Approaching the Transition State in the Crystal Structure of a Phosphoribosyltransferase[†,‡]

Pamela J. Focia, Sydney P. Craig, III, and Ann E. Eakin*

Laboratory of Molecular Parasitology & Drug Design, School of Pharmacy, University of North Carolina, Chapel Hill, North Carolina 27599-7360

Received September 4, 1998; Revised Manuscript Received October 16, 1998

ABSTRACT: Hypoxanthine phosphoribosyltransferase (HPRT) salvages 6-oxopurine bases in the nucleotide metabolic pathway. The 1.8 Å crystal structure of an asymmetric dimer of the HPRT from the protozoan parasite *Trypanosoma cruzi* was determined in a ternary complex with the primary substrate phosphoribosylpyrophosphate (PRPP) and an analogue of the substrate hypoxanthine, revealing both open and closed active site conformations. The ligands are positioned for in-line nucleophilic attack at the PRPP ribose C1′ by two metal ions which straddle the pyrophosphate leaving group. The structure provides the first evidence for the involvement of two metal ions in the HPRT-catalyzed reaction, and structural details further suggest the mechanism may proceed via S_N2-type chemistry. The closed conformation reveals the structural roles for invariant flexible loop residues Ser103 and Tyr104 and supports a role for the loop in the liberation of pyrophosphate. The pre-transition state structure is valuable for understanding the enzyme mechanism, as well as providing a foundation for antiparasite drug design efforts against *T. cruzi*, which causes Chagas' disease in humans. Additionally, the structure illuminates the molecular basis of three inherited mutations in the human HPRT leading to Lesch-Nyhan syndrome (D193N) or gout (S103R or S109L), as the homologous residues in the trypanosomal enzyme contribute to the previously unrecognized Mg^{2+} ion binding site and to the formation of the closed flexible loop, respectively.

We report the high-resolution crystal structure of a pre-transition state ternary complex of a dimer of the hypoxanthine phosphoribosyltransferase (HPRT)[1] from the protozoan parasite *Trypanosoma cruzi*, with both open and closed active site conformations. *T. cruzi* is the etiologic agent of Chagas' disease, affecting more than 24 million people in South and Central America (*1*). There is no satisfactory treatment for Chagas' disease, and the infection is lifelong. For many years, HPRTs have been proposed as chemotherapeutic targets for the treatment of diseases caused by parasites, including *T. cruzi*, which lack de novo pathways for purine nucleotide biosynthesis and thus must rely on scavenging purine bases from their hosts (*2*, *3*). Also, several inherited mutations in the human HPRT lead to severe metabolic diseases, such as Lesch-Nyhan syndrome or gout (*4*).

HPRTs recycle the purine base products of nucleic acid degradation that are eventually converted back to ATP or GTP. HPRT is a type I phosphoribosyltransferase (PRT) (*5*) of the PRT superfamily of enzymes that utilize phosphoribosylpyrophosphate (PRPP), a central metabolite in nucleotide biosynthesis (*6*). HPRTs catalyze a Mg^{2+}-dependent reversible ordered sequential reaction (*7*−*9*). In the forward direction, the enzyme binds PRPP first, followed by the purine (hypoxanthine or guanine). The proposed nucleophilic attack by the purine base at the PRPP ribose C1′ is followed first by the release of pyrophosphate and then by the release of the ribonucleotide (IMP or GMP) with inversion of the stereochemistry. If the reaction proceeds via S_N1-type nucleophilic displacement, the transition state is predicted to involve a charged oxycarbonium intermediate (*10*) which would require isolation from bulk solvent to prevent nonproductive PRPP hydrolysis. More recent studies have confirmed the presence of an oxycarbonium intermediate in the transition state for a type I PRT, orotate phosphoribosyltransferase (OPRT), consistent with an S_N1-type nucleophilic displacement mechanism (*11*, *12*). However, S_N2-type nucleophilic displacement has not been ruled out in HPRTs (*7*, *13*), and rapid phosphoribosyl transfer is thought to take place via a ternary complex (*8*). S_N2-type chemistry is associative rather than dissociative and would not be expected to result in formation of an oxycarbonium intermediate. Either mechanism would involve electrostatic rearrangement during the transition state as the PRPP C1′−O1 bond breaks and the C1′−purine N9 bond forms.

[†] This work was supported in part by NIH grants AI34326 (to S.P.C.) and AI38919 (to A.E.E.).

[‡] The atomic coordinates and observed structure factors have been deposited with the Protein Data Bank as 1tc2 and r1tc2sf, respectively.

* Corresponding author. Phone: (919) 966-6422. Fax (919) 966-6919. E-mail: eakin@unc.edu.

[1] Abbreviations: HPRT, hypoxanthine phosphoribosyltransferase; PRT, phosphoribosyltransferase; PRPP, 5-phospho-α-D-ribosyl 1-pyrophosphate or phosphoribosylpyrophosphate; HPP, 7-hydroxy [4,3-*d*]pyrazolopyrimidine; IMP, inosine 5′-monophosphate; GMP, guanosine 5′-monophosphate; PP_i, pyrophosphate; GPAT, glutamine phosphoribosylpyrophosphate amidotransferase; OPRT, orotate phosphoribosyltransferase; HGXPRT, hypoxanthine-guanine-xanthine phosphoribosyltransferase; XGPRT, xanthine-guanine phosphoribosyltransferase; PEG4K, poly(ethylene glycol) 4000; NCS, noncrystallographic symmetry; rmsd, root-mean-square deviation.

10.1021/bi9821465 CCC: $15.00
Published on Web 11/19/1998

Structures of HPRTs from several sources have been described (*13*−*16*), although none have been observed previously with two substrates bound in a closed conformation. Reported herein, for the first time, is the structure for a closed active site conformation of an HPRT in ternary complex with PRPP and a hypoxanthine analogue, 7-hydroxy[4,3-*d*]pyrazolopyrimidine (HPP). The 1.8 Å crystal structure of the HPRT dimer from *T. cruzi* shows two views of the substrates-bound active site, one open and solvent-exposed and the other closed and sequestered from solvent. The structure reveals the unexpected involvement of two divalent metal ions in substrate binding and positioning. The closed conformation ternary complex approximates the transition state for the enzyme-catalyzed reaction, provides evidence for the role of a long flexible loop in both active site protection and liberation of the pyrophosphate leaving group, and supports an S_N2-type nucleophilic displacement mechanism for HPRTs.

