Structural Analysis of the Active Site Geometry of N^5-Carboxyaminoimidazole Ribonucleotide Synthetase from *Escherichia coli*[†,‡]

James B. Thoden,[§] Hazel M. Holden,*[,§] and Steven M. Firestine*[,||]

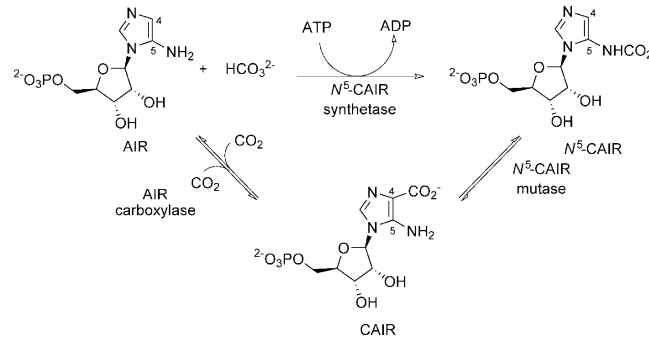
Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706, and Department of Pharmaceutical Sciences, Eugene Applebaum College of Pharmacy and Health Sciences, Wayne State University, Detroit, Michigan 48201

Received September 10, 2008; Revised Manuscript Received October 17, 2008

ABSTRACT: N^5-Carboxyaminoimidazole ribonucleotide synthetase (N^5-CAIR synthetase) converts 5-aminoimidazole ribonucleotide (AIR), MgATP, and bicarbonate into N^5-CAIR, MgADP, and P_i. The enzyme is required for *de novo* purine biosynthesis in microbes yet is not found in humans suggesting that it represents an ideal and unexplored target for antimicrobial drug design. Here we report the X-ray structures of N^5-CAIR synthetase from *Escherichia coli* with either MgATP or MgADP/P_i bound in the active site cleft. These structures, determined to 1.6-Å resolution, provide detailed information regarding the active site geometry before and after ATP hydrolysis. In both structures, two magnesium ions are observed. Each of these is octahedrally coordinated, and the carboxylate side chain of Glu238 bridges them. For the structure of the MgADP/P_i complex, crystals were grown in the presence of AIR and MgATP. No electron density was observed for AIR, and the electron density corresponding to the nucleotide clearly revealed the presence of ADP and P_i rather than ATP. The bound P_i shifts by approximately 3 Å relative to the γ-phosphoryl group of ATP and forms electrostatic interactions with the side chains of Arg242 and His244. Since the reaction mechanism of N^5-CAIR synthetase is believed to proceed via a carboxyphosphate intermediate, we propose that the location of the inorganic phosphate represents the binding site for stabilization of this reactive species. Using the information derived from the two structures reported here, coupled with molecular modeling, we propose a catalytic mechanism for N^5-CAIR synthetase.

The original pathway proposed for *de novo* purine biosynthesis by Buchanan in the late 1950s involved the 10-step conversion of phosphoribosyl pyrophosphate into the common branch-point intermediate inosine monophosphate (*1*). This view of purine biosynthesis existed well into the mid-1990s until it was discovered that the 10-step pathway was not universal across all organisms (*2*−*4*). Indeed, *de novo* purine biosynthesis is different in microbes versus humans, and the difference centers around the synthesis of the key intermediate, 4-carboxy-5-aminoimidazole ribonucleotide (CAIR).[1] In bacteria and yeast, CAIR is produced from 5-aminoimidazole ribonucleotide (AIR) by the action of two enzymes rather than one as in higher eukaryotes (Scheme 1). The first of these, N^5-carboxyaminoimidazole ribonucleotide synthetase, hereafter referred to as N^5-CAIR synthetase, converts AIR, MgATP, and bicarbonate into an unstable intermediate N^5-CAIR. The second enzyme, N^5-carboxyaminoimidazole ribonucleotide mutase or N^5-CAIR mutase, catalyzes a rearrangement in which the CO_2 group from N5 of the substrate is moved to the C4 position thereby generating CAIR. Strikingly, humans contain neither enzyme. Instead, AIR is converted directly to CAIR via the action of AIR carboxylase (*5*). Although amino acid sequence and structural comparisons have shown that AIR carboxylase is evolutionarily related to N^5-CAIR mutase, each enzyme is specific for its own substrate (*5*).

Scheme 1



Clearly, *de novo* purine biosynthesis has evolutionarily diverged among lower and higher organisms. This divergent

[†] This research was supported in part by an NIH grant (DK47814 to H.M.H.) and by Wayne State University (S.M.F.). Results in this report were derived from work performed at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source. Argonne is operated by the University of Chicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357.

[‡] X-ray coordinates have been deposited in the Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (accession nos. 3ETH and 3ETJ).

* To whom correspondence should be addressed. E-mail addresses: sfirestine@wayne.edu or Hazel_Holden@biochem.wisc.edu. Fax (S.M.F.): 313-577-2033. Phone (S.M.F.): 313-577-0455.

