Structural Insights into the Avian AICAR Transformylase Mechanism†

Dennis W. Wolan,‡ Samantha E. Greasley,‡,§ G. Peter Beardsley,|| 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, and Department of Pharmacology and Pediatrics, Yale University School of Medicine, New Haven, Connecticut 06520

Received August 1, 2002; Revised Manuscript Received September 25, 2002

ABSTRACT: ATIC encompasses both AICAR transformylase and IMP cyclohydrolase activities that are responsible for the catalysis of the penultimate and final steps of the purine de novo synthesis pathway. The formyl transfer reaction catalyzed by the AICAR Tfase domain is substantially more demanding than that catalyzed by the other folate-dependent enzyme of the purine biosynthesis pathway, GAR transformylase. Identification of the AICAR Tfase active site and key catalytic residues is essential to elucidate how the non-nucleophilic AICAR amino group is activated for formyl transfer. Hence, the crystal structure of dimeric avian ATIC was determined as a complex with the AICAR Tfase substrate AICAR, as well as with an IMP cyclohydrolase inhibitor, XMP, to 1.93 Å resolution. AICAR is bound at the dimer interface of the transformylase domains and forms an extensive hydrogen bonding network with a multitude of active site residues. The crystal structure suggests that the conformation of the 4-carboxamide of AICAR is poised to increase the nucleophilicity of the C5 amine, while proton abstraction occurs via His^{268} concomitant with formyl transfer. Lys^{267} is likely to be involved in the stabilization of the anionic formyl transfer transition state and in subsequent protonation of the THF leaving group.

Aminoimidazole-4-carboxamide ribonucleotide (AICAR)[1] transformylase/inosine monophosphate (IMP) cyclohydrolase (ATIC) (*1*) is a bifunctional enzyme responsible for the final two catalytic steps of the de novo purine biosynthesis pathway. AICAR transformylase (AICAR Tfase) catalyzes the penultimate step by transferring a formyl group from the cofactor *N*-10-formyl-tetrahydrofolate (10-f-THF) to the substrate AICAR. This reaction produces the stable intermediate 5-formyl-AICAR (FAICAR) and the byproduct THF (Figure 1A). IMP cyclohydrolase (IMPCH) performs the cyclization of FAICAR to the final product of the pathway IMP with the elimination of water.

Enzymes in this pathway have been targeted for development of anticancer compounds as many tumor cells rely more heavily on de novo synthesis rather than the salvage pathway for their production of purines (*2*). Glycinamide ribonucleotide (GAR) transformylase (GAR Tfase) also requires the same 10-f-THF cofactor for the formyl transfer in the third reaction of the pathway. Extensive crystallographic and kinetic studies of GAR Tfase have resulted in the design of several folate-based analogues as potential chemotherapeutic agents (*3*−*7*). Correspondingly, ATIC also poses as a potential target for the development of antineoplastic drugs. Despite some local sequence correlation that initially suggested some close similarities in the GAR Tfase and AICAR Tfase active sites, there is no significant overall sequence or structure homology between ATIC and GAR Tfase; residues in GAR Tfase critical for formyl transfer do not have any functional relevance in AICAR Tfase (*4*, *8*). The ATIC formylation reaction in itself is of major interest because of the inherent difficulty in the formyl transfer to the relatively non-nucleophilic amine on the AICAR substrate as compared to that of its GAR transformylase counterpart (*9*−*11*).

Development of therapeutic compounds has focused on design of novel antifolate inhibitors that are either tailored for the inhibition of a specific target or designed for general inhibition of all folate-dependent enzymes. Examples include 6*R*-dideazatetrahydrofolate (DDATHF, Lometrexol) (*12*, *13*), which specifically targets GAR Tfase and the multitargeted antifolate LY231514 (*14*), which inhibits enzymes including GAR Tfase, dihydrofolate reductase (DHFR), and thymidylate synthase (TS). Methotrexate (MTX) is the only clinically available antifolate and has been widely used in treatment of various diseases for over half a century. MTX is a specific inhibitor of DHFR and results in the accumulation of cellular dihydrofolate polyglutamates. However, it also inhibits other folate-dependent enzymes in the methionine, thymidylate, and de novo purine biosynthesis pathways. More recently, the dihydrofolate polyglutamates were shown to be potent

† This work was supported by the National Institute of Health Grant PO1 CA63536 (I.A.W. and G.P.B.) and National Science Foundation and Skaggs predoctoral fellowships (D.W.W.). This is publication 15188-MB from The Scripps Research Institute.

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

‡ The Scripps Research Institute.

§ Present address: Pfizer La Jolla, 10770 Science Center Drive, San Diego, CA 92121.

|| Yale University School of Medicine.

[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-formyl-tetrahydrofolate; FAICAR, 5-formyl-AICAR; IMPCH, IMP cyclohydrolase; GAR, glycinamide ribonucleotide; GAR Tfase, GAR transformylase; DDATHF, 6*R*-dideazatetrahydrofolate; DHFR, dihydrofolate reductase; TS, thymidylate synthase; MTX, methotrexate; XMP, xanthosine 5′-monophosphate; GMP, guanosine monophosphate.