EXPERIMENTAL PROCEDURES

The recombinant HPRT of *T. cruzi* was overexpressed in *Escherichia coli* and purified (*17*). The ternary complex crystals were grown by hanging drop vapor diffusion at 4 °C. The protein (5 mg/mL) was preincubated with PRPP and HPP (each at 5 mM) in 20 mM Bis-Tris (pH 6.8) and 6 mM $MgCl_2$. Drops containing equal volumes of the protein solution and precipitating solution were equilibrated against reservoirs of the precipitating solution which contained 0.1 M sodium acetate (pH 4.6), 0.2 M ammonium acetate, and 30% poly(ethylene glycol) 4000, as found in condition 10 of the Crystal Screen from Hampton Research. Cryoprotection was achieved with similar precipitating solutions enhanced with 10% PEG400. The crystals were mounted in nylon loops, flash-frozen and stored in liquid nitrogen, and then transferred into the beam using cryotongs at the Stanford Synchrotron Radiation Laboratory beamline 7-1 (SSRL BL7-1). Data were measured in a single sweep of 180° at a wavelength of 1.08 Å, using the 18 cm Mar image plate. Diffraction data were processed using the HKL package (*18*). The crystallographic data are summarized in Table 1. The unit cell volume indicated a dimer was expected to occupy the asymmetric unit, similar to a previous crystal form (*16*).

The structure was solved by molecular replacement using a monomer from the 1.4 Å crystal structure of the trypanosomal HPRT as the search probe (*16*) without including the bound ligands or any residues expected to be in the long flexible loop. The solution obtained with AMoRe (*19*) using X-ray data from 15 to 3.5 Å resolution had a correlation coefficient of 56.8 and an *R*-factor of 38.0% and placed the functional dimer in the asymmetric unit. The model was subjected to rigid body refinement using X-PLOR (*20*), and the noncrystallographic symmetry (NCS) operator determined using O (*21*), and then refined by positional and simulated annealing protocols against data from 6.0 to 2.5 Å resolution, imposing strict NCS constraints. Two-fold NCS averaging using RAVE (*22*) improved the initial $2F_o - F_c$ electron density maps, and clear $F_o - F_c$ electron density also facilitated modeling both ligands (PRPP and HPP) and two octahedrally coordinated Mg^{2+} ions (M1 and M2) and their associated water molecules into each active site of the dimer.

After the NCS constraints were released, the electron density for the long flexible loop in both the open and closed subunits was clearly interpretable, and permitted complete tracing of the loop in both the open and closed conformations. Also, eight additional residues which form a C-terminal helix could be modeled into the closed subunit. The subsequent refinement incorporated data to the 1.81 Å resolution limit and included refinement of individual isotropic temperature factors. During the early NCS-free refinement, a positive density peak was located at the position of one of the metal sites (M1) in the closed subunit, which led us to investigate the presence of a larger ion preferring octahedral geometry. It was deduced by the strength of the $F_o - F_c$ difference peak in omit maps (26σ vs $12-13\sigma$ for the other three ions), the bond lengths, and the behavior in the positional and temperature factor refinement that M1 in the closed subunit could be a manganese ion (discussed below). The temperature factor for this atom modeled as magnesium refined to 2.0 Å^2; after the atom was modeled as manganese, the temperature factor refined to 16 Å^2, similar to the refined temperature factors of the coordinated waters, coordinated PRPP oxygens, and neighboring protein atoms (12−16 Å^2).

The asymmetric unit of the crystal of the HPRT from *T. cruzi* contains a dimer of the enzyme, with one active site in the open conformation and the other closed. In the open subunit, flexible loop II is stabilized by packing contacts with other dimers in the crystal lattice, as also observed in the initial structure (*16*). However, in the closed subunit, the conformation of the flexible loop is not influenced by crystal packing and its position is not stabilized by lattice contacts but by specific interactions with the bound ligands. All atoms in the closed flexible loop are well-defined in the electron density maps, and refined temperature factors range between 20 and 30 Å^2 for main chain atoms in contact with the bound ligands and between 20 and 55 Å^2 for all loop atoms.

Table 1: Crystallographic Summary[a]

Diffraction Data	
space group	$P2_1$
a (Å)	39.4
b (Å)	101.8
c (Å)	51.9
β (deg)	94.2
resolution range (Å)	20.6−1.81 (1.85−1.81)
unique reflections (no.)	33026
completeness (%)[b]	88.8 (80.1)
R_{sym}[c]	0.051 (0.32)
$\langle I/\sigma(I)\rangle$	20.8 (3.8)
redundancy	3.6
Refinement	
resolution range (Å)	6.0−1.81
R_{cryst}[d]	0.191
R_{free}[e]	0.235
protein atoms (no.)	2998
ligand atoms (no.)	78
solvent atoms (no.)	119
average *B*-factor (Å^2)	
protein	22.8
ligands (open, closed)	23.02, 17.30
solvent	27.8
rmsd from target values	
bond lengths (Å)	0.010
bond angles (deg)	1.702

[a] Data were measured at SSRL BL 7−1. Values in parentheses are for the highest-resolution bin. [b] Completeness is the fraction of theoretically possible reflections observed at least once. [c] $R_{sym} = \sum|I_h - \langle I_h\rangle|/\sum I_h$, where $\langle I_h\rangle$ is the average intensity over symmetry equivalents. [d] $R_{cryst} = \sum|F_o - F_c|/\sum F_o$. [e] $R_{free} = R_{cryst}$ calculated for 6% of the reflections omitted from the refinement.

Temperature factors for all corresponding ligand atoms were lower by 2−20 Å^2 in the closed subunit than in the open subunit, consistent with their being solvent inaccessible. Ninety-three percent of residues in the structure are in the most favored regions of the Ramachandran plot, and all are in the allowed regions as analyzed by PROCHECK (*23*). The coordinates for the ternary complex of the HPRT of *T. cruzi* and the structure factors have been deposited with the Protein Data Bank, identification codes 1tc2 and r1tc2sf, respectively. Color figures were prepared with SETOR (*24*).

RESULTS AND DISCUSSION

Structure of the HPRT of T. cruzi. Ternary complex crystals were obtained by cocrystallization of the *T. cruzi* enzyme with PRPP and a hypoxanthine analogue, HPP, in the presence of Mg^{2+}. The structure was solved by molecular replacement and refined at 1.8 Å resolution to conventional and free *R*-factors of 19.1 and 23.5%, respectively. There is a dimer in the asymmetric unit of the crystal (Figure 1, top), and the active site of each monomer contains bound substrate (PRPP), analogue (HPP), and two octahedrally coordinated metal ions (M1 and M2). The active site is located at the domain interface and is comprised of four loops, I−IV (Figure 1). The two subunits differ primarily in the conformation of the "flexible loop" (loop II) and reveal two new views for the HPRT active site in the ternary complex, one open and solvent-exposed and the other sequestered from the solvent (Figure 1, bottom). A closed conformation of the flexible loop in the HPRT of *T. cruzi* was achieved using the nonreactive purine analogue HPP, which lacks a purine N9 and is unable to fulfill the role of nucleophile in the forward reaction. Therefore, this structure represents two pre-catalytic substrates-bound states, one prior to loop closure and the other after closure but prior to nucleophile activation.

