Structural Insights into the Human and Avian IMP Cyclohydrolase Mechanism via Crystal Structures with the Bound XMP Inhibitor[†,‡]

Dennis W. Wolan,[§] Cheom-Gil Cheong,[§,||] Samantha E. Greasley,[⊥] and Ian A. Wilson*

Department of Molecular Biology and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received July 1, 2003; Revised Manuscript Received October 21, 2003

ABSTRACT: Within *de novo* purine biosynthesis, the AICAR transformylase and IMP cyclohydrolase activities of the bifunctional enzyme ATIC convert the intermediate AICAR to the final product of the pathway, IMP. Identification of the AICAR transformylase active site and a proposed formyl transfer mechanism have already resulted from analysis of crystal structures of avian ATIC in complex with substrate and/or inhibitors. Herein, we focus on the IMPCH active site and the cyclohydrolase mechanism through comparison of crystal structures of XMP inhibitor complexes of human ATIC at 1.9 Å resolution with the previously determined avian enzyme. This first human ATIC structure was also determined to ascertain whether any subtle structural differences, compared to the homologous avian enzyme, should be taken into account for structure-based inhibitor design. These structural comparisons, as well as comparative analyses with other IMP and XMP binding proteins, have enabled a catalytic mechanism to be formulated. The primary role of the IMPCH active site appears to be to induce a reconfiguration of the substrate FAICAR to a less energetically favorable, but more reactive, conformer. Backbone (Arg^{64} and Lys^{66}) and side chain interactions (Thr^{67}) in the IMPCH active site reorient the 4-carboxamide from the preferred conformer that binds to the AICAR Tfase active site to one that promotes intramolecular cyclization. Other backbone amides (Ile^{126} and Gly^{127}) create an oxyanion hole that helps orient the formyl group for nucleophilic attack by the 4-carboxamide amine and then stabilize the anionic intermediate. Several other residues, including Lys^{66}, Tyr^{104}, Asp^{125}, and $Lys^{137'}$, provide substrate specificity and likely enhance the catalytic rate through contributions to acid−base catalysis.

The purine and pyrimidine *de novo* synthesis and salvage pathways are two avenues by which cells produce their DNA and RNA nucleotide pools (*1*). Salvage synthesis provides an economical pathway in which degraded nucleic acids are reprocessed for cellular use. Despite the cellular cost of *de novo* synthesis, rapidly dividing cancer cells rely mainly on new synthesis of nucleotides (*2*). As a consequence, enzymes involved in the *de novo* production of purines and pyrimidines become potential targets for the development of anticancer therapeutics. Structure-based drug design provides a useful tool in the development of specific protein inhibitors and has been utilized for enzymes within these pathways.

The penultimate and final steps of *de novo* purine biosynthesis are catalyzed by the bifunctional homodimeric enzyme aminoimidazole-4-carboxamide ribonucleotide (AICAR)[1] transformylase/inosine 5′-monophosphate (IMP) cyclohydrolase (ATIC) which is highly conserved in prokaryotes and eukaryotes (*3*, *4*). AICAR transformylase (AICAR Tfase) catalyzes the transfer of a formyl group from the cofactor N_{10}-formyltetrahydrofolate (10-f-THF) to the substrate AICAR to produce the stable intermediate 5-formyl-AICAR (FAICAR) and the byproduct tetrahydrofolate (THF) (Figure 1A). FAICAR is subsequently cyclized, in the final step of *de novo* purine synthesis, to IMP by IMP cyclohydrolase (IMPCH) with the loss of a water molecule. The AICAR Tfase and IMPCH activities reside in different domains of the same polypeptide chain (ATIC); each domain can be individually expressed in the laboratory as an enzymatically active fragment (*4*).

Two avian ATIC crystal structures have been previously reported in which primary emphasis was placed on elucidation of the AICAR Tfase mechanism. These structures

[†] This work was supported by the National Institutes of Health Grant PO1 CA63536 (I.A.W.) and National Science Foundation and Skaggs predoctoral fellowships (D.W.W.). This is publication 15882-MB from The Scripps Research Institute.

[‡] The coordinates and structure factors have been deposited with the Protein Data Bank as entry 1PKX.

* To whom correspondence should be addressed. E-mail: wilson@scripps.edu. Phone: (858) 784-9706. Fax: (858) 784-2980.

[§] These authors contributed equally to this work.

[||] Present address: Laboratory of Structural Biology, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709.

[⊥] Present address: Agouron Pharmaceuticals, Inc., 10777 Science Center Dr., San Diego, CA 92121.

[1] Abbreviations: AICAR, aminoimidazole-4-carboxamide ribonucleotide; IMP, inosine monophosphate; ATIC, aminoimidazole-4-carboxamide ribonucleotide transformylase/inosine monophosphate cyclohydrolase; AICAR Tfase, AICAR transformylase; 10-f-THF, N_{10}-formyltetrahydrofolate; FAICAR, 5-formyl-AICAR; THF, tetrahydrofolate; IMPCH, IMP cyclohydrolase; XMP, xanthosine 5′-monophosphate; GAR, glycinamide ribonucleotide; GAR Tfase, GAR transformylase; GMP, guanosine 5′-monophosphate; CPS, carbamoyl phosphate synthetase; IMPDH, IMP dehydrogenase; IPTG, isopropyl β-D-thiogalactopyranoside; *CC*, correlation coefficient; NCS, noncrystallographic symmetry; FAICA, 4-formyl-5-imidazolecarboxamide; PRPP, phosphoribosylpyrophosphate.

10.1021/bi030162i CCC: $27.50
Published on Web 01/14/2004

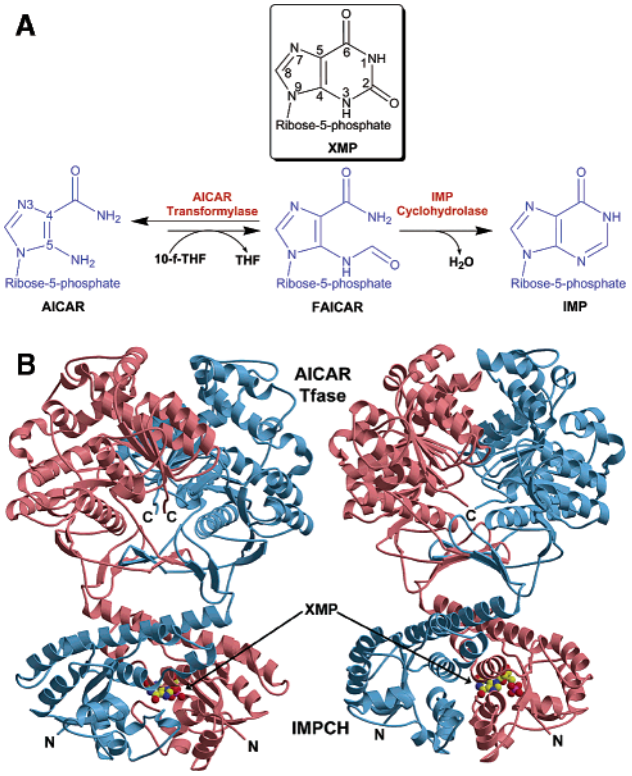

FIGURE 1: Reaction and crystal structure of ATIC. (A) Formyl transfer and cyclohydrolase reactions catalyzed by ATIC. The structure of inhibitor XMP is shown in the black box. (B) Two rotated views of the crystal structure of the apo human ATIC dimer. Monomers C and D that compose one of the dimers in the crystal asymmetric unit are colored rose and blue, respectively. The bound XMP molecule is depicted in a cpk representation (yellow for carbon, red for oxygen, blue for nitrogen, and magenta for phosphate). The N- and C-termini are labeled accordingly.

include ATIC in complex with both the AICAR Tfase substrate AICAR and an IMPCH inhibitor xanthosine 5′-monophosphate (XMP) (*5*), and a second complex with an AICAR Tfase multisubstrate adduct inhibitor β-DADF which coalesces the AICAR and folate moieties into one covalently linked entity (*6*). The combination of these structural data with previous mutational analyses has identified key residues involved in AICAR Tfase substrate binding and catalysis (*4*−*9*).

The proximity of the independent AICAR Tfase and IMPCH active sites on ATIC appears to be required as the formyl transfer reaction favors the starting products AICAR and 10-f-THF (*7*, *8*, *10*); the essentially irreversible cyclization of FAICAR (*10*) ensures formyl transfer in the forward direction and, hence, completion of purine biosynthesis. The bifunctionality of ATIC is conserved from *Escherichia coli* to humans, but the possibility of intermediate tunneling and channeling has been negated by ATIC structure (*11*) and intermediate single-turnover kinetic (*10*) analyses, respectively. A recently identified new class of monofunctional IMP cyclohydrolases in *Methanococcus jannaschii* has been identified, and its sequence is not at all homologous to that of the IMPCH domain of eukaryotic ATIC (*12*). In Archaea, FAICAR formation is independent of folate and is synthesized by an ATP-driven process (*13*), suggesting that the proximity of IMP cyclohydrolase activity is not required in this situation. These recent findings further support the notion that folate-dependent formyl transfer of AICAR Tfase requires close coupling with the cyclohydrolase activity within the bifunctional ATIC to ensure IMP production.

Herein, we describe the IMPCH active site and propose a cyclohydrolase mechanism based on the crystal structure of apo human ATIC and the previously reported structure of avian ATIC in complex with AICAR and XMP (*5*). Human and avian ATIC are highly homologous in both sequence (83% identical) and structure. The human ATIC structure was also determined to reveal whether any slight differences arise in the active site, or in the binding of ligands, that could facilitate design of more potent and selective human enzyme inhibitors, as was the case when human glycinamide ribonucleotide (GAR) transformylase (GAR Tfase) was compared with *E. coli* GAR Tfase (*14*, *15*).

Two isoforms (a and b) of ATIC have been reported in yeast (*16*) and humans (*4*, *17*). ATIC_b is the predominant form in humans (*17*) and, consequently, has been used here for structure determination. For comparison, avian ATIC has an additional alanine residue at the N-terminus (e.g., avian His^{268} is equivalent to human His^{267}).

The IMPCH active site was previously identified in the initial crystal structure of apo avian ATIC (*11*) due to the presence of a bound purine nucleotide ligand acquired during protein production or purification, which was later identified as XMP by HPLC analysis (*18*). XMP is a naturally occurring intermediate in subsequent processing of purine nucleotides in which IMP is converted to guanosine 5′-monophosphate (GMP). XMP is a potential feedback inhibitor of the IMPCH domain ($K_i = 0.12$ μM) when low levels of L-glutamine limit conversion of XMP to GMP (*19*, *20*).