[§] University of Wisconsin, Madison.

[||] Wayne State University.

[1] Abbreviations: ADP, adenosine diphosphate; AIR, 5-aminoimidazole ribonucleotide; AMPPNP, adenosine 5′-(β,γ-imido) triphosphate; ATP, adenosine triphosphate; CAIR, 4-carboxy-5-aminoimidazole ribonucleotide; IPTG, isopropyl-β-D-thiogalactopyranoside; Tris, tris-(hydroxymethyl)aminomethane.

10.1021/bi801734z CCC: $40.75
Published on Web 11/20/2008

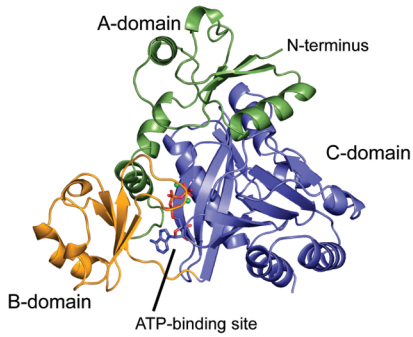


FIGURE 1: Molecular architecture of the N^5-CAIR synthetase subunit. The A-, B-, and C-domains, color-coded in green, yellow, and blue, respectively, are defined by Met1 to Ala95, Pro96 to Phe158, and Ser159 to Gly355. The ATP-binding pocket is located between the B- and C-domains.

nature of a primary metabolic pathway has significant ramifications for the design of novel antimicrobial agents. Genetic studies have shown that both N^5-CAIR synthetase and N^5-CAIR mutase are required for microbial growth (*6*−*9*). Deletion of either enzyme results in a purine auxotroph that is unable to propagate in human or mouse serum and is not viable in animal models predictive of disease (*6, 10*−*12*). These results support the conclusion that inhibitors of these enzymes could be potential antibacterial and antifungal agents.

The first structural analysis of an N^5-CAIR synthetase, namely, that from *Escherichia coli*, was reported in 1999, and it demonstrated that the enzyme contained a three-domain architecture similar to those proteins belonging to the ATP-grasp superfamily as indicated in Figure 1 (*13*). Members of the ATP-grasp superfamily catalyze a wide range of ligation reactions (*14, 15*). Their reactions are thought to proceed through the formation of an acyl phosphate intermediate, which is attacked by a nucleophile leading to a tetrahedral intermediate that ultimately collapses to form the product. The bound nucleotides in ATP-grasp superfamily members are invariably located wedged between the B- and C-domains (Figure 1). For the initial structural analysis of N^5-CAIR synthetase, two complexes were reported, a sulfate-liganded structure determined to 2.1-Å resolution and a MgADP-liganded structure solved to 2.5-Å resolution (*13*). The sulfate in the former structure was thought to mimic where the carboxyphosphate intermediate is stabilized within the active site region. Due to the limited resolution, however, it was not possible to either describe in detail the hydrogen-bonding pattern between the nucleotide and the protein or discuss the coordination geometry surrounding the magnesium ions. Such information is critical for structure-guided drug design.

Given the importance of N^5-CAIR synthetase for microbial growth, we recently initiated a project aimed at both identifying suitable inhibitors of the enzyme and determining higher resolution X-ray structures. High-throughput screening has already revealed 14 compounds that inhibit the *E. coli* enzyme to varying degrees (unpublished results). Some of these compounds display noncompetitive kinetics with respect to both AIR and ATP, thus suggesting that they bind at locations distinct from the active site cleft. To gain additional information, we crystallized N^5-CAIR synthetase from *E. coli* complexed with either AIR and MgATP or aminopyrazole ribonucleotide (an inhibitor) and MgATP. The 1.6-Å resolution structure of the enzyme complexed with the inhibitor and MgATP revealed only the presence of MgATP in the active site. Interestingly, for the structure of the enzyme complexed with MgATP and AIR, there was no density for the substrate and the MgATP ligand had been hydrolyzed to ADP and P_i. As such, these two structures provide snapshots of the active site geometry before and after ATP hydrolysis. In addition, the location of the inorganic phosphate is different than that observed for sulfate in the original analysis of N^5-CAIR synthetase, and thus it may represent a better mimic for where the carboxyphosphate intermediate is stabilized. On the basis of these two structures and model-building AIR into the active site cleft of N^5-CAIR synthetase, we propose a reaction mechanism for the enzyme and discuss the possible roles of conserved amino acid residues in catalysis.