Flexible Loop in Previous PRT Structures. The flexible loop (loop II) was generally characterized by poor electron density in previous HPRT structures (*13*−*16*), and the active sites in these structures are open, allowing for substrate binding or product release. To explain the proposed shielding of the active site chemistry, a closed conformation of the disordered flexible loop has been hypothesized for type I PRTs (*13*, *25*). Because of its importance for understanding enzyme mechanisms, the structure of a closed active site has long been a goal. A crystal structure for glutamine PRPP amidotransferase (GPAT), a related type I PRT, has been observed in a closed conformation in which the flexible loop completely buries the bound PRPP analogue (*26*).

Two other crystal structures have been interpreted as shedding light on a closed active site conformation for type I PRTs: a sulfate-bound OPRT from *E. coli* (*27*) and a sulfate-bound HPRT (HGXPRT) from *Toxoplasma gondii* (*15*). However, these structures did not provide information regarding interactions in the transition state because substrates for neither the forward nor reverse reactions were present as ligands in the enzyme active site. Also, the observed conformations of the flexible loop in both structures may have been affected by crystallographic artifacts, so the functional importance of loop residues is obscured (*15*, *27*). In the structure of the OPRT from *E. coli*, the closed flexible loop contributed an active site lysine residue to, and partially covered, the active site of the second subunit in the dimer

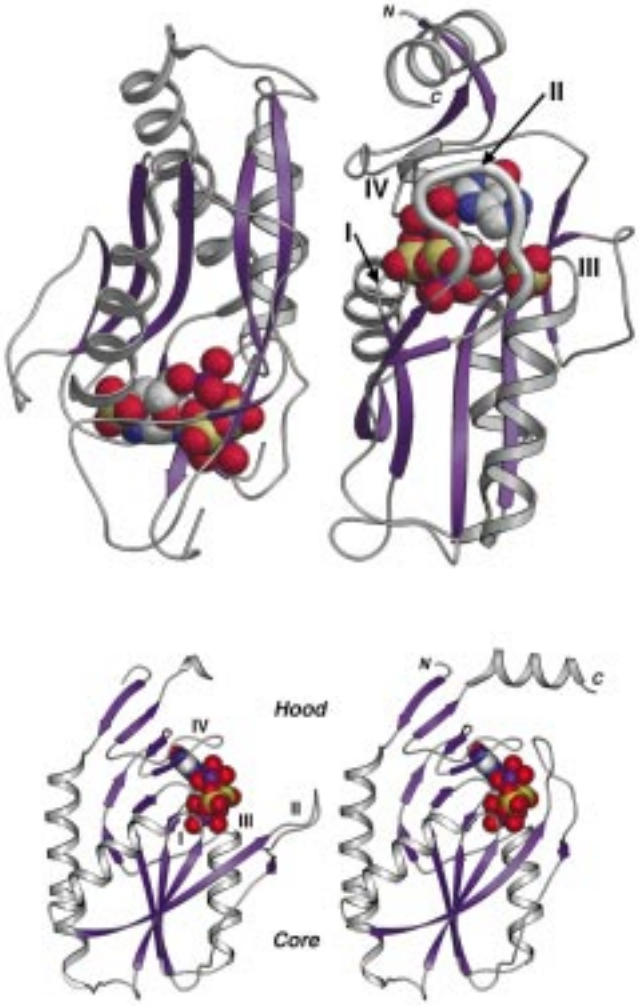

FIGURE 1: (Top) The HPRT from *T. cruzi* crystallized with a dimer in the asymmetric unit, containing one open (left) and one closed (right) subunit. The closed flexible loop (loop II), highlighted as a thicker rendition in the monomer on the right, buries the substrates. (Bottom) The monomers of the HPRT from *T. cruzi* in identical orientations are shown below the corresponding monomer in the dimer (above). The active site is located between the core and hood domains of the protein and is surrounded by four loops (I−IV). PRPP, HPP and hydrated metal ions are shown as van der Waals representations using standard atom coloring. Eight additional residues revealed in the electron density map for the closed subunit form a C-terminal helix above the closed flexible loop (right). Numbering of Atoms and Residues: The numbering system for amino acids used throughout the text and figures follows that for structurally homologous residues in the human HPRT (*13*, *16*). Atoms in HPP are identified according to the purine ring numbering system to facilitate discussion and comparisons with the natural substrate.

(*27*); however, it may have been stabilized in this conformation by crystal lattice contacts. In the structure of the HPRT from *T. gondii*, the flexible loop was only partly resolved, and the crystallographic temperature factors of loop residues were extremely high (80−100 Å^2 for most main chain atoms and 100 Å^2 for all side chain atoms). Although an invariant Ser-Tyr dipeptide was not modeled in this structure, a catalytic role for it was proposed, presumably deduced from the locations of adjacent residues in the loop which could be visualized in the electron density maps (*15*).

Substrates-bound type I PRTs with open active sites also have been reported. A recent crystal structure of the XGPRT