10.1021/bi020505x CCC: $22.00
Published on Web 12/10/2002

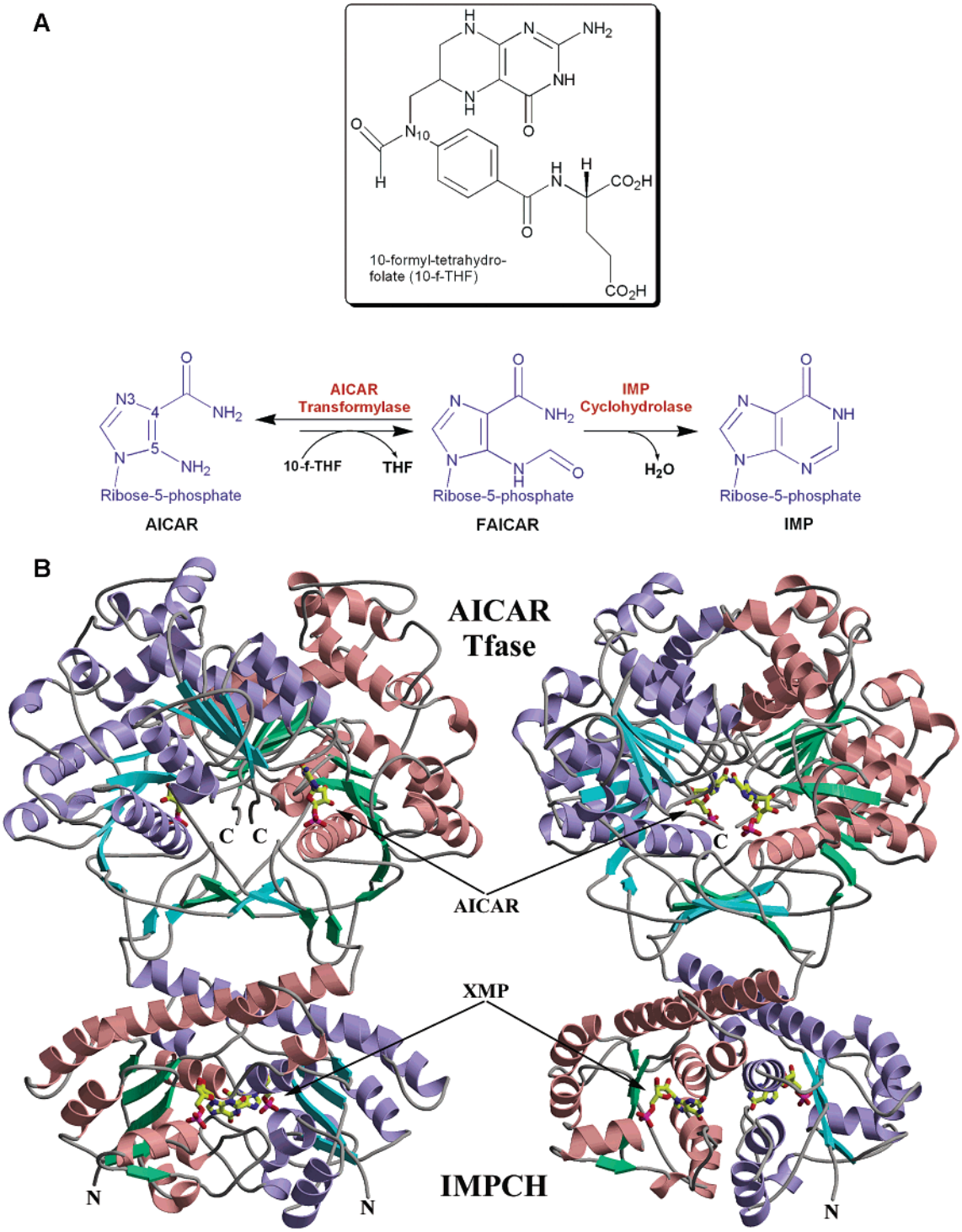

FIGURE 1: Reaction and crystal structure of ATIC. (A) Formyl transfer reaction by ATIC and structure of the folate cofactor (10-f-THF). (B) Two rotated views of the ATIC dimer. Monomer A is in blue helices (ribbons) and cyan β-strands (arrows) with monomer B in rose helices and green β-strands. Connecting residues for both monomers are gray. The bound AICAR and XMP molecules are depicted in a ball-and-stick representation (carbon, yellow; oxygen, red; nitrogen, blue; and phosphate, purple). The N- and C-termini are labeled accordingly.

inhibitors of AICAR Tfase (*15*) resulting in an anti-inflammatory response (*16*, *17*) and implicating ATIC as a target for development of nonsteroidal anti-inflammatory drugs (NSAIDs). Thus, purine nucleotide-based inhibitor design toward the inhibition of either or both the AICAR Tfase and the IMPCH activities of ATIC would provide potential leads for antineoplastic and anti-inflammatory agents.

The crystal structure of ATIC in complex with the AICAR Tfase substrate AICAR and the IMPCH purine-based inhibitor xanthosine 5′-monophosphate (XMP) (K_i = 0.12 μM) was determined to identify the active site location and the key catalytic residues (Figure 1B). Such structural insights could provide a significant advance in unraveling the mechanism of catalysis and aid in the iterative design of potent inhibitors. Heightened interest in the elucidation of the ATIC mechanism stems from the low nucleophilicity of the 5-amino group of the AICAR substrate and that formylation by AICAR Tfase favors the AICAR and 10-f-THF substrates rather than the FAICAR and THF products (*10*,

11) (Figure 1A). XMP is a naturally occurring intermediate in the reaction sequence following ATIC in which IMP is converted to guanosine monophosphate (GMP). XMP is a potential feedback inhibitor of ATIC, thus reducing IMP production when low levels of L-glutamine limit conversion of XMP to GMP (*15*). The IMPCH active site was previously identified in the apo avian ATIC crystal structure (*18*) because of the presence of a bound purine nucleotide ligand acquired during enzyme purification, which we later identified as XMP (Vergis et al., unpublished results). Thus, in this paper, we will focus only on the identification and structural analysis of the AICAR Tfase active site. Several key residues appear to be involved in the conversion of AICAR to FAICAR, and the AICAR-bound complex provides a structural platform to decipher the mechanism of substrate activation and formyl transfer.

Previous mechanistic studies have proposed that the 4-carboxamide oxygen of AICAR is essential for formyl transfer and is directly involved in the deprotonation of the 5-amino group (*10*, *19*) (Figure 1A). Bulky substitutions at the 4-carboxamide nitrogen inhibit binding to AICAR Tfase, suggesting that the AICAR 4-carboxamide oxygen is oriented toward its 5-amino group (Figure 1A). Also, substitution of the 4-carboxamide oxygen by a sulfur reduced the k_{cat} by 4-fold (*10*), whereas other substitutions result in a complete loss of activity. The potential for 4-carboxamide to act as a hydrogen acceptor was also proposed to be enhanced by the proximity to Lys^{267} and His^{268}, where His^{268} would facilitate the role of Lys^{267} as a general acid catalyst to protonate the N3 ($pK_a = 3.23$) position of AICAR. Nucleophilic attack on N3 was proposed to enable the 4-carboxamide oxygen to simultaneously deprotonate the 5-amino group upon its nucleophilic attack on the 10-f-THF. Mutational analysis on Lys^{267} and His^{268} have confirmed the absolute requirement for these active site residues in catalysis (*8*, *19*, *20*). Mutation of His^{268} to Ala or Gln results in complete loss of activity (*8*, *19*), while replacement of Lys^{267} to Arg or Ala results in either a 150-fold decrease in k_{cat} (*19*) or in complete ablation of activity (*20*), while maintaining wild-type substrate binding ($K_m = 16.8$ μM for human ATIC). However, the crystal structure of the AICAR Tfase active site bound with AICAR presented here structurally corroborates the proposed 4-carboxamide orientation but disagrees with several other aspects of previously proposed mechanisms for AICAR Tfase. The structural data then provide a new framework onto which new mechanistic insights into the AICAR formyl transfer reaction can be explored.

MATERIALS AND METHODS

Protein Expression and Purification. Avian ATIC cDNA was cloned into a pET28a vector (Novagen, Inc.) and transformed into *Escherichia coli* BL21.DE3 cells (Novagen, Inc.) for overexpression and was a kind gift from Dr. Stephen Benkovic as previously described (*18*). The *E. coli* transformant was grown in NZCYM media (Gibco BRL) at 37 °C to an OD_{600} of 0.6−0.7 and induced with 1 mM isopropyl-β-D-thiogalactopyranoside (Gibco BRL) for 4 h. The cells were harvested by centrifugation at 3500 × g for 15 min at 4 °C. Pellets were used immediately or frozen at −80 °C.

