

Domain Organization of *Salmonella typhimurium* Formylglycinamide Ribonucleotide Amidotransferase Revealed by X-ray Crystallography^{†,‡}

Ruchi Anand,[§] Aaron A. Hoskins,^{||} JoAnne Stubbe,[⊥] and Steven E. Ealick^{*,§}

Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853, and Departments of Chemistry and Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received April 29, 2004; Revised Manuscript Received June 14, 2004

ABSTRACT: Formylglycinamide ribonucleotide amidotransferase (FGAR-AT) catalyzes the ATP-dependent conversion of formylglycinamide ribonucleotide (FGAR) and glutamine to formylglycinamide ribonucleotide (FGAM), ADP, P_i, and glutamate in the fourth step of the purine biosynthetic pathway. In eukaryotes and Gram-negative bacteria, FGAR-AT is encoded by the *purL* gene as a multidomain protein with a molecular mass of about 140 kDa. In Gram-positive bacteria and archaeobacteria FGAR-AT is a complex of three proteins: PurS, PurL, and PurQ. We have determined the structure of FGAR-AT (PurL) from *Salmonella typhimurium* at 1.9 Å resolution using X-ray crystallography. PurL is the last remaining enzyme in the purine biosynthetic pathway to have its structure determined. The structure reveals four domains: an N-terminal domain structurally homologous to a PurS dimer, a linker region, an FGAM synthetase domain homologous to an aminoimidazole ribonucleotide synthetase (PurM) dimer, and a triad glutaminase domain. The domains are intricately linked by interdomain interactions and peptide connectors. The fold common to PurM and the central region of PurL represents a superfamily for which HypE, SelD, and ThiL are predicted to be members. A structural ADP molecule was found bound to a site related to the putative active site by pseudo-2-fold symmetry and two sulfate ions were found at the putative active site. These observations and the structural similarities between PurM and StPurL were used to model the substrates FGAR and ATP in the StPurL active site. A glutamylthioester intermediate was found in the glutaminase domain at Cys1135. The N-terminal (PurS-like) domain is hypothesized to form the putative channel through which ammonia passes from the glutaminase domain to the FGAM synthetase domain.

Inosine 5'-monophosphate (IMP)¹ is synthesized from phosphoribosyl pyrophosphate (PRPP) and small molecule precursors via the purine biosynthetic pathway. The details of the pathway were first worked out by Buchanan and co-workers in the 1950s using chicken and pigeon livers (1). This pathway requires 10 enzymatic steps while consuming four ATP molecules. In bacteria, an additional enzyme, N²-carboxyaminoimidazole ribonucleotide synthetase (PurK),

is required, resulting in 11 enzymatic steps and consuming five ATP molecules (2). Other variations to the basic pathway for vertebrates have been observed. *Escherichia coli* contain both a tetrahydrofolate (THF)-dependent glycinamide ribonucleotide (GAR) transformylase (PurN) and an ATP-dependent formylglycinamide ribonucleotide (FGAR) synthetase (PurT) (3), and in *Methanococcus jannaschii* the final step is catalyzed by a unique IMP cyclohydrolase (PurO) (4).

Representative enzyme structures for all of the steps in the purine biosynthetic pathway have been determined except formylglycinamide amidotransferase (FGAR-AT; PurL), which is the fourth step in this pathway. Four of the structures belong to the ATP-grasp superfamily (5). These are PurT (6), GAR synthetase (PurD) (7), PurK (8), and N-succinocarboxamide-5-aminoimidazole ribonucleotide synthetase (PurC) (9). These four enzymes catalyze the coupling of a carboxylate group and an amino group utilizing an ATP to form an acyl phosphate intermediate. The ATP-grasp fold is found in many proteins including D-alanine/D-alanine ligase (10) and biotin carboxylase (11). The remaining two ATP-dependent enzymes are PurL and aminoimidazole ribonucleotide (AIR) synthetase (PurM). PurL and PurM have both been the subject of extensive studies (2, 12–15).

PurL and PurM catalyze consecutive steps in the pathway and likely both utilize ATP to activate an amide of their

[†] This work was supported by NIH Grants RR15301 (S.E.E.) and GM32191 (J.S.). SEE is indebted to the W. M. Keck Foundation and the Lucille P. Markey Charitable Trust. A.A.H. was supported by an NSF predoctoral fellowship.

[‡] The coordinates of the StPurL structure have been deposited in the Protein Data Bank under accession number 1T3T.

* To whom correspondence should be addressed. Telephone: (607) 255-7961. Fax: (607) 255-1227. E-mail: see3@cornell.edu.

[§] Cornell University.

^{||} Department of Chemistry, Massachusetts Institute of Technology.

[⊥] Departments of Chemistry and Biology, Massachusetts Institute of Technology.

¹ Abbreviations: IMP, inosine monophosphate; FGAR, formylglycinamide ribonucleotide; FGAM, formylglycinamide ribonucleotide; FGAR-AT, formylglycinamide amidotransferase; ATP, adenosine 5'-triphosphate; ADP, adenosine 5'-diphosphate; GAR, glycinamide ribonucleotide; CPS, carbamoyl phosphate synthetase; AS, anthranilate synthase; SelD, selenophosphate synthetase; PRPP, phosphoribosyl pyrophosphate; AIR, aminoimidazole ribonucleotide; SeMet, selenomethionine; GMP, guanosine monophosphate; THF, tetrahydrofolate; St, *Salmonella typhimurium*; Bs, *Bacillus subtilis*; Mt, *Methanobacterium thermoautotrophicum*.

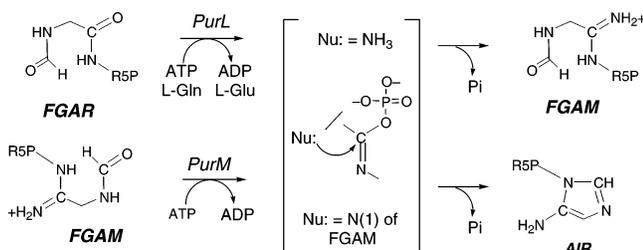


FIGURE 1: Reactions catalyzed by (a) PurL and (b) PurM. Both reactions proceed via a proposed iminophosphate intermediate. R5P is ribose 5'-phosphate.

substrate for nucleophilic attack (Figure 1). The conversion of FGAR to FGAM catalyzed by PurL utilizes ammonia derived from glutamine as its nucleophile. On the basis of a variety of glutamine-requiring amidotransferases (16), the ammonia is generated within a triad glutaminase domain and is likely channeled through a tunnel to the FGAR/ATP binding site. In the case of PurM, the nucleophile attacking the putative iminophosphate intermediate is the N(1) of FGAM. The structure of PurM revealed a novel fold and an ATP binding motif different from the ATP-grasp enzymes (14). Similarity in the functions and catalytic mechanisms of PurL and PurM, as well as sequence similarities identified through iterative BLAST (17) searches led us to propose a new structural superfamily (14) containing PurM and PurL, together with the ATP-dependent enzymes HypE (an NiFe-hydrogenase maturation protein) (18), selenophosphate synthetase (SeID) (19), and thiamine phosphate kinase (ThiL) (20), as members.

We undertook structural studies of PurL to further investigate the relationship between PurL and PurM, to understand the ATP binding domain of the PurM superfamily and to understand the domain organization within PurL. Two types of PurLs have been detected. The first type consists of a single polypeptide chain of about 1300 amino acids and is found in eukaryotes and Gram-negative bacteria. This type is designated large PurL (lgPurL) and can be divided into an N-terminal domain of unknown function, an FGAM synthetase domain homologous to a PurM dimer, and a C-terminal glutaminase domain. The second type of PurL consists of about 800 amino acids and is found in Gram-positive bacteria and archaeobacteria. This type is designated small PurL (smPurL) and requires two additional gene products, a glutaminase, PurQ, and an additional protein, PurS, of unknown function (21). smPurL is homologous to the central portion of lgPurL and the PurM dimer.

The role of the N-terminal domain of lgPurL and the requirement of PurS for reconstituting smPurL activity created a puzzle because no corresponding PurS domain was found within the lgPurL by sequence alignments. In addition, PurM catalyzes a similar reaction to PurL but is only half the size of smPurL. In this paper, we report the structure of lgPurL from *Salmonella typhimurium* (St). StPurL contains four distinct domains: an N-terminal domain that is structurally homologous to a PurS dimer (22), a linker domain, an FGAM synthetase domain, and a glutaminase domain. The N-terminal and linker domains align the FGAM synthetase and the glutaminase domains and potentially participate in the formation of an ammonia channel. The FGAM synthetase domain has pseudo-2-fold symmetry with one active site and one remnant active site with a bound ADP molecule. The

domains assemble such that the glutaminase domain aligns to the FGAM synthetase domain, thus positioning the ammonia for release.

EXPERIMENTAL PROCEDURES

Expression and Purification of Native PurL. PurL was cloned from *S. typhimurium* genomic DNA with the primers 5'-TAGTAGCATATGGAAATTCTGCGTGGTTTCG-3' and 5'-TAGTAGCTCGAGTTAACCCAAGCTGCTTACGCGC-3' (*NdeI* and *XhoI* sites underlined) using standard PCR techniques. The product was then ligated into the pET28a vector (Novagen) at the *NdeI* and *XhoI* sites, which adds an N-terminal (His)₆ tag. The expression construct was subsequently transformed into *E. coli* B834(DE3) cells (Novagen Inc.), which are auxotrophic for methionine. Cultures containing 1 L of LB media and 35 $\mu\text{g}/\text{mL}$ of kanamycin were inoculated with 1 mL of saturated starter culture and incubated at 37 °C. When the culture reached an OD₆₀₀ of approximately 0.6, the cells were induced with 500 μM isopropyl- β -D-thiogalactopyranoside for 5 h at 30 °C. Cells (4.5 g) were harvested by centrifugation and resuspended in 30 mL of ice cold binding buffer (2 mM imidazole, 50 mM KH₂PO₄, 300 mM NaCl, 1 mM glutamine, 5 mM MgCl₂, pH 7.2). All subsequent protein purification steps were carried out at 4 °C. Cells were lysed by two passes through a French press at 15 000 psi, and cell debris was removed by high-speed centrifugation. The clarified cell extract was mixed with a 3 mL slurry of Ni-nitrilotriacetic acid resin (Qiagen), which was preequilibrated in binding buffer, and gently stirred for 1 h. The resin was centrifuged for 10 min at slow speed, and the supernatant was decanted. The resin was resuspended in wash buffer (5 mM imidazole, 50 mM KH₂PO₄, 300 mM NaCl, 1 mM glutamine, 5 mM MgCl₂, pH 7.2) and loaded into a column. The column was washed with wash buffer until OD₂₈₀ < 0.005. The His-tagged PurL was eluted with 50 mM KH₂PO₄, 500 mM NaCl, 1 mM glutamine, 5 mM MgCl₂, and 100 mM imidazole, pH 7.4. The eluted protein was dialyzed against 1 mM glutamine, 5 mM MgCl₂, and 20 mM HEPES, pH 7.2. The protein was concentrated to 15 mg/mL as determined using the Bradford assay with bovine serum albumin as a standard (23). Purity was verified by running a 12% polyacrylamide gel followed by Coomassie staining (data not shown). Typically 50 mg of protein was obtained from 4.5 g of cells.