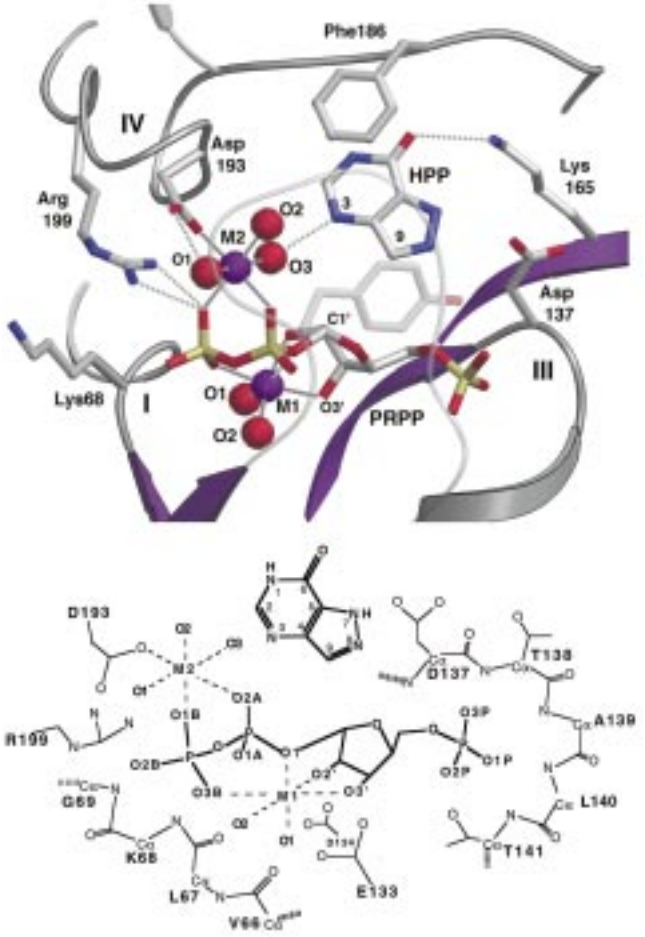

FIGURE 2: (Top) The active site of HPRT with bound PRPP, HPP and hydrated metal ions viewed through the closed flexible loop (transparent). Non-main chain protein−ligand hydrogen bonds are displayed, as well as metal (purple) coordination interactions with both substrates and the protein. On the left, the loop I conserved cis peptide interacts with both M1 O2 and PRPP O3B while the adjacent side chain, Lys68, projects across the dimer interface and interacts with the second subunit in the dimer. On the right, Lys165 forms a hydrogen bond with the exocyclic oxygen of HPP, and probably acts to discriminate for 6-oxopurine bases. Asp137 OD2 is within van der Waals distance of HPP and is 3.18 Å from the Tyr104 OH (transparent) in the closed active site. These three residues, as well as Asp193 and Arg199, are invariant in HPRTs. (Bottom) A schematic of the active site environment (see Table 2).

Table 2: Interatomic Distances between Ligand Atoms and Other Atoms in the Open and Closed HPRT Active Sites[a]

ligand atom	second atom	open	closed
M1	M1 O1	1.95	2.08
	M1 O2	1.99	2.14
	PRPP O1	2.52	2.62
	PRPP O2′	2.16	2.28
	PRPP O3′	2.30	2.30
	PRPP O3B	2.21	2.22
M1 O1	Val66 O	2.91	2.51
	Glu133 OE2	2.65	2.73
M1 O2	Leu67 O	2.93	
	Val101 O		2.58
M2	M2 O1	2.12	2.13
	M2 O2	2.17	2.10
	M2 O3	2.03	1.98
	PRPP O1B	2.17	2.22
	PRPP O2A	2.10	2.00
	Asp193 OD2	2.15	2.16
M2 O1	Asp193 OD2	2.98	2.74
	Gly105 O		2.64
M2 O2	Asp193 O	2.90	2.78
	water A	2.97	2.85
M2 O3	HPP N3	3.01	2.94
HPP			
N1	Val187 O	2.63	2.73
N3	M2 O3	3.00	2.95
O6	Lys165 NZ	2.87	2.93
N7	Asp137 OD2	4.05	3.27
N8	Asp137 OD2	4.10	3.31
PRPP			
O1B	Arg199 NH1	3.07	3.13
	Arg199 NH2	2.91	2.88
O2B	Lys68 N	2.99	3.09
O3B	Gly69 N	2.98	2.89
	Lys68 N	3.14	
O1A	Ser103 N		2.78
	Tyr104 N		2.84
O2′	Asp134 OD1	3.01	2.63
	Asp134 OD2	3.02	
O3′	Glu133 OE1	2.67	2.57
	Glu133 OE2	3.05	
O1P	Asp137 N	2.80	2.89
	Thr138 N	3.11	
	Ala139 N	2.93	2.95
O2P	Thr141 N		3.01
	Thr141 OG		2.65
	water C		2.88
O3P	Tyr104 OH		2.45
	Thr138 N	3.01	3.04
	Thr138 OG	2.85	2.63

[a] Refer to Figure 2 (bottom) for atomic identification; distances are in Angstroms.

from *E. coli* has been solved in a ternary complex with xanthine or guanine and a nonhydrolyzable PRPP analogue, but the flexible loop was involved in crystal lattice contacts and did not cover the active site (*28*). In the structure of an OPRT from *Salmonella typhimurium* with both substrates bound (orotate and PRPP), the flexible loop was partially disordered and did not bury the substrates, although the loop contributed an arginine residue that forms hydrogen bonds with the PRPP bound in the active site of the second subunit of the dimer (*29*). Thus far, the substrate analogues-bound GPAT structure has been the only example of a closed conformation for a type I PRT (*26*). The ternary complex in a closed conformation reported here is the first structure of an HPRT to provide details of the interactions between residues of the flexible loop and the ligands in the active site of the enzyme.

Substrate Binding Interactions in the Ternary Complex. In each active site of the trypanosomal HPRT, the PRPP is cradled between active site loops I and III (Figure 2) which form extensive networks of hydrogen bonds with the phosphate and pyrophosphate groups, primarily through main chain atoms (Table 2). One metal ion binding site is located on either side of the PRPP pyrophosphate group (Figure 2). The HPP purine-like ring binds above PRPP, in a hydrophobic pocket beneath Phe186 of the hood domain, and is positioned by main chain hydrogen bonds from loop IV residues and invariant Lys165 NZ. Ten of the 13 residues which are invariant among the amino acid sequences of the HPRTs from eukaryotes and prokaryotes (*16*) make contact with the bound ligands in the closed subunit. In both the open and closed subunits, eight invariant residues (Leu67, Gly69, Glu133, Asp134, Asp137, Lys165, Asp193, and Arg199) form hydrogen bonds with PRPP, HPP, or ion-coordinated water molecules, and Asp193 additionally

