Biochemistry

 Volume 33, Number 6 February 15, 1994

Accelerated Publications

Crystal Structure of Orotate Phosphoribosyltransferase[†,‡]

Giovanna Scapin,[§] Charles Grubmeyer,[∥] and James C. Sacchettini*[,§]

Department of Biochemistry, Albert Einstein College of Medicine, Bronx, New York 10461, and Department of Biochemistry, Temple University School of Medicine, Philadelphia, Pennsylvania 19140

Received November 22, 1993; Revised Manuscript Received December 15, 1993[⊛]

ABSTRACT: Phosphoribosyltransferases (PRTases) are enzymes involved in the synthesis of purine, pyrimidine, and pyridine nucleotides. They utilize α-D-5-phosphoribosyl-1-pyrophosphate (PRPP) and a nitrogenous base to form a β-*N*-riboside monophosphate and pyrophosphate (PP_i), and their functional significance in nucleotide homeostasis is evidenced by the devastating effects of inherited diseases associated with the decreased activity and/or stability of these enzymes. The 2.6-Å structure of the *Salmonella typhimurium* orotate phosphoribosyltransferase (OPRTase) complexed with its product orotidine monophosphate (OMP) provides the first detailed image of a member of this group of enzymes. The OPRTase three-dimensional structure was solved using multiple isomorphous replacement methods and reveals two major features: a core five-stranded α/β twisted sheet and an N-terminal region that partially covers the C-terminal portion of the core. PRTases show a very high degree of base specificity. In OPRTase, this is determined by steric constraints and the position of hydrogen bond donors/acceptors of a solvent-inaccessible crevice where the orotate ring of bound OMP resides. Crystalline OPRTase is a dimer, with catalytically important residues from each subunit available to the neighboring subunit, suggesting that oligomerization is necessary for its activity. On the basis of the presence of a common PRPP binding motif among PRTases and the similar chemistry these enzymes perform, we propose that the α/β core found in OPRTase will represent a common feature for PRTases. This generality is demonstrated by construction of a model of the human hypoxanthine–guanine phosphoribosyltransferase (HGPRTase) from secondary structure predictions for HGPRTase and the three-dimensional structure of OPRTase.

Nucleotide synthesis and salvage depend on the action of phosphoribosyltransferases, a group of enzymes that utilize the ribose 5-phosphate donor α-D-5-phosphoribosyl-1-pyrophosphate (PRPP) and a nitrogenous, generally aromatic base, in the presence of a divalent metal ion, to form an *N*-riboside monophosphate and pyrophosphate (PP_i). Members of this group play an essential role in both *de novo* and recycling pathways of purine, pyrimidine, and pyridine metabolism, as well as in the biosynthesis of histidine and tryptophan (Musick, 1981). The importance of these enzymes in nucleotide homeostasis is apparent from the clinical consequences of PRTase malfunction. Congenital deficiency of hypoxanthine–guanine PRTase results in hereditary gout or, in cases of severe deficiency, in mental retardation associated with Lesch–Nyhan syndrome (Stout & Caskey, 1989). A rare inborn defect of the orotate OPRTase domain of UMP synthase is linked to one form of hereditary orotic aciduria that may also result in mental retardation (Suttle et al., 1989).

Comparisons of the primary structures of the PRTases have revealed the conservation of only a short stretch of 12–13 residues. These similar sequences in most cases contain two

[†] This work was supported by a grant from the National Institutes of Health (GM48623) to C.G. and J.C.S.

[‡] Coordinates for the crystal structure of orotate phosphoribosyltransferase have been deposited with the Brookhaven Protein Data Bank under the file name 1STO.

* Address correspondence to this author at the Department of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461 [phone, (718) 430-2741; Fax, (718) 597-5692; E-mail, sacchett@scud.bioc.aecom.yu.edu].

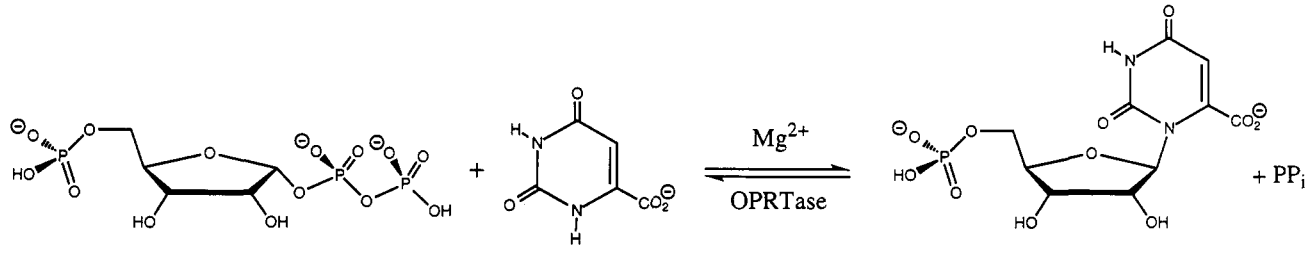
[§] Albert Einstein College of Medicine.

[∥] Temple University School of Medicine.

[⊛] Abstract published in *Advance ACS Abstracts,* January 15, 1994.

0006-2960/94/0433-1287$04.50/0

Scheme 1



PRPP Orotate OMP

adjacent carboxylate residues, flanked by hydrophobic residues and followed by four residues containing mostly small side chains, of which one or more is a glycine (Hersey & Taylor, 1986). In OPRTase from *Salmonella typhimurium*, the sequence is ^{120}VMLVDDVITAGT^{131}. Interestingly, a similar stretch of amino acids has also been identified in the primary structure of PRPP synthetase (Hove-Jensen et al., 1986). On the basis of its presence in both PRPP synthetase and the PRTases, this sequence has been proposed to represent a PRPP binding motif (Hersey & Taylor, 1986; Hove-Jensen et al., 1986). The need for structural information on this group of enzymes has led several groups to calculate secondary structure predictions and to hypothesize that PRTases have a common tertiary structure that includes a region of repeating α-helices and β-strands at the amino terminal end of the protein (Argos et al., 1983). On the basis of this theory, a three-dimensional structure has been predicted for human HGPRTase which describes sites for substrate binding and clinically important mutations (Wilson et al., 1983; Stout & Caskey, 1989).

