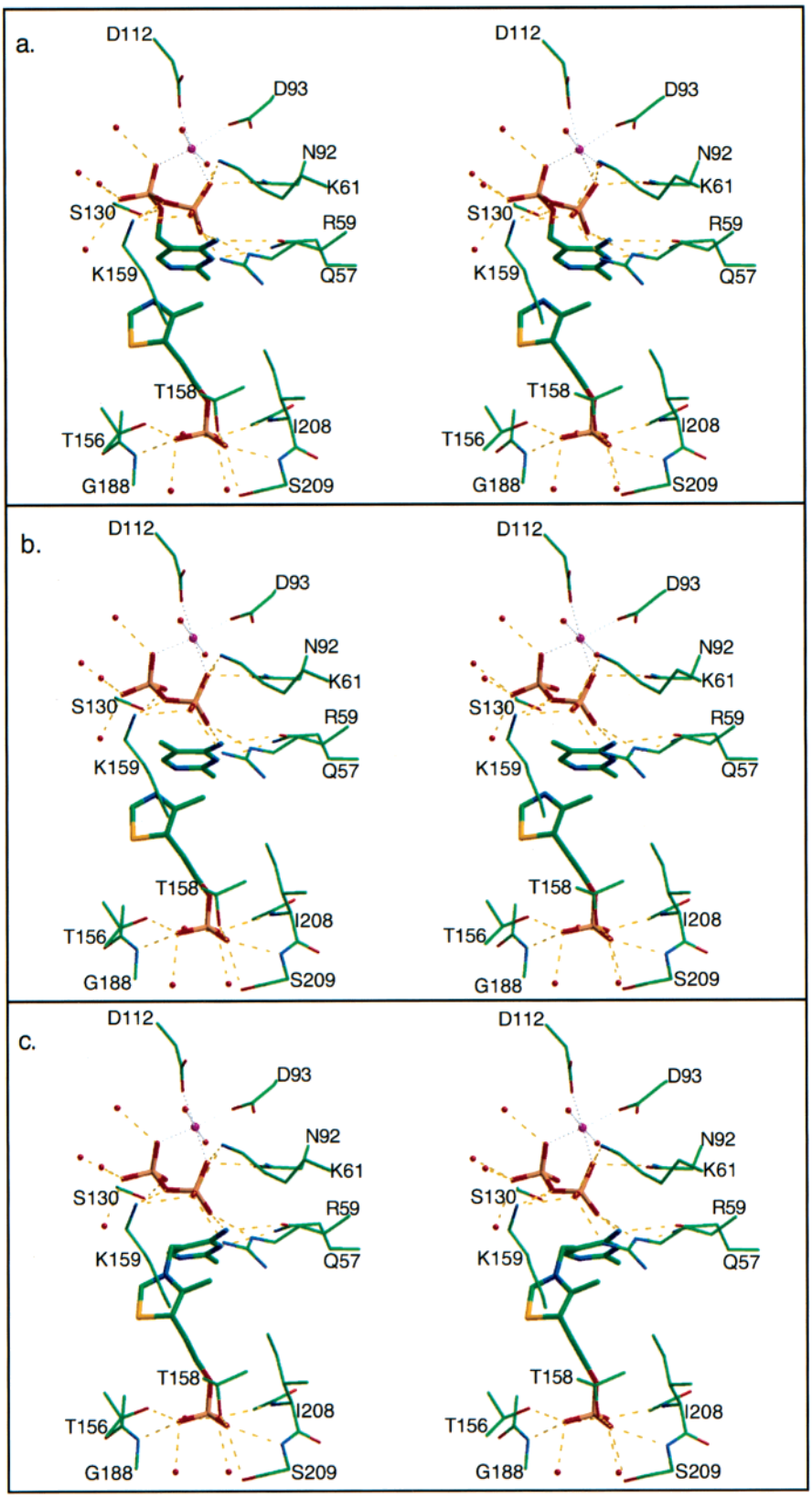


FIGURE 6: Stereoviews of the TP synthase active site. (a) Model of substrates with key active site residues. (b) Model of intermediates with key active site residues. (c) Model of products with key active site residues. Figures were made using Rasmol (*20*), Bobscript, (*21*), Raster3d (*22*), and GraphicConverter (*23*).