Cell pellets (approximately 15 g wet weight) were resuspended in 100 mL of ice cold 25 mM sodium phosphate (pH 7.5), 150 mM NaCl, 5 mM EDTA, and 10% glycerol plus two Complete Protease Inhibitor tablets (Boerhinger Mannheim) to which 1 mg/mL lysozyme was added. Cells were then sonicated at 4 °C at 20% power for three cycles of 1 min using a Virtis 475 Cell Disrupter. The cell lysate was clarified by centrifugation at 20 000 × g at 4 °C for 60 min, then dialyzed overnight at 4 °C against 4 L of buffer A (50 mM sodium phosphate (pH 8.0), 150 mM NaCl, 10 mM imidazole, and 5 mM β-mercaptoethanol) in Snakeskin Dialysis Tubing (Pierce) with a molecular weight cutoff (MWCO) of 10 000 Da. The dialyzed material was transferred to a 250 mL flask containing 20 mL of a 1:1 Ni−NTA (Qiagen) slurry and placed on a rotator for 2 h at 4 °C. The Ni−NTA beads were poured into a 2.5 × 10 cm column and washed with 10 column volumes of buffer A. ATIC was eluted using buffer B (50 mM sodium phosphate (pH 8.0), 150 mM NaCl, 250 mM imidazole, and 5 mM β-mercaptoethanol). Fractions were analyzed by SDS−PAGE. Fractions containing ATIC were pooled and concentrated using Millipore Ultrafree-15 devices with a MWCO of 10 000 Da. ATIC was next passed over a 2.6 × 90 cm Superdex 200HR column (Amersham Pharmacia Biotech) to remove any impurities that coeluted in the Ni−NTA purification. The column was equilibrated with buffer C (25 mM Tris (pH 7.4), 150 mM NaCl, 50 mM KCl, 5 mM DTT, and 5 mM EDTA) prior to loading the sample. Peak fractions were again analyzed by SDS−PAGE, pooled, and concentrated using Millipore Ultrafree-15 devices to a final protein concentration of 0.1−0.4 mg/mL. The N-terminal His_6 tag was not removed via thrombin cleavage as the tagged ATIC crystallized readily. ATIC can be stored at 4 °C indefinitely.

XMP and AICAR were added to the ATIC protein prior to concentration and crystallization in a 10-fold molar excess, heated in a 37 °C water bath for 30 min to prevent precipitation of sample, and incubated overnight at 4 °C. All samples were subsequently concentrated to approximately 10 mg/mL using Millipore Ultrafree-15 devices and stored at 4 °C for crystallization experiments.

Crystallization and Data Collection. Protein crystals were grown at 22 °C by sitting drop-vapor diffusion method by mixing equal volumes of avian ATIC (10 mg/mL) with a reservoir solution consisting of 14% (w/v) PEG 8000, 0.2 M imidazole (pH 7.2), and 5 mM DTT. Data were collected to 1.93 Å resolution on a single, flash-cooled crystal at −173 °C in a cryoprotectant that consisted of well buffer and 15% MPD on beamline 14-BM-C at The Advanced Photon Source (APS) (Argonne, IL). Data were processed to 1.93 Å with the DENZO−SCALEPACK suite (*21*).

Structure Determination and Refinement. Molecular replacement was performed with AMoRe (*22*) using apo avian ATIC (*18*) (PDB entry 1G8M) as the search model. The rotation search (15−4 Å) identified two peaks with correlation coefficients of 20.5 and 13.7 (noise peak = 6.9), consistent with an ATIC dimer in the asymmetric unit ($V_m = 2.4$ Å^3 Da^{-1}) (*23*). Following the translational search, the correlation coefficient was 55.3, and the R_{value} was 38.8%. Further rigid body refinement against the maximum likelihood target (*24*) in CNS (*25*) yielded an R_{cryst} of 33.7% (R_{free} of 35.3%).

The initial electron density maps clearly identified AICAR and XMP within their respective active sites. Model inter-

Table 1: Data Processing and Refinement Statistics

Data Processing	
resolution range (Å) (outer shell)	30−1.93 (1.96−1.93)
unique reflections	91184 (4619)
completeness (%)	98.2 (99.6)
redundancy	3.9 (3.4)
R_{sym} (%)[a]	4.4 (41.9)
average $I/\sigma(I)$	25.3 (2.4)
Refinement	
resolution range (Å) (outer shell)	29.2−1.93(1.98−1.93)
no. of reflections[b]	74522 (4809)
R_{cryst} (%)[c]	20.6 (29.4)
R_{free} (%)[c]	24.4 (33.9)
protein atoms/waters	9020/513
CV coordinate error[d] (Å)	0.18
rmsd bonds (Å)/angles (deg)	0.009/1.3
$\langle B \rangle$ all atoms/waters/substrates (Å^2)	51.9/49.1/51.4
Ramachandran Statistics (%)	
most favored	92.4
additional allowed	7.6
generously allowed	0.0
disallowed	0.0

[a] $R_{sym} = 100\sum_h\sum_i|I(h)_i - \langle I(h)\rangle|/\sum_h\sum_i I(h)_i$, where $I(h)_i$ is the ith measurement of reflection h, and $\langle I(h)\rangle$ is the average measurement value. [b] Reflections with $I > 0$ 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 10% test set structure factors. [d] Cross-validated (CV) coordinate errors from σ_A plots.

pretation involved 18 cycles of rebuilding and refinement including simulated annealing, conjugate-gradient minimization, restrained individual temperature factors, and a bulk solvent correction. Two-fold noncrystallographic symmetry (NCS) was applied for the initial four rounds of refinement then released to accommodate differences within the individual monomers. Water molecules were automatically positioned by CNS using a 3σ cutoff in $F_o - F_c$ maps and manually inspected. After convergence in CNS, as judged by monitoring R_{free}, the final three rounds of refinement were performed using REFMAC5, including TLS refinement of four domains (*22*) that decreased R_{free} by 3%, to account for anisotropic displacement of the individual monomers. The model was manually fitted to σ_A-weighted $3F_o - 2F_c$ maps (*26*) with O (*27*). The final refined structure (Table 1) consists of all residues in each monomer, except 1−3 and the 20 additional N-terminal His_6 tag and thrombin cleavage site.

The model was analyzed with PROCHECK (*28*), WHATCHECK (*29*), CNS (*25*), and CCP4 (*22*). S_c coefficients (*30*) and buried surface areas were calculated with SC (*22*) and MS (*31*) using 1.7 and 1.4 Å probes, respectively. Hydrogen bonds and van der Waals interactions were identified with CONTACSYM (*32*) with default parameters. Root mean square deviation (rmsd) calculations on the complexed and apo avian ATICs were carried out with PROFIT (*43*). Superposition of monomer A onto monomer B gave an rmsd of 0.44 Å for the main chain and 0.67 Å for all atoms, whereas the corresponding values were 0.31 and 0.60 Å for the IMPCH domain (residues 4−199) and 0.37 and 0.63 Å for the AICAR Tfase domain (residues 200−593). Figures 1−4 and 6 were created with Bobscript (*33*) and rendered with Raster3D (*34*). Coordinates and structure factors have been deposited without hold in the PDB (*35*), with accession code 1M9N[2].