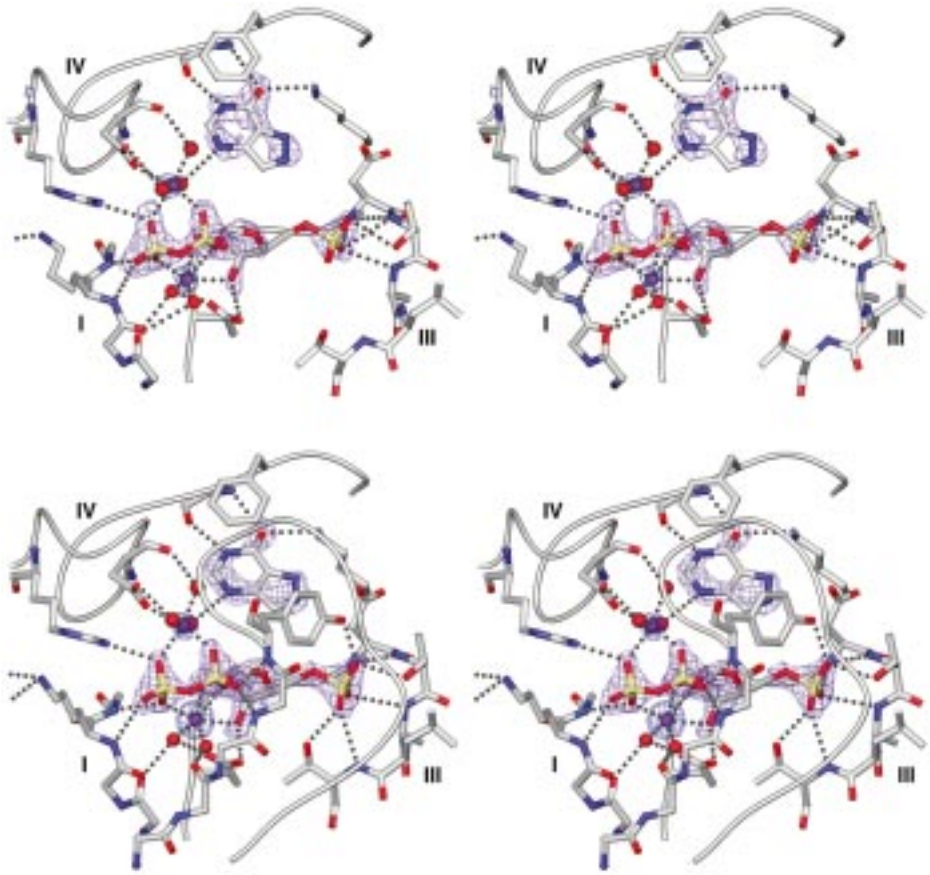

FIGURE 3: A stereoview of the open (Top) and closed (Bottom) active sites. Interactions between the protein and the HPP, PRPP and two divalent metal ions and coordinated water molecules are shown. Hydrogen bonds between the PRPP O2′ and Asp 134 (located behind the PRPP) are omitted for clarity, as are a few ordered water molecules which contribute to a network of hydrogen bonds in both active sites. For some atoms participating in multiple hydrogen bonds, not all are shown. Side chains displayed are, from loop I − Lys68 (plus adjacent mainchain atoms), from loop III − Glu133 (bottom center in each panel), Asp137, Thr138 Ala139, Leu140 and Thr141, and from loop IV − Phe186 (plus mainchain atoms of Val187), Asp193 and Arg199. At the upper right, Lys165 forms a hydrogen bond with the HPP exocyclic oxygen. Metal M1 is below the PRPP and M2 is above. Omit maps ($F_o - F_c$) for the bound ligands are shown superimposed, in purple for PRPP and HPP (contoured at 3.9σ), and in blue for the metal ions (contoured at 8σ). The density for the PRPP and HPP is consistent with the refined temperature factors which are 2−20 Å^2 lower in the closed active site for the equivalent ligand atoms. The peak height of three of the four metal ions is 12−13σ, while the peak height of M1 in the closed active site (Bottom) is 26 σ. In the closed active site (Bottom) only the position of the main chain of the closed flexible loop is shown, with the exception of invariant Tyr104.

provides the only direct protein−metal bond (Figure 2). The side chain of Arg199 interacts with the PRPP pyrophosphate group, as predicted by Craig et al. (*30*), and also interacts (not shown) with the side chain of Asp193, possibly serving to position both side chains. Main chain atoms from conserved loop I residues form hydrogen bonds with M1 ion-coordinated waters and two PRPP pyrophosphate oxygens (Table 2 and Figure 3), similar to the interactions observed in GPAT (*26*). A conserved non-proline cis peptide in loop I exposes the amide nitrogen to the active site so that the peptide bond between Leu67 and Lys68 contributes two adjacent hydrogen bonds to the PRPP−metal complex in the open active site and one in the closed active site (Figure 3 and Table 2). The PRPP binding motif, also conserved among type I PRTs, contributes active site loop III which encircles the ribose−phosphate group (Figure 2, top); its members Glu133 and Asp134 (invariant in HPRTs and conserved in type I PRTs) form hydrogen bonds with both PRPP hydroxyls, with one of the M1 ion-coordinated waters (M1 O1) (Figure 2, bottom, Figure 3 and Table 2), and with one or two additional ordered active site waters (not shown). These two acidic side chains bind the PRPP−metal complex, but they do not participate in any direct bonds with the metal ion. They were predicted to participate in stabilizing the transition state of the reaction (*13*, *25*), and they may be responsible for networking other water molecules bound in the active site to facilitate further hydration of the pyrophosphate leaving group during or following the reaction.

Two Divalent Metal Ions. Phosphate transfer reactions nearly always require a metal ion (*31*). The involvement of divalent ion pairs in catalysis has been documented in several enzymes that catalyze phosphotransfer reactions, such as DNA polymerase (*32*), phosphoenolpyruvate carboxykinase (*33*), pyruvate kinase (*34*), and glutamine synthetase (*35*), some utilizing a combination of different ions. The possibility that two divalent metal ions are involved in the mechanism for an HPRT was not anticipated. The two metal ions in this structure, located on opposite sides of the PRPP pyrophosphate group (Figure 2), are separated by 5.57 and 5.43 Å in the open and closed subunits, respectively, and each interacts with two pyrophosphate oxygens (Figure 2 and Table 2). The electron density for all ligand atoms, metal ions, and their coordinated waters was clearly interpretable in this structure (Figure 3). Mg^{2+} was present throughout the purification and crystallization of the protein, and three