In comparison to other cyclohydrolases, the IMPCH domain of ATIC involves only one catalytic step. On the other hand, other cyclohydrolase enzymes, such as methylenetetrahydrofolate dehydrogenase-cyclohydrolase (*21*) and GTP cyclohydrolases I (*22*, *23*) and II (*24*, *25*), catalyze multiple transformations via detectable intermediates within their cyclohydrolase active site. Key binding and catalytic residues have been elucidated here by comparison of the human and avian XMP-bound IMPCH active sites with the apo avian ATIC structure (*11*) and through analysis of analogous IMP and XMP binding proteins, including carbamoyl phosphate synthetase (CPS) (*26*, *27*) and IMP dehydrogenase (IMPDH) (*28*, *29*). Identification of unique interactions of XMP within the human and avian IMPCH active sites has allowed formulation of the cyclohydrolase catalytic mechanism that is consistent with and complements mutagenesis studies of human ATIC (*18*).

MATERIALS AND METHODS

Protein Expression and Purification. Human ATIC cDNA was cloned into a pET28a vector (Novagen, Inc.) and transformed into *E. coli* BL21.DE3 cells (Novagen, Inc.) for overexpression and was a kind gift from G. Peter Beardsley. The *E. coli* transformant was grown in 2YT medium (Gibco BRL) at 37 °C to an OD_{600} of 0.6. Isopropyl β-D-thiogalactopyranoside (IPTG) (Gibco BRL) was added to a final concentration of 0.4 mM, and the mixture was incubated for an additional 3 h at 30 °C. Cells were harvested by centrifugation at 3500*g* for 15 min at 4 °C. Pellets were subsequently washed with 0.85% NaCl and used immediately or frozen at −80 °C. Cell pellets (approximately 3 g wet

weight) were resuspended in 100 mL of ice-cold buffer A [50 mM sodium phosphate, 300 mM NaCl, and 20 mM imidazole (pH 8.0)]. Lysozyme was added to the pellets to a final concentration of 1 mg/mL and sonicated at 4 °C. Insoluble material was removed by centrifugation at 20000*g* at 4 °C for 60 min. The resulting supernatant was mixed with 10 mL of a 1:1 slurry of Ni−NTA Superflow beads and incubated on a rotator overnight at 4 °C. The Ni−NTA Superflow beads were transferred to a 1 cm × 10 cm column and washed with 200 mL of buffer A and 100 mL of 99% buffer A and 1% buffer B [50 mM sodium phosphate, 300 mM NaCl, and 500 mM imidazole (pH 8.0)]. Human ATIC was eluted using a 1 to 50% buffer B gradient, and fractions were analyzed by SDS−PAGE. Fractions containing ATIC were pooled and concentrated using Millipore Ultrafree-15 filters (molecular mass cutoff of 10 kDa). The concentrated protein sample was loaded onto a 2.6 cm × 90 cm Superdex 200HR column (Amersham Biotech) that was pre-equilibrated with buffer C [20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 50 mM KCl, 5 mM EDTA, and 5 mM DTT]. Peak fractions were analyzed by SDS−PAGE, pooled, and concentrated using Millipore Ultrafree-15 filters to a final protein concentration of 0.1 mg/mL and stored at 4 °C. Prior to crystallization, human ATIC was concentrated to 10 mg/mL using Millipore Ultrafree-15 filters.

Crystallization, Data Collection, and Structure Determination. Crystals of human ATIC (10 mg/mL) were grown at 22 °C by sitting-drop vapor diffusion by mixing equal volumes of protein and a reservoir solution of 15−17% PEG 3000, 0.1 M Tris-HCl (pH 7.5−8.0), 5% MPD, 6 mM DTT, and 0.2 M NaCl. Data were collected to 1.90 Å resolution on a single, flash-cooled crystal in a cryoprotectant of the reservoir solution and 10% MPD, and processed with HKL2000 (*30*) in monoclinic space group $P2_1$ ($a = 77.52$ Å, $b = 93.56$ Å, $c = 179.88$ Å, and $\beta = 91.1°$). The Matthews coefficient ($V_m = 2.5$ Å^3/Da) (*31*) suggested four monomers per asymmetric unit (solvent content of 51%).

The human structure was determined by molecular replacement using the apo avian ATIC structure (1G8M) as the search model with the program AMoRe (*32*). The rotation search incorporating data from 10 to 3.5 Å identified four major peaks with *CC*s of 10.9−10.4 (first noise peak at 7.1). The subsequent translation search yielded a final *CC* and an R_{cryst} of 67.1 and 35.5%, respectively. The four monomers in the asymmetric unit comprise two biologically relevant homodimers of ATIC. Data statistics are outlined in Table 1.

Structure Refinement. The human ATIC structure was refined using the program CNS (*33*) with four cycles of conventional positional refinement and simulated annealing (MLF target). Approximate 2-fold NCS restraints were applied during initial rounds of refinement and subsequently released in later rounds. Water molecules were added when they satisfied the electron density and geometric criteria using CNS and manual inspection in O (*34*) and TURBO-FRODO (*35*). The final R_{cryst} and R_{free} are 21.1 and 24.9%, respectively, with 91.9% of the residues in the most favored region of the Ramachandran plot (Table 1).

S_c coefficients (*36*) and buried surface areas for both human and avian ATIC were calculated with SC (*32*) and MS (*37*) using 1.7 and 1.4 Å probes, respectively. Hydrogen bonds and van der Waals interactions were identified with default parameters in CONTACSYM (*38*). The models were analyzed with CNS (*33*), CCP4 (*32*), PROCHECK (*39*), and WHATCHECK (*40*). Root-mean-square deviation (rmsd) comparisons and calculations (Table 2) were carried out with PROFIT (Martin, A. C. R., SciTech Software, Chico, CA). Figures 1−5 were created with Bobscript (*41*) and rendered with Raster3D (*42*). Coordinates and structure factors for avian ATIC in complex with AICAR and XMP have been previously deposited in the Protein Data Bank (*43*) as entry 1M9N. Coordinates and structure factors for apo human ATIC have been deposited in the Protein Data Bank as entry 1PKX.

Table 1: Human ATIC Data Processing and Refinement Statistics

data processing	
resolution range (Å) (outer shell)	50−1.9 (1.97−1.90)
no. of unique reflections	177173 (9598)
completeness (%)	87.7 (47.8)
redundancy	2.1 (1.8)
R_{sym} (%)[a]	5.3 (39.4)
average $I/\sigma(I)$	15.0 (2.1)
refinement	
resolution range (Å)	50.0−1.90
no. of reflections[b] (test set)	160871 (8496)
R_{cryst} (%)[c]	21.1
R_{free} (%)[c]	24.9
no. of protein atoms/no. of waters	17762/977
CV[d] coordinate error (Å)	0.30
rmsd for bonds (Å)	0.009
rmsd for angles (deg)	1.39
$\langle B \rangle$ for subunits A−D (Å^2)	33.1, 28.1, 46.5, 26.2
$\langle B \rangle$ for waters and ligands (Å^2)	31.1, 34.8
Ramachandran statistics (%)	
most favored	91.9
additionally allowed	7.8
generously allowed	0.2
disallowed	0.0

[a] $R_{sym} = 100 \times \sum_h \sum_i |I(h)_i - \langle I(h) \rangle| / \sum_h I(h)$, where $I(h)_i$ is the ith measurement of reflection h and $\langle I(h) \rangle$ is the average measurement value. [b] Reflections with positive I values were used for refinement. [c] $R_{cryst} = \sum_h ||F_o| - |F_c|| / \sum_h |F_o|$, where F_o and F_c are the structure factor amplitudes from the data and the model, respectively. R_{free} is R_{cryst} with 5.2% of the test set structure factors. [d] Cross-validated (CV) Luzzati coordinate errors.

RESULTS

Comparison of Human versus Avian ATIC. Full-length avian ATIC (residues 1−593) and human ATIC (residues 1−592) when overexpressed in *E. coli* can be purified by gel filtration as noncovalent dimers. The apo human and avian AICAR−XMP complex (*5*) ATIC structures were both determined via molecular replacement using the native ATIC structure (1G8M) (*11*) as the initial search model. The final R_{cryst} and R_{free} values for the AICAR−XMP-bound avian ATIC structure were 20.6 and 24.4%, respectively, with 92.4% of the residues in the most favored region of the Ramachandran plot (*5*). The structure includes residues 3−592 of both monomers, AICAR and XMP bound within their respective active sites, two potassium ions, and 513 water molecules. The corresponding final R_{cryst} and R_{free} values are 21.1 and 24.9%, respectively, for the apo human ATIC structure with 91.9% of the residues in the most favored region of the Ramachandran plot. The final human model contains two dimers (monomers A−D, 2353 residues) and 977 water molecules within the asymmetric unit. Monomers A and B comprise one dimer; monomer A

Table 2: Rmsd Comparisons within and between Human and Avian ATICs for Corresponding α Carbons

domain	rmsd (Å)
avian ATIC (AICAR and XMP)	
monomer A vs monomer B	
overall	0.44
IMPCH (4−199)	0.31
AICAR Tfase (200−593)	0.37
human ATIC	
monomer A vs monomer B	
overall	0.90
IMPCH (4−198)	1.09
AICAR Tfase (199−592)	0.54
monomer C vs monomer D	
overall	0.81
IMPCH (4−198)	1.06
AICAR Tfase (199−592)	0.37
human ATIC vs avian ATIC	
monomer C vs monomer A	
overall	0.61
IMPCH	0.42
AICAR Tfase	0.56
monomer C vs monomer B	
overall	0.71
IMPCH	0.46
AICAR Tfase	0.65
monomer D vs monomer A	
overall	0.97
IMPCH	1.19
AICAR Tfase	0.66
monomer D vs monomer B	
overall	1.06
IMPCH	1.20
AICAR Tfase	0.77

contains residues 3−592, but with no electron density, and hence presumed disorder, between residues 483 and 484, whereas monomer B comprises all residues from residue 4 to 592. Monomers C and D comprise the second dimer; monomer C is made up of residues 4−592 with disorder at residues 481 and 482, whereas monomer D contains all residues from 4 to 592.