We have used the 213-residue *S. typhimurium* OPRTase as a model to understand the structural basis for catalysis by this group of enzymes. Orotate phosphoribosyltransferase catalyzes the OMP-forming step in *de novo* pyrimidine nucleotide biosynthesis (Scheme 1). A survey of sequences in GenBank, EMBL, and the Swiss Protein data bases reveals a total of 13 OPRTase sequences, which represent divergent species, including Gram-negative and -positive bacteria, fungi, insects, and mammals, with sequence identities ranging between 30% and 95%. *S. typhimurium* OPRTase has been sequenced (Scapin et al., 1993), overexpressed in *E. coli*, and purified to homogeneity. Its kinetic mechanism (Bhatia et al., 1990) and crystallization (Scapin et al., 1993) have been reported. The enzyme catalyzes the formation of OMP and PP_i from orotate and PRPP, in the presence of Mg^{2+}, in a freely reversible reaction. We now report the three-dimensional structure of *S. typhimurium* OPRTase with bound OMP to 2.6-Å resolution. These studies reveal the functional enzyme to be a symmetric dimer. The core structure of OPRTase has a fold similar to the α/β nucleotide binding domain of many dehydrogenases. A crevice, formed between the C-terminal end of this α/β core and an additional structure located at the protein's N-terminus, serves to bind the orotate ring of OMP. The location of the secondary structural elements found for OPRTase is different from what has been previously predicted (Argos et al., 1983), in that the α/β fold begins nearer to the middle and not at the N-terminal end of the protein. We have used the structure of OPRTase along with a recently reported algorithm for predicting secondary structure (Rost & Sander, 1992) to build a new model of human HGPRTase.

EXPERIMENTAL PROCEDURES

Crystallization. Crystals of OPRTase were grown using the hanging drop vapor diffusion method, as previously described (Scapin et al., 1993). The crystals are tetragonal, space group $P4_12_12$, with unit cell parameters $a = b = 48.4$ Å, $c = 210.50$ Å, and $\alpha = \beta = \gamma = 90.0°$. There is one monomer per asymmetric unit. The solvent content of these crystals is ~53%, and they diffract to better than 2.3-Å resolution. Two derivatives [K_2PtCl_4 and *p*-(chloromercuri)-benzenesulfonate, PCMBS] were used to solve the structure. The platinum derivative was collected after a native crystal was soaked overnight in 5 mM K_2PtCl_4 in 30% PEG 4000, 100 mM Hepes, pH = 7.5, 10 mM OMP, and 0.05% NaN_3. The mercury derivative was obtained by pre-reacting the protein (15 mg/mL in 100 mM Hepes, pH = 7.5, and 0.05% NaN_3) with a 4-fold molar excess of PCMBS for about 30 min at 4 °C and then crystallizing the complex under the same conditions that gave native crystals. Crystals of the OPRTase–PCMBS complex were tetragonal and isomorphous ($a = b = 48.2$ Å, $c = 210.2$ Å) with the native form and could then be used in the MIR procedures.

Data Collection and Processing. All data sets were collected on a Siemens multiwire area detector, using a Rigaku RU-200 rotating anode X-ray source operating at 55 kV and 85 mA. Data were reduced using the Siemens package XENGEN (Siemens Analytical X-ray Instruments, Inc., Madison, WI) on a Silicon Graphics Iris computer. For the native data set, the *R*-merge on intensities was 6.2% to 2.5-Å resolution for 6389 reflections (~60% complete). The platinum derivative had an *R*-merge on intensities of 7.2% for 6630 reflections to 2.7-Å resolution. For the OPRTase–PCMBS data set, the *R*-merge on intensities for 8648 reflections to 2.4-Å resolution was 8.4%.

Structure Determination. The crystal structure of OPRTase was determined using multiple isomorphous replacement (MIR) data collected from the two derivatives described above. Table 1 summarizes the statistics for phase determination. The heavy atom positions (as calculated from difference Patterson maps) were refined by an iterative series of phase refinement, using the package PHASES (W. Furey, VA Medical School and University of Pittsburgh, Pittsburgh, PA), running on a Silicon Graphics Iris computer. The anomalous scattering contribution of the mercury derivative was used to assign the absolute configuration of the molecule (Blundell & Johnson, 1976). These data produced a mean figure of merit of 0.550 for 6280 phased reflections with $F > 1\sigma$ (including anomalous scattering in the phase calculation for the mercury data set). Solvent flattening (Wang, 1985) procedures, as implemented in PHASES, were used to further improve the MIR phases. From the resulting electron density map, a partial polyalanine model was built using the program TOM, a derivative of FRODO (Jones, 1985), displayed on a Iris Graphics workstation. An extensive series of "combined" electron density maps (i.e., maps obtained combining MIR phases and model-based phases) were used to trace the complete model, with exclusion of five residues located in a

Table 1: Heavy Atom Derivatives Used in the OPRTase Structure Determination

derivative	site	x	y	z	occ[a]	B[b]	R_c[c]	R_k[d]	no. of reflections	phasing[e] power	D (Å)[f]
PCMBS	HG1	0.697	0.398	0.102	0.78	30.15	0.588	0.176	5566	1.61	2.70
	HG2	0.734	0.362	0.016	0.93	20.50					
K_2PtCl_4	PT1	0.397	0.120	0.098	0.91	20.48	0.633	0.194	5639	1.12	2.70
	PT2	0.674	0.382	0.099	0.66	41.46					

[a] Relative occupancy. [b] Isotropic temperature factor. [c] Centric R-factor (R-Cullis): $\sum\|F_{PH}|_o \pm |F_P|_o| - |F_H|_c|/\sum\|F_{PH}|_o \pm |F_P|_c|$, where F_P, F_{PH}, and F_H are the native structure factor, the protein + heavy atom structure factor, and the heavy atom structure factor, respectively. [d] Acentric R-factor (R-Kraut): $\sum\|F_{PH}|_o - |F_{PH}|_c|/\sum|F_{PH}|_o$. [e] The phasing power is defined as the mean value of the heavy atom structure factor divided by the residual lack-of-closure error. [f] Maximum resolution.

surface-exposed loop, and to incorporate the complete amino acid sequence.

Structure Refinement. The data set used for final refinement was collected a second time, after a native crystal was soaked overnight in 10 mM OMP. The final R-merge on intensities was 11.7% (for 7610 reflections between 28.0- and 2.6-Å resolution, ~92% complete). This crystal was not completely isomorphous, showing 2–3% variation in the unit cell parameters ($a = b = 47.2$ Å, $c = 216.7$ Å). Rigid body refinement was then used to adjust the MIR model to the new data set. This model was then refined using molecular dynamics and energy minimization (Brünger et al., 1987). In the first step of the computer-based refinement, the simulated annealing procedure, "slow cooling" (as described in the X-PLOR manual; Brünger, 1992), was used without modification. Electron density maps [both ($2|F_o| - |F_c|$) and ($|F_o| - |F_c|$)] calculated using the refined atomic coordinates were of high quality, allowing us to improve the model and to include the bound OMP and water molecules. In the last stage, the least-squares positional and temperature factor refinement as implemented in TNT (Tronrud et al., 1988) was used, with reflections between 15.0- and 2.6-Å resolution. The current refined model includes 1617 of the 1660 non-hydrogen protein atoms, 77 solvent molecules, and the bound substrate OMP. The refined structure has a conventional R-factor of 17.6%, using data between 15.0 and 2.6 Å, with $F > 1\sigma$. The rms deviations from ideality for the bond lengths and bond angles are 0.02 Å and 2.5°, respectively.