RESULTS

Avian ATIC comprising residues 1−593 was expressed in *E. coli* and was purified by gel filtration as a noncovalent dimer. XMP and AICAR were added in 10-fold molar excess to ATIC overnight and concentrated to approximately 10 mg/mL for crystallization. The ATIC complex crystallized in monoclinic space group $P2_1$ (a = 56.48 Å, b = 107.88 Å, c = 103.86 Å, and β = 91.2°), and the structure was determined to 1.93 Å resolution by molecular replacement (MR) with apo avian ATIC (PDB entry 1G8M) as the initial MR model (*18*). The final refined model of dimeric ATIC in complex with AICAR and XMP includes residues 4−593 of both monomers, AICAR and XMP bound in their respective active sites, two potassium ions, and 513 water molecules. The final R_{cryst} and R_{free} values are 20.6 and 24.4%, respectively, with 92.4% of the residues in the most favored region of the Ramachandran plot (Table 1).

The ATIC monomer is composed of the IMPCH (residues 4−199) and AICAR Tfase (residues 200−593) domains. The domain fold topology of the complexed ATIC is conserved with the previously described apo ATIC (*18*). The IMPCH domain, responsible for the cyclohydrolase activity, is composed of a Rossmann-like fold consisting of a five-stranded parallel β-sheet surrounded by α-helices. The AICAR Tfase domain encompasses three structural subdomains with the two major domains termed domain 2 (residues 200−374) and domain 4 (residues 375−593) and a third smaller domain of 64 residues (residues 469−532) inserted into domain 4 (*18*). Domains 2 and 4 both have a primarily parallel five-stranded β-sheet with one antiparallel strand (strand 1) surrounded by three α-helices (B, C, and D) on one side and helix A on the other (Figure 2); a potassium binding site is located within the loop between helix A and strand 1 of domain 4 (*18*) (Figure 2). Domain 3 is composed of four α-helices and two short β-strands and maintains a unique topology (*18*). The ATIC dimer is assembled from two highly intertwined monomers with an extensive buried surface area of 4956 Å^2 (using a 1.4 Å probe). An eight-stranded β-sheet bridges the IMPCH and AICAR Tfase domains, which negates any possibility of FAICAR tunneling between active sites.

The crystal structure of the complexed avian ATIC varies little from the apo avian ATIC structure (rmsd of 0.59 and 0.88 Å for monomers A and B). The equivalent comparison for the individual IMPCH domains and the AICAR Tfase domains (IMPCH monomer A = 0.38 Å, IMPCH monomer B = 1.03 Å, AICAR Tfase monomer A = 0.56 Å, and AICAR Tfase monomer B = 0.67 Å) indicate that the largest real difference occurs in IMPCH monomer B; the apo avian ATIC monomer already has a bound nucleotide, similar to XMP bound here in monomer A but not monomer B. Thus, significant side-chain and main-chain rearrangements occur in the IMPCH domain upon binding XMP. However, very little rearrangements are observed in the AICAR Tfase domain upon binding AICAR, where the largest main-chain deviation (rmsd 3.0 Å) is within subdomain 3, which is displaced as a rigid body and has an overall increased thermal displacement as evidenced by slightly increased B values

[2] The coordinates have been deposited with the Protein Data Bank with ID code 1M9N.

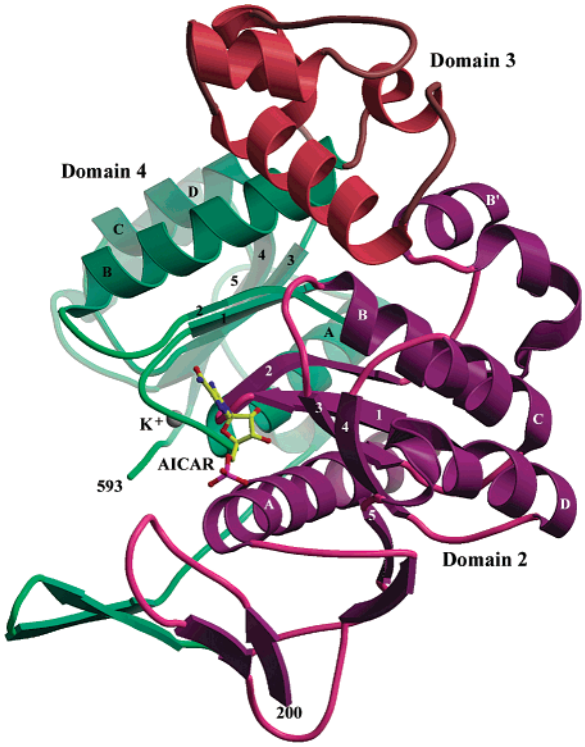

FIGURE 2: AICAR Tfase monomer active site location and relation to subdomains. Domain 2 (residues 200−374) is purple, domain 3 (residues 469−532) is red, and domain 4 (residues 375−468 and 533−593) is green. The bound AICAR molecule is represented as a ball-and-stick (coloring as in Figure 1B). Helices and β-strands are labeled as previously described (*18*).

(60.1 vs 51.9 Å^2 for ATIC). The positional variation of subdomain 3 in AICAR Tfase is likely related to crystal packing as crystal contacts surround the entire subdomain.

Initial $F_o - F_c$ maps revealed clearly interpretable electron density within the AICAR Tfase active sites depicting a bound nucleotide. AICAR precisely fit the electron density with a 3′-endo sugar pucker conformation (Figure 3A). Residues from both monomers compose each active site supporting biochemical data that show that the AICAR Tfase domain is only active in the dimeric form (*36*). Superposition of the bound monomers revealed no significant differences in the location, position, and conformation of AICAR or in the active site residue conformations. Water molecules that directly interact with the bound substrate AICAR are conserved in both monomers. Furthermore, the AICAR location and interaction with active site residues are maintained in the folate-bound complexes that are currently in refinement in our laboratory (Wolan et al., unpublished results and Cheong et al., unpublished results).

The AICAR molecule is shielded from the solvent with 84% of the molecule being buried by ATIC side chains. AICAR's phosphate group is heavily anchored to one side of the N-terminus of α-helix A (Figures 2 and 3A). In this orientation, the phosphate does not take full advantage of the dipole interaction from the helix, as frequently seen in phosphate binding sites where the phosphate is directly positioned atop the N-terminus of the helix. Side-chain interactions are provided by Tyr^{209} and Arg^{208}, as well as by Arg^{589} from the adjacent subunit of the dimer (Figure 3B), whereas main-chain hydrogen bonds from Ile^{239}, Gly^{210}, and Asn^{240} to the phosphate are mediated via ordered water molecules. Side-chain and main-chain contacts from Asp^{340} and Gly^{317}, respectively, provide specificity for ribose ring oxygens at positions 2 and 3. Asp^{243} and Asn^{240} position an ordered water molecule that also interacts with O2 of the ribose sugar. The AICAR nucleotide moiety is encased by a shell of water molecules ordered by the side chains of Arg^{452} and Asn^{432} and the main chain of Phe^{542} of the opposing subunit (Figure 3A). The aromatic ring of Phe^{542} also provides a perpendicular π hydrogen bond acceptor for the amino group of the carboxamide attached to the C4 position of AICAR (*37*). The His^{268} imidazole forms the only direct hydrogen bond interaction with the AICAR nucleotide through close interaction with the reactive amino group attached to C5. Density for Lys^{267} is well-defined within the active site with the side-chain position facing away from the AICAR substrate but adjacent to the folate binding site (Wolan et al., unpublished results and Cheong et al., unpublished results). The Lys^{267} side chain does not appear to be sterically hindered from being able to adopt other rotamer conformations in the active site when folate is bound.