Expression and Purification of SeMet-PurL. Production of selenomethionine (SeMet)-incorporated PurL followed the same protocol as above with a few modifications. The cells were grown in 1 L cultures containing M9 media supplemented with 40 $\mu\text{g}/\text{mL}$ L-amino acids (excluding methionine), 1 \times BME vitamin solution (GibcoBRL), 0.4% (w/v) glucose, 2 mM MgSO₄, 25 $\mu\text{g}/\text{mL}$ FeSO₄·7H₂O, 35 $\mu\text{g}/\text{mL}$ kanamycin, 0.1 mM CaCl₂, and 40 $\mu\text{g}/\text{mL}$ L-SeMet. A 1 mL starter culture containing the above medium with L-methionine in place of L-SeMet was grown. Before inoculation into 1 L of media, the starter culture was pelleted and washed in the M9 media to remove all traces of L-methionine. SeMet-StPurL was purified in the presence of 1 mM β -mercaptoethanol to protect against oxidation and typically gave 35 mg of SeMet-PurL per 4.0 g of cells.

Crystallization of PurL. Initial crystallization conditions for native PurL were determined using the sparse matrix

Table 1: Summary of Data Collection and Processing Statistics^a

	SeMet StPurL	native StPurL
resolution (Å)	2.4	1.85
wavelength (Å)	0.979	0.964
space group	<i>P</i> 6 ₅	<i>P</i> 6 ₅
no. of reflns	817 526	747 468
no. of unique reflns	65 930	144 593
redundancy	12.4 (12.5)	5.0 (4.9)
completeness	100.0 (100.0)	95.7 (95.7)
<i>R</i> _{sym} (%) ^b	11.5 (33.2)	10.2 (28.8)
<i>I</i> / σ	6 (2.3)	9.7 (3.1)

^a Values for the outer resolution shell are given in parentheses. ^b $R_{\text{sym}} = \sum_i |I_i - \langle I \rangle| / \sum_i \langle I \rangle$, where $\langle I \rangle$ is the mean intensity of the N reflections with intensities I_i and common indices h, k, l for the native and derivative crystals, respectively.

screens Crystal Screen 1 and Crystal Screen 2 (Hampton Research) using 15 mg/mL of PurL in 25 mM HEPES (pH 7.1), 1 mM glutamine. The crystals were grown at room temperature using the hanging-drop vapor diffusion technique. Drops (6 μ L) containing a 2:1 mixture of protein and reservoir solutions were optimal for crystal growth. The optimized condition was found to be 2.0 M ammonium sulfate for the starting reservoir solution. Diffraction quality crystals grew over a period of 20–25 days. Under these conditions, StPurL crystallizes in the hexagonal space group *P*6₅ with unit cell dimensions of $a = 145.3$ and $c = 140.9$ Å. The asymmetric unit contains one monomer, corresponding to a calculated solvent content of 57%. Crystals of SeMet-StPurL were also grown under the conditions described above.

Data Collection and Processing. A single-wavelength data set was collected on a frozen SeMet-StPurL crystal on beamline F-2 at the Cornell High Energy Synchrotron Source (CHESS). A cryoprotectant solution of 20% glycerol in the mother liquor was used to prevent damage during freezing. An X-ray absorption spectrum in the vicinity of the Se K-absorption edge was determined for the SeMet-PurL crystal by recording X-ray fluorescence as a function of wavelength. Diffraction data were then collected to 2.4 Å resolution at a wavelength of 0.9791 Å, corresponding to the peak in the spectrum. The data were measured in 0.5° oscillation steps with 20 s exposure times using a Quantum-210 CCD detector (Area Detector Systems Corporation) with a crystal to detector distance of 221.7 mm. To minimize systematic errors, Bijvoet pairs were acquired close in time by collecting the data in 5° wedges followed by a wedge having inverse beam geometry ($\phi + 180^\circ$). A total of 90° of data was collected. The data were processed with MOSFLM (24) and scaled using SCALA (25). Final data processing statistics are shown in Table 1.

Single-Wavelength Anomalous Diffraction (SAD) Phasing. The selenium atom positions were determined using direct methods as implemented in the Shake-and-Bake procedure of Hauptman and co-workers (26). The DREAR (27) suite of programs was used to calculate the normalized anomalous differences (ΔE) to 2.4 Å resolution, which were then input into the computer program SnB (28). A total of 1000 random trials were used with the 990 largest E values and 9900 triple phase relationships. For each trial, phases were refined for 66 cycles. The results of the 1000 trials showed a bimodal distribution of the minimal function and indicated that 89 of the random trials had resulted in solutions. Twenty-nine

Table 2: Refinement Statistics

	SeMet StPurL	native StPurL
resolution (Å)	25–2.4	25–1.85
total no. of non-hydrogen atoms	10 551	11 024
no. of protein atoms	9820	9885
no. of ligand atoms	40	60
no. of water atoms	691	1079
no. of reflns in refinement	126 688	127 665
no. of reflns in test set	12 420	10 822
<i>R</i> factor (%) ^a	18.9	18.7
<i>R</i> _{free} (%) ^b	21.5	20.4
rms deviation from ideal geometry		
bonds (Å)	0.007	0.004
angles (deg)	1.54	1.32
average <i>B</i> -factor (Å ²)	21.3	20.2
Ramachandran plot		
most favored region (%)	89.1	
additional allowed region (%)	10.1	
generously allowed region (%)	0.3	
disallowed region (%)	0.5	

^a R factor = $\sum_{hk} | |F_{\text{obsd}}| - k|F_{\text{calcd}}| | / \sum_{hk} |F_{\text{obsd}}|$, where F_{obsd} and F_{calcd} are observed and calculated structure factors, respectively. ^b For R_{free} , the sum is extended over a subset of reflections (8%) that were excluded from all stages of refinement.

of the thirty-three Se atom positions found were correct. The Se atom positions were refined and phases were calculated to 2.4 Å resolution using the CNS suite of programs (29) yielding experimental phases with an overall figure of merit (FOM) of 0.30. The phases were further improved to a final FOM of 0.89 using density modification functions in CNS (29). The resulting map was readily interpretable and showed all of the secondary structural elements.

Model Building and Refinement. The C α trace was built through clear stretches of electron density at 2.4 Å resolution using the computer program O (30) followed by addition of side chains. The 29 SeMet residues, along with several aromatic residues and well-defined isoleucine residues, served as markers and aided in sequence alignment. The first round of tracing allowed the fitting of 1253 out of 1295 residues. Density for residues 27–32, 118–125, 605–612, and 786–793, which were missing in the SAD map, appeared after the first round of refinement. Refinement of the initial model was carried out using the CNS suite of programs. Several rounds of simulated annealing minimization followed by manual map refitting were performed. Water molecules were included in subsequent rounds of refinement based on the criteria that the peak in difference electron density maps was greater than 3 σ and the water molecule formed at least one hydrogen bond with a protein, ligand, or solvent atom. In addition two sulfate molecules, three magnesium ions, an ADP molecule, and a covalently bound glutamylthioester at Cys1135 were added to the model. The loop spanning residues 448–466 is disordered and hence missing from the final model. The final data refinement statistics are presented in Table 2.

ATP and FGAR Modeling. To construct a model of ATP and FGAR in the active site cleft of the FGAM synthetase domain, a starting position of ATP was obtained by using the pseudo-2-fold axis of StPurL to first generate an active site ADP molecule from the auxiliary ADP molecule found in the crystal structure. The position for the γ -phosphate of ATP was then modeled using the crystallographic position of sulfate 2. FGAR was then modeled into the active site by using the crystallographic position of sulfate 1 to model the

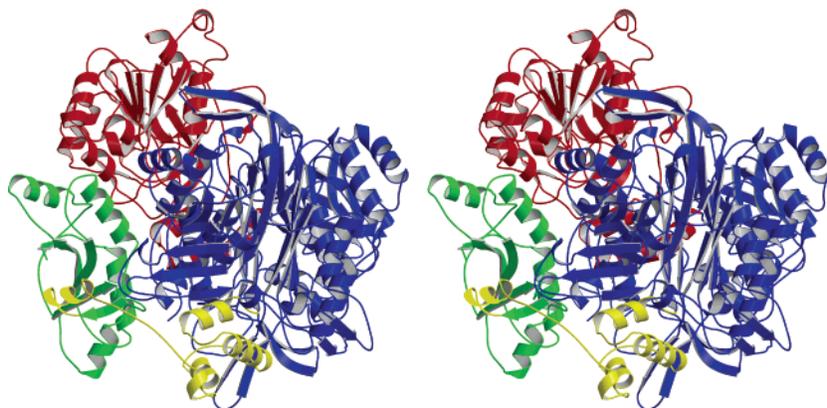


FIGURE 2: Structure of StPurL. A stereoview of the StPurL monomer is color-coded according to domain architecture. The structure shows four major domains: the N-terminal domain (1–140) is shown in green, the linker domain (141–214) is shown in yellow, the FGAM synthetase domain (215–979) is color-coded in blue, and the glutaminase domain (980–1295) is shown in red. The figure was prepared using MOLSCRIPT (51) and RASTER3D (52).

5'-phosphate of FGAR. The docking studies used a GB/SA solvation model (31) and AMBER* force field as implemented in version 7.2 of the program MacroModel (32). Several generalized ligand parameters were adjusted using quantum mechanics calculation at the 6-31G** (geometry optimization) and LMP2/cc-pvtz (-f) (single-point energies) levels.

Molecular mechanics simulations were based on constructing a 12 Å shell around the active site, and then a conformational search was performed using the I-LMOD (33) and MCMM (34) methods to generate trial structures. Residues 216, 219, 259, 295, 296, 297, 298, 292, 299, 318, 502, 776, 778, the ATP, and the FGAR were allowed to move freely, while all other atoms were frozen. The search was performed using a block of at least 7500 trial steps resulting in no new structures within 25 kJ/mol of the global minimum. Structures were minimized using a gradient of 0.05 kJ/(mol·Å).

RESULTS

Protein Expression. StPurL was first purified to homogeneity by French et al. directly from the organism in 1963 (13). His₆-tagged PurL from *S. typhimurium* was purified in 300 mM NaCl, 50 mM potassium phosphate, pH 7.0, and 1 mM glutamine using nickel affinity chromatography and eluted at 100 mM imidazole. The PurL was >95% pure with a specific activity of 4.4 U/mg at 37 °C. This protein was then dialyzed into 25 mM HEPES, 1 mM glutamine, pH 7.1, for crystallization.

Overall Structure of StPurL. StPurL was crystallized in the presence of glutamine, and the crystal structure shows a covalently bound glutamylthioester. Previous studies showed that *E. coli* PurL treated with glutamine in the absence of FGAR and ATP forms a 1:1 protein/glutamine complex that is catalytically incompetent (12). The StPurL structure reported here most likely will require conformational changes within or between domains to achieve a catalytically competent state.