The ATIC monomer maintains two activities on a single polypeptide chain in which the IMP cyclohydrolase and AICAR transformylase activities reside on residues 1−198 and 199−592, respectively (*4*). Consistent with their 83% level of sequence identity, the human and avian ATIC structures are highly conserved. The topology of the AICAR Tfase domain of avian ATIC has been previously described for the native ATIC (PDB entry 1G8M) (*11*), for ATIC in complex with AICAR and XMP (1M9N) (*5*), and for the *β*-DADF complex (1O20) (*6*). The IMPCH domain is composed of a parallel five-stranded *β*-sheet (5−4−1−2−3 strand order) with three α-helices on one side of the sheet and seven on the other (Figure 2). The IMPCH domain does not contain any common mononucleotide or phosphate binding motifs and does not contain signature fingerprint sequences, such as GXXGXGK or GXGXXG (X represents any residue), despite containing a Rossmann-like fold.

After initial refinement, the human apo ATIC $F_o - F_c$ electron density maps revealed unambiguous electron density for a purine nucleotide bound to monomers A and C of the independent dimers within the asymmetric unit, despite purification and crystallization in the absence of any purine nucleotides, whereas no ligands were observed in monomers B and D. Similarly, monomer A of the apo avian ATIC structure also contained a bound nucleotide carried throughout purification and crystallization (*11*). As this nucleotide was later identified as XMP through HPLC analysis (*18*), the bound ligand in the IMPCH domains of monomers A and C of the apo human structure was modeled as XMP.

Identification of the IMPCH active sites in the avian AICAR−XMP-bound ATIC structure was also facilitated by clearly interpretable initial $F_o - F_c$ maps (*5*). Here, we describe for the first time a detailed interpretation of the additional density in both IMPCH active sites that could be readily modeled as XMP (Figure 3A,B). The XMP nucleotides maintain an approximate 3′-endo sugar pucker which is more pronounced when monophosphate nucleotides [AICAR (*5*) and the multi-substrate adduct inhibitor *β*-DADF (*6*)] are bound to the AICAR Tfase active site; however, the electron density for the sugar in one of the XMPs is not as well defined (Figure 3A) and may represent some minor fluctuations around the 3′-endo conformation. Slight, but important, differences arise in the conformation of the two independent XMP molecules, as electron density in one IMPCH active site reveals a planar xanthosine purine ring, while in the opposing active site the C2 carbonyl is bent out of the plane by approximately 20° (Figure 3A,B). The "bent" XMP molecule is similar to the conformation of the endogenous purine ligand found selectively in one IMPCH active site of the native avian ATIC structure (1G8M) (*11*). Superposition of the XMP-bound IMPCH monomers oth-

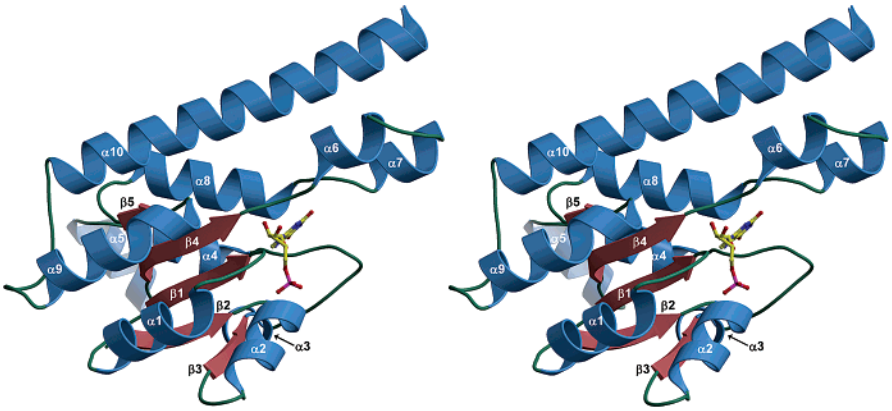

FIGURE 2: Stereoview of the IMP cyclohydrolase domain with bound XMP. IMP cyclohydrolase contains a typical Rossmann fold within its parallel *β*-sheet with a 5−4−1−2−3 strand order. The phosphate is positioned at the N-terminus of helix α2, and the XMP C2 carbonyl is situated directly above the amino end of helix α8. Helices and *β*-sheets are colored blue and rose, respectively. The bound XMP molecule is depicted in a ball-and-stick representation with atoms colored as in Figure 1B.

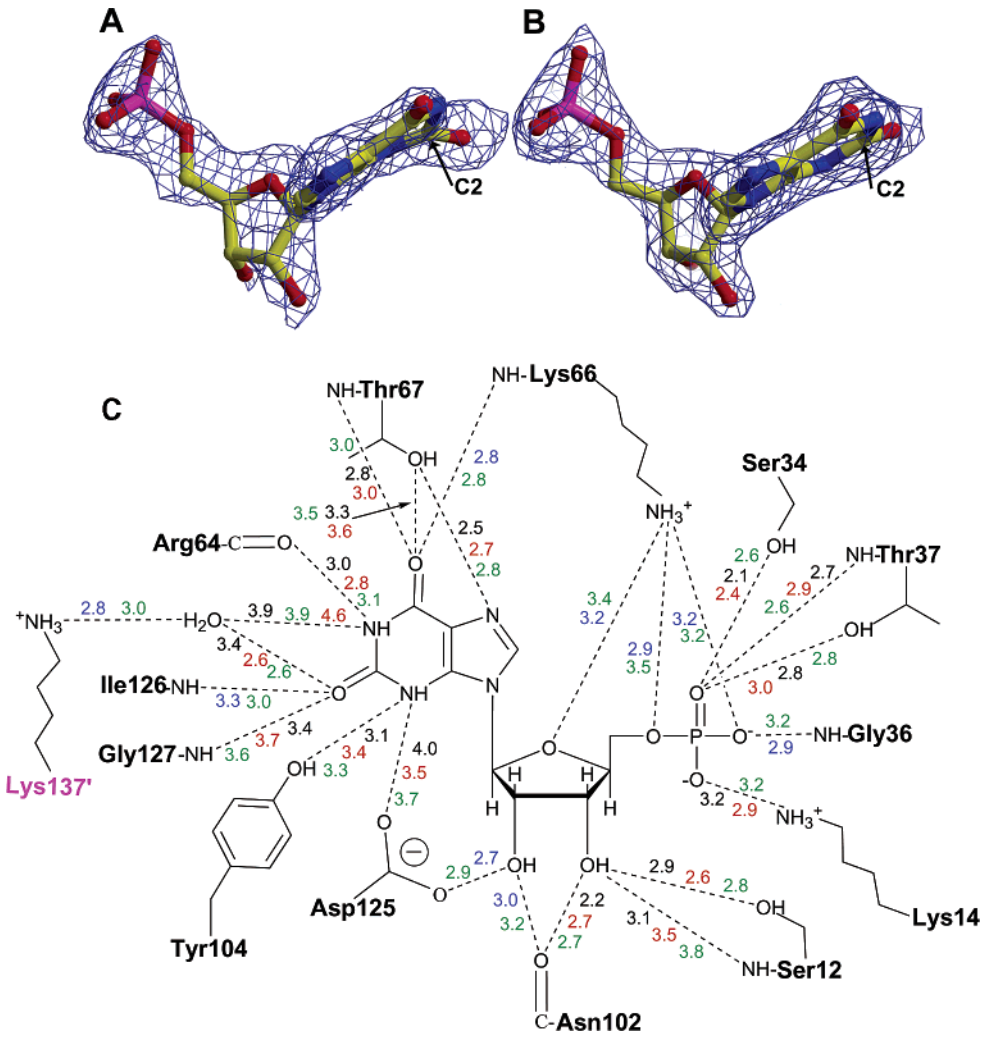

FIGURE 3: Comparison of the conformation of XMP bound to the two IMPCH active sites of the avian ATIC dimer and the associated hydrogen bonding network within the active site. (A) XMP bound to monomer A. (B) XMP bound to monomer B. The electron density for XMP is shown as a blue $2F_o - F_c$ map contoured at 1.0σ. Atoms are depicted in ball-and-stick representation and colored as in Figure 1B. The C2 carbonyl is labeled. (C) Schematic representation of the hydrogen bonding network and corresponding distances within the active site when XMP is bound. Differences arise in various subunits due to the bent (avian B) vus planar conformations of the bound XMP (avian A and human A and C). Hydrogen bond distances in the AICAR−XMP-bound avian ATIC structure are shown in angstroms and labeled in black (monomer A distances), red (monomer B distances), and blue (equivalent distances in monomers A and B). Corresponding distances in XMP-bound monomer C of the human apo structure are labeled in green. Lys$^{137'}$ from the opposite subunit is labeled in pink and denoted with a prime (′). Side chains are depicted for only those residues that provide direct interactions with XMP.

erwise reveals no significant differences in the active site residues. Similarly, superposition of the human XMP-bound monomer A onto the avian XMP-bound monomer B resulted in an rmsd of only 0.46 Å (Figure 4B). However, certain regions of the XMP-bound IMPCH domains of the avian AICAR−XMP and apo human structures vary significantly from the native and β-DADF complex avian ATIC subunits in which no endogenous ligands were found or when no exogenous XMP was added. Unlike the relatively static, solvent-rich AICAR Tfase active site which forms at the dimer interface, the IMPCH active site (Figure 4A) is comprised mainly from residues in one subunit and undergoes substantial conformational rearrangements when it binds XMP (Figure 4C), and excludes all but one bound water molecule (Figure 4A,B).

XMP is almost completely shielded from the bulk solvent with 95% of its surface buried by IMPCH active site residues. XMP is firmly anchored in the active site and takes advantage of dipole moments supplied by the proximity of two α-helices (Figures 2 and 4A). The XMP phosphate is positioned at the N-terminus of α-helix 2, and the C2 exocyclic carbonyl interacts with the amino end of α-helix 8. Side chain interactions with the phosphate are provided by Lys^{14}, Ser^{34}, Gly^{36}, Thr^{37}, and Lys^{66}, while the main chain of Thr^{37} also supplies a hydrogen bond (Figures 3C and 4A). Main chain and side chain contacts with Asn^{102} and Asp^{125}, respectively, as well as main chain and side chain interactions with Ser^{12}, anchor the XMP ribose ring oxygens at positions 2 and 3. The ribose ring is further secured by an interaction between O1 and the side chain of Lys^{66}. The N1 position of the xanthosine ring contacts only the main chain carbonyl of Arg^{64}. The exocyclic C2 carbonyl hydrogen bonds to the main chain amides of Ile^{126} and Gly^{127}, which are both situated at the N-terminus of α-helix 8 (Figures 2, 3C, and 4A). Lys$^{137'}$ provides the only interaction from the opposing subunit where it donates an indirect hydrogen bond to the C2 carbonyl via the only structural water molecule within the active site.