RESULTS AND DISCUSSION

Description of the Structure. Recombinant OPRTase from *S. typhimurium* is a single-domain enzyme of about 51 × 33 × 27 Å, shown as a ribbon diagram in Figure 1. Three substructures can be identified in the protein: (1) a core α/β structure composed of five parallel β-strands (B1–B5) surrounded by three α-helices (A3–A5); (2) an N-terminal region which partially covers the C-terminal side of the core; and (3) a C-terminal region consisting of two antiparallel α-helices. The N-terminal substructure (Met 1–Ser 60) is formed by two antiparallel α-helices (A1, Met 1–Leu 13, and A2, Asn 41–Val 58) connected by three short antiparallel β-strands. The β-strands are arranged such that they partly envelop the C-terminal side of β-strands B-3 and B-5, resembling a "hood" over the core. A five-residue turn (Gly 61–Asp 65) joins this N-terminal hood to the α/β protein core. The topology of the core (residues Leu 66–Ile 181) is similar to the dinucleotide binding fold of many dehydrogenases (Branden & Tooze, 1991): it contains a twisted parallel β-sheet in the middle, surrounded by α-helices. The core can be divided into two similar halves, which wind in a right-handed fashion from the inside to the outside of the protein. The first half contains two parallel β-strands (B-1, Leu 66–Ala 71, and B-2, Lys 91–Glu 101) and one α-helix (A-3, Gly 74–Glu 87) and is connected to the second half by a long loop, Ala 102–Arg 119.

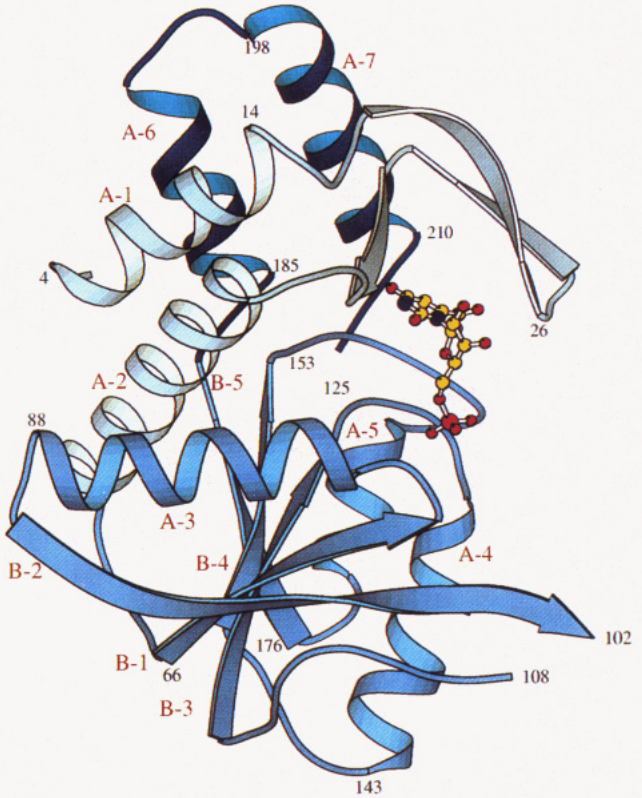


FIGURE 1: Ribbon diagram of the three-dimensional structure of *S. typhimurium* orotate phosphoribosyltransferase. The N-terminal hood domain (light blue), including helices A-1 and A-2, extends across the upper portion of the figure. The C-terminal helices A-6 and A-7 (dark blue) are seen behind the hood structure. The α/β nucleotide binding fold (medium blue) occupies the lower portion of the figure. The bound OMP substrate is shown as a ball-and-stick model. Figures 1–3 and 5 were prepared with the molecular graphics program MOLSCRIPT (Kraulis, 1991).

Five of the 19 residues of this surface-exposed loop (residues 103–107) are located in a region of the map that showed very weak electron density, and a satisfactory fitting was not achieved. These five residues are omitted from the current structure. The second half of this α/β structure is formed of the remaining three parallel β-strands (B-3, Val 120–Asp 125, B-4, Leu 147–Ile 152, and B-5, Cys 176–Ile 181) and two α-helices (A-4, Arg 134–Gln 141, and A-5, Ser 165–Asp 173). Ile 182 through Leu 184 connect the α/β structure to the C-terminal portion of the protein. This region contains two antiparallel α-helices (A-6, Lys 185–Glu 192, and A-7, Met 197–Glu 210), connected by a four-residue loop (Glu 193–Asp 196). A-6 and A-7 constitute the last two helices of a five-helix (A-3, A-2, A-1, A-6, and A-7, as seen in Figure 1) layer that surrounds the C-terminal portion of the β-sheet and the hood. The location of these last two helices suggests that they serve to stabilize the conformation of the hood; A-1 and A-6 run perpendicular to one another, and favorable van der Waals interactions exist between Met 1 and Ile 9 of A-1