StPurL is a 140 kDa protein and can be divided into four domains (Figure 2). The N-terminal domain consists of residues 1–140 (Figure 3a) and is structurally homologous to the PurS dimers observed for *Methanobacterium thermoautotrophicum* (22) and *Bacillus subtilis* (35). A linker

domain consisting of residues 141–214 (Figure 3b) is made up of a long hydrophilic belt with an extended conformation. The linker is initiated with helix $\alpha 5$, which is associated with the N-terminal domain, and terminates with a three-helix bundle (helices $\alpha 6$ – $\alpha 8$). The central region, consisting of residues 215–980 (Figure 3d), is the main FGAM synthetase domain where FGAR in the presence of ammonia is converted to FGAM. The C-terminal portion, consisting of residues 1016–1295 (Figure 3c), forms the glutaminase domain. The FGAM synthetase domain can be further divided into four subdomains. Subdomains A1 and B1 (Figure 3e) are related to subdomains A2 and B2 (Figure 3f) by 2-fold pseudosymmetry.

The N-Terminal Domain. The N-terminal domain consists of a five-stranded antiparallel β -sheet flanked by four α -helices and has the appearance of a half-barrel (Figure 3a). The β -sheet topology is $\beta 1 \uparrow \beta 2 \downarrow \beta 4 \uparrow \beta 3 \downarrow \beta 5 \uparrow$ (Figure 4a). Strands $\beta 1$ – $\beta 4$ and the four helices of the N-terminal domain display a pseudo-2-fold axis perpendicular to the sheet; however, the sequence identity between the two halves based on structural alignment is only 19%. The central β -strands, $\beta 2$ and $\beta 4$, contain 11 and 12 residues, respectively, and form 12 interstrand hydrogen bonds. Strand $\beta 2$ forms an additional six hydrogen bonds to strand $\beta 1$ and strand $\beta 4$ forms six hydrogen bonds to strand $\beta 3$. The N-terminal domain is structurally homologous to the central four β -strands and four α -helices of PurS (22, 35) but shows no detectable homology with PurS sequences. The four helices of the N-terminal domain pack against the central FGAM synthetase and glutaminase domains, while the concave surface of the sheet faces outward. Contacts with the FGAM synthetase domain bury 1033 Å² of surface area, while contacts with the glutaminase domain bury 1919 Å². The structure and position of the N-terminal domain suggest that it could form an ammonia channel and that it plays a role in aligning the FGAM synthetase and glutaminase domains.

Linker Domain. The linker domain (Figures 3b and 4a) begins with an α -helix ($\alpha 5$) that packs against the concave surface of the N-terminal domain β -sheet. This helix is followed by an 18-residue hydrophilic loop (151–168) that is in a mostly extended conformation and is exposed to the solvent. Although the electron density for the loop backbone was clear, the side chain density was poor, suggesting that

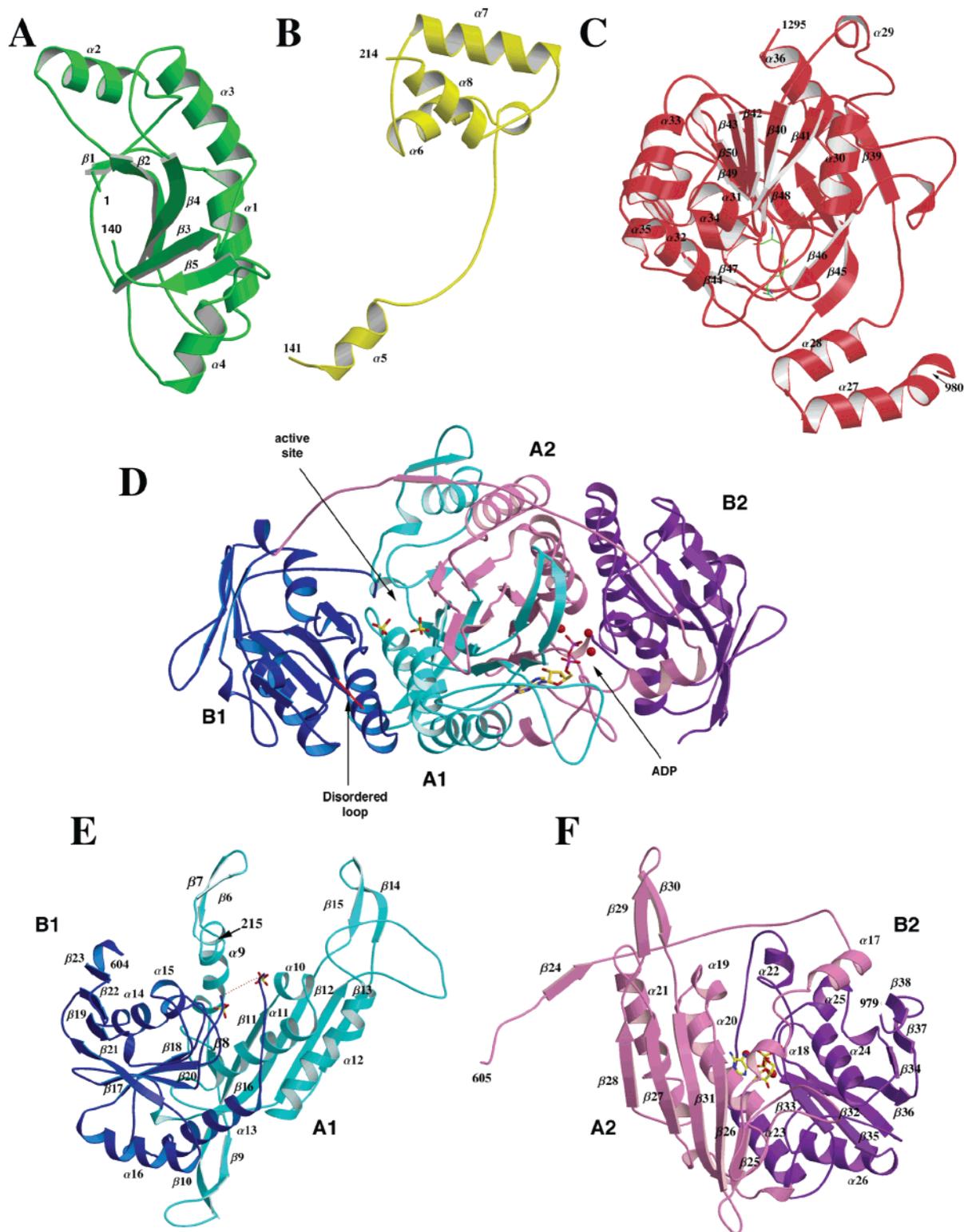


FIGURE 3: Structures of the domains of StPurL: (A) N-terminal domain (1–140); (B) linker domain (151–214); (C) glutaminase domain (1016–1295) with the stick model showing the glutamylthioester; (D) FGAM synthetase domain showing the arrangement of subdomains A1 (cyan), A2 (pink), B1 (blue), and B2 (magenta), the location of the active site and ADP binding site, and a disordered loop (448–466) near the active site, shown as a red line (the pseudo-2-fold axis is perpendicular to the plane of the figure); (E) FGAM synthetase subdomains A1 (225–427) and B1 (428–602) with stick models showing the crystallographic sulfates; (F) FGAM synthetase subdomains A2 (615–810) and B2 (811–1015) with stick models showing the structural ADP molecule and sulfate ions. The pseudo-2-fold axis is vertical in panels E and F. The figure was prepared using MOLSCRIPT (51) and RASTER3D (52).

the loop is flexible. The linker region terminates with an antiparallel three-helix bundle. Most of the linker region contacts, with the exception of the first helix, occur with the FGAM synthetase domain.

FGAM Synthetase Domain. The FGAM synthetase domain (Figure 3d) accounts for about half the molecular mass of the StPurL. The FGAM synthetase domain can be divided into two halves (Figures 3e and 3f) that are related by 2-fold

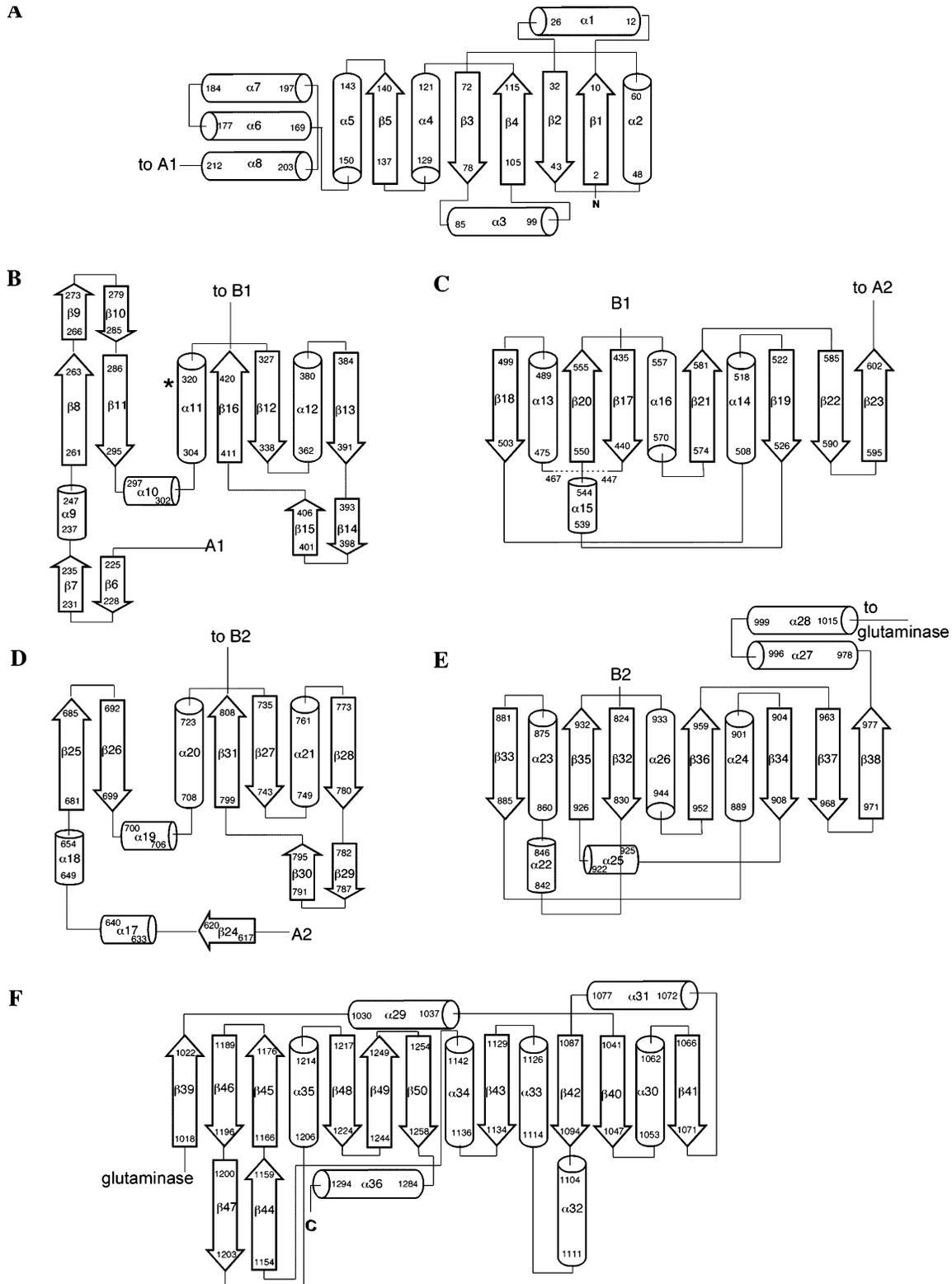


FIGURE 4: Topology diagram for StPurL: (A) N-terminal domain (1–150) and linker domain (151–214); (B) A1 subdomain (225–427) with the * denoting the location of the Dx₄GAXP signature sequence; (C) B1 subdomain (428–602); (D) A2 subdomain (615–810); (E) B2 subdomain (811–1015); (F) C-terminal glutaminase domain (1016–1295). The α -helices and β -strands are labeled. The dotted line in the B1 subdomain corresponds to a disordered loop near the FGAM synthetase active site. The first and the last residue numbers for each secondary structural element are shown.