MATERIALS AND METHODS

Cloning, Expression, and Purification. Genomic DNA from *E. coli* w3110 (ATCC 39936) was isolated by standard procedures. The *purE/purK* genes that encode N^5-CAIR mutase and N^5-CAIR synthetase, respectively, were PCR-amplified using primers that introduced a 5′ *Nde*I site and a 3′ *Xho*I site. The purified PCR product was A-tailed and ligated into the pGEM-T (Promega) vector for screening and sequencing. A PurE/PurK-pGEM-T vector construct of the correct sequence was then appropriately digested and ligated into a pET31b(+) (Novagen) plasmid for protein expression.

The PurE/PurK-pET31 plasmid was used to transform HMS174 (DE3) *E. coli* cells (Novagen). The culture in LB media was grown at 37 °C with shaking until an optical density of 0.7 was reached at 600 nm, which was then followed by induction with 1 mM IPTG. These cells were subsequently allowed to express protein at 37 °C for 6 hours. N^5-CAIR synthetase was purified according to previously published procedures (*2*). The final sample was dialyzed against 10 mM Tris-HCl and 200 mM NaCl at pH 8.0 and concentrated to 17 mg/mL.

Structural Analysis of N^5-CAIR Synthetase. Crystallization conditions were initially surveyed by the hanging drop method of vapor diffusion with a sparse matrix screen developed in the laboratory. The protein sample contained 10 mM MgATP and either 10 mM AIR or 10 mM aminopyrazole (gift of Dr. V. Jo Davisson). Larger crystals were subsequently grown via hanging drop vapor diffusion against 20−25% monomethylether poly(ethylene glycol) 5000 at pH 7.5. Crystals grown in the presence of the substrate belonged to the space group *C*2 with unit cell dimensions of $a = 128.8$ Å, $b = 70.5$ Å, $c = 88.7$ Å, and $\beta = 124.3°$, whereas those grown in the presence of the inhibitor belonged to the space group *P*1 with unit cell dimensions of $a = 57.1$ Å, $b = 57.1$ Å, $c = 59.2$ Å, $\alpha = 77.5°$, $\beta = 82.3°$, and $\gamma = 77.3°$. Both crystal forms contained one dimer per asymmetric unit.

X-ray data from both crystal forms were measured at 100 K at SBC-CAT Beamline 19-ID using the SBC-3 CCD-detector. These data were processed and scaled with HKL2000 (*16*). Prior to X-ray data collection, the crystals were stabilized by transfer into a 30% poly(ethylene glycol) solution containing 400 mM NaCl, 10 mM MgATP, 15% ethylene glycol, and 10 mM of either AIR or aminopyrazole

Table 1: X-ray Data Collection Statistics

	enzyme complexed with MgATP	enzyme complexed with MgADP and P_i
space group	*P*1	*C*2
resolution limits	30.0−1.6 (1.66−1.6)[b]	30.0−1.6 (1.66−1.6)[b]
no. independent reflns	87837 (7649)	80957 (6164)
completeness (%)	93.8 (81.9)	93.6 (71.1)
redundancy	2.3 (1.8)	3.8 (2.6)
avg *I* (avg $\sigma(I)$)	29.5 (6.4)	28.6 (2.6)
R_{sym} (%)[a]	3.3 (9.8)	5.0 (20.7)

[a] $R_{sym} = (\Sigma|\Sigma I - \bar{I}|/\Sigma I) \times 100$. [b] Statistics for the highest resolution bin.

Table 2: Least-Squares Refinement Statistics

	enzyme complexed with MgATP	enzyme complexed with MgADP and P_i
resolution limits (Å)	30.0−1.6	30.0−1.6
R-factor[a] (overall), %/no. reflns	17.8/87820	16.0/80952
R-factor (working), %/no. reflns	17.5/79043	15.8/72840
R-factor (free), %/no. reflns	22.8/8777	21.5/8085
no. protein atoms	5590[b]	5555[c]
no. heteroatoms	756[d]	750[e]
Average *B* Values		
protein atoms (Å^2)	26.6	27.2
ligands (Å^2)	20.3	24.0
solvent (Å^2)	35.1	37.0
Weighted rms Deviations from Ideality		
bond lengths (Å)	0.013	0.013
bond angles (deg)	2.25	2.45
trigonal planes (Å)	0.008	0.007
general planes (Å)	0.015	0.014
torsional angles[f] (deg)	16.7	16.7

[a] *R*-factor $= (\Sigma|F_o - F_c|/\Sigma|F_o|) \times 100$ where F_o is the observed structure-factor amplitude and F_c is the calculated structure-factor amplitude. [b] These include multiple conformations for Glu55, Arg135, Ser173, and Ser345 in subunit 1 and Ser173 in subunit 2. [c] These include multiple conformations for Glu140 and Ser352 in subunit 1 and Arg135, Ser173, and Ser326 in subunit 2. [d] Heteroatoms include 2 ATP, 4 Mg^{2+}, and 690 waters. [e] Heteroatoms include 2 ADP, 2 P_i, 4 Mg^{2+}, 1 Cl^-, and 681 waters. [f] The torsional angles were not restrained during the refinement.

ribonucleotide. Relevant X-ray data collection statistics are presented in Table 1.