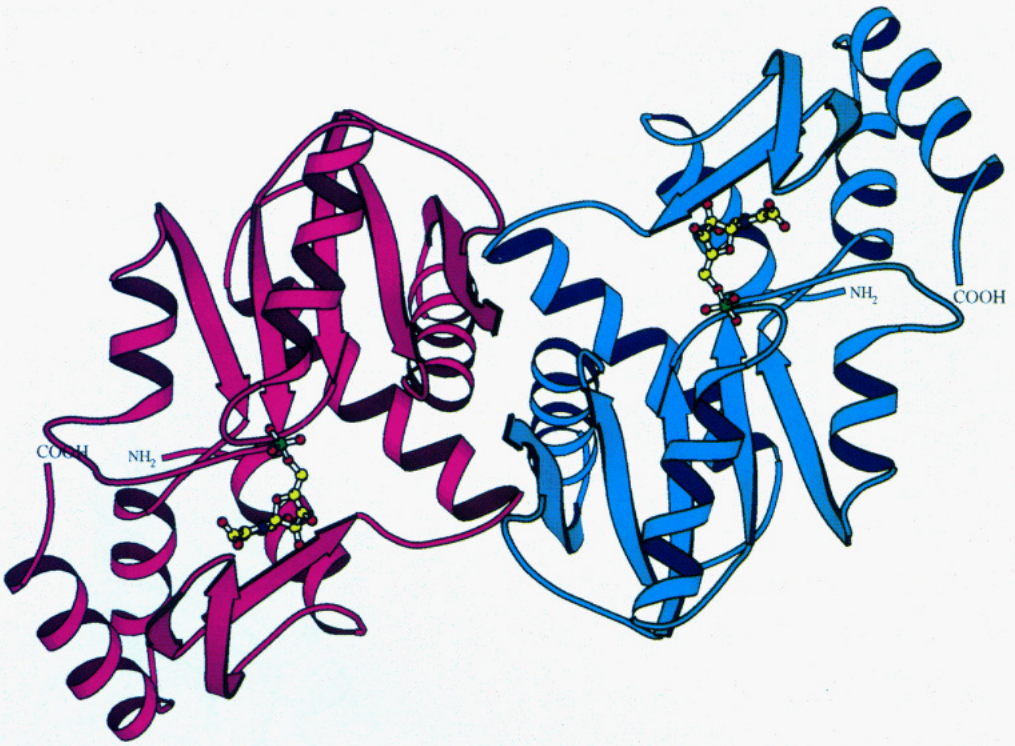


FIGURE 2: Ribbon diagram of the OPRTase dimer, as generated in the $P4_12_12$ space group. The view is orthogonal to that in Figure 1. OMP is seen from the solvent-accessible side of the active site.

and Leu 187 and Tyr 190 of A-6. In addition, the side chains of Gln 5 and Arg 6 from A-1 and Asp 186 and Glu 193 from A-6, respectively, are within hydrogen bond distance. The C-terminal end of A-7 (residues 200–207) is packed against residues 29–33 of the hood, and again, van der Waals interactions between these regions may help the positioning of the hood.

The OPRTase Dimer Is Formed on a Perfect 2-fold Axis. Gel-filtration studies have demonstrated that bacterial OPRTase exists as a dimer in solution (Bhatia et al., 1990). Analysis of crystal packing of the OPRTase monomer in the $P4_12_12$ space group revealed the presence of this dimer, with the two subunits related by a 2-fold symmetry axis, corresponding to a crystallographic axis (Figure 2). The formation of the dimer buries about 700 Å^2 of the total 9600-Å^2 surface area of a monomer. The interface consists of two regions of the protein: (i) the helix A-3 (Gly 74–His 88), the loop, and the first six residues of the following β-strand (Lys 73–Tyr 95); and (ii) the four residues (Asn 41–Asp 45) at the beginning of α-helix A-2. Helix A-3 and its symmetry mate run parallel, with an interhelix angle of about −60° [class II helix–helix interaction (Richardson, 1981)]. Of the 27 residues from each monomer that contribute atoms to the dimer interface, 4 have small polar side chains, 11 have hydrophobic side chains, 9 are ionizable, and 1 is aromatic. The remaining 2 residues are glycines. In the midsection of the interface zone, 7 residues of one subunit are within van der Waals contact distance (<4.5 Å) of atoms from the other subunit. The absence of residues with large or bulky side chains in A-3 allows for the tight packing of the two helices. The dimer is also stabilized by seven hydrogen bonds involving side-chain and main-chain atoms present near the edge of the interface. Although residues 99–102 do not contribute to the interactions between subunits, they are located near the interface, and their side chains extend into the active site of the adjacent protein.

The Active Site of OPRTase. The active site of OPRTase appears as a broad irregular cavity in the enzyme surface, located on the C-terminal side of the β-sheet. This location was defined according to three criteria: (i) One is the presence of the bound substrate within the cavity. OMP is bound in its *syn* conformation, suggested to be the low-energy state for this 6-substituted pyrimidine (Davies et al., 1985). The ribose is in the 2′-*endo* conformation, as observed for crystalline UMP (Seshadri et al., 1980). (ii) Another is analogy to other nucleotide binding proteins of similar structure, where the active site is universally found on the C-terminal side of the β-sheet (Branden & Tooze, 1991). (iii) Finally, the C-terminal six residues of the proposed PRPP binding motif (VMLVD-DVITAGT) are found within the bounds of the cavity.

The orotate group of bound OMP is located in a solvent-inaccessible crevice formed by residues of the core and the hood (Figure 3). Surface accessibility calculations (GRASP, A. Nicholls and B. Honig, Columbia University, New York, NY) showed that only 10 Å^2 out of the 250 Å^2 of the orotate total surface are solvent accessible when OMP is bound to OPRTase. The orotate ring is stacked under the benzene ring of Phe 34 (as viewed in Figure 3). This residue is aromatic in all sequenced OPRTases. The average distance between the two aromatic rings is 4.1 Å, indicating strong van der Waals interactions. A network of four hydrogen bonds between orotate and atoms of the protein appears to be important in providing the base specificity of the enzyme. Orotate–N3 forms a hydrogen bond (3.1 Å) to the main-chain carbonyl oxygen of Phe 35; orotate–O4 forms two hydrogen bonds (3.2 and 2.9 Å, respectively) with the amide nitrogen of Phe 35 and the side-chain nitrogen of Arg 156; the orotate carboxylate forms the fourth hydrogen bond (3.0 Å) with the main-chain nitrogen of Lys 26. Base specificity is well documented for the mammalian OPRTase enzyme (Niedzwicki et al., 1984) but is less fully characterized for the yeast and bacterial OPRTases (Flaks, 1963). However, both the mammalian and bacterial enzymes strongly discriminate against pyridine bases and bulky, 5-substituted pyrimidines and against uracil, which is recognized specifically by a