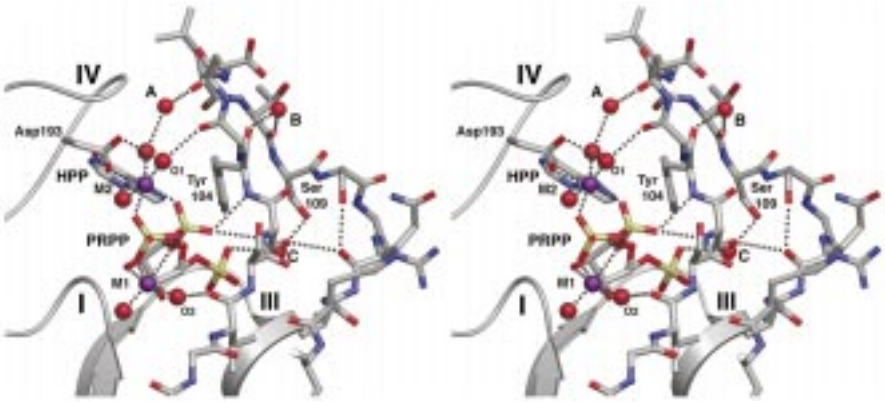

FIGURE 4: Interactions of the closed flexible loop with the ligands bound to the HPRT active site. This stereoview is rotated approximately ninety degrees on the vertical axis relative to those shown in Figures 2 and 3. Metal-protein and metal-water interactions are displayed, as well as hydrogen bonds formed within the loop, and between the loop and the bound ligands. Some main chain hydrogen bonds are omitted for clarity. Active site loops I, III and IV are shown as ribbons and labeled. All non-hydrogen atoms are drawn for the long flexible loop II (not labeled) which begins near the bottom center and ends near the bottom right of the figure. Also shown are the non-hydrogen atoms of the bound ligands (PRPP, HPP) and hydrated metal ions (M1, M2), Asp193, and three additional bound water molecules (labeled A, B and C) discussed in the text. Tyr104 is located near the center of the figure and projects away from the viewer. Ser103 (not labeled) precedes Tyr104 in the loop and forms a hydrogen bond with water C, which also shares a hydrogen bond with Ser109 and the PRPP 5′-phosphate group.

of the four metal sites (in the dimer) appear to be occupied by magnesium ions. However, the ion located in site M1 in the closed active site exhibits much stronger difference electron density than the ions at the other three sites (Figure 3), and the bond lengths, geometry, and behavior of the temperature factors during crystallographic refinement are all consistent with the ion being Mn^{2+} (which could be present only as a trace contaminant). Interestingly, Mn^{2+} supports catalysis of the trypanosomal HPRT with a slightly higher specific activity than with Mg^{2+}, although no synergistic effect (*33*) was noted when a combination of the two divalent ions was used (data not shown).

A single Mg^{2+} or Mn^{2+} ion (analogous to M1) has been observed in related type I PRT structures with PRPP or a PRPP analogue bound (*26*, *28*, *29*). This ion (shown below the PRPP in Figures 2 and 3) forms no direct interactions with the protein and is likely responsible for maintaining a conformation of the PRPP that facilitates binding to the active site (*26*), and may also mask the negative charges on some PRPP oxygens. M1 forms a slightly distorted octahedral coordination geometry with three bonds to PRPP via O3B and the 2′- and 3′-hydroxyls, and two bonds with water molecules (M1 O1, and M1 O2) (Figure 2). The metal−oxygen distance to the PRPP O1 is too long to optimally satisfy the metal coordination sphere in either active site (Table 2) (*31*), raising the possibility that formation of the sixth metal bond with O1 during cleavage of the PRPP may be a driving force in the reaction. A Mn^{2+} ion, which has an additional electron shell and a slightly larger ionic radius compared to Mg^{2+}, might better tolerate this distorted geometry.

The second metal ion (M2) (above and to the left of PRPP in Figures 2 and 3) was unanticipated, and clearly can be interpreted as Mg^{2+} in both active sites in this structure, with average metal−oxygen bond distances of 2.1 Å (see Table 2) and less distorted octahedral geometry. The metal is coordinated by invariant residue Asp193 OD1, three water molecules (M2 O1−O3), and two PRPP pyrophosphate oxygens (O1B and O2A) (Figure 2 and Table 2). These PRPP pyrophosphate oxygens are in a high-energy eclipsed conformation due to the metal ligation, which may facilitate activation of the substrate for catalysis. Metal ion-coordinated water molecules M2 O1 and M2 O2 form additional hydrogen bonds with the Asp193 OD2 and carbonyl group, respectively, while water M2 O3 forms a hydrogen bond to the nitrogen at purine position N3 of HPP (Figures 2 and 3). The interaction of the 'N3' of the hypoxanthine analogue with the metal's water ligand (M2 O3) suggests how it could similarly position a bound purine base (via N3) for in-line nucleophilic attack at the PRPP ribose C1′ (Table 2), a geometry which would be anticipated if the proposed catalytic mechanism indeed utilizes S_N2-type nucleophilic displacement. This Mg^{2+} ion (M2) both binds PRPP and, through its coordinated waters, appears to stabilize the relative positions of PRPP and HPP, while firmly anchored to the protein scaffold via Asp193. Other PRPP-bound type I PRT structures revealed only a single metal site (M1) (*26*, *28*, *29*). Asp193 is invariant in HPRTs; thus, all HPRTs are likely to utilize two divalent metal ions for catalysis.

Closed HPRT Flexible Loop. With both substrates bound, the active site architecture is completed by the closure of flexible loop II, which forms additional and extensive interactions with the bound PRPP and hydrated metal ions. In the closed active site, the main chain atoms of invariant residues Ser103 and Tyr104 form hydrogen bonds to the PRPP O1A (Figure 3, bottom, and Figure 4), while the tyrosine hydroxyl group forms a strong (2.45 Å) hydrogen bond to the 5′-phosphate oxygen O3P (Figure 3, bottom). Tyr104 forms interactions which span the PRPP molecule. The plane of its aromatic ring is nearly perpendicular to the PRPP ribose ring, positioned above the O4′, and may provide electrostatic stabilization during the transition state. When loop II closes, one hydrogen bond is lost between the loop I cis peptide carbonyl group and M1 O2, another is lost between the cis peptide amide nitrogen and the PRPP β-phosphate oxygen O3B, and the interaction between the cis peptide amide nitrogen, and the PRPP β-phosphate oxygen O2B lengthens (Figure 3 and Table 2). An additional hydrogen bond in the closed subunit forms across the dimer interface, involving the lysine side chain adjacent to the cis

peptide (Lys68) (Figure 3, bottom), possibly providing a mechanism for communication between the active sites of the dimer as formerly suggested (*16*). The closed flexible loop II forms a new β-sheet interaction with the loop I cis peptide carbonyl group previously bound to M1 O2 and replaces the M1 O2 hydrogen bond lost by loop I with a hydrogen bond contributed by the Val101 carbonyl group (Figure 3 and Figure 4). Also, the closed loop makes four additional hydrogen bonds with pyrophosphate oxygens and metal-coordinated waters, providing a possible mechanism for liberation of the pyrophosphate leaving group (Figure 4). Although Mg^{2+} was located in the crystal structure of a product-bound GPAT (*26*), no metal ions have been observed in product-bound HPRT structures (*8*, *13*, *14*). In addition, PRPP binding to the human HPRT is not detectable in the absence of Mg^{2+} (*8*). Thus, both metal ions are likely to enter and leave the active site in association with the negatively charged PRPP or pyrophosphate.