The structures were solved via molecular replacement with Phaser (*17*) using the original structure of N^5-CAIR synthetase solved in this laboratory as the search model (PDB accession code 1b6s). Subsequent least-squares refinement with the software package TNT (*18*) reduced the *R*-factors to 17.8% for the MgATP/aminopyrazole ribonucleotide cocrystals and 16.0% for the MgATP/AIR cocrystals for all measured X-ray data from 30 to 1.6 Å resolution. No electron density was observed for the aminopyrazole ribonucleotide ligand, and thus this complex shows the structure of the enzyme before ATP hydrolysis. The ATP/AIR complex did not show electron density for the substrate, AIR, but only for MgADP and P_i. As such this complex represents the posthydrolysis structure. Relevant refinement statistics are presented in Table 2.

RESULTS AND DISCUSSION

Structure of N^5-CAIR Synthetase Complexed with MgATP. Crystals employed for the structural analysis of this complex contained one dimer per asymmetric unit, and the X-ray data extended to a nominal resolution of 1.6 Å. The overall quality of the dimer model is excellent with 94.4% and 5.6% of the polypeptide dihedral angles lying within the core and allowed regions of the Ramachandran plot, respectively. The two subunits correspond to a root-mean-square deviation of 0.16 Å, and thus for the sake of clarity the following discussion refers only to subunit 1 in the X-ray coordinate file.

A superposition of the N^5-CAIR synthetase/MgATP complex model with that of the apo-form of the enzyme shows that the A- and C-domains adopt similar positions, but the B-domains are in different orientations. Specifically, the α-carbons for the A- and C-domains superimpose with root-mean-square deviations of 0.45 and 0.31 Å respectively, whereas the B-domains correspond with a root-mean-square deviation of 0.83 Å. The loop extending from Thr123 to Gly130, which is disordered in the apoenzyme structure, becomes ordered with Gly125, Tyr126, Asp127, and Gly128 forming part of the binding pocket for the γ-phosphoryl group of ATP.

In the original structural analysis of N^5-CAIR synthetase, it was suggested that there might not be a second magnesium-binding site associated with the nucleotide (*13*). As can be seen in Figure 2a, however, there are clearly two magnesium ions accompanying the ATP ligand. Each of these magnesiums is octahedrally coordinated with metal−ligand bond lengths ranging from 1.8 to 2.3 Å. A close-up view of the active site pocket is presented in Figure 2b. The backbone amide group of Ile156, the carbonyl group of Gln154, and the side chains of Glu153 and Lys120 form hydrogen-bonding interactions with the adenine ring of the nucleotide. In addition to its interaction with the adenine ring, the side chain of Lys120 lies within 2.7 Å of an α-phosphoryl oxygen. The nucleotide ribose adopts the $C_{3'}$-*endo* pucker with the C-2 and C-3 hydroxyl groups anchored to the protein via the carboxylate side chain of Glu161. In addition to Lys120, there is a second positively charged residue, Arg80, that lies within 2.8 Å of the β-phosphoryl oxygen that is involved in metal ligation. One of the magnesiums, labeled "A" in Figure 2b, is ligated by an α-phosphoryl oxygen, a γ-phosphoryl oxygen, a water molecule, the side chain of Glu226, which functions as a bidentate ligand, and finally an oxygen contributed by Glu238. The second magnesium, labeled "B" in Figure 2b, is surrounded by a β-phosphoryl oxygen, a γ-phosphoryl oxygen, two water molecules, and the carboxylate oxygens of Glu238. These two metals are separated by 3.7 Å with Glu238 serving as the bridging ligand. This arrangement of the magnesiums and their accompanying coordination spheres is virtually identical to that observed in *E. coli purT*-encoded glycinamide ribonucleotide transformylase (PurT), another enzyme involved in *de novo* purine biosynthesis (*19*). On the basis of both structural and sequence alignments, it was postulated that Asp127 in N^5-CAIR synthetase might function as a metal ligand (*13*). In fact, the carboxylate group of Asp127 is situated over 7 Å from the metals, but its backbone amide nitrogen does participate in a hydrogen-bonding interaction with one of γ-phosphoryl oxygens of ATP.