pseudosymmetry, consistent with a gene duplication and fusion event. Each half of the FGAM synthetase domain can be further divided into A and B subdomains each with an α/β fold giving an overall A1B1A2B2 organization. The A1 subdomain (Figure 4b) and A2 subdomain (Figure 4d) each

consist of a four-stranded antiparallel β -sheet flanked by two long α -helices and one additional short β -strand. The core β -strands are unusually long with 10–11 residues each. The A1 and A2 subdomains join to form a central eight-stranded β -barrel flanked on each side by four α -helices. This

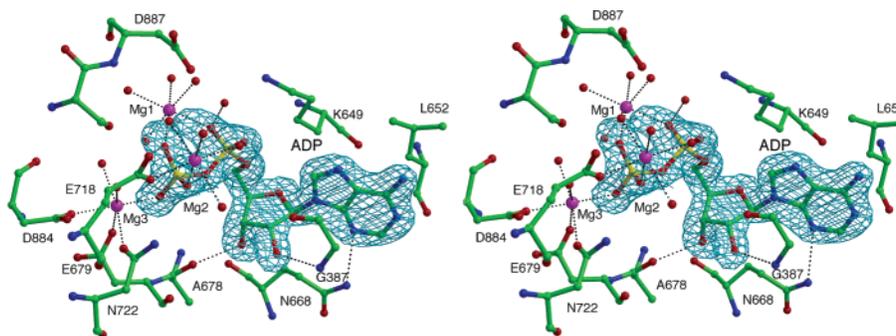


FIGURE 5: Stereoview of the auxiliary ADP site. The ball-and-stick representation is color-coded by atom type with the three magnesium ions colored in magenta. The dashed lines represent metal ligand bonds or hydrogen bonds. A section of the $F_o - F_c$ electron density showing the ADP molecular boundary contoured at 3.56σ is shown in blue. The figure was prepared using BOBSCRIPT (53) and RASTER3D (52).

distinctive barrel forms the core of the FGAM synthetase domain. The helices of the A1 half of the barrel pack against the N-terminal and glutaminase domains, while the helices of the A2 half of the barrel are mostly solvent-exposed.

The B1 (Figure 4c) and B2 (Figure 4e) subdomains show an α/β fold consisting of a seven-stranded mixed β -sheet with topology as $\beta 18\downarrow\beta 20\uparrow\beta 17\downarrow\beta 21\uparrow\beta 19\downarrow\beta 22\downarrow\beta 23\uparrow$ for the B1 subdomain and $\beta 33\downarrow\beta 35\uparrow\beta 32\downarrow\beta 36\uparrow\beta 34\downarrow\beta 37\downarrow\beta 38\uparrow$ for the B2 subdomain. The α -helices and β -strands are shorter than those of the A1 subdomain. The B1 and B2 subdomains flank the barrel on opposite sides. The B1 subdomain is near the N-terminal domain, the glutaminase domain, and the three-helix bundle of the linker domain. The B2 subdomain is largely solvent-exposed except for its contacts with the A2 subdomain. The B1 and A2 subdomains are joined by a long extended linker containing 27 residues (604–630) and terminated by helix $\alpha 17$. Both the main chain and side chain density were clear in the electron density map. The linker traverses the entire link of the FGAM synthetase domain and is anchored near its midpoint (612–616; $\beta 24$) by four hydrogen bonds to strand $\beta 7$ from the A1 subdomain. Strands $\beta 6$, $\beta 7$, and $\beta 24$ form a small antiparallel β -sheet. Helix $\alpha 17$ at the C-terminal end of the linker inserts into a cleft between the A2 and B2 subdomains.

A DALI (36) search using the FGAM synthetase domain identified PurM (14) with a Z-score of 7.8 as the only known structural relative. PurM is a homodimer, and the PurM monomer aligns with either the N-terminal half (subdomains A1 and B1) or C-terminal half (subdomains A2 and B2) of StPurL. A superposition of the FGAM synthetase domain onto the PurM dimer results in an rms deviation of 3.4 Å for 666 C α positions.

Auxiliary ADP Binding Site. Although no ADP was added during purification or crystallization, a deeply buried ADP molecule with three magnesium ions and extensive hydrogen bonding was found in a cleft formed by the central barrel (A1 and A2 subdomains) and the B2 subdomain of the FGAM synthetase domain (Figures 3d and 5). The entrance to the ADP binding cleft is blocked by helix $\alpha 17$ at the end of the linker that connects the B1 and A2 subdomains. A conserved Lys649 located at the C-terminus of this helix hydrogen bonds to the ADP α -phosphate via a water molecule. The ADP binding site includes several highly conserved residues (Lys649, Glu718, Asn722, Asp884, and Asp887) that make hydrogen bonds to ADP, either directly or through water molecules, or serve as ligands for the

magnesium ions. The ADP binding site also contains 11 well-ordered water molecules, 10 of which serve as ligands to a magnesium ion.

The adenine base of ADP is located in a mostly hydrophobic pocket formed by residues Val333, Phe335, Met366, Pro370, Phe389, Phe651, and Leu652. The adenine base forms two hydrogen bonds. The N3 nitrogen atom accepts a hydrogen bond from the Gln668 side chain and the N1 nitrogen accepts a hydrogen bond from a well-ordered water molecule, which in turn hydrogen bonds to the backbone carbonyl of Leu652. The ribose sugar in ADP adopts a C2'-endo pucker. The sugar O2' oxygen atom hydrogen bonds to the backbone amide nitrogen atom of Gly387 and the O3' oxygen atom hydrogen bonds to the carbonyl oxygen atom of Ala678. The α -phosphate provides a ligand to Mg1 and hydrogen bonds to a water molecule, which in turn hydrogen bonds to Asp884. Each oxygen atom of the β -phosphate provides a ligand for a different magnesium ion. One of the β -phosphate oxygen atoms hydrogen bonds to the side chain of Asn722.

Each magnesium ion is octahedrally coordinated (Figure 5). Mg1 bridges the α - and β -phosphates and the coordination sphere is completed by four water molecules, which in turn hydrogen bond to main chain atoms of Val646, Thr645, and Ser886 and side chain atoms of Asp887 and Glu896. Glu718 and the β -phosphate provide ligand atoms for Mg2, and the coordination sphere is completed by four water molecules, which in turn hydrogen bond to main chain atoms of Ile697 and Glu699 and to side chain atoms of Glu699, Glu718, and Asp887. Ligands for Mg3 are provided by oxygen atoms from the side chains of Asp679, Asn722, and Asp884, the β -phosphate group, and two water molecules. The water molecules are in turn hydrogen bonded to main chain atoms of Val677 and Asp884 and the side chains of Glu718, Asn722, His883, and Thr893.

Sulfate Binding Sites. Crystals of StPurL were grown from 2 M ammonium sulfate solutions, which resulted in the binding of two sulfate ions in a cleft formed at the A1 and B1 subdomain interface (Figure 6). One of the sulfate ions (sulfate 1) is hydrogen bonded to the side chain of Lys776, the amide nitrogen atom of Ala505, and a water molecule. The water molecule hydrogen bonds with conserved residues Glu294 and Asp318. The other sulfate ion (sulfate 2) forms hydrogen bonds with the side chains of His216, Thr295, His298, and Ser778 and the amide nitrogen atom of Asn297. The cleft in which the sulfates are bound is lined by

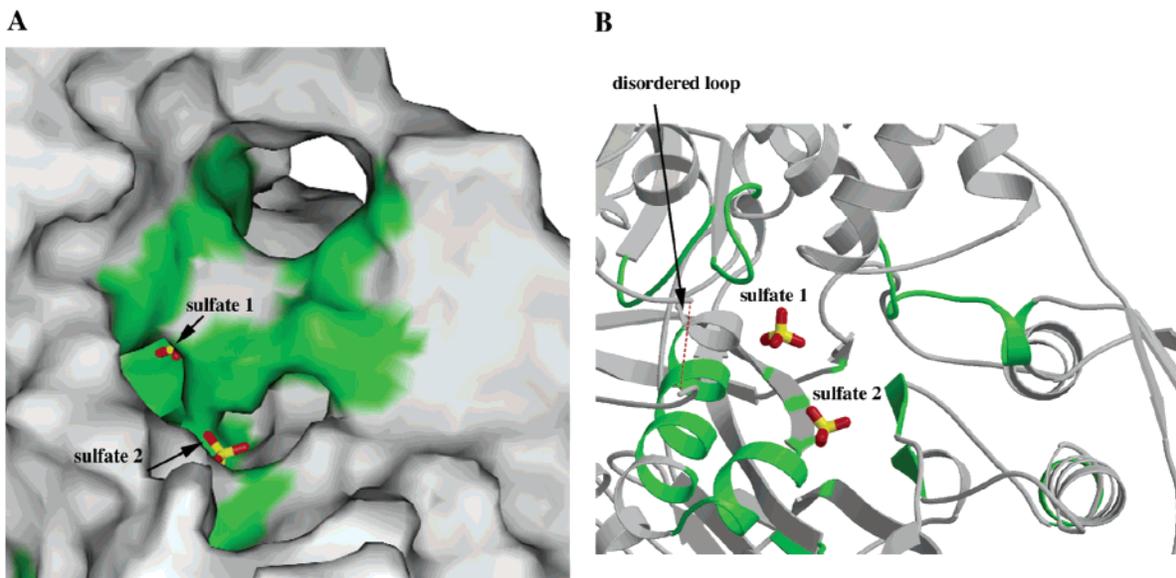


FIGURE 6: Proposed FGAM synthetase active site. Panel A presents a surface representation of the putative active site showing crystallographic sulfate sites in ball-and-stick representation. The conserved residues are highlighted in green. Panel B presents a ribbon diagram in the same orientation as panel A. The dashed line indicates the only disordered loop in StPurL spanning residues 448–466. The figure was prepared using SPOCK (54), MOLSCRIPT (51), and RASTER3D (52).

conserved residues and corresponds to the proposed active site cleft of PurM. Therefore, we hypothesize that the sulfate ions occupy the phosphate binding sites of FGAR or ATP.