DISCUSSION

The crystal structure of avian ATIC in complex with AICAR and XMP provides substantial new insights into the overall mechanism of this bifunctional enzyme. In this analysis, we focus on the AICAR Tfase domain as the substrate-bound structure provides a first glimpse of how AICAR is bound in the active site. Despite performing similar reactions, the AICAR Tfase domain of ATIC has no significant sequence or structural homology to GAR Tfase and is consistent with a previous mutational analysis (*8*) that showed that a catalytic trio of residues responsible for formylation in GAR Tfase (*4*, *38*) does not perform the equivalent role in AICAR Tfase. Therefore, ATIC has a novel site for formyl transfer. The substrate-bound avian ATIC crystal structure now reveals that several residues are likely to play a key role in the activation of the 5-amino group of AICAR for formyl transfer from 10-f-THF. Although the folate cofactor is not present in this structure, it is unlikely that the AICAR location and orientation will change significantly upon folate binding based on other folate-bound structures currently in refinement (Wolan et al., unpublished results and Cheong et al., unpublished results). Thus, here we will discuss the similarity and differences with previously proposed mechanisms for the formyl transfer reaction based on these new structural data that provide a clear view of the substrate-bound conformation and potential roles for key active site residues.

Orientation of the AICAR Carboxamide. The orientation of the 4-carboxamide of AICAR cannot be unambiguously determined based on its electron density alone even at 1.93 Å resolution. Only two possible conformers are consistent with the density with either the amino nitrogen or carbonyl oxygen of the 4-carboxamide pointing upward away from the imidazole ring (Figure 3A). As expected, refinement also results in this ambiguity in the orientation for the carboxamide. However, the environment around the carboxamide favors one conformer over the other because of potential hydrogen bonding interactions of AICAR with the ATIC. Combined with the small molecule crystal structure of AICAR (*39*) and the relative activity of 4-carboxamide substitutions in the AICAR transformylase reaction (*10*), the

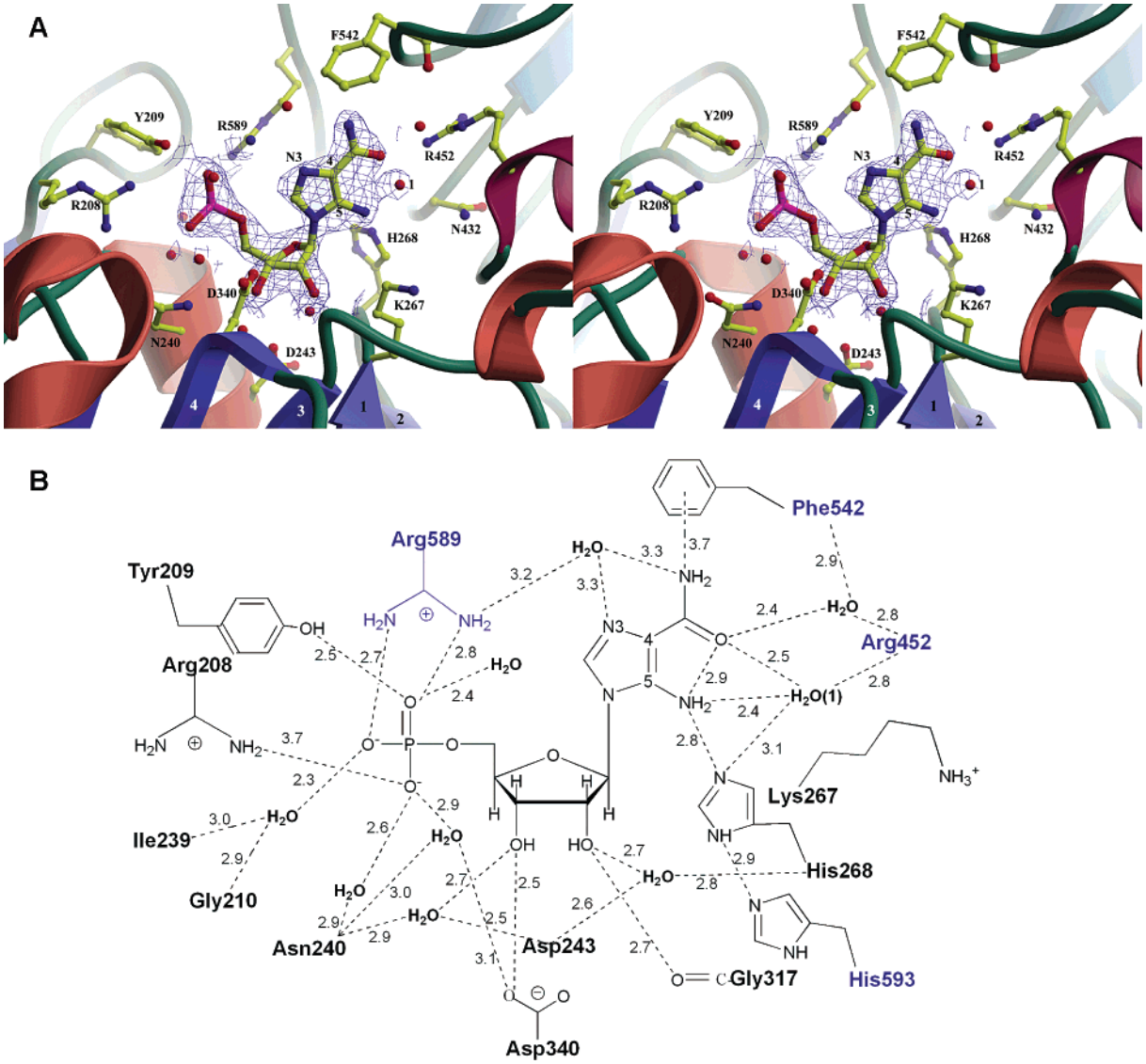

FIGURE 3: AICAR Tfase active site. (A) Stereoview of the residues that make up the AICAR binding site with electron density surrounding AICAR. Residues are represented in ball-and-stick with atoms colored according to Figure 1B. AICAR is surrounded by a blue $2F_o - F_c$ density map contoured at 1.2σ. (B) Schematic representation of the hydrogen bonding network and corresponding distances within the active site between AICAR and AICAR Tfase. Interacting residues from the opposing subunit are labeled in blue. Side chains are depicted for only those residues that provide direct interactions to AICAR.