Structure of N^5-CAIR Synthetase Complexed with MgADP and P_i. Crystals used for this structural analysis were grown in the presence of MgATP and AIR. They also contained a dimer in the asymmetric unit, and the X-ray data extended to 1.6 Å resolution. A Ramachandran analysis demonstrates

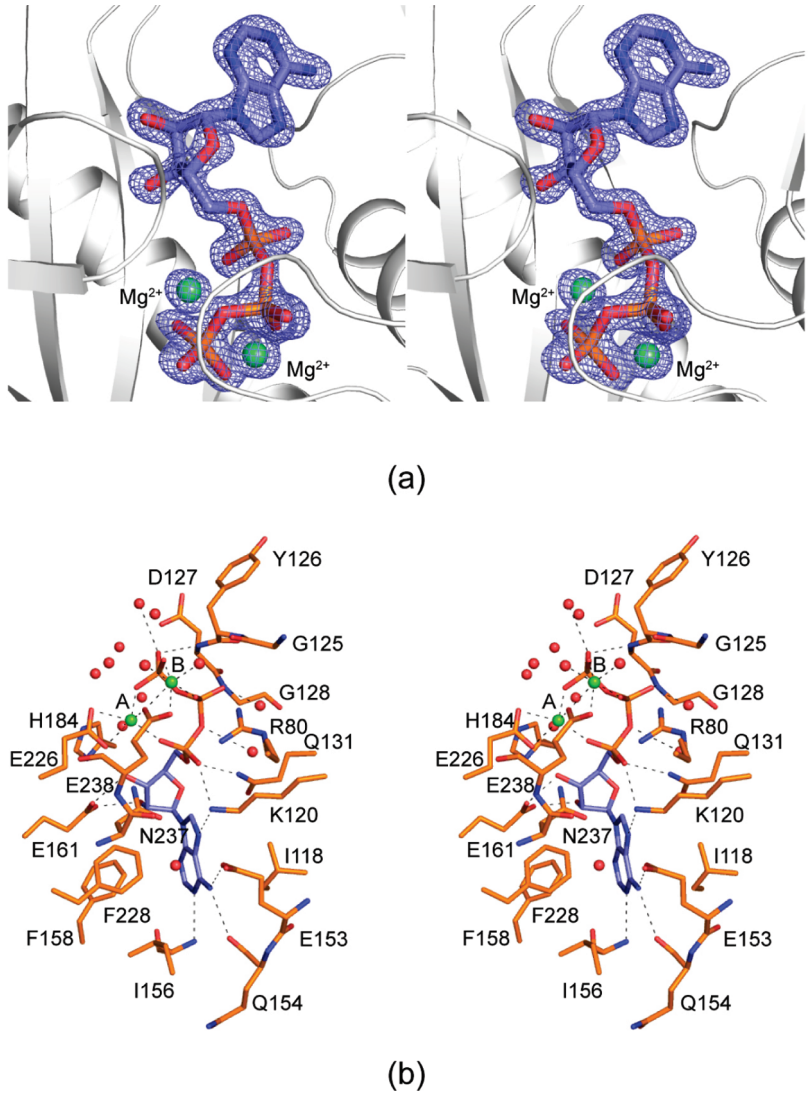


FIGURE 2: The active site for the N^5-CAIR synthetase/MgATP complex. (a) Electron density corresponding to the two magnesium ions and the ATP ligand is shown. The map was contoured at $\sim 4\sigma$ and calculated with coefficients of the form $(F_o - F_c)$, where F_o was the native structure-factor amplitude and F_c was the calculated structure-factor amplitude. (b) Close-up view of the active site within 3.6 Å of the MgATP is displayed. The magnesium ions are highlighted in green, whereas the nucleotide is displayed in blue-filled bonds. The octahedral coordination spheres surrounding the magnesium ions are indicated by the dashed lines. Additional hydrogen-bonding interactions within 3.2 Å between ATP and the protein are also displayed as dashed lines.

that 94.3% and 5.7% of the backbone dihedral angles fall within the core and allowed regions, respectively. The α-carbon positions for the two subunits superimpose with a root-mean-square deviation of 0.5 Å, which is higher than that observed for the enzyme/MgATP complex model. This higher value is due to differences in the B-domains between the two subunits in the dimer. In subunit 1, the loop between Gly125 and Gly130 is ordered, albeit the electron density is weaker suggesting flexibility. The corresponding loop in subunit 2 is disordered as observed in the apoenzyme structure. In light of this disorder, the following discussion refers only to subunit 1.

The initial analysis of the map revealed no electron density for AIR, and in addition, the electron density corresponding to the MgATP clearly revealed that it had been hydrolyzed to MgADP and P_i as shown in Figure 3a. Again there are two magnesium ions present in the electron density. A close-up view of the active site with bound MgADP and P_i is presented in Figure 3b. The β- and γ-phosphorus atoms are now separated by 4 Å. As can be seen, the adenine and ribose moieties are accommodated within the active site cleft in nearly an identical fashion as that observed for MgATP. The main differences between the pre- and posthydrolysis states occur around the magnesiums. First, Glu226 is no longer a bidentate but rather a monodentate ligand to the "A" metal. The other ligands to this metal include a water molecule, a phosphoryl oxygen donated by the inorganic phosphate, an α- and a β-phosphoryl oxygen from the ADP, and an oxygen contributed by Glu238, which still serves to bridge the two metals. The inorganic phosphate shifts in the active site by ~3 Å relative to the γ-phosphoryl group of ATP thereby allowing its electrostatic interactions with the side chains of Arg242 and His244. Neither of these two residues is strictly conserved among members of the ATP-grasp superfamily (*15*). In addition to these side chain interactions, the inorganic phosphate lies within hydrogen-bonding distance of the backbone amide nitrogen of Asp127. In the original structure of N^5-CAIR synthetase solved in the presence of MgADP (*13*), there was no inorganic phosphate observed in the active site. As a consequence, the loop delineated by Gly125 to