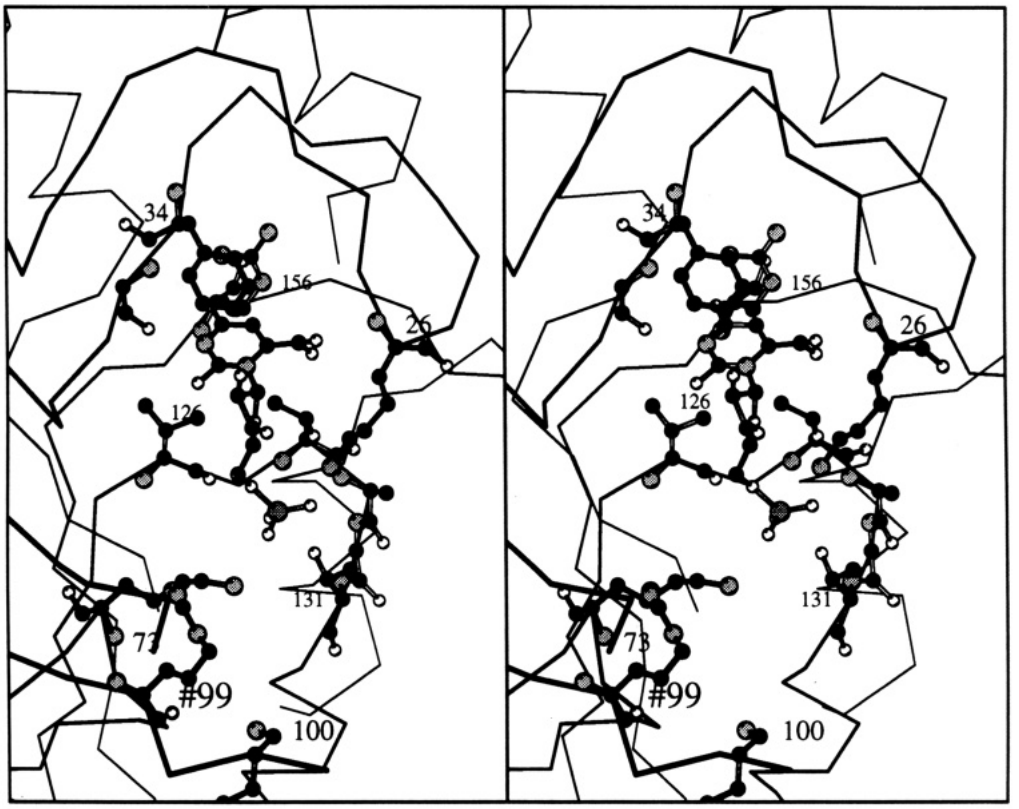


FIGURE 3: Stereo diagram of the OMP binding site in OPRTase. Residues that are involved in the OMP binding are displayed. The residue labeled #99 is from the adjacent subunit of the dimer.

separate enzyme, UPRTase. The hydrogen bond network involving orotate–N3 and –O4 provides interactions that serve to enhance binding of pyrimidine over pyridine nucleotides. Additional substrate discrimination comes from steric constraints imposed by the shape of the crevice that are best visualized in a surface representation of the binding site (Figure 4). The location of the orotate–C5 in the tightly packed crevice, near Arg 156 and the protein backbone, precludes the binding of bulky 5-substituted bases. The specificity of the *S. typhimurium* enzyme produces a 10-fold higher catalytic efficiency for orotate over its methyl ester, arising largely from an increase in K_m, with only a minor reduction in k_{cat} (Bhatia & Grubmeyer, 1993). This result is consistent with steric constraints and the presence of hydrogen bond interactions between the orotate carboxylate and the main-chain amide nitrogen of Lys 26. Uracil, which lacks a C6 substituent, exhibits a k_{cat}/K_m value for *S. typhimurium* OPRTase 10^6-fold lower than orotate and exhibits a steady-state K_m value 100-fold higher than orotate. It seems likely that these large changes are the result of unproductive binding modes of the uracil ring associated with rotational and lateral movement permitted in the absence of the anchoring interactions involving the carboxylate of the base and residues of the crevice.

The O2′ and O3′ hydroxyl groups of the bound OMP are located in a solvent-accessible region of the binding site. The two ribose hydroxyls hydrogen bond only with the ϵ-amino group of Lys 26, a highly conserved residue. This interaction was previously postulated from chemical modification studies (Grubmeyer et al., 1993). The ribosyl ring oxygen and the 5′ oxygen are within hydrogen-bonding distance of the side-chain oxygen of Thr 128, which is also conserved among OPRTases. The 5′-phosphate is located within a pocket formed by residues comprising the C-terminal portion of the PRPP binding motif (Val 126–Thr 131). Thr 128 and Thr 131 interact with the 5′-phosphate through their side-chain oxygens. Additional hydrogen bonds occur between the backbone nitrogens of the six residues of the pocket and the oxygens of the phosphate. The side chain of Lys 73, a residue conserved among bacterial and fungal OPRTases, extends across the active site depression such that the ϵ-amino group is within hydrogen-bonding distance (3.4 Å) of the phosphate moiety. An additional favorable binding interaction may be provided by the positive dipole of α-helix A-4 (Arg 134–Gln 141), which is in close proximity to (~3 Å) and pointed toward the 5′-phosphate. Despite the fact that the nucleoside orotidine, which lacks this phosphate, binds with only a 100-fold higher K_d than OMP, it fails to serve as a substrate for the enzyme (Dessen, 1993; C. Grubmeyer, unpublished material): this finding suggests that the 5′-phosphate may play a critical role in catalysis. There is still uncertainty about the functional role of the two conserved aspartic acids, Asp 124 and Asp 125. In the structure of the OPRTase–OMP complex, the two residues do not interact with groups of the bound substrate. In the electrostatic potential surface representation of the active site (Figure 4) the side chain of the two aspartic acids clearly forms a large, negatively charged area, located below the bound OMP. It is possible that their negative charge may help to stabilize a positive charge developed during the enzyme-catalyzed reaction, as discussed below.

OPRTase binds OMP and PP_i randomly before catalysis of the reverse reaction (Bhatia et al., 1990). The OMP–enzyme complex was crystallized in the absence of PP_i. Inspection of the structure does not reveal any likely PP_i binding sites in the immediate region of the OMP. Two residues, Lys 100 and Lys 103, have been shown by chemical modification studies to be involved in PP_i group binding (Grubmeyer et al., 1993). Lys 100 is located at the end of β-strand B-2, several angstroms from the OMP binding site, as seen in Figure 3. The lack of order in the surface-exposed loop containing Lys