Previous studies of the monomer and dimer have shown that human AICAR Tfase activity requires assembly of the

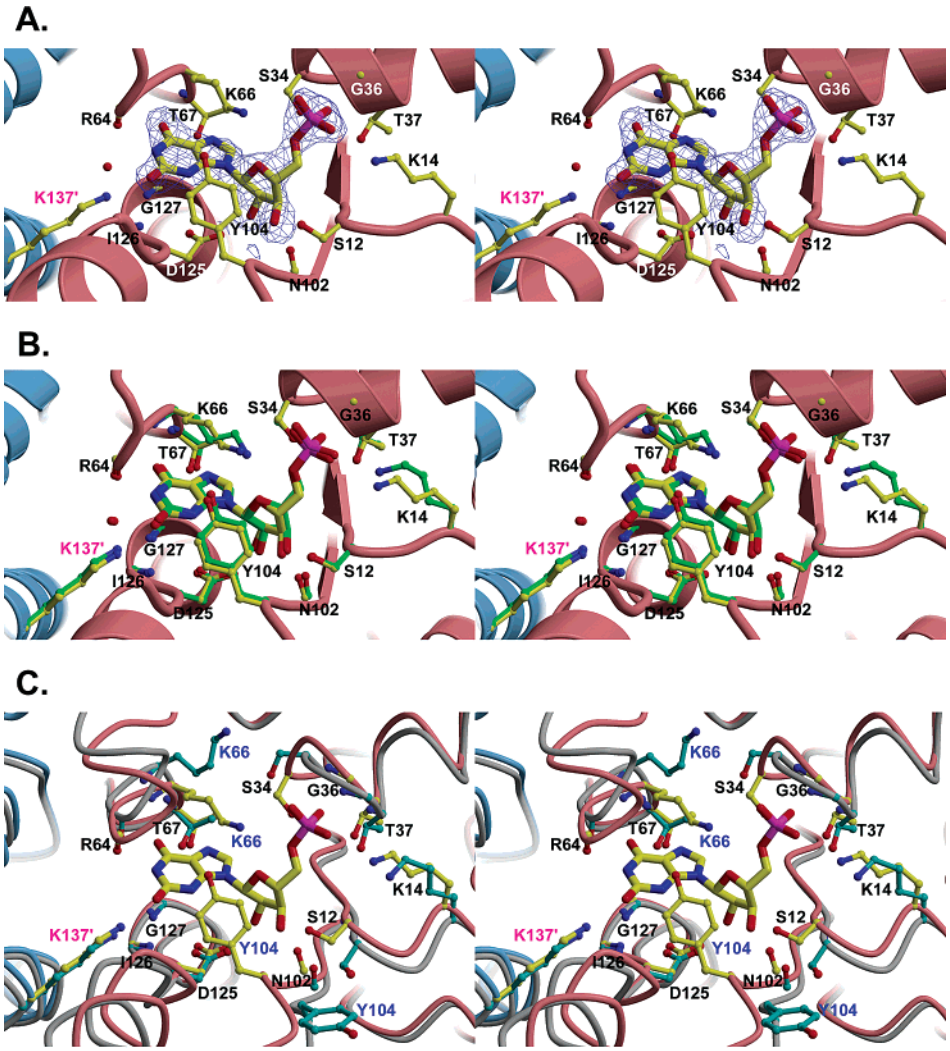

FIGURE 4: Apo human IMPCH active site with bound XMP, and comparisons of XMP bound and unbound human IMPCH active sites. (A) Stereoview of the residues that make up the IMPCH active site. XMP, active site residues, and the water molecule are represented as balls and sticks with atoms colored as in Figure 1B. XMP is surrounded by its corresponding $F_o - F_c$ electron density omit map contoured at 2.0σ. (B) Stereoview of the superposition of XMP-bound human (monomer C) and avian (monomer B) structures. XMP, active site residues, and labels for the apo human structure are represented as in panel A, while the AICAR−XMP-bound avian structure has its XMP and active site residues colored green. Only small deviations are seen in some of the active site residues. (C) Stereoview of the conformational changes in the active site that occur upon ligand interaction. The Cα traces for the XMP-bound monomer C and unliganded monomer D (colored gray) are shown. XMP, active site residues, and labels for the XMP-bound structure are as in panel A. The unliganded active site residues of monomer D are colored in cyan (side chains). Note the conformational changes in Lys[66] and Tyr[104] (labeled in blue).

dimer, while the IMPCH domain can function as a monomer, albeit it is 6-fold less active than the naturally occurring homodimer (*44*). This observation also supports a role for Lys[137′] in facilitating catalysis and/or in binding of the natural substrate FAICAR via a long-range electrostatic effect on N1 mediated through the bound water molecule. N3 of the purine ring interacts with the side chain hydroxyl group of Tyr[104]. Similarly, the carboxyl side chain of Asp[125] may contribute electrostatic effects on N3 of XMP. Asp[125] is not sterically hindered in the binding site and could adopt either of the two alternative rotamers that would provide closer interactions with N3. The C6 exocyclic carbonyl has three potential contacts with the main chain and side chain of Thr[67], as well as an amide backbone interaction with Lys[66].

Conformational and Orientational Differences of Bound XMP. Superposition of the two independent IMPCH active sites of the avian ATIC dimer with bound AICAR and XMP reveals differences in the orientation and conformation of their respective bound XMP molecules that result in, or are a consequence of, variation of hydrogen bond interactions with the enzyme (Figure 3C). Slight differences exist in the ribose ring sugar pucker, in the torsion of the xanthosine moiety, and in the configuration of the C2 carbonyl (Figure 3A,B). Refinement of the XMP molecules revealed that the ligand bound to monomer A adopts a more pronounced 3′-endo sugar pucker that results in differential hydrogen bond interactions with residues Ser[12] and Asn[102] within the two IMPCH active sites (Figure 3C). As the electron density for the ribose 3′-hydroxyl was lacking in monomer A, compared to monomer B, the XMP ribose bound to monomer A was weighted toward geometry rather than data in the structure refinement, resulting in a preference for the energetically favored 3′-endo pucker. Second, as a consequence of the torsional variation of the xanthosine moiety relative to the ribose ring between the two bound XMP molecules, slight differences arise in the hydrogen bond distances between XMP and residues Tyr[104], Asp[125], and Gly[127] (Figure 3C). However, the major difference between these bound XMP

molecules is the conformation of the C2 carbonyl, which is bent out of the plane of the purine ring by approximately 20° in monomer A (Figure 3A,B). As a consequence, the hydrogen bond distance from the C2 carbonyl to the only active site water deviates by 0.8 Å in the two independent avian IMPCH active sites.

However, the XMP molecules selectively bound to monomers A and C (but not to monomers B and D) of the apo human structure reveal the purine rings and the C2 carbonyl are in plane. Thus, the bent endogenous nucleotide ligand bound to monomer A of the AICAR−XMP-bound avian structure either was not displaced, despite the presence of a 10-fold molar excess of XMP prior to crystallization, or indeed constitutes a different XMP conformation compared to the human or second avian IMPCH binding sites. The former situation could be explained if (1) the bound ligand in the crystal is not XMP, as identified by HPLC analysis (*18*), and either has a higher affinity than XMP or is conformationally trapped in the binding site and does not exchange or (2) a real difference exists in binding of XMP to avian and human ATIC in which XMP and IMPCH both subtly alter their conformation in avian ATIC by induced fit so that the affinity for the second site is reduced and not occupied even in the presence of relatively higher concentrations of added XMP. However, the electron density for the bent ligand is well resolved, and its interpretation is consistent with a xanthosine moiety, although minor substitutions in the ring cannot obviously be differentiated at this resolution (such as the nitrogen at position 3 of the ring for a carbon atom). Normally, due to the favorable conjugation with the aromatic imidazole moiety of xanthosine, N3 would be expected to be in plane with the imidazole. Nevertheless, the six-membered xanthosine moiety may not be as rigid as most aromatic rings as the nitrogen atoms can maintain an sp^3 hybridization with the lone pair of electrons, which would abolish the planarity of the ring (Figure 1A).

Other protein crystal structures with a bound XMP, including *Methanobacterium thermoautotrophicum* orotidine-5′-monophosphate decarboxylase (*45*) (PDB entry 1LOL) and *Toxoplasma gondii* hypoxanthine-guanine phosphoribosyltransferase (*46*, *47*) (PDB entry 1QK5), as well as the small molecule crystal structures of anhydrous xanthosine (*48*) and xanthosine dihydrate (*49*), depict a planar purine ring. Despite similar distances, the bent conformation does in fact make better interactions between the C2 carbonyl's lone pairs of electrons with the N-terminus of α-helix 8 and the backbone amides of Ile^{126} and Gly^{127}. Therefore, the bent conformation may have a higher affinity when bound first to one of the two IMPCH binding sites of the dimer in the avian enzyme.

Various crystal structure determinations have shown that nucleotides, which have not been added at any stage in protein expression, can be carried throughout purification and crystallization, as for GTP cyclohydrolase I which had a mixture of GTP and ATP bound to the 10 active sites of the decamer (*23*). In the ATIC case, selective binding of the bent XMP molecule to only one of the identical IMPCH active sites in the apo avian ATIC crystal structure, as well as the identification of XMP bound to only one of the IMPCH actives sites within the human dimers, suggests the possibility of half-the-sites reactivity. Despite conformational changes upon ligand binding, no clear structural evidence indicates that induced conformational rearrangements incapacitate the opposing IMPCH active site from binding substrate, as clearly shown in thymidylate synthase (*50*). However, many other crystal structures have been determined for which structural evidence fails to support induced conformational changes as the main factor in explaining negative cooperativity, despite strong biochemical evidence. For example, the crystal structure of dihydropteroate synthase depicts a ligand bound to only one subunit of the dimer, as in ATIC, but no significant conformational changes occur within the binding active site upon interaction, to explain why the other subunit does not bind ligand (*51*). Similarly, tyrosyl-tRNA synthetase is a negatively cooperative protein (*52*); in this case, the crystal structure revealed the intermediate tyrosyl adenylate in both subunits of the dimer (*53*). However, in this experiment, the native crystals were soaked under saturating conditions of tyrosine and 12 mM ATP, thereby forcing both active sites to bind tyrosyl adenylate. As avian ATIC was also cocrystallized with a high molar excess (10-fold) of AICAR and XMP, it appears that this excess was sufficient to populate the second binding site and, hence, suggests that first IMPCH site has a higher affinity and that a higher concentration of ligand is required to fully occupy the second site. Thus, under physiological conditions, half-the-sites reactivity may need to be considered for ATIC.