Modeling of Substrates ATP and FGAR. Multiple attempts to soak FGAR and ATP analogues into the StPurL crystals were unsuccessful. Analysis of 24 PurL sequences, the arrangement of IgPurL domains, the structural homology with PurM, and the location of sulfate ions suggested that the FGAM synthetase active site is located in the cleft between the A1 and B1 subdomains. Interestingly, this cleft is related by 2-fold pseudosymmetry to the binding site of the auxiliary ADP molecule. Therefore, we used the 2-fold axis of the FGAM synthetase domain to generate an ADP molecule in the putative FGAM synthetase active site. Several lines of evidence support this as the binding site for the product ADP. The active site ADP molecule sits in a cleft surrounded by conserved residues (His219, Phe222, Glu294, Asp318, Asp502, and Lys776). The ADP molecule has many favorable contacts and no steric clashes. Sulfate 2 (Figure 6) observed in the crystal structure of StPurL is consistent with the predicted position of the ATP γ -phosphate, which allowed us to generate an ATP model by replacing the sulfate ion with phosphate and joining it to the ADP β -phosphate. Finally, the ATP γ -phosphate is positioned near a cavity that is the right size and shape to accommodate a molecule of FGAR. This cavity also contains the second sulfate ion.

The adenine ring of the modeled ATP sits in a hydrophobic pocket lined by residues Phe222 and Phe238. The N6 nitrogen atom hydrogen bonds to Asp777. The three phosphate groups of the ATP molecule are stabilized by interactions with positively charged residues Lys776 and Lys292. Three conserved negatively charged residues, Asp318, Asp294, and Asp502, are positioned to bind magnesium ions. The equivalent residues in the auxiliary ADP binding site, Asn718, Glu722, and Asp 884, provide ligand atoms for Mg3, which would bridge the β - and γ -phosphates of the modeled ATP.

The FGAR was modeled using the crystallographic position of the bound sulfate as the 5'-phosphate site. With the phosphate group anchored, the FGAR molecule was positioned to best fit the cavity. In this model, the hydroxyl oxygen atoms of the ribose are hydrogen bonded to backbone atoms of Thr295 and Asp777. The extended glycinamide moiety is sandwiched in a hydrophilic pocket lined by conserved residues His216, His219, His296, His298, and Pro299. His216 and His296 are positioned near the C2 carbonyl oxygen atom and the N4 nitrogen atom of the glycinamide moiety. FGAR and ATP are ideally positioned for nucleophilic attack with a C2–O2–P angle of 108°. The only disordered loop in the entire StPurL structure (448–466) is positioned to cover the active site after substrate binding (Figure 6b). Figure 7 shows the modeled ATP and FGAR in the conserved cleft. Mutation of either His216 or His296, two conserved residues implicated by the model to be involved in FGAR interactions, resulted in complete loss of activity (unpublished results).

Glutaminase Domain. The glutaminase domain (Figure 3c) consists of a ten-stranded mixed β -sheet flanked on one side by two α -helices and on the other by three α -helices (Figure 4f). A DALI (36) search (Table 3) showed that the StPurL glutaminase domain is most similar to anthranilate synthase (37), HisH from *Thermotoga maritima* (38), and yeast (39), and GMP synthetase (40). The glutaminase domain of StPurL contains the classic catalytic triad Cys1135-His1260-Glu1261. Cys1135 occurs in a tight turn between strand β 43 in the central β -sheet and helix α 34; this structural motif is known as the “nucleophilic elbow” (41). The glutaminase domain makes contacts with both the N-terminal domain and the FGAM synthetase domain. The opening of the active site of the glutaminase domain is directed toward the opening of the FGAM synthetase active site; the two active sites are separated by about 30 Å.

Glutaminase Active Site. A glutamylthioester intermediate was observed at the StPuL active site at Cys1135. The density for the ester intermediate was well defined and not

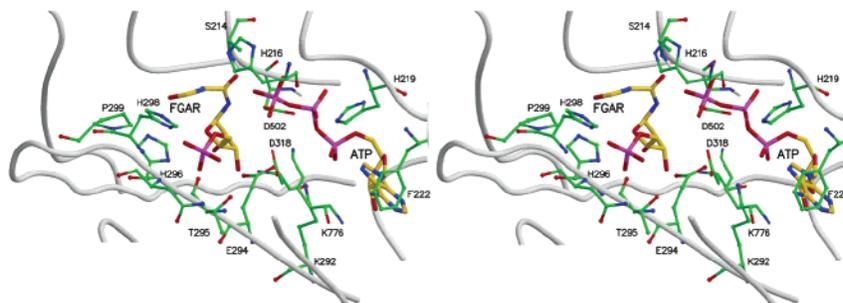


FIGURE 7: Stereoview of modeled FGAR and ATP in the active site. The conserved residues lining the active site cleft along with the modeled ATP and FGAR are shown in a ball-and-stick representation. The figure was prepared using MOLSCRIPT (51) and RASTER3D (52).

Table 3: Structural Similarity of the Glutaminase Domain of StPurL with Homologous Proteins

	rmsd (Å)	no. of C α atoms aligned	identity (%)	PDB code
AS	2.5	175	19	1qdl
HisH	2.8	178	15	1k9v
GMP	2.9	180	18	1gpm
bifunctional HisH/F	3.1	199	14	1jvn
CPS	4.3	194	12	1c3o

unexpected because the crystals were grown in the presence of glutamine. The glutamylthioester and surrounding residues are shown in Figure 8. The binding pocket is mostly hydrophilic and contains a number of ordered water molecules. The terminal amino group of the glutamyl moiety hydrogen bonds with the carbonyl oxygen atoms of Gly1093 and Asp1098. The terminal carboxylate group hydrogen bonds to the side chain of Gln1139 and the backbone amide of Glu1198. The sulfhydryl group of Cys1135 is covalently attached to the CD carbon atom of the glutamyl residue and OE1 oxygen atom of the glutamyl residue is hydrogen bonded to the amide nitrogen atom of Gly1093. His1260 of the catalytic triad and His1196 are located near the thioester group, but neither makes direct interactions with the glutamylthioester. Electron density that could be attributed to either water or ammonia is located 3.2 Å away from the CD1 carbon atom of the glutamyl residue. This molecule is stabilized by hydrogen bonds with the carbonyl group of Ser1195 and the side chain of Arg1263.

Protein Domain Interfaces. The interface between the glutaminase domain and the FGAM synthetase domain is extensive and has a buried surface area of 6280 Å² (Figure 2). Residues from the glutaminase and the FGAM synthetase domain interact across one entire face of both the FGAM synthetase and the glutaminase domains. The interface contains both hydrophobic and electrostatic interactions. Interface residues Phe1094, Arg1263, Glu657, Glu476, Arg658, and Glu1055 are mostly conserved. In the A1 subdomain, residues Leu363, Thr367, Leu371, Asn378, and Glu379, lying on the long extended helix α 12, pack against loop residues Arg1164, Phe1165, Phe1094, Arg996, and Leu995. Residues 980–993 in the linker helix α 27, which connects the FGAM synthetase and glutaminase domains, pack against residues 648–654 in the A2 subdomain helix α 18. Helix α 27 also interacts with the preceding helix α 17, which bridges the B1 and the A2 subdomains. Residues in strands β 44 and β 46 in the glutaminase domain are also involved in interactions with helix α 18 in the A2 subdomain. Helix α 32 in the glutaminase domain further stabilizes the

binding by packing against helices α 12 and α 13 lying in the A1 and B1 subdomains.

In addition to interactions of the core α -helices, the FGAM synthetase domain also interacts with the glutaminase domain via long loops and short two-stranded β -sheet-like structures that protrude out from the main core structure. These extra secondary structural elements are unique to PurL and are missing in PurM, which does not require additional domains to carry out catalysis. Extended loop regions 660–667 and 687–690 in the A2 subdomain interact with the extended loop region 1263–1273 lying between the last strand, β 50, and α 36 in the glutaminase domain. Two other short strands, β 9 and β 10, interact with 1066–1068 on β 41 in the glutaminase domain. The β -turn in residues 1047–1050, preceding helix α 30, tethers the domains together by packing with helices α 12 and α 13 in the A1 and B1 subdomains, respectively.

The N-terminal domain (1–140) spans both the FGAM synthetase domain and the glutaminase domain (Figure 2). The buried surface area between the glutaminase and the N-terminal domain is 1919 Å², whereas that between the FGAM synthetase domain and N-terminal domain is only 1033 Å². There is a noticeable gap of a few angstroms between the N-terminal domain and the rest of the structure, suggesting the possibility of a tighter complex under different conditions. The N-terminal domain interacts with the glutaminase domain mainly via helices α 1 and α 3. The glutaminase domain extensively interacts via residues in the range 1096–1114, which encompasses helix α 32. In addition to hydrophobic interactions, the two domains are held in position by hydrogen bonding interactions of Asn97 with Asn1113 and Gln1096 and also by interaction of Asp93 with Gln1048. Residues 13–18 in helix α 1 also tether the glutaminase domain by packing against residues 995–999, which serve as connecting residues between helices α 27 and α 28 in the glutaminase domain. This two-helix bundle links the glutaminase domain to the FGAM synthetase domain. Helix α 3 of the N-terminal domain interacts with both the glutaminase domain and the FGAM synthetase domain. The FGAM synthetase domain interacts via residues 470–475, which follow the only disordered loop in the entire structure, 448–466, and precede the B1 subdomain helix α 13. This suggests that when substrate binds the disordered loop might become ordered resulting in an inward movement of the N-terminal domain such that helix α 4 caps the opening to the active site cleft.

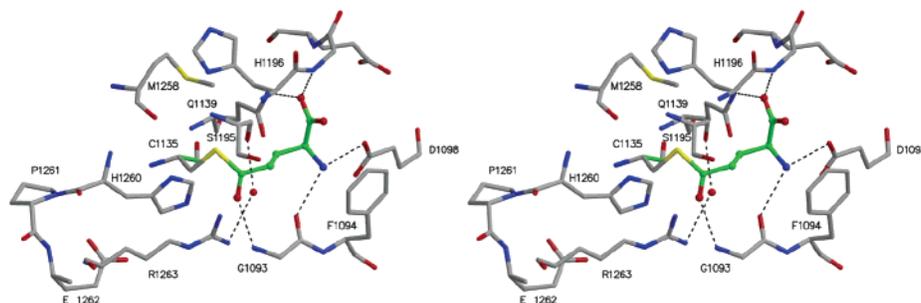


FIGURE 8: Stereodiagram of the glutaminase active site showing the environment around the glutamylthioester intermediate. The glutamyl carbon atoms are highlighted in green to distinguish them from the other amino acid carbon atoms, which are color-coded in gray. The figure was prepared using MOLSCRIPT (51) and RASTER3D (52).