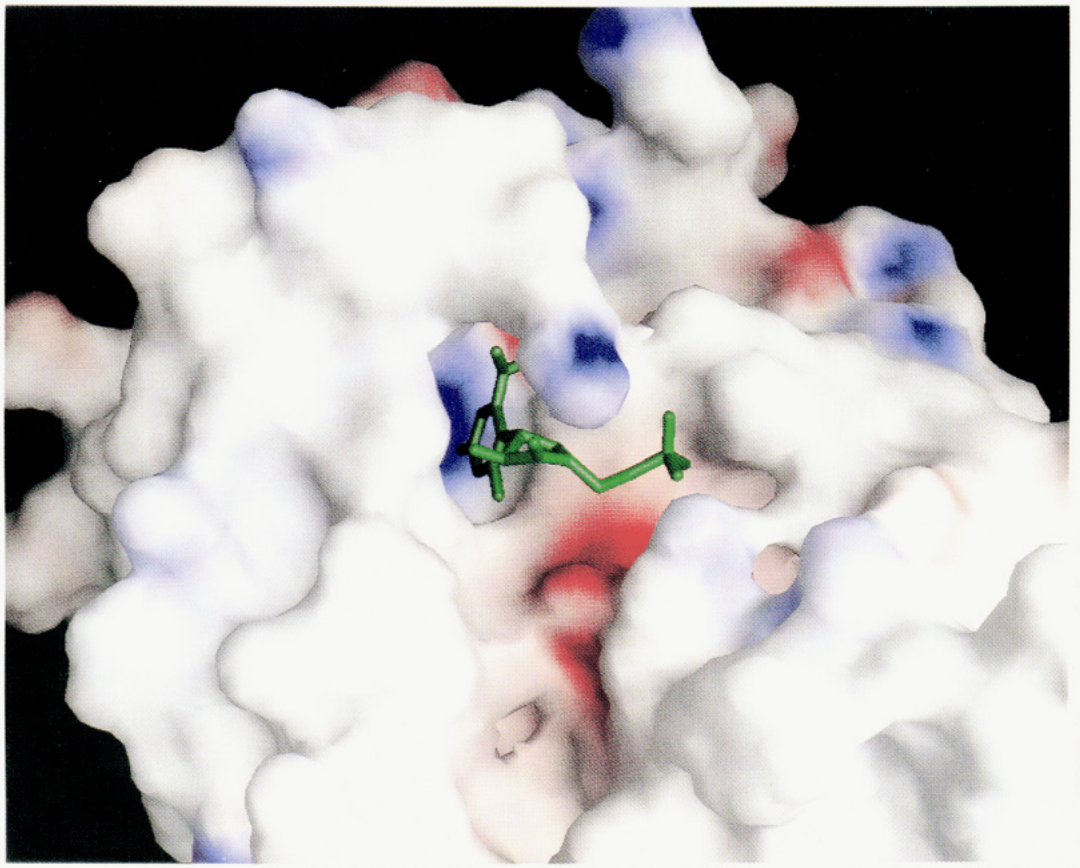


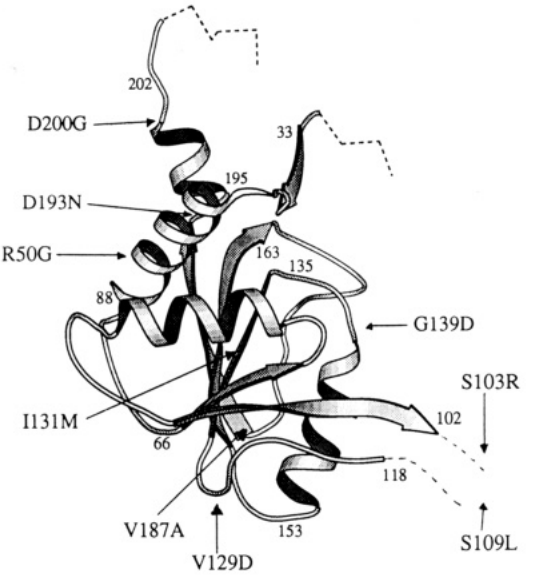
FIGURE 4: Electrostatic potential surface in the region of the OPRTase active site. The figure was generated using the program GRASP (A. Nicholls and B. Honig, Columbia University, New York, NY). Blue and red coloration represent regions of positive and negative potential, respectively. The view is from the same side of the binding site as in Figure 2. The close fit of the orotate group in the crevice that provides base specificity is readily seen. Below the ribose ring is a region of negative potential formed by Asp 124 and Asp 125 of the PRPP binding motif.

103 may be due to the absence of PP_i or PRPP in the crystal. The location of Lys 103 in the region of the active site at the dimer interface could make it available in the active site of either subunit. A third basic residue, Arg 99, which is conserved among OPRTases, is also located at the beginning of this loop. This residue is well ordered, and its side chain extends into the active site of the adjacent subunit of the dimer. It appears likely that the PP_i binding site will be found in the active site cavity near these basic residues. An additional consideration is that the PP_i probably binds to OPRTase as its monomagnesium complex (Bhatia & Grubmeyer, 1993). It is possible that the nearby aspartate residues at positions 124 and 125 form part of a Mg^{2+}–pyrophosphate binding site. This would position PP_i at least 7 Å from the bound OMP. Such a configuration would require movement of the protein or of the ribose 5′-phosphate itself to position the ribosyl C1′ near the PP_i group, a prerequisite for catalysis. A similar situation has been reported in the three-dimensional structure of an analogous complex of purine nucleoside phosphorylase (Ealick et al., 1990), where the distance between a proposed phosphate binding site and the nucleoside appears to be greater than expected for nucleophilic displacement.

Although the chemical mechanisms of PRTases are poorly understood, it is known that the reactions proceed with inversion of stereochemistry at the anomeric carbon (Chelsky & Parsons, 1975), providing evidence against a covalent phosphoribosylated enzyme intermediate. Reports have appeared suggesting that PRTase reactions involve an intermediate or transition state with substantial oxocarbonium character (Goitein et al., 1978; Victor et al., 1979). Work on *S. typhimurium* OPRTase (Bhatia et al., 1990) has favored the latter proposal. The long PP_i–OMP distance, which we propose, may be partially accounted for by the substantial bond lengthening between the ribose group and the incoming nucleophile or leaving base in transition states for (ribosyl) group transfer, as recently described (Horenstein & Schramm, 1993).

The nucleotide binding core of OPRTase is similar to that of the classic nucleotide binding fold (Branden & Tooze, 1991). Features which differentiate the OPRTase core from the typical dinucleotide binding fold include the following: First, the β-sheet is composed of five β-strands rather than six, two in the N-terminal domain and three in the C-terminal domain. OPRTase is more closely related to flavodoxin in this respect (Branden & Tooze, 1991). Second, the commonly observed dinucleotide binding motif GXGXXG is absent in the OPRTase structure, as is the GXXXXGKT sequence commonly found in ATP binding sites. Third, the loops at the C-terminal sides of β-strands B-3 and B-4 are flattened in OPRTase to form a broad platform associated with OMP binding. Fourth, the manner in which OMP binds to the protein is quite distinct from that found for NAD, which would lie in an extended form along the top of the C-terminal side of the β-sheet. Rather, OMP is above and nearly perpendicular to the platform, interacting with residues of the α/β structure mainly through the 5′-phosphate. These differences may reflect the novel chemistry performed by OPRTase, which occurs at the C1′–N1 bond rather than at the base or the phosphate moieties as seen with other enzymes of similar α/β structure.