Conformational Changes upon Ligand Binding. Comparison of the unbound and bound IMPCH active sites in the apo human structure, as well as the apo and AICAR−XMP-bound avian ATIC structures, reveals several conformational changes upon ligand binding. The most significant movement is for the loop of residues 103−108, where its primary role appears to be to flip the side chain of Tyr^{104} into the active site and sequester the bound ligand from bulk solvent (Figure 4C). This conformational change, with an rmsd of approximately 4 Å for main chain Cα atoms, results in the XMP molecule being 95% buried in the active site (using a 1.4 Å probe). Other significant movements include the N-terminus of α-helix 2, which is displaced by approximately 3 Å, as measured for main chain Cα atoms. Similarly, rearrangement of side chains occurs for those residues in α-helix 2 and others that hydrogen bond to the phosphate and ribose ring, including Ser^{12}, Lys^{14}, Ser^{34}, Thr^{37}, and Lys^{66} (Figure 4C). The loop region of residues 64−66 has an approximate displacement of its main chain toward XMP of 3.8 Å upon ligand binding. Lys^{66} has an extreme side chain rearrangement as it orients to within 3.4 Å of the hydroxyl group of Tyr^{104}. In comparison, in the unbound monomer D of the apo human structure, the distance between these two groups is approximately 21 Å (Figure 4C). Surprisingly, with the exception of Tyr^{104} and Arg^{64}, those residues (including $Lys^{137'}$ from the opposite subunit) that interact with the xanthosine moiety are primarily fixed in conformation after introduction of the XMP molecule. This observation suggests these residues are already poised and ready to contribute to the catalytic transformation once substrate FAICAR has bound and been enveloped by Tyr^{104}.

Active Site Comparisons with CPS and IMPDH.. The XMP-bound IMPCH domains of both the apo human and AICAR−XMP-bound avian structures were compared to all structures within the Protein Data Bank using the program Dali (*54*). The two top *Z*-scores for human and avian were 9.8 and 9.5 for CPS (PDB entry 1C3O) (*26*, *27*) and 8.5 and

8.3 for methylglyoxal synthase (PDB entry 1B93) (*55*), respectively (a Z of <2.0 means structural dissimilarity, XMP-bound IMPCH monomer B to itself is 37.5). Sequence alignments revealed conservation of IMPCH active site residues with CPS in the phosphate binding loop; however, sequence conservation with methylglyoxal synthase was not located near the IMPCH active site. CPS is a multidomain heterodimeric protein in which IMP binds as an effector molecule to the subdomain of residues $Ser^{937}-Lys^{1073}$ of the large subunit (*27*). Sequence alignment of this region with the human IMPCH domain of ATIC indicates an overall level of identity of 28% (Figure 5) and reveals significant similarities in residues that interact with the purine ligand (*27*). Superposition of the IMP-binding domain of CPS and XMP-bound human IMPCH gave an rmsd of 2.5 Å for 112 equivalent Cα atoms of the corresponding domains containing a Rossmann fold. Superimposition of XMP in the IMPCH active site of ATIC with IMP of the CPS binding domain reveals a high level of sequence and three-dimensional conservation of residues that interact with the phosphate group. Strikingly, the conserved interactions between the two proteins occur in the same order in their respective sequences and relatively similar hydrogen bond interactions (Figure 5). On the other hand, IMPCH active site residues that interact with the ribose ring and xanthosine moiety are not conserved. However, one main chain interaction is conserved with the C6 exocyclic carbonyl of both XMP and IMP by IMPCH residue Thr^{68} and by CPS residue Val^{994}, respectively (Figure 5A,B). Thr^{68} of the IMPCH domain has an additional side chain interaction with the C6 carbonyl via its hydroxyl group (Figures 3C and 4A).

IMPDH is involved in the catalysis of IMP to the intermediate XMP prior to conversion to GMP. Sequence alignment with the IMPCH domain of human ATIC reveals no significant homology between the two proteins. Second, no structural similarities are found with IMPDH (PDB entries 1EEP and 1JR1) which is a β/α-barrel (*28*, *29*). However, closer examination of the IMP binding site of IMPDH with those corresponding residues involved in XMP interaction in the IMPCH active site reveals some similarity in interactions with the phosphate group. Otherwise, the lack of three-dimensional similarity or in conservation of hydrogen bond donors and/or acceptors suggests that the IMPCH and IMPDH active sites bind FAICAR and IMP, respectively, in unique ways.

DISCUSSION

The crystal structure of the bifunctional homodimeric apo human ATIC complexed with endogenous XMP when compared to the avian ATIC structure in complex with AICAR Tfase substrate AICAR and IMPCH inhibitor XMP (*5*) now provides major insights into the cyclization reaction performed by the IMPCH active site. These structural results can then be combined with mutagenesis data (*18*) to elucidate potential catalytic mechanisms for IMPCH.

Key Residues of the IMPCH Active Site. Comparison of the unbound and XMP-bound forms of the IMPCH domain of both human and avian ATIC, as well as superposition of the IMPCH active site with proteins similar in structure and function, allows us to differentiate those residues that are involved in substrate binding versus those likely to play a role in the cyclization reaction of FAICAR. Conserved active site residues correspond to those that provide direct interaction to the bound nucleotide, including those residues responsible for phosphate positioning in IMPDH and CPS. Residues to consider in any IMPCH mechanistic proposal must include the active site water; Tyr^{104}, as it is involved in a large conformational shift upon ligand interaction; Lys^{66}, when it orients toward the hydroxyl group of Tyr^{104}; Asp^{125}, whose pK_a could be perturbed on ligand binding and subsequent active site side chain rearrangements; $Lys^{137'}$ from the opposing subunit; the main chain interaction to N1 by Arg^{64} and the hydroxyl side chain interaction to the C6 carbonyl by Thr^{67}. Tyr^{104}, Asp^{125}, and $Lys^{137'}$ have been shown to be important for catalysis by mutational analysis (*18*). However, Ile^{126} and Gly^{127} both provide key main chain interactions (that cannot be probed by mutagenesis) with the C2 exocyclic carbonyl of xanthosine that may orient the 5-formylamine of FAICAR for nucleophilic attack by the 4-carboxamide. These hydrogen bonds, provided by Ile^{126} and Gly^{127}, may account for potential preferential binding of substrate FAICAR and XMP over product IMP which lacks the C2 carbonyl.

Several enzyme structures have revealed that large conformational changes of an active site tyrosine side chain accompany ligand interaction. Tyr^{248} of carboxypeptidase A, and its associated polypeptide loop, enclose the active site upon ligand binding and sequester the substrate from bulk solvent. The tyrosine hydroxyl is displaced by approximately 12 Å, enabling a hydrogen bond to form between the tyrosine side chain and the terminal carboxylate of the substrate (*56*−*59*). Similarly, Tyr^{291} of the bifunctional enzyme 4-hydroxy-2-ketovalerate aldolase/acylating acetaldehyde dehydrogenase is responsible for controlling the access to a tunnel separating the two activities for the intermediary product (*60*). Conformational flipping of the side chain allows Tyr^{291} to play this regulatory role.

The consequences of Tyr^{104} and Lys^{66} side chain movements may be 3-fold, as they not only sequester XMP and orient the phosphate, respectively, but also bring the side chains of these two groups together such that they fully bury Asp^{125} (Figure 4C). The side chain rearrangement brings the positive charge of the Lys^{66} amino group into the proximity of Tyr^{104} such that it could perturb the pK_a of Tyr^{104} for a potential role in catalysis. Asp^{125} is approximately 84% buried from bulk solvent within the unbound IMPCH active site and approximately 97% buried in the XMP-bound form (using a 1.4 Å probe), suggesting that the pK_a of Asp^{125} may also be perturbed upon ligand binding. Asp^{125} hydrogen bonds to the ribose ring O2 hydroxyl group; however, it also maintains a long-range electrostatic interaction with N3 of the xanthosine moiety (Figures 3C and 4A). Despite retention of its side chain conformation upon ligand binding, Asp^{125} does not appear to be conformationally restricted within the active site, and the two other possible rotamers could bring it to within 2.8 Å of N3 without steric hindrance from other active site residues. Therefore, the potential of pK_a perturbation and close interactions with XMP suggest Asp^{125} may play a role in FAICAR cyclization. Asp^{125} and Tyr^{104} are then the two candidates most likely to act as catalytic bases due to their structural juxtaposition to the N3 of the purine ring and their potential pK_a perturbation on ligand binding. Mutational analysis indicates both residues contribute to