The lack of strong interactions between residues in the flexible loop and the remainder of the protein may minimize the energy costs of closing and opening the loop. The closed loop predominantly interacts with the bound ligands and metal-coordinated water molecules. These interactions appear to be necessary for forming the closed active site. There are also specific interactions among loop residues that may internally stabilize the formation of its closed conformation (Figure 4). The loop is anchored to the core domain parallel β-sheet by a short antiparallel strand (Figure 1, bottom). Two conserved serine residues near the β-anchor (S103 and S109) both interact with the same water molecule which also forms a hydrogen bond with PRPP 5′-phosphate oxygen O2P (water C in Figure 4 and Table 2). The two serines, together with Ser110 and Gln112 form a network of hydrogen bonds near the β-anchor, while the omega-shaped turn of the loop is stabilized by several interactions between its main chain atoms, one mediated by a bridging water molecule (water B in Figure 4). The sequence of the loop is not well-conserved in HPRTs following the Ser103-Tyr104 dipeptide but generally contains polar side chains, all of which are solvent-exposed in this structure (Figure 4). The main chain atoms of the loop bind PRPP and the hydrated metal ions, including one interaction via a water molecule which is in the second solvation shell of metal M2 (water molecule A in Figure 4). There is one hydrophobic residue in the *T. cruzi* HPRT loop II sequence which forms van der Waals contacts with, and is located between, Phe186 (Figure 2) and the C-terminal helix (Figure 1).

Active Site and Ligand Comparisons. In addition to the differences in the location of the flexible loop in the two subunits, the relative positions of the ligands have shifted within the two active sites, and a comparison illustrates some details consistent with the proposed catalytic mechanism. PRPP and HPP are farther apart than would be expected for the two substrates during the transition state (C1′ to ′C9′ distances are 3.8 Å), and they are prevented from closer approach probably due to the presence of a hydrogen on the carbon at purine position 9 that would be absent on the N9 of a purine substrate. Nevertheless, concerted differences in hydrogen bond lengths in the two active sites indicate that HPP may be released somewhat from the purine binding site following loop II closure, allowing it to more closely approach PRPP and the candidate general base, Asp137 (*13*, *36*) (Figure 2 and Figure 3). Specifically, the bond lengths between Lys165NZ and the exocyclic oxygen of HPP and between the carbonyl group of Val187 and HPP ′N1′ lengthen somewhat, while the bond between HPP ′N3′ and the substrates-orienting water M2 O3 becomes shorter and the position of HPP shifts nearly 0.8 Å to bring ′N7′ and ′N8′ within van der Waals contact with Asp137 OD2 (Figure 3 and Table 2). Furthermore, loop III tightens around the PRPP 5′-phosphate group in the closed subunit, forming two additional hydrogen bonds with the substrate (Figure 3, bottom, and Table 2) that contribute to the displacement of two water molecules observed only in the open active site (not shown).

The conformations of the two bound PRPP molecules differ as well, indicating that the closed loop may be associated with a more tightly bound, and possibly more strained substrate, as would be expected in the transition state (*37*). The two metal ions move closer together in the closed subunit (5.57 Å vs 5.43 Å), and there is a slight reduction in the angle between the PRPP pyrophosphate and 5′-phosphate groups, which may promote the formation of a more planar ribose ring, and may also result in increased strain on the scissile bond between C1′−O1 (*26*). This structure does not support the conclusion that the enzyme plays a secondary role in a primarily 'substrate-assisted' catalytic mechanism, as was proposed by Krahn et al. for GPAT (*26*), nor does it support a mechanism which would necessarily result in a positively charged intermediate indicative of S_N1-type nucleophilic displacement (*10*). If the reaction were driven by hydrolysis of the highly unstable PRPP substrate and the development of an oxycarbonium intermediate, the presence of the intact PRPP molecule in the active site would not be expected. Thus, the structure reported here is consistent with an S_N2-type nucleophilic displacement mechanism leading to the formation of the transition state, with nucleophilic attack by the purine base occurring simultaneously with departure of the pyrophosphate leaving group via a ternary complex. The PRPP is poised for pyrophosphate removal by contacts with the metal ions and the closed flexible loop (Figure 3, bottom, and Figure 4), but this step is apparently dependent on nucleophile activation, prevented in our HPP-bound structure by the presence of the carbon at purine position 9 (Figure 2). Thus, the structure represents a catalytically trapped active site awaiting activation of the nucleophile.

The long flexible loop may function not simply to protect the active site chemistry from solvent. Instead it appears to be important for binding the PRPP pyrophosphate group and associated hydrated metals, and may be responsible for facilitating the removal of an activated pyrophosphate leaving group as well, with invariant loop member Tyr104 possibly playing a role in electrostatic stabilization of the transition state (Figure 2, top, and Figure 3, bottom). Due to their electron withdrawing propensity, the two metal ions probably have a role in activating the pyrophosphate leaving group, as is observed in other phospho-transfer enzymes utilizing two metal ions (*38*). The relative positions of the two metal ions and their ligands change between the open and closed subunits, consistent with their roles in substrate positioning and activation in the HPRT-catalyzed reaction.

HPRT Structure in Human Disease. Mutations in the X-linked gene encoding the human HPRT are associated with

enzyme deficiencies leading to gouty arthritis or Lesch-Nyhan syndrome (*4*, *39*). Because of considerable homology between the structures of the human and trypanosomal enzymes, the structure reported here provides new evidence for the molecular basis of three point mutations in the human HPRT that result in disease syndromes: S103R and S109L, which cause gout, and D193N, which causes Lesch-Nyhan syndrome. Both the S103R and S109L mutations would result in the destabilization of the hydrogen bond network found near the β-anchor of the flexible loop II (Figure 4), and would presumably interfere with correct formation of the closed loop and active site. The mutation of Asp193 to asparagine severely impairs catalytic activity (*39*), probably because the substitution of a nitrogen for the oxygen in the M2 binding site impairs binding to both the metal ion and its coordinated water (Figure 2), energetically destabilizing the only direct interaction between the enzyme and the metal ion which appears to be important for maintaining the relative orientations of the two substrates. The mutation could also disrupt the hydrogen bonding interactions between Asp193 and Arg199 that may be essential for proper positioning of both side chains as necessary for substrate binding.