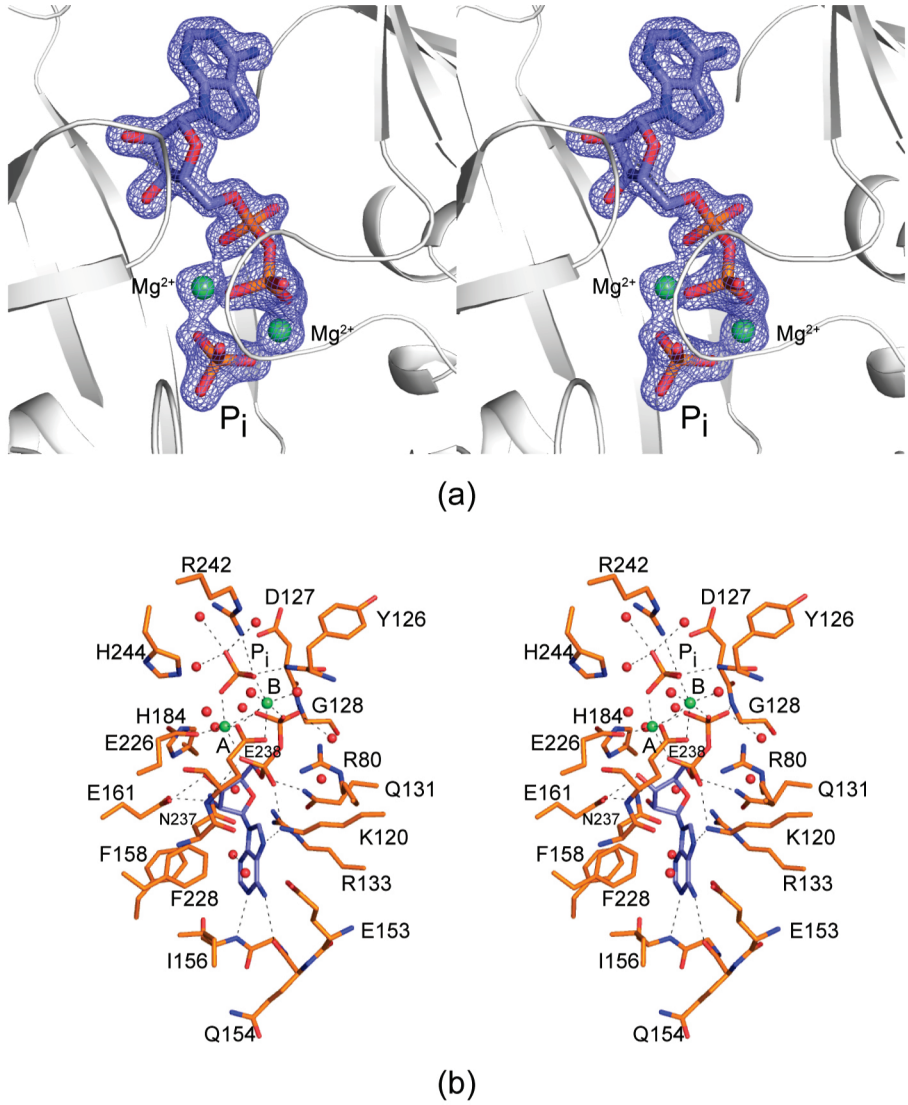


FIGURE 3: The active site for N^5-CAIR synthetase with bound MgADP and P_i. (a) Electron density corresponding to the two magnesium ions, the ADP, and the inorganic phosphate is shown. The map was contoured at $\sim 4\sigma$ and calculated with coefficients of the form $(F_o - F_c)$, where F_o was the native structure-factor amplitude and F_c was the calculated structure-factor amplitude. (b) Close-up view of the active site within 3.6 Å of the MgADP and P_i ligands is displayed. The magnesium ions are highlighted in green, whereas the nucleotide is shown in blue-filled bonds. The octahedral coordination spheres surrounding the magnesium ions are indicated by the dashed lines. Additional hydrogen-bonding interactions within 3.2 Å between ADP, P_i, and the protein are indicated by the dashed lines.