where that residue is not ligated to Mg^{2+}. Backbone flexibility is also seen at the Gly62-Gly63 region, which could allow peptide flipping of one or both of those residues. In the high-resolution structures of TP synthase, only a few surface side chains are poorly defined and less than 5% of all side chains appear to be disordered. One residue that is consistently present in two different conformations and also appears near the active site is His132. A disordered Arg163 lies over His132, and an ordered Glu136 is to one side. For one conformation of His132 in the S130A−intermediate

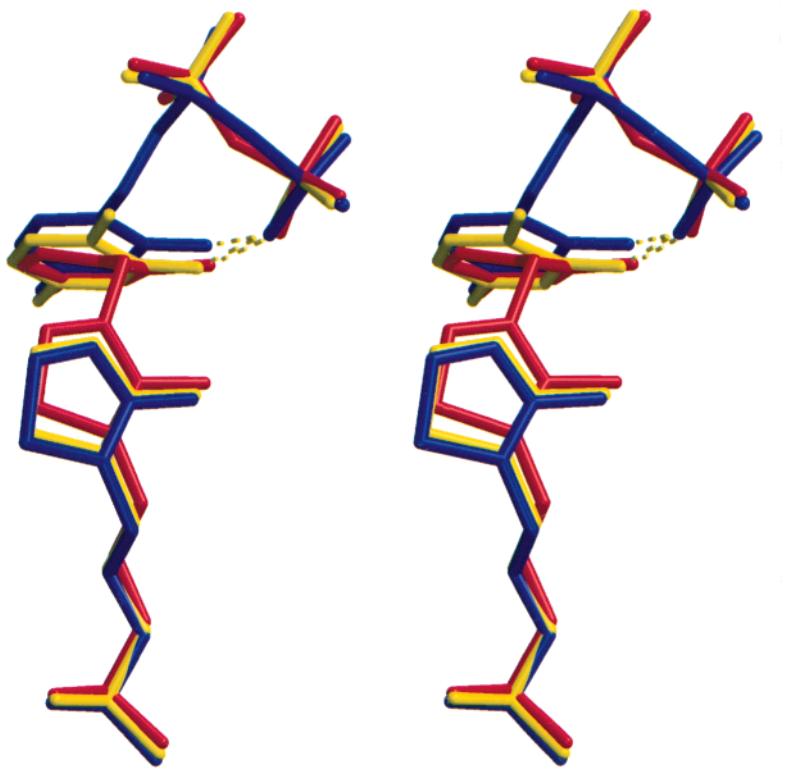


FIGURE 7: Stereoview showing a superposition of substrates (blue), intermediates (yellow), and products (red). The comparison illustrates the movements of the thizaole and pyrimidine rings that occur during catalysis. The hydrogen bonds between the pyrimidine 4-amino group and the PP_i oxygen atom are shown with dashed lines.

complex, Nϵ is hydrogen bonded to a conserved water molecule that in turn is hydrogen bonded to PP_i while Nδ is hydrogen bonded to Glu136 Oϵ, thus creating a possible charge relay system. For the second conformation, the His132 side chain is solvent exposed. The possibility of a charge relay triad is not supported by sequence comparisons, which show that His132 is not conserved in TP synthases of other species.

Structural Characterization of the TP Synthase Reaction Coordinate. The structural characterization of the free enzyme and the enzyme−substrate, enzyme−intermediate, and enzyme−product complexes allows us to construct a clear picture of the reaction coordinate for TP synthase (Figure 6). After binding HMP-PP and Thz-P, loop 6 closes down on the active site to give the ternary enzyme−substrate complex. The ionization of the pyrimidine pyrophosphate is facilitated by the stabilization of the PP_i leaving group and by the delocalization of the positive charge into the electron rich pyrimidine. The carbenium ion is further stabilized by proton transfer to the PP_i to give the pyrimidine iminemethide intermediate. This proton transfer is not concerted with carbenium ion formation because the C−O bond and the N−H bonds are orthogonal in the transition state that leads to the carbenium ion. The structure of the pyrimidine iminemethide complex demonstrates that at this point along the reaction coordinate C7′ of the pyrimidine has moved ∼3.5 Å away from the PP_i oxygen by a pivotal movement about the methyl group. To form the product, C7′ of the pyrimidine continues its pivotal movement for an additional 1 Å and the thiazole moves a little less than 1 Å toward C7′ of the pyrimidine (Figure 7). During the course of the reaction, the length of the hydrogen bond between the pyrimidine amino group and the closest PP_i oxygen changes from ∼2.6 Å (by inference) in the enzyme−substrate complex to ∼3.0 Å in the enzyme−iminemethide complex and then to ∼2.9 Å in the enzyme−product complex.