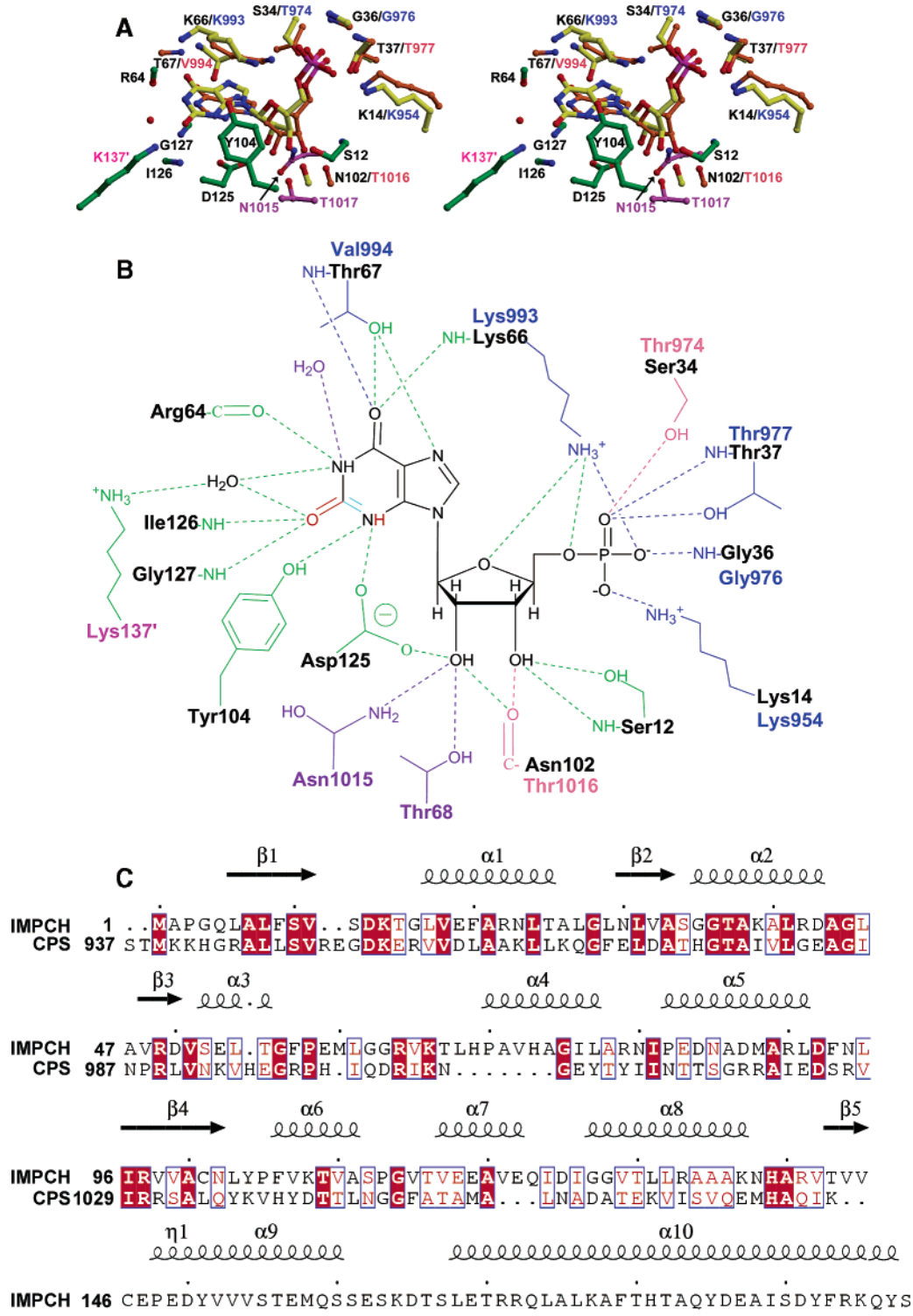

FIGURE 5: Comparison of the IMP-bound CPS and XMP-bound human apo IMPCH structures. (A) Stereoview of the superposition of the IMP-binding domain of CPS (PDB entry 1C3O) and the IMPCH active site. Identical IMPCH residues and CPS residues are labeled in black and blue, respectively, while functionally homologous residues in CPS are labeled in rose. XMP and active site residues of IMPCH conserved with CPS have yellow carbon atoms, as in Figure 1B. IMP and interacting CPS residues conserved with those of IMPCH have orange carbon atoms. Interactions unique between XMP and IMPCH are represented by green carbon atoms. Interactions unique to IMP and CPS are highlighted with purple carbon atoms. (B) Comparative schematic representation of the IMPCH active site and IMP-binding domain of CPS. Conserved interactions in CPS are labeled in blue. Homologous interactions are labeled in rose. Interactions unique to the IMPCH are labeled in green. Interactions unique to the CPS structure are labeled in purple. The XMP C2 carbonyl is colored red, and the double bond between N3 and C2 in IMP is colored cyan. (C) Sequence alignment of the human IMPCH domain (residues 1−198) with CPS (residues 937−1073) using the GCG program Gap (Genetics Computer Group, Madison, WI). Secondary structure assignments (α and *β*) correspond to those of human ATIC. Identical residues are denoted with white letters on a red background, while homologous residues are denoted with red letters on a white background.

catalysis as substitution of Tyr^{104} with an alanine or phenylalanine has little effect on the K_m, but reduces the k_{cat} by at least 1 order of magnitude ($k_{cat} = 8.6\ s^{-1}$ for the wild type, $0.136\ s^{-1}$ for the $Tyr^{104}Ala$ mutant, and $0.036\ s^{-1}$ for the

Tyr[104]Phe mutant) (*18*). Similarly, mutation of Asp[125] to alanine reduces the k_{cat} by 2 orders of magnitude with a K_m 10-fold greater than the wild-type K_m ($K_m = 0.9\ \mu M$ for the wild type, 9.1 μM for the Asp[125]Ala mutant, and 0.014 s^{-1} for the Asp[125]Ala mutant). Nevertheless, the double mutation consisting of alanine substitution for both Tyr[104] and Asp[125] still maintains a detectable k_{cat} (*18*).

Proposed Mechanism. Previous experiments on inhibitor design of the IMPCH domain and on IMPCH activity measurements, as well as consideration of the relatively facile chemical synthesis of hypoxanthine, provide important guidelines for proposing an IMPCH active site mechanism. Derivatives of AICAR with bulky substitutions at the 4-carboxamide nitrogen do not bind to the AICAR Tfase active site, but do inhibit IMPCH (*7*). These results suggest that the 4-carboxamide of AICAR adopts a different conformation in the AICAR Tfase and IMPCH active sites. AICAR, in the AICAR Tfase active site of the AICAR−XMP-bound avian ATIC structure, orients its 4-carboxamide with the nitrogen in the favored "up" position, where Phe[541] can provide a perpendicular π-hydrogen bond acceptor (*5*). Therefore, bulky additions of the 4-carboxamide nitrogen would result in steric clashes with active site residues. However, these same 4-substituted AICAR derivatives can bind to the IMPCH active site and inhibit cyclohydrolase activity, suggesting that the 4-carboxamide flips to a less favorable conformer with the oxygen in the up position so that its nitrogen is now less sterically restricted (*7*). This carboxamide reorientation is confirmed by the structural results presented here and is a consequence of the configuration of the IMPCH binding site, which provides a hydrogen bond donor (Thr[67]) and main chain hydrogen bond acceptor (Arg[64]) to the oxygen and nitrogen of the 4-carboxamide, respectively. If the 4-carboxamide were to flip so that the nitrogen would now be in the up position, unfavorable van der Waals interactions between the lone pairs of electrons from the Arg[64] main chain carbonyl and the oxygen of the 4-carboxamide would result. Similarly, the hydrogens from the Thr[67] side chain hydroxyl would sterically clash with the 4-carboxamide amino group (Figures 3C and 4A). Therefore, the IMPCH active site reorients the 4-carboxamide of FAICAR from its favored conformation in solution so that its amino nitrogen is correctly poised for nucleophilic attack on the 5-formyl group. Cyclization of 4-formyl-5-imidazolecarboxamide (FAICA) to hypoxanthine has been shown to be a facile reaction which can occur readily under weak alkali solutions, such as 0.05 N potassium bicarbonate (*61*). Thus, once the substrate is appropriately induced into a reactive configuration, the role of other residues in the IMPCH active site during catalysis would appear then to provide substrate specificity and to aid in any acid−base catalysis that would enhance IMP product formation through water elimination from the proposed intermediate (Figure 6B).

Thus, the primary role of the IMPCH active site would then appear to be to correctly orient both the FAICAR 4-carboxamide and the 5-formylamine to promote intramolecular nucleophilic attack and subsequent cyclization, and to aid in elimination of the water molecule from the proposed transitory reaction intermediate. The IMPCH active site induces a conformational rearrangement in the substrate from its favored conformation that binds to the AICAR Tfase active site to a more disfavored conformation that promotes intramolecular cyclization. Other enzymes proposed to act in a similar manner include scytalone dehydratase (*62*), glycosidases (*63*), and serine proteases (*64*). Upon binding FAICAR in the appropriate conformation, loops of residues 102−107 and 64−66 conformationally rearrange such that the side chain of Tyr[104] flips into the active site to sequester the substrate and Asp[125] from bulk solvent, and Lys[66] then comes within hydrogen bonding distance of the hydroxyl group of Tyr[104]. Reorientation of the 4-carboxamide by Arg[64] and Thr[67] most likely overcomes the primary barrier to catalysis (Figure 6A). Similarly, Ile[126] and Gly[127] position the 5-formylamine relative to the 4-carboxamide nitrogen at the appropriate tetrahedral angle for optimal nucleophilic attack. Both Ile[126] and Gly[127] are located at the N-terminus of helix α8 (Figure 4A), which can also aid in polarization of the formyl group. Once the substrate is rearranged, intramolecular nucleophilic attack on the 5-formyl group by the 4-carboxamide nitrogen is facilitated.

Competitive inhibitors of the IMPCH active site prefer electronegative substituents, such as sulfur, fluorine, or oxygen in the C2 position of the purine ring (*20*), suggesting that the reaction intermediate may have a negative charge on its C2 carbonyl during ring closure. The positioning of the C2 carbonyl at the N-terminus of helix α8 would also assist in stabilizing any negatively charged oxygen within the transition state. Similarly, the backbone amide interactions between Ile[126] and Gly[127] and the C2 carbonyl emulate the oxyanion hole of chymotrypsin-like enzymes (*65*−*67*).

Within the active site, no side chains are in the vicinity to act as a catalytic base to deprotonate the 4-carboxamide nitrogen of FAICAR prior to, or concomitant with, ring closure. The only polar moieties located within the vicinity of N1 of the purine ring in the XMP-bound IMPCH active sites are the main chain carbonyl oxygen of Arg[64] and the bound water molecule forming a bridge to Lys[137′]. The extremely weak basicity of backbone carbonyl oxygens rules it out as a potential proton acceptor, and hence, water is most likely involved in this role.