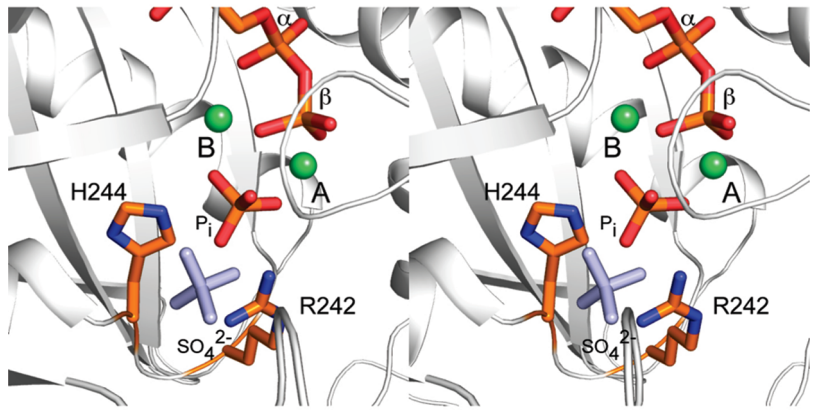


FIGURE 4: Comparison of the observed positions for sulfate versus inorganic phosphate in the N^5-CAIR synthetase active site. The sulfate anion is colored in gray bonds.

Gly130, which includes Asp127, is shifted slightly away from the active site. If this loop is excluded, however, the α-carbons for the two models are very similar and superimpose with a root-mean-square deviation of 0.48 Å. Metal "B" in the present structure is coordinated by an oxygen from the inorganic phosphate, two waters, a β-phosphoryl oxygen,