Substrate-Induced Product Release. The TP synthase−product ternary complex is stable and requires the addition of HMP-PP to drive product dissociation. One possible mechanism for substrate-assisted product release in the case of TP synthase is dissociation of the PP_i from the enzyme product complex allowing the pyrimidine to migrate into the adjacent solvent-filled cavity. This migration into a less favorable site may be driven by binding of a subsequent molecule of the pyrimidine pyrophosphate substrate to the catalytic site. If loop 6 then moves back from the active site, the thiazolium moiety can dissociate from the enzyme surface followed by the pyrimidine. The observation of a stable enzyme−product complex for TP synthase while unusual is not without precedent. Stable enzyme−product complexes have previously been detected for protein farnesyltransferase (*16*), MTA phosphorylase (*17*), phosphopantothenate adenylyl transferase (*18*), and dihydrofolate reductase (*19*). The physiological role of this tight TP synthase−product complex is not yet known. One hypothesis is that it may play a role in the regulation of thiamin biosynthesis.

ACKNOWLEDGMENT

We thank Ms. Leslie Kinsland for assistance in the preparation of the manuscript. We thank CHESS and the Structural Biology Center at the APS for providing beam time.

REFERENCES

1. Backstrom, A. D., McMordie, A., and Begley, T. P. (1995) *J. Am. Chem. Soc. 117*, 2351−2351.
2. Zhang, Y., Taylor, S. V., Chiu, H. J., and Begley, T. P. (1997) *J. Bacteriol. 179*, 3030−3035.
3. Nosaka, K., Nishimura, H., Kawasaki, Y., Tsujihara, T., and Iwashima, A. (1994) *J. Biol. Chem. 269*, 30510−30516.
4. Levin, J. Z., Potter, S. L., and Williams, M. (2000) *PCT Int. Appl.*, Patent WO 00/00623.
5. Chiu, H., Reddick, J., Begley, T., and Ealick, S. (1999) *Biochemistry 38*, 6460−6470.
6. Reddick, J. J., Nicewonger, R., and Begley, T. P. (2001) *Biochemistry 40*, 10095−10102.
7. Leslie, A. G. W. (1990) in *Chemistry to Biology* (Moras, D., Podjarny, A. D., and Thierry, J. C., Eds.) pp 50−61, Oxford University Press, Oxford, U.K.
8. Collaborative Computational Project (Number 4) (1994) *Acta Crystallogr. D50*, 760−763.
9. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
10. 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) *Acta Crystallogr. D54*, 905−921.
11. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
12. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
13. Martin, A. C. (1998) *ProFit*, version 1.8, SciTech Software, London.
14. Luzzati, P. V. (1952) *Acta Crystallogr. 5*, 802−810.
15. Ramachandran, S. (1968) *Adv. Protein Chem. 23*, 283−437.
16. Tschantz, W. R., Furfine, E. S., and Casey, P. J. (1997) *J. Biol. Chem. 272*, 9989−9993.
17. Appleby, T. C., Erion, M. D., and Ealick, S. E. (1999) *Struct. Folding Des. 7*, 629−641.
18. Geerlof, A., Lewendon, A., and Shaw, W. V. (1999) *J. Biol. Chem. 274*, 27105−27111.
19. Fierke, C. A., Johnson, K. A., and Benkovic, S. J. (1987) *Biochemistry 26*, 4085−4092.
20. Sayle, R. A., and Milner-White, E. J. (1995) *Trends Biochem. Sci. 20*, 374−376.
21. Esnouf, R. (1997) *J. Mol. Graphics 15*, 132−134.
22. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505−524.
23. Lemke, T. (2000) *GraphicConverter 4.0*, Lemke Software, Peine, Germany.

BI0104726