A Model for the HGPRTase Structure Suggests Generality of the PRTase Fold. Like the NAD-linked dehydrogenases,




```
                10        20        30        40        50        60        70
OPRT            MKPYQRQFIE FALNKQVLKF GEFTLKSGRK SPYFFNAGLF NTGRDLALLG RFYAEALVDS GIEFDLLFGP
Pred:             HHHHHHHH HHHHHHHHEE   EEE         EE      HHHHHHHHHH HHHHHHHHH     EEEEEE
Found:          HHHHHHHHHH HHH         EEEEE EE E EEEEEEE    HHHHHHHHHH HHHHHHHH        EEEEE

HG              ATRSPGVVIS DDEPGYDLDL FCIPNHYAED LERVFIPHGL IMDRTERLAR DVMKEMGGHH IVALCVLKGG
Pred:                EEEEE          E EEE HHHHHH HEEEEE     HHHHHHHHHH HHHHH     E EEEEEEE

                80        90        100                   110       120       130       140
Oprt            AYKGIPIATT TAVALAEHHD KDLPYCFNRK EAK------- ---DHGEGGS LVGSALQGRV MLVDDVITAG TAIRESMEII-
Pred:                HHHHH HHHHHHHH       EE     E                      EEE     EE EEEEE         HHHHHHHH
Found:          E  HHHHHHH HHHHHHH     EEEEEEEEEE E                        E EEEEE          HHHHHHH

Hg:             YKFFADLLDY IKALNRNSDR SIPMTVDFIR LKSYCNDQST GDIKVIGGDD -LSTLTGKNV LIVEDIIDTG KTMQTLLSLVR
Pred:           HHHHHHHHHH HHHH          EEEEEE EEEE         EEEEE             EE EEEEE      HHHHHHHHHH

                150       160       170       180       190       200       210
OPRT            QAHGATLAGV LISLDRQERG RGEISAIQEV ERDYGCKVIS IITLKDLIAY LEEKPDMAEH LAAVRAYREE FGV
Pred:           HH    EEEE EEEE            EEEEE EE    EEEE EEEEHHHHHH HH    HHHH HHHHHHHHH H
Found:          H     EEEE EE              HHHHHH HHH   EEEE EHHHHHHHHH HH    HHHH HHHHHHHHH

HG              QYNPKMVKVA SLLVKRTPRS VGYKPDFVGF EIPDKFVVGY ALDYNEYFRD LNHVCVISET GKAKYKA
Pred:           HH    EEEE EEEEEE            EE E    EEEEE EEEHHHHHH      EEEEE
```

FIGURE 5: Comparison of the location of secondary structure elements for OPRTase and HGPRTase and prediction of the three-dimensional structure for HGPRTase. Lower panel: secondary structure elements for OPRTase and HGPRTase. The alignment is based on secondary structure element prediction (see text for a detailed description). Upper panel: ribbon diagram of the proposed HGPRTase structure, with the location, indicated by arrows, of HGPRTase mutations. Many of the known HGPRTase mutations related to Lesch–Nyhan syndrome or hereditary gout have been identified, and in some cases a preliminary characterization of the enzyme has been attempted. Localization of these mutations on a three-dimensional structure provides some insights about the structural basis of these diseases. The mutations $HPRT_{Toronto}$, R50G, and $HPRT_{Ann\ Arbor}$, I131M (Wilson et al., 1983), $HPRT_{Midland}$, V129D (Davidson et al., 1988), and V187A (Yamada et al., 1991) produce unstable enzyme forms. In our predicted three-dimensional structure, all of the above-mentioned mutations are located in regions of the protein that would appear to be important in stabilizing the α/β core of the enzyme. Other reported clinical mutations that have been shown to affect the activity of the enzyme are located in regions involved in substrate binding. $HPRT_{Munich}$ S103R (Wilson et al., 1983) is at the COOH terminal of the second β-strand of the core nucleotide domain. This area of OPRTase is near the active site where the conformation of the protein might be very sensitive to mutation. $HPRT_{London}$, S109L (Wilson et al., 1983), is in the loop connecting the two halves of the nucleotide fold. The mutation $HPRT_{Tokyo}$, G139D (Fujimori et al.,1991), affects the highly conserved glycine of the phosphate binding loop of the enzymes (TAGT of OPRTase and DTGK of HGPRTase), and the additional atoms and negative charge from this substitution might hinder binding of the ribose 5′-phosphate group. $HPRT_{Kinston}$, D193N (Wilson et al., 1983), at the end of the final β-strand of the core and $HPRT_{Ashville}$, D200G (Davidson et al., 1989), located in the helix following the final β-strand, are sites where the N-terminal domain interacts with the core, and their modification might disrupt this interaction and affect substrate binding.

the members of the PRTase family of enzymes do not show substantial primary sequence similarity. However, the existence of a common PRPP binding motif and similar chemistry suggests that, as with many of the NAD-binding enzymes, the PRTases might have a common fold. This possibility was explored through the construction of a three-dimensional structural model for another PRTase.

HGPRTase activity is responsible for the recycling of purine bases, and its malfunction causes symptoms of gout and mental retardation. Because of the clinical relevance of this enzyme, a three-dimensional model was previously constructed for the human protein on the basis of a template nucleotide binding fold (Wilson et al., 1983). We examined the possibility of building a new model for the 217-residue HGPRTase by first predicting secondary structure using a recently developed homology-assisted computer program (Rost & Sander, 1992). We tested this approach using the primary sequence of OPRTase from *S. typhimurium* as the test sequence and employing all 13 available OPRTase sequences to strengthen the predictive power (Rost & Sander, 1992). The program was successful in identifying most of the secondary structure elements of OPRTase (Figure 5, lower panel), correctly predicting 12 of the 13 regions of α- and β-structure. When applied to HGPRTase, using primary structures for HGPRTases from eight species, the program calculated the secondary structure for the human enzyme shown in Figure 5, lower panel. Alignments of the structural features for the two proteins was done so as to minimize introduction of deletions and was assisted by the positional correspondence of the sequence discussed below. The extent of similarity of

predicted HGPRTase secondary structure with the known OPRTase secondary structure is quite good. Each of the five β-strands and the first two of the three helices of the nucleotide binding core are readily found in the HGPRTase prediction, permitting us to identify this domain with a high degree of certainty. Helix A-5 is not predicted for either the OPRTase or HGPRTase structures, and β-structure is predicted for residues 113–117 of HGPRTase, which do not have homologs in the OPRTase sequence. The short β-strand and the α-helix (residues 33–55) connecting the hood to the nucleotide binding domain of OPRTase are also readily found in the proposed HGPRTase structure, although the β-structure of the hood and its N-terminal α-helix are not predicted. This finding is not unexpected in that the hood is clearly responsible for much of the enzyme's specificity and might be expected to vary among PRTases. These observations support the proposal that a similar pattern of secondary structure will be followed by all PRTases. We next modeled human HGPRTase residues 33–202 using the OPRTase three-dimensional structure and the secondary structure prediction. On examination, many residues of importance in OPRTase have positional homologs in the predicted HGPRTase structure. For example, the PRPP binding motif VLIVEDIIDTGK of HGPRTase is found at the top of B-3, as for OPRTase, and could be proposed to make similar interactions with the nucleotide. The sequence ^{99}RKEAK^{103} of OPRTase, which is at the top of B-2 and whose basic residues are thought to be important in the binding/stabilization of the substrate PP_i, is aligned with a sequence of similarly basic composition, ^{99}IRLKS^{103}. Three residues (Phe 34, Lys 73, and Arg 156) that are involved in the binding of nucleotide in OPRTase also have positional homologs in HGPRTase (Phe 35, Lys 72, and Arg 166), which may play similar roles in this enzyme.