most likely conformation is with the carboxamide nitrogen pointing upward away from the imidazole ring (Figure 3A). The small molecule crystal structure of AICAR depicts the amino nitrogen of the 4-carboxamide in the favored up position (*39*), as the 4-carboxamide amino hydrogens can interact with the lone pairs of electrons of N3 in the imidazole ring, whereas the carboxamide oxygen can hydrogen bond to the C5 amino group. Another option to alleviate potential destabilizing interactions within the substrate would be if the 4-carboxamide were to rotate around the C5−C bond by 90° so that it was now perpendicular rather than parallel to the imidazole ring. However, this orientation would eliminate energetically favorable conjugation with the imidazole ring and is definitely not consistent with the electron density that clearly shows a planar molecule, and therefore, an inplane carboxamide. Despite a paucity of interacting residues with the up group of the 4-carboxamide, Phe[542] potentially provides a perpendicular π hydrogen bond acceptor for the 4-carboxamide nitrogen of AICAR (*37*) that would favor the preferred nitrogen up orientation. In the opposite conformer, the Phe[542] interaction with the oxygen lone pairs of electrons would not provide a favorable electrostatic interaction. Thus, we concur that the enzyme itself helps orient the carboxamide such that it can take advantage of internal hydrogen bonding stabilization of the possible tautomeric forms at the substrate itself to promote amine activation and formyl transfer (*10*, *19*).

Activation of the AICAR Amino Group. The free 5-amino group of AICAR is extremely non-nucleophilic because of conjugation and delocalization of the lone pairs of electrons, suggesting that it must be activated in order for the formyl transfer reaction to proceed. Partial charge redistribution or hydrogen bonding that would lead eventually to complete removal of a proton would make the amino group more nucleophilic for attack on the 10-f-THF. The substrate-bound ATIC structure suggests that two potential key players could facilitate hydrogen sharing, and subsequent removal during formyl transfer, are the 4-carboxamide of AICAR and His[268] (Figure 3B). The 4-carboxamide being only 2.9 Å from the 5-amino group could certainly increase the amine nucleo-

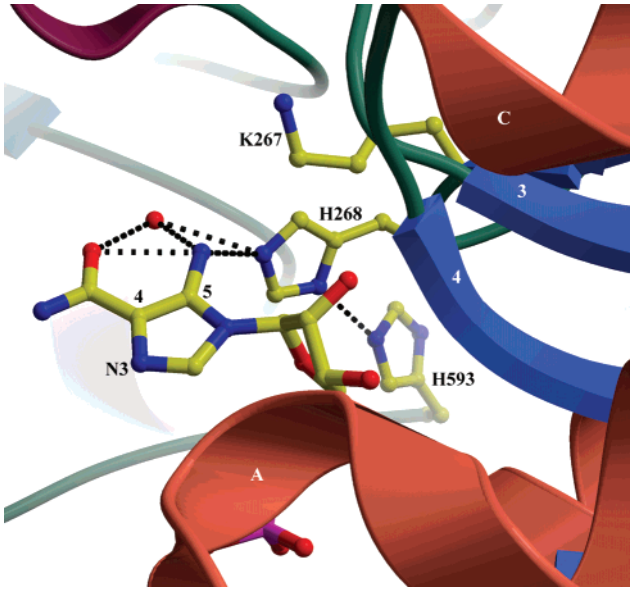

FIGURE 4: Potential role of His[268] as a base. The hydrogen bonding network around His[268] and AICAR. Depicts the catalytic residues potentially responsible for the activation of the 5-amino group of AICAR. Hydrogen bonds are shown in black dashed lines. This orientation illustrates the potential of proton shuttling to solvent via His[268] and His[593].

philicity by the sharing of a proton (*19*). Upon formylation, the proton could be accepted by the 4-carboxamide. The Benkovic group has further proposed that this same proton could be transferred to the N10 leaving group of the folate (*19*). However, the proposed transition state would proceed via an energetically unfavorable four-member ring. Instead, we favor a major role for His[268] in hydrogen bonding to the amine and for subsequent proton abstraction.

Shim et al. incorrectly reported that Lys[267] and His[268] are located above the imidazole ring, where they could assist in reducing the energy barrier of the reaction by protonating N3 (pK_a = 3.23), thus assisting the simultaneous deprotonation of the 5-amino group by the 4-carboxamide oxygen and nucleophilic attack on the 10-f-THF. Instead, His[268] is located on the opposite side of the substrate within hydrogen bonding distance of the 5-amino group of AICAR (Figure 3A). The only possible ATIC interaction that could occur with N3 is with the side chain of Arg[589] from the opposite subunit, but this distance is 4.2 Å, too long for a hydrogen bond, but within a reasonable distance for a longer range electrostatic effect. However, the primary role of Arg[589] would appear to be as a ligand to orient the phosphate group of AICAR.

His[268] and Lys[267] are conserved in ATIC throughout various species and are essential for catalysis (*8*, *19*, *20*). On the basis of pH profile kinetics that provides more definitive evidence for direct involvement of either His[268] or Lys[267] in the rate-limiting step (*19*), there is no pH dependence in the range from 6 to 10.5. However, the close vicinity of a tightly bound water (#1), His[593], and the positive charge of Lys[267] could suppress the pK_a of His[268] and aid in the eventual deprotonation of the 5-amino by His[268] concomitant with nucleophilic attack of the 5-amino group on 10-f-THF. The abstracted proton on His[268] could then easily be shuttled out to the solvent via the interaction with His[593] (Figures 4 and 5); an ordered water molecule (#1) within the active site (Figure 3A,B) bound to the 5-amino group could also facilitate rapid proton shuttling with His[268].

Role of Lys[267] in the Formyl Transfer Reaction. Lys[267], in an extended conformation, is likely to be involved in the stabilization of the oxyanion transition state during formyl transfer. It is also possible that the Lys[267] side chain could also transfer a proton to the N10 of THF; rapid proton exchange with Lys[267] could proceed via either His[268] or from a water molecule (Figure 5). Formylation of Lys[267] by 10-f-THF as an intermediary state and subsequent transfer onto AICAR is also unlikely as evidenced by previous experiments (*10*, *40*). Thus, Lys[267] is well-poised to aid in the formyl transfer reaction by stabilizing the transition state, as well as providing a possible conduit for proton transfer to the N10 leaving group, either directly or via water molecules or His[268].