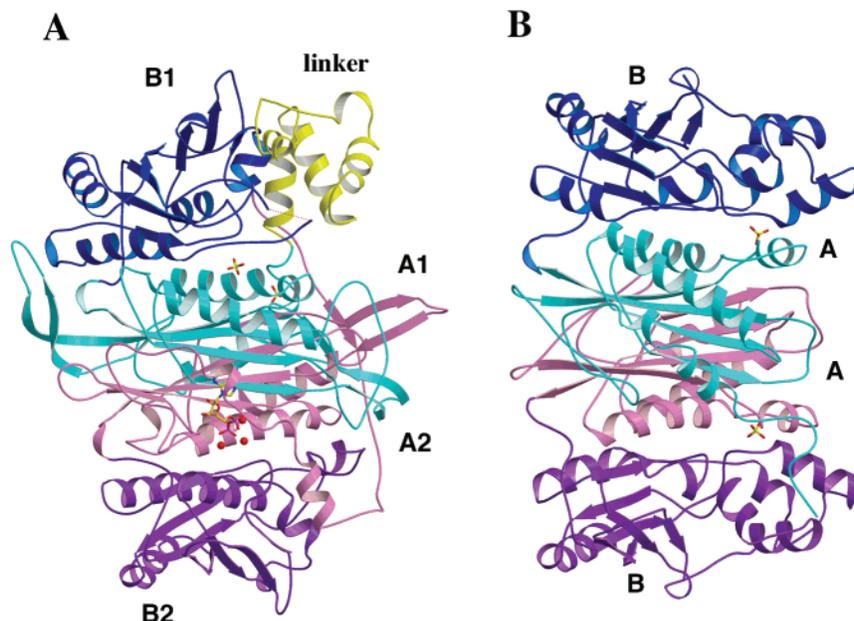


FIGURE 9: Comparison of (A) the StPurL FGAM synthetase domain and (B) the *E. coli* PurM homodimer. The pseudo-2-fold axis of FGAM synthetase and the 2-fold axis of PurM are horizontal. Ligand molecules are shown as stick models. The linker domain of StPurL is shown in yellow. The three helix bundle of the linker domain, which is near the FGAM synthetase active site, corresponds to a structurally homologous three helix bundle in PurM. The structural ADP binding site of FGAM synthetase lacks this structural feature. The color coding is the same as that in Figure 3. The figure was prepared using MOLSCRIPT (51) and RASTER3D (52).

DISCUSSION

The IgPurL N-Terminal Domain Is Homologous to PurS. The N-terminal domain of IgPurL forms a five-stranded half-barrel flanked by four α -helices on the outer periphery of the barrel. The interior of the barrel is mostly hydrophilic and solvent-exposed. This N-terminal domain is structurally homologous to the MtPurS dimer (22) and the BsPurS dimer (35). This suggests that the N-terminal domain is the result of an internal gene duplication; however, there is only 18% sequence identity between the two halves of the N-terminal domain. Structural superposition of the two halves shows an rms deviation of 2.2 Å for 56 C α positions. Superposition of BsPurS and MtPurS on the N-terminal domain using DALI (36) shows rms deviations of 2.8 Å for 129 C α positions and 2.6 Å for 130 C α positions, respectively. The sequence identity based on the structural alignments is 10% for BsPurS and 9% for MtPurS.

The PurM/PurL Superfamily. The FGAM synthetase domain shows pseudo-2-fold symmetry and the topology of each of the halves is similar to that of PurM, which catalyzes the next step of the purine biosynthetic pathway (Figure 1). Therefore the structure of the FGAM synthetase domain

resembles the biologically active PurM homodimer (Figure 9). The A1 subdomain of StPurL has two short extra strands at the start of the domain, spanning residues 225–228 (β 6) and 231–235 (β 7), respectively. These strands make hydrogen bonding interactions with β 24, which is part of the extended loop linking the two halves of the FGAM synthetase domain.

Structural comparison of the A subdomain in PurM with the A1 and A2 subdomains in StPurL reveals that the loops connecting the central β -barrel are substantially longer in StPurL. Loop 339–352 between β 12 and helix α 12 of the A1 subdomain is missing in both the StPurL A2 subdomain and the PurM A subdomain. This loop interacts with strands β 29 and β 30 present only in the StPurL A2 subdomain. At the start of the A1 subdomain, residues 266–285 form a pair of β -strands that interacts with strand β 41 in the glutaminase domain.

The structural similarity between the FGAM synthetase domain of StPurL and the PurM homodimer is striking since there is only 11% sequence identity between the two. Comparison of residues in the two active sites also showed few identities. A comparison of 26 active site residues from

PurM with structurally equivalent residues in StPurL showed only five identities. Lys239 in StPurL corresponds to Lys27 in PurM. This residue is proposed to be near the γ -phosphate of ATP and was shown by labeling studies to be important for catalysis (2). Ala261 in StPurL is predicted to form a hydrophobic pocket for the adenine base and corresponds to Ala47 in PurM. Asp318 in StPurL (Asp94 in PurM) provides a magnesium ion ligand and is part of the Dx₄GxP signature sequence of the PurM/PurL superfamily. Residues Gly447 and Ser469 in StPurL are part of the FGAR/FGAM binding site and align structurally with residues Gly193 and Ser195, respectively, in PurM. In the case of StPurL these two residues are joined by a disordered loop (448–466). The remaining 21 residues show very little similarity suggesting that the two active sites are highly divergent. One exception is Glu294 (Asp65 in PurM), which is predicted to provide a magnesium ion ligand in each case.

Both PurM and StPurL catalyze the ATP-dependent conversion of an amide to an amidine, the principal difference being that the nitrogen atom for PurL is derived from ammonia and for PurM from the N(1) atom of FGAM (Figure 1). The structural similarities between PurL and PurM, the similar mechanistic strategies, with similar proposed iminophosphate intermediates, and their juxtaposition in the purine biosynthetic pathway suggest that these two enzymes are members of a superfamily.

Iterative BLAST searches identified three other enzymes, HypE, SelD, and ThiL, as possible members of the PurM/PurL superfamily (14). All five enzymes utilize a common structural motif for binding ATP. The common signature sequence for ATP binding in these five enzymes is Dx₄GxP. In the inactive half of StPurL, this position corresponds to Asn722, which is a ligand for the magnesium ion associated with the auxiliary ADP molecule.

HypE is required for the biosynthesis of NiFe-hydrogenase cyanide ligands and catalyzes the ATP-dependent dehydration of an *S*-carbamoyl moiety to give the enzyme thiocyanate, which in turn donates cyanide to the iron atom (18). The proposed iminophosphate intermediate for HypE is similar to those of PurM and PurL.

SelD (selenophosphate synthetase) catalyzes the ATP-dependent conversion of selenide to selenophosphate, the selenium donor for the biosynthesis of selenium-dependent enzymes and seleno-tRNAs (19, 42). In the case of SelD, isotopic labeling and kinetic studies show that in the absence of selenide, there is an initial reaction of the enzyme with the γ -phosphoryl group of ATP to form an enzyme phosphoryl intermediate and ADP (43). Because SelD has the same ATP binding signature as the other three members of the family, it is possible that the phosphorylated intermediate resembles an iminophosphate-like intermediate, which could then undergo a nucleophilic attack by the selenide moiety resulting in the formation of selenophosphate. This phosphorylated enzyme intermediate could potentially form at either a backbone amide nitrogen atom or the amide group of an asparagine or glutamine residue. However, because the enzyme phosphorylated intermediate has only been observed in the absence of selenide and at a high concentration of ATP, it is also possible that the enzyme was nonspecifically phosphorylated and that the phosphoryl-enzyme intermediate is an artifact.

The last superfamily member thus far identified, ThiL (thiamine phosphate kinase), catalyzes the ATP-dependent phosphorylation of thiamin phosphate to thiamin pyrophosphate, the active form of vitamin B₁ (20). Because iminophosphate intermediates have been proposed for PurL, PurM, HypE, and SelD, it is possible that the phosphorylation reaction in the case of ThiL may follow similar chemistry as the other members of the superfamily and hence may also proceed via an initial enzyme phosphorylation.

Role of the Auxiliary ADP Molecule. ADP was not added during crystallization or purification of StPurL, so we were surprised to discover a bound ADP deeply buried between the A2 and B2 subdomains, which have no known catalytic function. This ADP molecule can only be removed by protein denaturation and cannot be exchanged with [³H]-ADP or be removed by extensive dialysis against 10 mM EDTA (unpublished results). The ADP binding site is related by pseudo-2-fold symmetry to the proposed FGAM synthetase active site, suggesting that after gene duplication the catalytic activity was lost during evolution but that the ADP binding site was retained. Given the number of contacts, either directly or through water, between ADP and the A2 and B2 subdomains and the blocking of the active site by helix α 17, it is likely that the ADP molecule is structurally important. Helix α 17 is also at the termination of the long linker that connects the B1 and A2 subdomains; therefore, positioning of this helix by the ADP molecule could indirectly affect the FGAM synthetase active site. The importance of the ADP binding site extends to the PurSLQ complex found in some organisms. Several ADP binding residues are conserved between the IgPurL and smPurL families, and recent biochemical studies on *B. subtilis* PurSLQ have shown that ADP is required to reconstitute the complex (44). In an earlier study, an ADP molecule of unknown function was found bound to trimethylamine dehydrogenase (45). Trimethylamine dehydrogenase is a three-domain iron–sulfur-containing protein that catalyzes the oxidative demethylation of trimethylamine to yield dimethylamine and formaldehyde. The ADP site is bound to the C-terminal domain of trimethylamine dehydrogenase and occupies a position equivalent to the ADP moiety of FAD in the FAD-binding domain of glutathione reductase (46). The structural similarity of the C-terminal domain of trimethylamine dehydrogenase to the dinucleotide-binding domain of glutathione reductase has led to the proposal that the C-terminal domain is a vestigial remnant of an ancestral FAD-binding domain. A similar scenario also exists in the case of StPurL, where the C-terminal portion of the FGAM synthetase domain has an ADP molecule of unknown function occupying the predicted ATP binding position.

A structurally bound tryptophan was observed in aminodeoxychorismate (ADC) synthase (47). This enzyme is a part of a heterodimeric complex in which the first step, generation of ammonia, is encoded by the *pabA* glutaminase gene. The ammonia is then channeled to the second enzyme encoded by *pabB*, which catalyzes the conversion of chorismate to aminodeoxychorismate. While ADC synthase has a structural tryptophan bound, a similar enzyme, anthranilate synthase (AS), which provides a branch of chorismate utilization in the biosynthesis of tryptophan, is allosterically regulated by tryptophan (48). The striking similarity between PurL and ADC synthase is that both these enzymes utilize ammonia,

which is channeled from one active site to another, and also bind structural small molecules that play an essential structural role. In the case of ADC synthase and AS, it has been suggested that ADC synthase arose via a gene duplication of an ancestral AS followed by divergent evolution resulting in the loss of regulation by tryptophan but increased binding affinity (48). In the case of PurL, gene duplication followed by fusion resulted in a catalytically active subunit and an inactive subunit containing a structural ADP moiety.

Implications for Protein Evolution. PurM and PurL are members of the same structural superfamily, catalyze similar chemical reactions, utilize similar iminophosphate intermediates, catalyze consecutive steps in the purine biosynthetic pathway, and thus both have FGAM binding sites. Therefore it is reasonable to hypothesize a common ancestor capable of catalyzing both the synthesis of FGAM and the cyclization of FGAM to AIR. The FGAM synthetase activity would be facilitated by a glutaminase to provide the ammonia molecule. Regulation of the ancestral enzyme might have occurred through a glutaminase recruitment protein similar to PurS. A product ADP molecule in the 2-fold related active site would also be positioned to stabilize the glutaminase domain. A separate recruitment protein would prevent the preferential evolution of strong glutaminase/FGAM synthetase interactions, which would compete out the AIR synthetase activity. At some point PurM and PurL diverged, resulting in separate AIR synthetase and FGAR-AT enzymes. In the case of FGAR-AT, the ADP molecule in the second active site corresponds to the auxiliary ADP of StPurL and the regulatory ADP of smPurL. Interestingly, the K_i of ADP for PurM is 11 μM , while the K_m for ATP is 60 μM , suggesting that ADP may play a role in regulating the activity of PurM (15). IgPurL most likely resulted from fusion of the PurS, smPurL, and PurQ genes.