In summary, the secondary structure predictions and model building studies of HGPRTase, and the similar chemistry carried out by PRTases, are convincing evidence that members of this group will maintain a fold analogous to what we report for OPRTase. The three-dimensional structure of OPRTase has a core similar in conformation to that of dinucleotide binding domains of dehydrogenases and adenylate kinase, though functionally OPRTase requires distinct base specificity and carries out catalysis at the C1′–N1 bond. The N-terminal hood region is clearly responsible for the protein's base specificity. However, it is not yet obvious how the chemistry of phosphoribosyl group transfer is accomplished. Basic residues that have been implicated in PP_i and PRPP binding from mutagenesis and chemical modification studies are far removed from the C1′–N1 bond of bound OMP. Substantial conformational changes are therefore necessary in OPRTase during a catalytic cycle. Analysis of the structure of OPRTase in complexes containing PRPP, PP_i, and Mg^{2+} will elucidate additional details about the structural basis of the enzyme mechanism.

ACKNOWLEDGMENT

We thank S. Patel, M. Ahkami, and R. Cohen for their technical assistance, M. Bhatia, A. Dessen, and R. Dorfman for preparing the enzyme, and E. Hehre, J. Blanchard, and V. L. Schramm for many helpful discussions.

REFERENCES

Argos, P., Hanei, M., Wilson, J. M., & Kelley, W. N. (1983) *J. Biol. Chem. 258*, 6450–6457.

Bhatia, M. B., & Grubmeyer, C. (1993) *Arch. Biochem. Biophys. 303*, 321–325.

Bhatia, M. B., Vinitsky, A., & Grubmeyer, C. (1990) *Biochemistry 29*, 10480–10487.

Blundell, T. L., & Johnson, L. N. (1976) *Protein Crystallography* (Harcourt Brace Jovanovich Publishers, Eds.) pp 373–375, Academic Press Ltd., London.

Branden, C., & Tooze, J. (1991) *Introduction to Protein Structure*, pp 141–159, Garland, New York, NY.

Brünger, A. T. (1992) *X-PLOR Version 3.0 Manual: a System for Crystallography and NMR*, Yale University, New Haven, CT.

Brünger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 223*, 458–460.

Chelsky, D., & Parsons, S. (1975) *J. Biol. Chem. 250*, 5669–5673.

Davidson, B. L., Palella, T. D., & Kelley, W. N. (1988) *Gene 68*, 85–91.

Davidson, B. L., Pashmforoush, M., Kelley, W. N., & Palella, T. D. (1989) *J. Biol. Chem. 264*, 520–525.

Davies, D. B., Rajani, P., & Sadikot, H. (1985) *J. Chem. Soc., Perkin Trans. 2*, 279–285.

Dessen, A. (1993) Ph.D. Thesis, New York University, New York.

Ealick, S. E., Rule, S. A., Carter, D. C., Greenhough, T. J., Sudhakar Babu, Y., Cook, W. J., Habash, J., Helliwell, J. R., Stoeckler, J. D., Parks, R. E., Jr., Chen, S.-F., & Bugg, C. E. (1990) *J. Biol. Chem. 265*, 1812–1820.

Flaks, J. G. (1963) *Methods Enzymol. 6*, 136–157.

Fujimori, S., Tagaya, T., Yamaoka, N., Kamatani, N., & Akaoka, I. (1991) in *Purine and Pyrimidine Metabolism in Man* (Harkness, R. A., et al., Eds.) Vol. 8, Part B, pp 101–104, Plenum Press, New York.

Goitein, R. K., Chelsky, D., & Parsons, S. M. (1978) *J. Biol. Chem. 253*, 2963–2971.

Grubmeyer, C., Segura, E., & Dorfman, R. (1993) *J. Biol. Chem. 268*, 20299–20304.

Hershey, H. V., & Taylor, M. (1986) *Gene 43*, 287–293.

Horenstein, B. A., & Schramm, V. L. (1993) *Biochemistry 32*, 7089–7097.

Hove-Jensen, B., Harlow, K. W., King, C. J., & Switzer, R. L. (1986) *J. Biol. Chem. 261*, 6765–6771.

Jones, T. A. (1985) *Methods Enzymol. 115*, 157–171.

Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946–950.

Musick, D. L. (1981) *CRC Crit. Rev. Biochem. 11*, 1–34.

Niedzwicki, J. G., Iltzch, M. H., El Kouni, M. H., & Cha, S. (1984) *Biochem. Pharmacol. 33*, 2383–2395.

Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167–339.

Rost, B., & Sander, C. (1992) *Nature 360*, 540.

Scapin, G., Sacchettini, J. C., Dessen, A., Bhatia, M. B., & Grubmeyer, C. (1993) *J. Mol. Biol. 230*, 1304–1308.

Seshadri, T. P., Viswamitra, M. A., & Kartha, G. (1980) *Acta Crystallogr., Sect. B 36*, 925.

Stout, J. T., & Caskey, C. T. (1989) in *The Metabolic Basis of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S., & Valle, D., Eds.) 6th ed., pp 1007–1028, McGraw-Hill, New York.

Suttle, D. P., Becroft, D. M. O., & Webster, D. R. (1989) in *The Metabolic Basis of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S., & Valle, D., Eds.) 6th ed., pp 1195–1226, McGraw-Hill, New York.

Tronrud, D. E., Ten Eyck, L. F., & Mathews, B. W. (1988) *Acta Crystallogr., Sect. A 43*, 489–501.

Victor, J., Greenderg, L. B., & Sloan, D. L. (1979) *J. Biol. Chem. 254*, 2647–2655.

Wang, Bi-C. (1985) *Methods Enzymol. 115*, 90–112.

Wilson, J. M., Young, A. B., & Kelley, W. N. (1983) *N. Engl. J. Med. 309*, 900–910.

Yamada, Y., Goto, H., & Ogasawara, N. (1991) in *Purine and Pyrimidine Metabolism in Man* (Harkness, R. A., et al., Eds.) Vol. 8, Part B, pp 121–124, Plenum Press, New York.