FAICAR Intermediate. The absence of 10-f-THF throughout purification and crystallization obviates the possibility of the intermediate FAICAR being present within either of the ATIC active sites. However, FAICAR could be surprisingly easily modeled into the AICAR Tfase active site using density for the substrate and bound water to obtain a possible indication of how the product may be located after formylation of AICAR (Figure 6). Replacement of the tightly bound water molecule (#1) hydrogen bonded to the 5-amino position of AICAR with the formyl group transferred from 10-f-THF leads to an anticipated nonplanar FAICAR conformation with

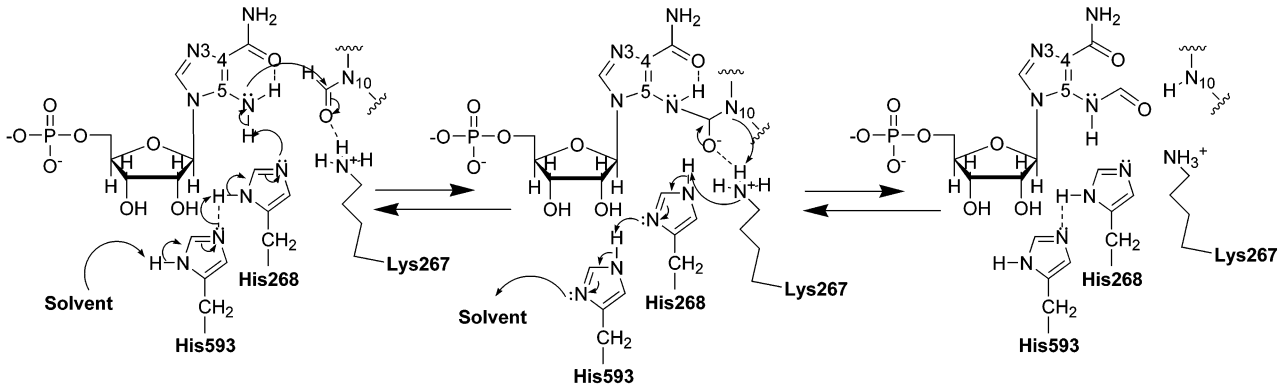

FIGURE 5: Proposed mechanism of formyl transfer. Potential mechanism for removal of a proton from the 5-amino group of AICAR during formyl transfer and nucleophilic attack on 10-f-THF are depicted. The substrate proton abstracted by His[268] can be shuttled out via the interaction with His[593] to the bulk solvent. Lys[267] could then protonate the N10 leaving group of THF in the transition state through rapid proton exchange with solvent or with His[268].

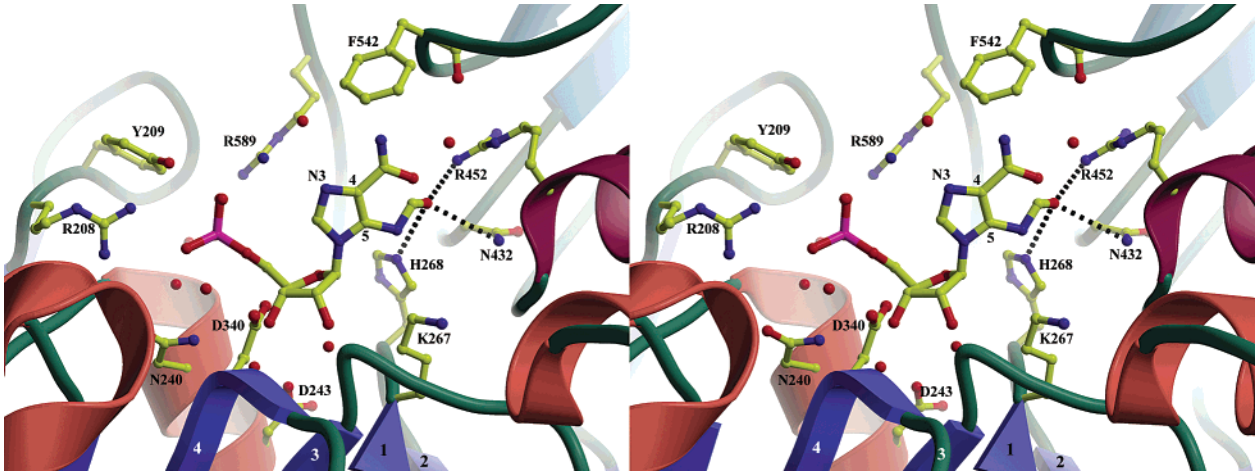

FIGURE 6: Stereo figure of FAICAR modeled into the AICAR active site. All labels and colors are conserved with Figure 3A. The formyl group of FAICAR has been modeled in by placing the formyl oxygen into the position of a highly conserved water molecule. Potential hydrogen bonds of active site residues to the formyl group are shown in black dashed lines. These additional hydrogen bonds may account for the increased affinity of FAICAR over AICAR within the AICAR Tfase active site.

the 5-formyl group out of the plane by approximately 30°. Asn[432], Arg[452], and His[268] are all within the vicinity of the 5-formyl group to provide stabilizing side-chain hydrogen bond interactions. These additional hydrogen bonds may account for the 2-fold higher affinity that is observed for FAICAR over AICAR (K_i = 0.4 μM) (*10*, *11*). In this conformation, the 5-formyl group is clearly not in a productive position for formyl transfer, but could explain product inhibition because of additional stabilization from active site residues.

Comparison to GAR Tfase. GAR Tfase also performs a similar transformylase reaction in the de novo purine synthesis pathway using the same 10-f-THF cofactor. However, there is no significant sequence or structural homology between the two proteins. GAR Tfase uses a catalytic trio composed of residues Asn[106], His[108], and Asp[144] for its transformylase reaction (*4*). His[108] has two roles in the GAR Tfase formyl transfer reaction; first, as a protonated imidazolium group to aid in nucleophilic attack by withdrawing electrons from the formyl group of 10-f-THF (*4*) and second, to stabilize the oxyanion of the transition state analogue. Asn[106] also helps stabilize the transition state, and the negative charge of Asp[144] serves to increase the pK_a of the His[108] imidazolium from 6.2 to 9.2 (*4*). Site-directed mutagenesis has shown that none of the catalytic trio residues in GAR Tfase are absolutely required for catalysis (*41*), although two or more mutations inactivate the enzyme. A similar constellation of residues was revealed in ATIC by sequence homology suggesting similarity in the active sites of GAR Tfase and ATIC that was later discounted from mutational analyses (*8*). The ATIC crystal structure further reveals that these residues are located far from the AICAR Tfase active site (*18*). Comparison of the crystal structures of GAR Tfase (*4*, *38*, *42*) with AICAR Tfase also does not support any similarity in the active sites, either for substrate or folate binding or in the mechanism of the transformylase reaction. The glycinamide ribonucleotide substrate amino group to which the formyl group is transferred is much more reactive in GAR than its AICAR counterpart, and consequently, its deprotonation during formyl transfer is suggested to occur via a structural water or bulk solvent molecule (*4*). The conversion of AICAR to FAICAR favors the starting substrates so that the driving force behind the reaction becomes the favorable cyclization of FAICAR to IMP (*10*, *11*). Thus, activation of the substrate amine is more of a consideration in AICAR Tfase where the substrate conformation itself, as well as His[268], appears to aid in its activation. As deprotonation of the amine is unlikely prior to nucleophilic attack on the formyl group, His[268] is positioned, therefore, to remove the proton coincident with bond formation in the covalent intermediate, similar to the proposed role of a structural water in GAR Tfase. Lys[267], like His[108] and Asn[106] of GAR Tfase, helps to stabilize the anionic transition state. Unlike the catalytic trio in GAR Tfase, His[268] and Lys[267] are absolutely required for the formyl transfer reaction to occur (*8*, *19*, *20*).