Comparison with Other Glutaminases. A DALI (36) search using the C-terminal portion (981–1295) of IgPurL showed that AS (37), HisH (the glutaminase subunit of imidazole glycerol phosphate synthase) (38, 39), GMP synthetase (40), and CPS (carbamoyl phosphate synthetase) (49) are the closest structural homologues to IgPurL. Like all other members of this family, the glutaminase domain of StPurL has a central ten-stranded mixed β -sheet flanked by α helices and variable length loops on either side. Table 3 compares the StPurL glutaminase domain to five known structures. The key mechanistic feature for this family is a classic catalytic triad motif, Glu-His-Cys, with the active site cysteine residue occurring in a nucleophilic elbow motif (41). A glutamylthioester intermediate is observed in the IgPurL glutaminase active site.

A glutamylthioester intermediate was also observed in the H353N mutant of CPS synthetase (50) and in wild-type AS from *Serratia marcescens* (37). The binding geometry of the thioester intermediate is similar in all three enzymes. However, there are some significant differences in the vicinity of the intermediates. Residues 240–245 in CPS stabilize the N1 nitrogen atom in the ester intermediate via interactions with the backbone oxygen atoms of Gly241 and Gly243 and the amide nitrogen atom of Gly313. These residues correspond to residues 58–63 in *S. marcescens* AS, which have a similar geometry. Both enzymes have a conserved Gly-Pro-Gly, a common oxyanion motif found in class I amidotransferases. However, in the case of StPurL,

this six residue loop is replaced by an 11 residue loop spanning 1093–1102. Due to this insertion, there are conformational differences in this region between the two structures, causing the region around the ester intermediate to be more open. This loop region in StPurL contains a five residue turn involved in interactions with the N-terminal domain. This could provide a mechanism for regulating the glutaminase activity by the N-terminal domain. Since there is a conformational difference introduced in this region, effective stabilization of the N1 nitrogen atom of the ester intermediate occurs by hydrogen bonding interactions of the side chain of Asp1098 lying in this region.

The StPurL glutaminase domain also has a conserved Arg1263, which is replaced by Ala356 in CPS (50) and Ser175 in *S. marcescens* AS (37). Arg1263 is near strong density in the electron density map that is also near the CD1 carbon atom of the thioester intermediate (the glutamine amide carbon that was modified by attack of the sulfhydryl group during the formation of the thioester intermediate). This density was modeled as a water molecule that bridges the NHZ nitrogen of Arg1263 with the OE1 carbonyl oxygen of the ester. The presence of the glutamylthioester intermediate in native StPurL and the location of a water molecule positioned to hydrolyze the thioester intermediate suggests that completion of the reaction requires a conformational change. One possibility is that ammonia has been formed and sequestered. Hydrolysis of the thioester intermediate and release of ammonia could be triggered by conformation changes associated with FGAR and ATP binding. Although the interactions between the N-terminal and glutaminase domains are tight, there is an apparent gap between the N-terminal domain and the FGAM synthetase domain, consistent with such a conformational domain rearrangement.

Examination of the glutaminase active site of StPurL reveals that the environment is more hydrophobic than for other amidotransferases. Residues 1091–1095 form an oxyanion hole, which is a common feature in triad amidotransferases. The negative charge during the hydrolytic reaction is stabilized by hydrogen bonds from the amide nitrogen atoms of Gly1093 and Ser1095. This region precedes a conserved five residue turn that makes hydrophobic packing interactions with the N-terminal domain (residues 1–140).

The Role of Ammonia Channeling. Channeling is a widely accepted mechanism for delivery of an ammonia molecule from the glutaminase domain to a second active site. The ammonia channel of CPS covers a distance of 45 Å (49), while that of imidazole glycerol phosphate synthase covers 25 Å (38, 39). Channel architecture is highly variable among ammonia channeling enzymes and can either be preformed as is the case for CPS, be formed only after substrates are bound as is the case of PRPP glutamine AT, or require the formation of a complex between the glutaminase and a second protein (38, 39). In the case of StPurL the two active sites are separated by about 30 Å and two possible paths exist (Figure 10).

Both paths exit the glutaminase domain through a gate formed by Phe1094 and Phe1165. Upon entering the FGAM synthetase domain, one path passes along the edge of the central β -barrel and is formed in part by the auxiliary ADP molecule. This path is mostly hydrophobic and is lined by Val333, Phe335, Ala374, Met 412, and Ile655. The second

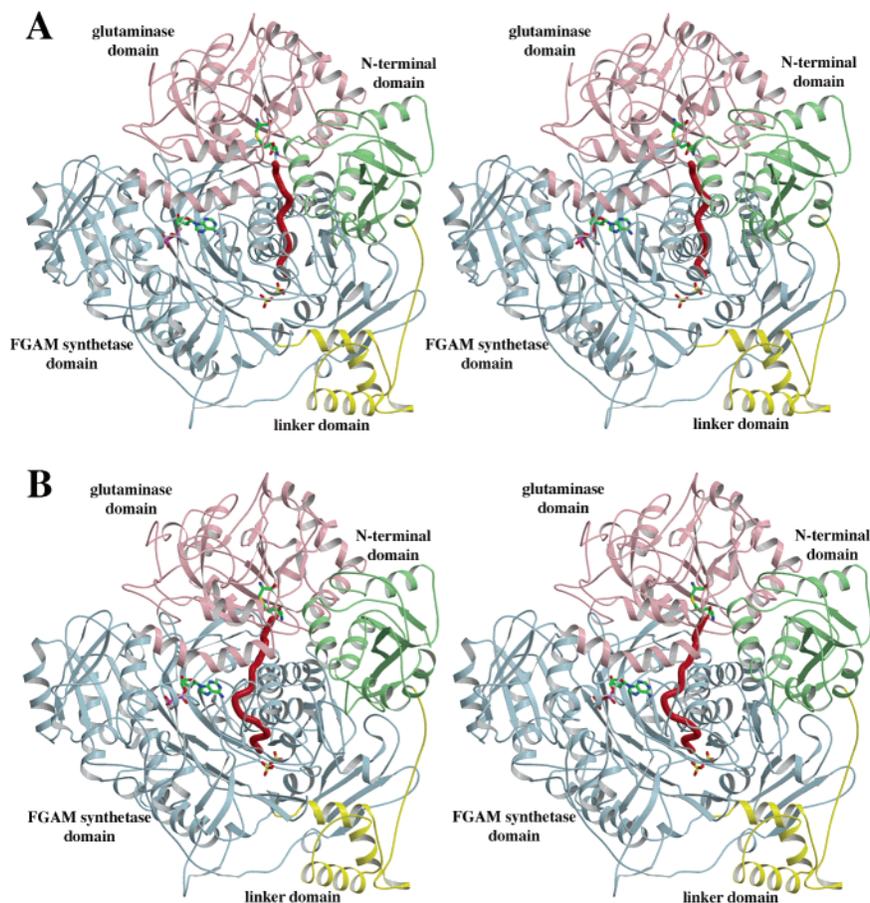


FIGURE 10: Stereoviews of the two putative ammonia channels: (A) favored path passing along the N-terminal domain; (B) alternate path passing near the ADP binding site. Red tubes are used to illustrate the paths of the channels.

possible path is more direct and contains a mixture of hydrophilic and hydrophobic residues. This path is lined by conserved residues Gly313, Ile316, Arg317, and Phe380 and is formed at the interface of the N-terminal and A1 and B1 subdomains of the FGAM synthetase domain. While the current evidence does not distinguish between these two paths, we favor the second, which is more direct, involves more conserved residues, and requires active participation of the StPurL N-terminal domain and, by analogy, the PurS domain of the PurS/smPurL/PurQ complex. Previous studies showed that *E. coli* PurL incubated with glutamine forms a stable 1:1 complex and that the isolated complex is not competent to generate FGAM from FGAR and ATP (12). In the structure of StPurL, glutamine is covalently bonded suggesting a loss of ammonia. Furthermore, because the N-terminal domain, which is proposed to align the FGAM synthetase and glutaminase domains, is loosely packed against the rest of the structure, it is likely that conformational changes in the active protein during catalysis may cause a more apparent channel to form transiently.

Conclusions. With the determination of the X-ray structure of StPurL, an enzyme structure is now available for every step of the purine biosynthetic pathway. The structure of StPurL confirms the existence of a PurM/PurL superfamily, which also includes HypE, SelD, and ThiL, and suggests a common ancestor for PurM and PurL. The structure of StPurL revealed two key residues, His216 and His296, that may play a role in formation of the proposed iminophosphate intermediate and several conserved residues in the ATP

binding site. A glutamylthioester was trapped in the glutaminase binding site, confirming key catalytic residues. The structure of StPurL shows that its N-terminal domain is homologous to a PurS dimer and that the domain architectures of IgPurL and smPurL/PurS/PurQ complexes are analogous (35). The StPurL structure revealed an unexpected structural ADP molecule and suggests a role for ADP in reconstitution of the PurSLQ complex, which has now been experimentally confirmed (44). Finally, the structure revealed a possible ammonia channel between the glutaminase active site and the FGAM synthetase active site and suggests that formation of the channel may be influenced by conformational changes in StPurL that take place during catalysis.

ACKNOWLEDGMENT

We thank Dr. Cynthia Kinsland for preparing the StPurL plasmid and Dr. Eric Bennett for assistance with the modeling studies. We thank Ms. Leslie Kinsland for assistance in the preparation of this manuscript. We thank the Cornell High Energy Synchrotron Source for provision of beam time.