Dimerization of ATIC promotes IMPCH activity, compared to the monomer which is 6-fold less active (*44*). A solid surface representation of the monomer reveals a tunnel from the outside surface of the active site to the C2 carbonyl of XMP in the absence of Lys[137′] from the neighboring subunit in the dimer, suggesting the hydroxyl group of the FAICAR intermediate can exchange protons with bulk solvent in the monomeric form. However, in the dimer, ordering of a buried water molecule by Lys[137′] would provide a more optimal situation (Figures 3C and 4A). Lys[137′] itself is completely buried in the dimer interface, suggesting a propensity for an uncharged amino group. The bound water molecule, 3.9 Å from the N3 group of XMP, can then remove the amide proton and transiently transfer it to the Lys[137′] side chain. After ring closure, the C2 carbonyl oxygen of the transition state must then become protonated. The proton removed from the attacking 4-carboxamide nitrogen by Lys[137′] via the bound structural water molecule can then be shuffled back to protonate the C2 carbonyl. The distance between the water molecule and the C2 carbonyl of XMP is 2.6 Å, which is an optimal distance for proton transfer (Figures 3C and 4A). Deprotonation of the attacking 4-car-

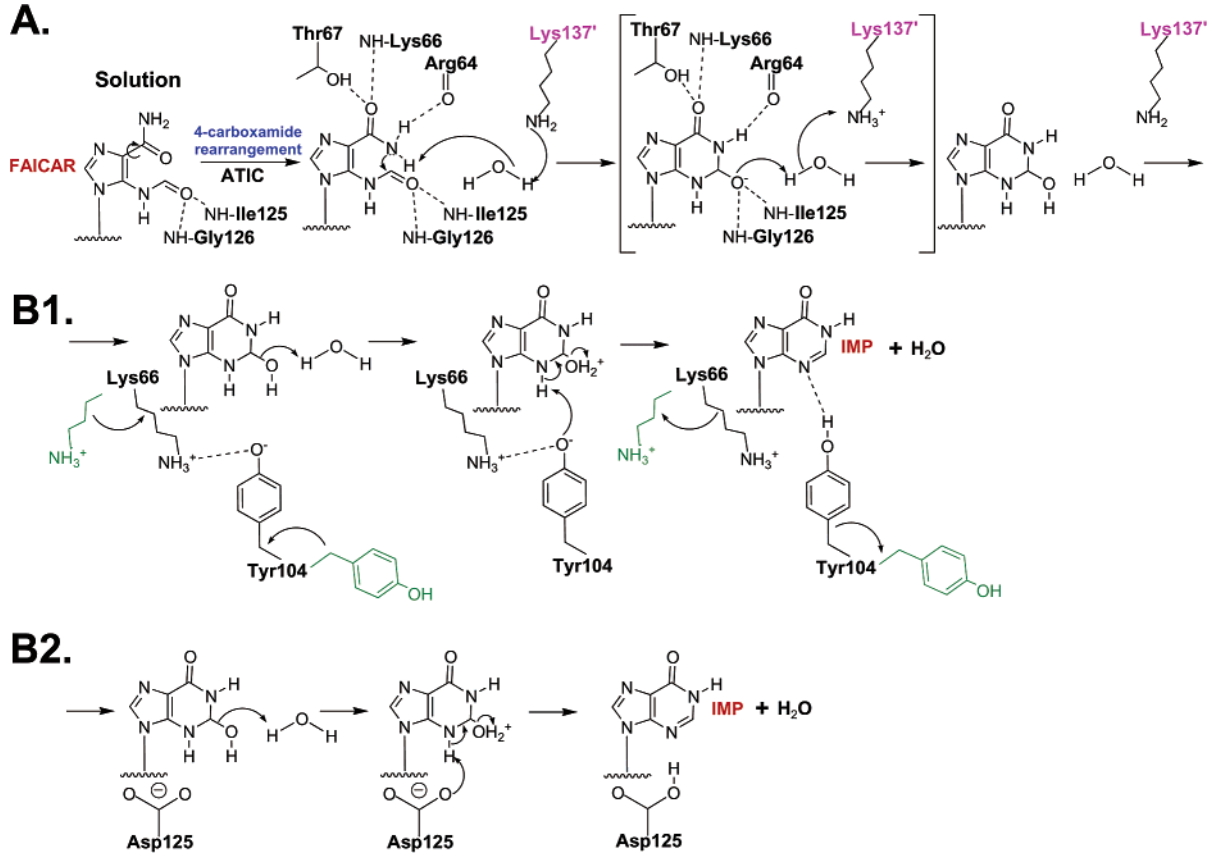

FIGURE 6: Proposed mechanism of IMP cyclization. Rearrangement of FAICAR 4-carboxamide occurs upon binding to the IMPCH active site and promotes intramolecular cyclization of FAICAR to IMP. (A) Ring closure and the first protonation step of the C2 carbonyl. Arg[64], Lys[66], and Thr[67] are responsible for the 4-carboxamide rearrangement to a disfavored conformation that promotes 4-carboxamide amino nucleophilic attack on the 5-formyl group. Backbone amides of Ile[126] and Gly[127] optimally orient the 5-formyl group for nucleophilic attack and provide an oxyanion hole to stabilize the negatively charged transition state (brackets). Lys[137′] may act as an electrophile to help remove the proton from the attacking 4-carboxamide nitrogen via the bound water molecule. The proposed first protonated intermediate (right) and second (see below) are likely to be transitory, as they involve only proton shuffling to and from a water molecule. (B1) Potential pathway for the elimination of water involving active site residues Tyr[104] and Lys[66]. Tyr[104] and Lys[66] are involved in large conformational rearrangements upon ligand binding. The positive amino group of Lys[66] is poised to stabilize the negatively charged Tyr[104] side chain in the bound form. Tyr[104] can then abstract a proton from the N3 position of the purine ring to complete IMP formation after a second protonation step that allows elimination of a water molecule. (B2) Alternate pathway for the elimination of the water leaving group involving Asp[125]. Asp[125] may act as a catalytic base and accept a proton from N3 of the intermediate. Both pathways (B1 and B2) could be utilized in the cyclization reaction and, hence, explain the mutagenesis data (*18*).

boxamide nitrogen and the first protonation of the C2 oxyanion are proposed to involve transient proton shuffling within the active site (Figure 6A). Subsequent protonation of the C2 hydroxyl reaction intermediate results in a better leaving group (water); deprotonation of N3 of the purine ring for water elimination completes the reaction sequence.

The proximity of the hydroxyl group of Tyr[104] to the amino group of Lys[66] in the ligand-bound conformation suggests the positive charge on the side chain of Lys[66] may lower the pK_a of the hydroxyl group of Tyr[104] and, hence, stabilize its deprotonated state. Tyr[104] is then poised as a catalytic base to accept the proton from N3 of the purine ring that would promote the water elimination leading to the formation of the purine double bond (Figure 6B1). Alternatively, deprotonation of the N3 group by Asp[125] and subsequent water elimination could occur if Asp[125] assumes a rotamer that would bring it closer to N3 (Figure 6B2). Two features of the Asp[125] location suggest it could be involved in the chemistry of cyclization of FAICAR; it is not conformationally restricted within the active site, and its pK_a is most likely perturbed during the movement of Tyr[104]. However, when the basicity of the side chains of Tyr[104] and Asp[125] is considered, the deprotonated form of Tyr[104] would be a stronger base than Asp[125] (Figure 6B1). The pK_a value of Tyr[104] would be decreased from its solution value of 10.5 due to the proximity of a positively charged Lys[66]. Therefore, in principle, its pK_a value could match the pK_a of 7.57 previously calculated for optimal IMPCH activity (*19*). These conclusions are supported by the mutagenesis studies in which mutation of Tyr[104] to alanine or phenylalanine greatly affects the k_{cat} (*18*).

Conversion of FAICAR to the product IMP would result in the loss of three potential hydrogen bonds, including main chain interactions with Ile[126] and Gly[127], as well as the Lys[137′] side chain interaction meditated through the active site water (Figures 3C and 4A). Loss of these hydrogen bonds would favor release of IMP from the active site and allow Tyr[104] to swing out in the open conformation of the loop of residues 102−107.

Implications of ATIC Bifunctionality. The question of why a separate IMPCH domain in ATIC is required at all arises, as the equivalent cyclization of FAICA to hypoxanthosine readily occurs in a 0.05 N bicarbonate solution to 85% completion (*61*). Many sophisticated enzymes, including

GTP cyclohydrolases I (*22*, *23*) and II (*24*, *25*), as well as methylenetetrahydrofolate dehydrogenase-cyclohydrolase (*21*), catalyze multistep reactions, including a cyclohydrolase activity, within one active site. Throughout evolutionary development why did the AICAR Tfase activity not expand to include the relatively simple cyclohydrolase ring closure transformation that completes the *de novo* purine biosynthesis pathway? The FAICAR intermediate does appear to dissociate from the ATIC surface (*10*), suggesting that the stable FAICAR intermediate may also be used in another pathway or used as a cell signaling molecule. However, no evidence currently supports such hypotheses. The IMPCH domain is most likely required to facilitate the rearrangement of the 4-carboxamide and help eliminate the water molecule. On the other hand, the bifunctionality of ATIC may simply be for regulatory purposes, as inhibition of the IMPCH activity allows catalysis of the reverse reaction of FAICAR and THF to regenerate AICAR and 10-f-THF within the AICAR Tfase domain (*7*). Several naturally occurring inhibitors of the IMPCH active site have been identified, including the feedback inhibitor XMP, suggesting that conversion of FAICAR to IMP is highly regulated. An oversupply of IMP could have some deleterious effects within a cell as it is both an inhibitor and effector molecule of several pathways. Phosphoribosylpyrophosphate (PRPP) synthetase produces PRPP, a precursor in purine, pyrimidine, and histidine biosyntheses, and is inhibited by IMP, AMP, and GMP (*68*). Similarly, IMP binds to CPS as an effector molecule in pyrimidine synthesis (*26*, *27*).

Conclusions. The crystal structures of human and avian ATIC have led to the identification of the key IMPCH active site residues that aid in catalysis of the essentially irreversible cyclization of FAICAR to IMP. A conformational change in the loop of residues 103−108 allows it to close over the active site and flip in the side chain of Tyr^{104} to sequester XMP from bulk solvent. This conformational change is coupled with a large conformational shift of the Lys^{66} side chain amino group that brings it into the proximity of the hydroxyl group of Tyr^{104}. However, the induction of a conformational change in the substrate by the enzyme appears to be probably the most important contribution of the IMPCH active site. The structure clearly shows that the backbone amide and carbonyl of Arg^{64} and Lys^{66}, respectively, as well as the side chain of Thr^{67}, orient the carboxamide of FAICAR to promote intramolecular cyclization. Similarly, backbone amides of Ile^{126} and Gly^{127} appropriately orient the 5-formyl group for attack and provide an oxyanion hole for stabilization of the transition state. The side chains of $Lys^{137'}$, Tyr^{104}, and Asp^{125}, as well as a bound water, facilitate catalysis, but may not be absolutely essential for this relatively facile cyclization reaction, consistent with a substantial diminution, but not abolishment, of the catalytic rate when these residues are mutated (*18*).