CONCLUSION

The AICAR-bound structure of the AICAR Tfase of ATIC has revealed potential key catalytic residues and suggested how the non-nucleophilic amino group of AICAR might become activated. On the basis of the substrate-bound ATIC structure, a potential mechanism for formyl transfer can be proposed that differs from previous proposals. Within the AICAR Tfase active site, the substrate is positioned such that its carboxamide and His[268] can provide stabilizing hydrogen bonds with the reactive amino group at C5. His[268] would act as a base, and its pK_a could be suppressed by the close interactions with His[593] and the positive charge of Lys[267], thereby facilitating the transfer of a proton from the 5-amino group of AICAR (Figure 5) during the formylation reaction. The positive charge on Lys[267] is likely to play a major role in the stabilization of the oxyanion transition state. The proton abstracted from the substrate by His[268] could then easily be shuttled out to the solvent via the interaction with His[593] or even shuffled through Lys[267] to protonate the leaving N10 of THF (Figure 5). A high-resolution AICAR Tfase structure with bound folate and a nonreactive substrate should clarify any remaining ambiguities in the formyl transfer mechanism of ATIC.

ACKNOWLEDGMENT

We thank T. Horton for excellent technical support in protein expression and purification; A. Heine and X. Dai

for data collection and processing; 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 beamline 14-BM-C.

REFERENCES

1. Baldwin, S. W., Tse, A., Gossett, L. S., Taylor, E. C., Rosowsky, A., Shih, C., and Moran, R. G. (1991) *Biochemistry 30*, 1997−2006.
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.
3. Varney, M. D., Palmer, C. L., Romines, W. H., III, Boritzki, T., Margosiak, S. A., Almassy, R., Janson, C. A., Bartlett, C., Howland, E. J., and Ferre, R. (1997) *J. Med. Chem. 40*, 2502−2524.
4. Su, Y., Yamashita, M. M., Greasley, S. E., Mullen, C. A., Jennings, P. A., Warren, M. S., Benkovic, S. J., and Wilson, I. A. (1998) *J. Mol. Biol. 281*, 485−499.
5. Greasley, S. E., Yamashita, M. M., Cai, H., Benkovic, S. J., Boger, D. L., and Wilson, I. A. (1999) *Biochemistry 38*, 16783−16793.
6. Greasley, S. E., Marsilje, T. H., Cai, H., Baker, S., Benkovic, S. J., Boger, D. L., and Wilson, I. A. (2001) *Biochemistry 40*, 13538−13547.
7. Marsilje, T. H., Labroli, M. A., Hedrick, M. P., Jin, Q., Desharnais, J., Baker, S. J., Gooljarsingh, L. T., Ramcharan, J., Tavassoli, A., Zhang, Y., Wilson, I. A., Beardsley, G. P., Benkovic, S. J., and Boger, D. L. (2002) *Bioorg. Med. Chem. 10*, 2739−2749.
8. Rayl, E. A., Moroson, B. A., and Beardsley, G. P. (1996) *J. Biol. Chem. 271*, 2225−2233.
9. Yamazaki, A., and Okutsu, M. (1978) *J. Heterocyclic Chem. 15*, 353−358.
10. Wall, M., Shim, J. H., and Benkovic, S. J. (2000) *Biochemistry 39*, 11303−11311.
11. Bulock, K. G., Beardsley, G. P., and Anderson, K. S. (2002) *J. Biol. Chem. 277*, 22168−22174.
12. Beardsley, G. P., Moroson, B. A., Taylor, E. C., and Moran, R. G. (1989) *J. Biol. Chem. 264*, 328−333.
13. Erba, E., Sen, S., Sessa, C., Vikhanskaya, F. L., and D'Incalci, M. (1994) *Br. J. Cancer 69*, 205−211.
14. Mendelsohn, L. G., Shih, C., Chen, V. J., Habeck, L. L., Gates, S. B., and Shackelford, K. A. (1999) *Semin. Oncol. 26*, 42−47.
15. Szabados, E., Hindmarsh, E. J., Phillips, L., Duggleby, R. G., and Christopherson, R. I. (1994) *Biochemistry 33*, 14237−14245.
16. Cronstein, B. N., Naime, D., and Ostad, E. (1993) *J. Clin. Invest. 92*, 2675−2682.
17. Gadangi, P., Longaker, M., Naime, D., Levin, R. I., Recht, P. A., Montesinos, M. C., Buckley, M. T., Carlin, G., and Cronstein, B. N. (1996) *J. Immunol. 156*, 1937−1941.
18. Greasley, S. E., Horton, P., Ramcharan, J., Beardsley, G. P., Benkovic, S. J., and Wilson, I. A. (2001) *Nat. Struct. Biol. 8*, 402−406.
19. 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.
20. 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 *In Purine and Pyrimidine Metabolism in Man* (Griesmacher, A., Ed.) pp 221−226, Plenum Press, New York.
21. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
22. *CCP4* (1994) *Acta Crystallogr. D50*, 760−763.
23. Matthews, B. W. (1968) *J. Mol. Biol. 33*, 491−497.
24. Pannu, N. S., and Read, R. J. (1996) *Acta Crystallogr. A52*, 659−668.
25. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
26. Read, R. J. (1986) *Acta Crystallogr. A42*, 140−149.
27. Jones, T. A., Cowan, S., Zou, J. Y., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
28. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
29. Hooft, R. W., Vriend, G., Sander, C., and Abola, E. E. (1996) *Nature 381*, 272.
30. Lawrence, M. C., and Colman, P. M. (1993) *J. Mol. Biol. 234*, 946−950.
31. Connolly, M. L. (1983) *Science 221*, 709−713.
32. Sheriff, S., Hendrickson, W. A., and Smith, J. L. (1987) *J. Mol. Biol. 197*, 273−296.
33. Esnouf, R. M. (1997) *J. Mol. Graphics 15*, 132−134.
34. Merritt, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr. D50*, 869−873.
35. 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.
36. Vergis, J. M., Bulock, K. G., Fleming, K. G., and Beardsley, G. P. (2001) *J. Biol. Chem. 276*, 7727−7733.
37. Levitt, M., and Perutz, M. F. (1988) *J. Mol. Biol. 201*, 751−754.
38. Klein, C., Chen, P., Arevalo, J. H., Stura, E. A., Marolewski, A., Warren, M. S., Benkovic, S. J., and Wilson, I. A. (1995) *J. Mol. Biol. 249*, 153−175.
39. Adamiak, D. A., Saenger, W., Kinas, R., and Stec, W. J. (1977) *Z. Naturforsch., C: Biosci. 32*, 672−677.
40. Mueller, W. T., and Benkovic, S. J. (1981) *Biochemistry 20*, 337−344.
41. Warren, M. S., Marolewski, A. E., and Benkovic, S. J. (1996) *Biochemistry 35*, 8855−8862.
42. Chen, P., Schulze-Gahmen, U., Stura, E. A., Inglese, J., Johnson, D. L., Marolewski, A., Benkovic, S. J., and Wilson, I. A. (1992) *J. Mol. Biol. 227*, 283−292.
43. Martin, A. C. R. *PROFIT*, version 1.7 (1996), SciTech Software, Chico, CA.

BI020505X