REFERENCES

1. Buchanan, J. M., and Hartman, S. C. (1959) Enzymic reactions in the synthesis of the purines, *Adv. Enzymol. Relat. Areas Mol. Biol.* 21, 199–261.
2. Mueller, E. J., Meyer, E., Rudolph, J., Davisson, V. J., and Stubbe, J. (1994) N⁵-Carboxyaminoimidazole ribonucleotide: evidence for a new intermediate and two new enzymatic activities in the de

- novo purine biosynthetic pathway of *Escherichia coli*, *Biochemistry* 33, 2269–2278.
- Marolewski, A. E., Mattia, K. M., Warren, M. S., and Benkovic, S. J. (1997) Formyl phosphate: a proposed intermediate in the reaction catalyzed by *Escherichia coli* PurT GAR transformylase, *Biochemistry* 36, 6709–6716.
 - Graupner, M., Xu, H., and White, R. H. (2002) New class of IMP cyclohydrolases in *Methanococcus jannaschii*, *J. Bacteriol.* 184, 1471–1473.
 - Galperin, M. Y., and Koonin, E. V. (1997) A diverse superfamily of enzymes with ATP-dependent carboxylate-amine/thiol ligase activity, *Protein Sci.* 6, 2639–2643.
 - Thoden, J. B., Firestone, S., Nixon, A., Benkovic, S. J., and Holden, H. M. (2000) Molecular structure of *Escherichia coli* PurT-encoded glycinamide ribonucleotide transformylase, *Biochemistry* 39, 8791–8802.
 - Wang, W., Kappock, T. J., Stubbe, J., and Ealick, S. E. (1998) X-ray crystal structure of glycinamide ribonucleotide synthetase from *Escherichia coli*, *Biochemistry* 37, 15647–15662.
 - Thoden, J. B., Kappock, T. J., Stubbe, J., and Holden, H. M. (1999) Three-dimensional structure of N⁵-carboxyaminoimidazole ribonucleotide synthetase: a member of the ATP grasp protein superfamily, *Biochemistry* 38, 15480–15492.
 - Levdikov, V. M., Barynin, V. V., Grebenko, A. I., Melik-Adamyanyan, W. R., Lamzin, V. S., and Wilson, K. S. (1998) The structure of SAICAR synthase: an enzyme in the *de novo* pathway of purine nucleotide biosynthesis, *Structure* 6, 363–376.
 - Fan, C., Moews, P. C., Walsh, C. T., and Knox, J. R. (1994) Vancomycin resistance: structure of D-alanine:D-alanine ligase at 2.3 Å resolution, *Science* 266, 439–443.
 - Waldrop, G. L., Rayment, I., and Holden, H. M. (1994) Three-dimensional structure of the biotin carboxylase subunit of acetyl-CoA carboxylase, *Biochemistry* 33, 10249–10256.
 - Schendel, F. J., Mueller, E., Stubbe, J., Shiau, A., and Smith, J. M. (1989) Formylglycinamide ribonucleotide synthetase from *Escherichia coli*: cloning, sequencing, overproduction, isolation, and characterization, *Biochemistry* 28, 2459–2471.
 - Dawid, I. B., French, T. C., and Buchanan, J. M. (1963) Azaserine-reactive sulphydryl group of 2-formamido-N-ribosylacetamide 5'-phosphate: L-glutamine amido-ligase (adenosine diphosphate). II. Degradation of azaserine-C-14-labeled enzyme, *J. Biol. Chem.* 238, 2178–2185.
 - Li, C., Kappock, T. J., Stubbe, J., Weaver, T. M., and Ealick, S. E. (1999) X-ray crystal structure of aminoimidazole ribonucleotide synthetase (PurM), from the *Escherichia coli* purine biosynthetic pathway at 2.5 Å resolution, *Structure* 7, 1155–1166.
 - Schrimsher, J. L., Schendel, F. J., Stubbe, J., and Smith, J. M. (1986) Purification and characterization of aminoimidazole ribonucleotide synthetase from *Escherichia coli*, *Biochemistry* 25, 4366–4371.
 - Rauschel, F. M., Thoden, J. B., and Holden, H. M. (2003) Enzymes with molecular tunnels, *Acc. Chem. Res.* 36, 539–548.
 - Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res.* 25, 3389–3402.
 - Reissmann, S., Hochleitner, E., Wang, H., Paschos, A., Lottspeich, F., Glass, R. S., and Bock, A. (2003) Taming of a poison: biosynthesis of the NiFe-hydrogenase cyanide ligands, *Science* 299, 1067–1070.
 - Veres, Z., Kim, I. Y., Scholz, T. D., and Stadtman, T. C. (1994) Selenophosphate synthetase. Enzyme properties and catalytic reaction, *J. Biol. Chem.* 269, 10597–10603.
 - Begley, T. P., Downs, D. M., Ealick, S. E., McLafferty, F. W., Van Loon, A. P., Taylor, S., Campobasso, N., Chiu, H. J., Kinsland, C., Reddick, J. J., and Xi, J. (1999) Thiamin biosynthesis in prokaryotes, *Arch. Microbiol.* 171, 293–300.
 - Saxild, H. H., and Nygaard, P. (2000) The yexA gene product is required for phosphoribosylformylglycinamide synthetase activity in *Bacillus subtilis*, *Microbiology* 146, 807–814.
 - Batra, R., Christendat, D., Edwards, A., Arrowsmith, C., and Tong, L. (2002) Crystal structure of MTH169, a crucial component of phosphoribosylformylglycinamide synthetase, *Proteins* 49, 285–288.
 - Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem.* 72, 248–254.
 - Leslie, A. G. W. *MOSFLM*, (1997) version 5.40, MRC Laboratory of Molecular Biology, Cambridge, U.K.
 - Evans, P. R. (1993) *SCALA*, version 3.3, MRC Laboratory of Molecular Biology, Cambridge, U.K.
 - Miller, R., DeTitta, G. T., Jones, R., Langs, D. A., Weeks, C. M., and Hauptman, H. A. (1993) On the application of the minimal principle to solve unknown structures, *Science* 259, 1430–1433.
 - Blessing, R. H. (1989) DREDD-data reduction and error analysis for single-crystal diffractometer data, *J. Appl. Crystallogr.* 22, 396–397.
 - Weeks, C. M., and Miller, R. (1999) The design and implementation of *SnB v2.0.*, *J. Appl. Crystallogr.* 32, 120–124.
 - Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D* 54, 905–921.
 - Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for the building of protein models in electron density maps and the location of errors in these models., *Acta Crystallogr. A* 47, 110–119.
 - Qiu, D., Shenkin, P. S., Hollinger, F. P., and Still, W. C. (1997) The GB/SA Continuum Model for Solvation. A Fast Analytical Method for the Calculation of Approximate Born Radii., *J. Phys. Chem. A* 101, 3005–3014.
 - Mohamadi, F., Richards, N. G. J., Guida, W. C., Liskamp, R., Lipton, M., Caufield, C., Chang, G., Hendrickson, T., and Still, W. C. (1990) MacroModel – An Integrated Software System for Modeling Organic and Bioorganic Molecules Using Molecular Mechanics, *J. Comput. Chem.* 11, 460–467.
 - Kolossvary, I., and Guida, W. C. (1999) Low-mode conformational search elucidated: application to C39H80 and flexible docking of 9-deazaguanine inhibitors into PNP, *J. Comput. Chem.* 20, 1671–1684.
 - Chang, G., Guida, W. C., and Still, W. C. (1989) An internal-coordinate Monte Carlo method for searching conformational space, *J. Am. Chem. Soc.* 111, 4379–4386.
 - Anand, R., Hoskins, A. A., Bennett, E. M., Stubbe, J., and Ealick, S. E. (2004) A Model for the *Bacillus subtilis* PurS/PurL/PurQ Formylglycinamide Ribonucleotide Amidotransferase Complex, *Biochemistry* 43, 10343–10352.
 - Holm, L., and Sander, C. (1998) Touring protein fold space with Dali/FSSP, *Nucleic Acids Res.* 26, 316–319.
 - Spraggon, G., Kim, C., Nguyen-Huu, X., Yee, M. C., Yanofsky, C., and Mills, S. E. (2001) The structures of anthranilate synthase of *Serratia marcescens* crystallized in the presence of (i) its substrates, chorismate and glutamine, and a product, glutamate, and (ii) its end-product inhibitor, L-tryptophan, *Proc. Natl. Acad. Sci. U.S.A.* 98, 6021–6026.
 - Douangamath, A., Walker, M., Beismann-Driemeyer, S., Vega-Fernandez, M. C., Sterner, R., and Wilmanns, M. (2002) Structural evidence for ammonia tunneling across the (β)α8 barrel of the imidazole glycerol phosphate synthase holoenzyme complex, *Structure* 10, 185–193.
 - Chaudhuri, B. N., Lange, S. C., Myers, R. S., Chittur, S. V., Davissou, V. J., and Smith, J. L. (2001) Crystal Structure of imidazole glycerol phosphate synthase: a tunnel through a (β)α8 barrel joins two active sites, *Structure* 9, 987–997.
 - Tesmer, J. J., Klem, T. J., Deras, M. L., Davissou, V. J., and Smith, J. L. (1996) The crystal structure of GMP synthetase reveals a novel catalytic triad and is a structural paradigm for two enzyme families, *Nat. Struct. Biol.* 3, 74–86.
 - Zalkin, H., and Smith, J. L. (1998) Enzymes utilizing glutamine as an amide donor, *Adv. Enzymol. Relat. Areas Mol. Biol.* 72, 87–144.
 - Lacourciere, G. M., Mihara, H., Kurihara, T., Esaki, N., and Stadtman, T. C. (2000) *Escherichia coli* NiFeS-like proteins provide selenium in the pathway for the biosynthesis of selenophosphate, *J. Biol. Chem.* 275, 23769–23773.
 - Walker, H., Ferretti, J. A., and Stadtman, T. C. (1998) Isotope exchange studies on the *Escherichia coli* selenophosphate synthetase mechanism, *Proc. Natl. Acad. Sci. U.S.A.* 95, 2180–2185.
 - Hoskins, A. A., Anand, R., Ealick, S. E., and Stubbe, J. (2004) Initial Purification and Characterization of the Formylglycinamide Ribonucleotide Amidotransferase Complex from *Bacillus Subtilis*: An Example of Metabolite-Mediated Complex Formation, *Biochemistry* 43, 10314–10327.
 - Lim, L. W., Mathews, F. S., and Steenkamp, D. J. (1988) Identification of ADP in the iron-sulfur flavoprotein trimethylamine dehydrogenase, *J. Biol. Chem.* 263, 3075–3078.

46. Thieme, R., Pai, E. F., Schirmer, R. H., and Schulz, G. E. (1981) Three-dimensional structure of glutathione reductase at 2 Å resolution, *J. Mol. Biol.* *152*, 763–782.
47. Parsons, J. F., Jensen, P. Y., Pachikara, A. S., Howard, A. J., Eisenstein, E., and Ladner, J. E. (2002) Structure of *Escherichia coli* aminodeoxychorismate synthase: architectural conservation and diversity in chorismate-utilizing enzymes, *Biochemistry* *41*, 2198–2208.
48. Morollo, A. A., and Eck, M. J. (2001) Structure of the cooperative allosteric anthranilate synthase from *Salmonella typhimurium*, *Nat. Struct. Biol.* *8*, 243–247.
49. Thoden, J. B., Holden, H. M., Wesenberg, G., Raushel, F. M., and Rayment, I. (1997) Structure of carbamoyl phosphate synthetase: a journey of 96 Å from substrate to product, *Biochemistry* *36*, 6305–6316.
50. Thoden, J. B., Miran, S. G., Phillips, J. C., Howard, A. J., Raushel, F. M., and Holden, H. M. (1998) Carbamoyl phosphate synthetase: caught in the act of glutamine hydrolysis, *Biochemistry* *37*, 8825–8831.
51. Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr.* *24*, 946–950.
52. Merritt, E. A., and Bacon, D. J. (1997) Raster3D: Photorealistic Molecular Graphics, *Methods Enzymol.* *277*, 505–524.
53. Esnouf, R. M. (1999) Further additions to MolScript version 1.4, including reading and contouring of electron-density maps, *Acta Crystallogr. D* *55*, 938–940.
54. Christopher, J. A. (1998) *SPOCK*, Texas A&M University, College Station, TX.

BI0491301