ACKNOWLEDGMENT

We thank T. Horton for excellent technical support in protein expression and purification; A. Heine and X. Dai for data collection and processing; G. P. Beardsley, D. Boger, F. Romesberg, J. Vergis, N. Larsen, R. Stanfield, and A. Heine for helpful suggestions and discussion; M. Elsliger and M. Rudolph for computational assistance; and the staff of APS 14-BM-C for beamline assistance.

REFERENCES

1. Smith, J. L. (1995) *Curr. Opin. Struct. Biol. 5*, 752−757.
2. Jackson, R. C., and Harkrader, R. J. (1981) in *Nucleosides and Cancer Treatment* (Tattersall, M. H. N., and Fox, R. M., Eds.) pp 18−31, Academic Press, Sydney, Australia.
3. Baldwin, S. W., Tse, A., Gossett, L. S., Taylor, E. C., Rosowsky, A., Shih, C., and Moran, R. G. (1991) *Biochemistry 30*, 1997−2006.
4. Rayl, E. A., Moroson, B. A., and Beardsley, G. P. (1996) *J. Biol. Chem. 271*, 2225−2233.
5. Wolan, D. W., Greasley, S. E., Beardsley, G. P., and Wilson, I. A. (2002) *Biochemistry 41*, 15505−15513.
6. Wolan, D. W., Greasley, S. E., Wall, M. J., Benkovic, S. J., and Wilson, I. A. (2003) *Biochemistry 42*, 10904−10914.
7. Wall, M., Shim, J. H., and Benkovic, S. J. (2000) *Biochemistry 39*, 11303−11311.
8. Shim, J. H., Wall, M., Benkovic, S. J., Diaz, N., Suarez, D., and Merz, K. M., Jr. (2001) *J. Am. Chem. Soc. 123*, 4687−4696.
9. Beardsley, G. P., Rayl, E. A., Gunn, K., Moroson, B. A., Seow, H., Anderson, K. S., Vergis, J., Fleming, K., Worland, S., Condon, B., and Davies, J. (1997) in *Purine and Pyrimidine Metabolism in Man* (Griesmacher, A., Ed.) pp 221−226, Plenum Press, New York.
10. Bulock, K. G., Beardsley, G. P., and Anderson, K. S. (2002) *J. Biol. Chem. 277*, 22168−22174.
11. Greasley, S. E., Horton, P., Ramcharan, J., Beardsley, G. P., Benkovic, S. J., and Wilson, I. A. (2001) *Nat. Struct. Biol. 8*, 402−406.
12. Graupner, M., Xu, H., and White, R. H. (2002) *J. Bacteriol. 184*, 1471−1473.
13. White, R. H. (1997) *J. Bacteriol. 179*, 3374−3377.
14. Zhang, Y., Desharnais, J., Greasley, S. E., Beardsley, G. P., Boger, D. L., and Wilson, I. A. (2002) *Biochemistry 41*, 14206−14215.
15. Zhang, Y., Desharnais, J., Marsilje, T. H., Li, C., Hedrick, M. P., Gooljarsingh, L. T., Tavassoli, A., Benkovic, S. J., Olson, A. J., Boger, D. L., and Wilson, I. A. (2003) *Biochemistry 42*, 6043−6056.
16. Tibbetts, A. S., and Appling, D. R. (1997) *Arch. Biochem. Biophys. 340*, 195−200.
17. Sugita, T., Aya, H., Ueno, M., Ishizuka, T., and Kawashima, K. (1997) *J. Biochem. 122*, 309−313.
18. Vergis, J., and Beardsley, G. P. (2004) *Biochemistry 43*, 1184−1192.
19. Szabados, E., Hindmarsh, E. J., Phillips, L., Duggleby, R. G., and Christopherson, R. I. (1994) *Biochemistry 33*, 14237−14245.
20. Szabados, E., Manthey, M. K., Wilson, P. K., and Christopherson, R. I. (1998) *Biochem. Mol. Biol. Int. 44*, 617−623.
21. Sundararajan, S., and MacKenzie, R. E. (2002) *J. Biol. Chem. 277*, 18703−18709.
22. Schramek, N., Bracher, A., Fischer, M., Auerbach, G., Nar, H., Huber, R., and Bacher, A. (2002) *J. Mol. Biol. 316*, 829−837.
23. Rebelo, J., Auerbach, G., Bader, G., Bracher, A., Nar, H., Hosl, C., Schramek, N., Kaiser, J., Bacher, A., Huber, R., and Fischer, M. (2003) *J. Mol. Biol. 326*, 503−516.
24. Ritz, H., Schramek, N., Bracher, A., Herz, S., Eisenreich, W., Richter, G., and Bacher, A. (2001) *J. Biol. Chem. 276*, 22273−22277.
25. Kaiser, J., Schramek, N., Eberhardt, S., Puttmer, S., Schuster, M., and Bacher, A. (2002) *Eur. J. Biochem. 269*, 5264−5270.
26. Thoden, J. B., Huang, X., Raushel, F. M., and Holden, H. M. (1999) *Biochemistry 38*, 16158−16166.
27. Thoden, J. B., Raushel, F. M., Wesenberg, G., and Holden, H. M. (1999) *J. Biol. Chem. 274*, 22502−22507.
28. Sintchak, M. D., Fleming, M. A., Futer, O., Raybuck, S. A., Chambers, S. P., Caron, P. R., Murcko, M. A., and Wilson, K. P. (1996) *Cell 85*, 921−930.
29. McMillan, F. M., Cahoon, M., White, A., Hedstrom, L., Petsko, G. A., and Ringe, D. (2000) *Biochemistry 39*, 4533−4542.
30. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
31. Matthews, B. W. (1968) *J. Mol. Biol. 33*, 491−497.
32. Collaborative Computational Project No. 4 (1994) *Acta Crystallogr. D50*, 760−763.

33. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
34. Jones, T. A., Cowan, S., Zou, J. Y., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
35. Roussel, A., and Cambillau, C. (1991) in *Silicon Graphics Geometry Partners Directory 86*, Silicon Graphics, Mountain View, CA.
36. Lawrence, M. C., and Colman, P. M. (1993) *J. Mol. Biol. 234*, 946−950.
37. Connolly, M. L. (1983) *Science 221*, 709−713.
38. Sheriff, S., Hendrickson, W. A., and Smith, J. L. (1987) *J. Mol. Biol. 197*, 273−296.
39. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
40. Hooft, R. W., Vriend, G., Sander, C., and Abola, E. E. (1996) *Nature 381*, 272.
41. Esnouf, R. M. (1997) *J. Mol. Graphics 15*, 132−134.
42. Merritt, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr. D50*, 869−873.
43. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) *Nucleic Acids Res. 28*, 235−242.
44. Vergis, J. M., Bulock, K. G., Fleming, K. G., and Beardsley, G. P. (2001) *J. Biol. Chem. 276*, 7727−7733.
45. Wu, N., and Pai, E. F. (2002) *J. Biol. Chem. 277*, 28080−28087.
46. Heroux, A., White, E. L., Ross, L. J., and Borhani, D. W. (1999) *Biochemistry 38*, 14485−14494.
47. Heroux, A., White, E. L., Ross, L. J., Davis, R. L., and Borhani, D. W. (1999) *Biochemistry 38*, 14495−14506.
48. Lesyng, B., Marck, C., and Saenger, W. (1984) *Z. Naturforsch. 39c*, 720−724.
49. Koyama, G., Nakamura, H., and Umezawa, H. (1976) *Acta Crystallogr. B32*, 969−972.
50. Anderson, A. C., O'Neil, R. H., DeLano, W. L., and Stroud, R. M. (1999) *Biochemistry 38*, 13829−13836.
51. Hampele, I. C., D'Arcy, A., Dale, G. E., Kostrewa, D., Nielsen, J., Oefner, C., Page, M. G., Schonfeld, H. J., Stuber, D., and Then, R. L. (1997) *J. Mol. Biol. 268*, 21−30.
52. Jakes, R., and Fersht, A. R. (1975) *Biochemistry 14*, 3344−3350.
53. Brick, P., Bhat, T. N., and Blow, D. M. (1989) *J. Mol. Biol. 208*, 83−98.
54. Holm, L., and Sander, C. (1993) *J. Mol. Biol. 233*, 123−138.
55. Saadat, D., and Harrison, D. H. (1999) *Struct. Folding Des. 7*, 309−317.
56. Christianson, D. W., and Lipscomb, W. N. (1989) *Acc. Chem. Res. 22*, 62−69.
57. Christianson, D. W., Mangani, S., Shoham, G., and Lipscomb, W. N. (1989) *J. Biol. Chem. 264*, 12849−12853.
58. Cho, J. H., Kim, D. H., Lee, K. J., and Choi, K. Y. (2001) *Biochemistry 40*, 10197−10203.
59. van Aalten, D. M., Chong, C. R., and Joshua-Tor, L. (2000) *Biochemistry 39*, 10082−10089.
60. Manjasetty, B. A., Powlowski, J., and Vrielink, A. (2003) *Proc. Natl. Acad. Sci. U.S.A. 100*, 6992−6997.
61. Shaw, E. (1950) *J. Biol. Chem. 185*, 439−447.
62. Zheng, Y. J., Basarab, G. S., and Jordan, D. B. (2002) *Biochemistry 41*, 820−826.
63. Vasella, A., Davies, G. J., and Bohm, M. (2002) *Curr. Opin. Chem. Biol. 6*, 619−629.
64. Case, A., and Stein, R. L. (2003) *Biochemistry 42*, 3335−3348.
65. Liao, D. I., Breddam, K., Sweet, R. M., Bullock, T., and Remington, S. J. (1992) *Biochemistry 31*, 9796−9812.
66. Barrette-Ng, I. H., Ng, K. K., Mark, B. L., Van Aken, D., Cherney, M. M., Garen, C., Kolodenko, Y., Gorbalenya, A. E., Snijder, E. J., and James, M. N. (2002) *J. Biol. Chem. 277*, 39960−39966.
67. Reiling, K. K., Krucinski, J., Miercke, L. J., Raymond, W. W., Caughey, G. H., and Stroud, R. M. (2003) *Biochemistry 42*, 2616−2624.
68. Stryer, L. (1995) *Biochemistry*, 4th ed., W. H. Freeman and Co., New York.

BI030162I