CONCLUDING REMARKS

The catalytically competent active site of the closed structure described here provides new information to further our understanding of the HPRT enzyme mechanism. It reveals the roles of a number of conserved amino acids in substrate binding and loop closure, provides the first evidence for the involvement of two metal ions in the enzyme-catalyzed reaction, and provides a glimpse of the complete HPRT active site just prior to nucleophile activation and formation of the transition state. In addition, the structure reported here provides the first insight for understanding disease-causing mutations which alter residues in the human HPRT important to the formation of the closed flexible loop, or which contribute to the previously unrecognized Mg^{2+} ion binding site (M2). Furthermore, this structure of the complete active site of the trypanosomal HPRT in a closed conformation will be of considerable value for the support of future efforts to discover chemotherapeutic agents for the treatment of diseases caused by parasites.

ACKNOWLEDGMENT

We are grateful to D. M. Freymann of Northwestern University for helpful suggestions in preparing the manuscript and R. J. Fletterick of the University of California at San Francisco for resources. We thank Olivier Froelich of Novartis Products, Inc., for synthesis of the HPP. Also, we thank T. W. Traut, R. V. Wolfenden, and E. J. Collins of the University of North Carolinia at Chapel Hill for critical reading of the manuscript. This work is based upon research conducted at the Stanford Synchrotron Radiation Laboratory. The SSRL is funded by the Department of Energy and the National Institutes of Health.

REFERENCES

1. Garcia-Zapata, M. T. A., McGreevy P. B. and Marsden, P. D. (1991) in *Hunter's Tropical Medicine* (Strickland, G. T., Ed.) pp 628−637, W. B. Saunders, Philadelphia.
2. Berens, R. L., Krug, E. C., and Marr, J. J. (1995) in *Biochemistry and Molecular Biology of Parasites* (Marr, J. J., and Müller, M., Eds.) pp 89−117, Academic Press Inc., San Diego.
3. Craig, S. P., III, and Eakin, A. E. (1997) *Parasitol. Today 13*, 238−241.
4. Sculley, D. G., Dawson, P. A., Emmerson, B. T., and Gordon, R. B. (1992) *Hum. Genet. 90*, 195−207.
5. Eads, J. C., Ozturk, D., Wexler, T. B., Grubmeyer, C., and Sacchettini, J. C. (1997) *Structure 5*, 47−58.
6. Smith, J. L. (1995) *Curr. Opin. Struct. Biol. 5*, 752−757.
7. Yuan, L., Craig, S. P., III, McKerrow, J. H., and Wang, C. C. (1992) *Biochemistry 31*, 806−810.
8. Xu, Y., Eads, J., Sacchettini, J. C., and Grubmeyer, C. (1997) *Biochemistry 36*, 3700−3712.
9. Munagala, N. R., Chin, M. S., and Wang, C. C. (1998) *Biochemistry 37*, 4045−4051.
10. Goitein, R. K., Chelsky, D., and Parsons, S. M. (1978) *J. Biol. Chem. 253*, 2963−2971.
11. Bhatia, M. B., and Grubmeyer, C. (1993) *Arch. Biochem. Biophys. 303*, 321−325.
12. Tao, W., Grubmeyer, C., and Blanchard, J. S. (1996) *Biochemistry 35*, 14−21.
13. Eads, J. C., Scapin, G., Xu, Y., Grubmeyer, C., and Sacchettini, J. C. (1994) *Cell 78*, 325−334.
14. Somoza, J. R., Chin, M. S., Focia, P. J., Wang, C. C., and Fletterick, R. J. (1996) *Biochemistry 35*, 7032−7040.
15. Schumacher, M. A., Carter, D., Roos, D. S., Ullman, B., and Brennan, R. G. (1996) *Nat. Struct. Biol. 3*, 881−887.
16. Focia, P. J., Craig, S. P., III, Nieves-Alicea, R., Fletterick, R. J., and Eakin, A. E. (1998) *Biochemistry 37*, 15066−15075.
17. Eakin, A. E., Guerra, A., Focia, P. J., Torres-Martinez, J., and Craig, S. P., III (1997) *Antimicrob. Agents Chemother. 41*, 1686−1692.
18. Otwinowski, Z. (1993) *Oscillation data reduction program*, SERC Daresbury Laboratory, Warrington.
19. Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
20. Brünger, A. T. (1992) *X-PLOR*: *A System for X-ray Crystallography and NMR*, Yale University Press, New Haven.
21. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
22. Kleywegt, G. J., and Jones, T. A. (1994) *Halloween....Masks and Bones*, SERC Daresbury Laboratory, Daresbury.
23. Laskowski, R. A., MacArthur, M. W., and Thorton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
24. Evans, S. V. (1993) *J. Mol. Graphics 11*, 134−138.
25. Scapin, G., Grubmeyer, C., and Sacchettini, J. C. (1994) *Biochemistry 33*, 1287−1294.
26. Krahn, J. M., Kim, J. H., Burns, M. R., Parry, R. J., Zalkin, H., and Smith, J. L. (1997) *Biochemistry 36*, 11061−10068.
27. Henriksen, A., Aghajari, N., Jensen, K. F., and Gajhede, M. (1996) *Biochemistry 35*, 3803−3809.
28. Vos, S., Parry, R. J., Burns, M. R., de Jersey, J., and Martin, J. L. (1998) *J. Mol. Biol. 282*, 875−889.
29. Scapin, G., Ozturk, D. H., Grubmeyer, C., and Sacchettini, J. C. (1995) *Biochemistry 34*, 10744−10754.
30. Craig, S. P., III, Focia, P. J., and Fletterick, R. J. (1997) *Biochim. Biophys. Acta 1339*, 1−3.
31. Glusker, J. P. (1991) *Adv. Protein Chem. 42*, 1−76.
32. Beese, L. S., and Steitz, T. A. (1991) *EMBO 10*, 25−33.
33. Tari, L. W., Matte, A., Goldie, H., and Delbaere, L. T. J. (1998) *Nat. Struct. Biol. 4*, 990−994.
34. Baek, Y. H., and Nowak, T. (1982) *Arch. Biochem. Biophys. 217*, 491−497.
35. Liaw, S.-H., and Eisenberg, D. (1994) *Biochemistry 17*, 675−681.
36. Xu, Y., and Grubmeyer, C. (1998) *Biochemistry 37*, 4114−4124.
37. Wolfenden, R. (1974) *Mol. Cell. Biochem. 3*, 207−211.
38. Dismukes, G. C. (1996) *Chem. Rev. 96*, 2909−2926.
39. Wilson, J. M., Young, A. B., and Kelley, W. N. (1983) *N. Engl. J. Med. 309*, 900−910.

BI9821465