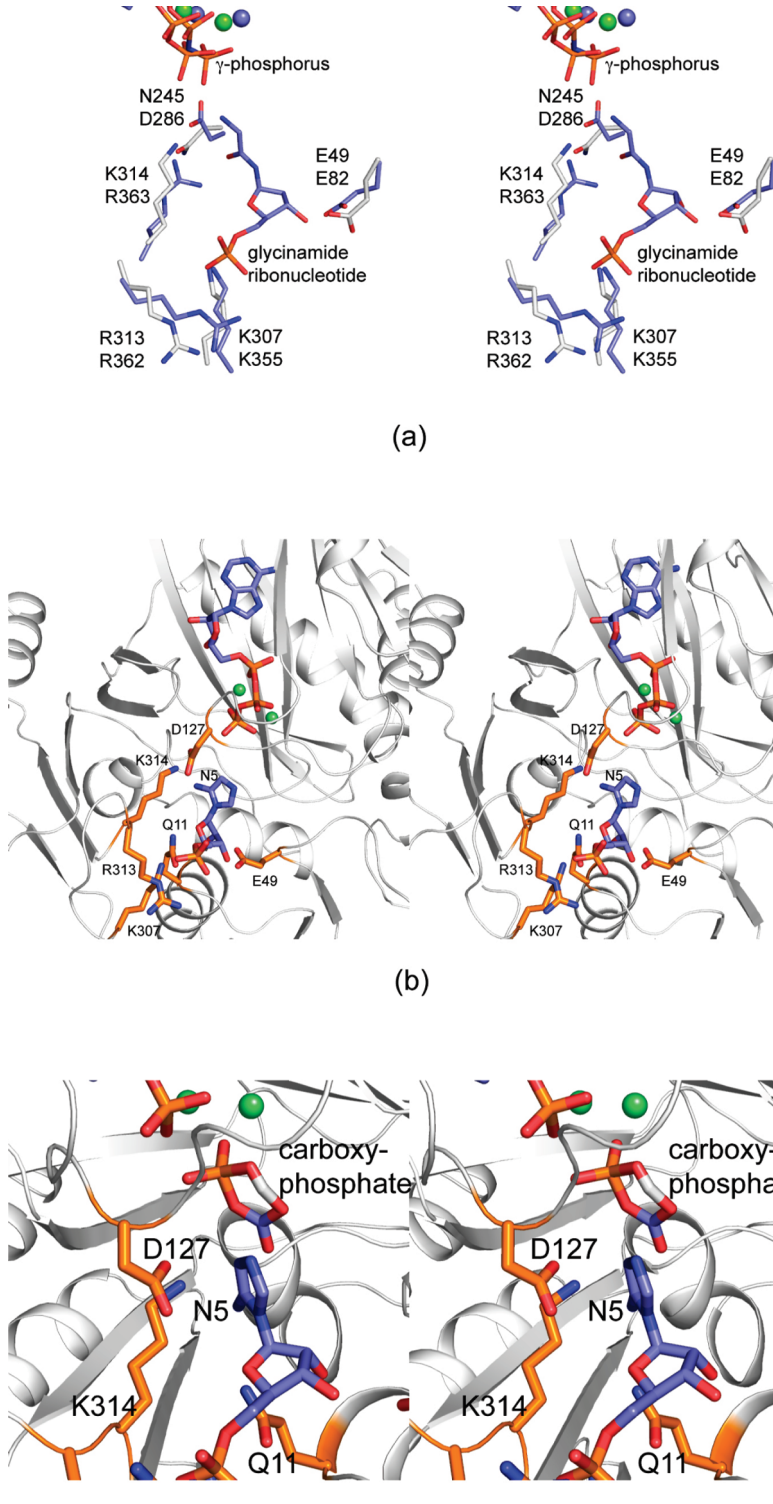


FIGURE 5: Models of the N^5-CAIR synthetase active site with bound ligands. (a) Superposition of the active site of N^5-CAIR synthetase onto the PurT glycinamide ribonucleotide binding pocket. Residues belonging to N^5-CAIR synthetase and PurT are displayed in white and slate, respectively. Residue labels at the top refer to N^5-CAIR synthetase whereas those at the bottom correspond to PurT. The blue and green spheres represent the positions of the magnesium ions in PurT and N^5-CAIR synthetase, respectively. (b) A possible binding mode for the substrate AIR based on that observed for glycinamide ribonucleotide in PurT is shown. Those residues thought to important for the binding of AIR are included. (c) A model of the carboxyphosphate intermediate in the pseudochair conformation bound in the active site is depicted. The shared hydrogen is displayed in white.

and the side chain of Glu238, which still functions as a bidentate ligand. Magnesium−ligand bond distances range from 1.9 to 2.4 Å. The overall coordination geometry surrounding the metal ions and the orientation of the inorganic phosphate relative to ADP in N^5-CAIR synthetase is strikingly similar to that observed for carbamoyl phosphate synthetase (*20*). Like N^5-CAIR synthetase, carbamoyl phosphate synthetase belongs to the ATP-grasp superfamily, and

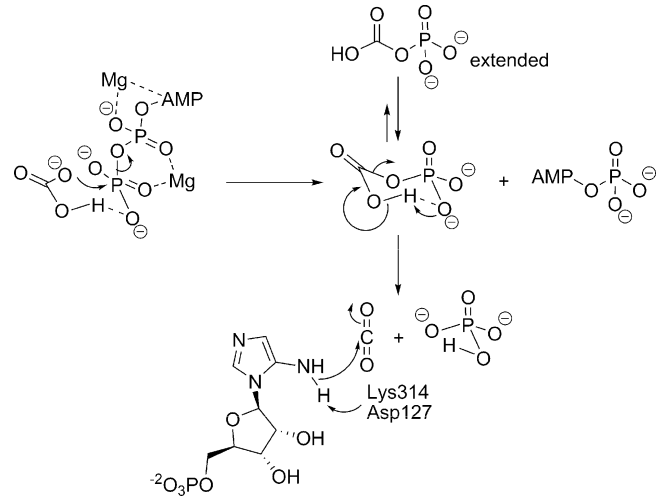
its reaction mechanism is thought to proceed through the formation of a carboxyphosphate intermediate.

One of the first structures to be solved of N^5-CAIR synthetase was done in the presence of sulfate, and it was postulated that the observed position of this anion represented the location of the carboxyphosphate intermediate. Figure 4 shows the positions adopted by the inorganic phosphate versus the sulfate ion in the N^5-CAIR synthetase active site. We suggest that the binding position for the inorganic phosphate, which was produced from the catalytically active synthetase crystal, represents a better mimic for the putative carboxyphosphate intermediate and its stabilized location on the protein.

Model of N^5-CAIR Synthetase with Bound MgATP and AIR. Attempts to crystallize a ternary complex of N^5-CAIR synthetase with ATP (or ATP-analog) and AIR have thus far been unsuccessful. However, the ternary structure of PurT bound with the nonhydrolyzable ATP analog AMPPNP (*21*) and glycinamide ribonucleotide is known (*19*). PurT catalyzes the formylation of glycinamide ribonucleotide using ATP and formate, and its reaction mechanism is thought to proceed via a formyl phosphate intermediate in a manner analogous to N^5-CAIR synthetase. These two enzymes demonstrate both a high degree of amino acid sequence homology (25% identity and 42% similarity based on 305 aligned residues) and structural homology (root-mean-square deviation of 1.7 Å for 285 α-carbons). A superposition of their active sites near the binding pocket for glycinamide ribonucleotide is presented in Figure 5a. In light of their high degree of three-dimensional similarities, it is possible to propose with reasonable confidence the location of the AIR-binding pocket in N^5-CAIR synthetase as shown in Figure 5b. The model for AIR was placed into the active site according to that observed for glycinamide ribonucleotide in PurT. According to this model, both Lys307 and Arg313 from the C-domain provide electrostatic interactions with the phosphoryl group of AIR. The side chains of Gln11 and Glu49 from the A-domain also contribute to the binding pocket. In particular, we propose that the carboxylate group of Glu49 hydrogen bonds with the 2- and 3-hydroxyl groups of the ribose. Note that in the initial structural analysis of N^5-CAIR synthetase, the phosphate of AIR was proposed to interact with a "P-loop" defined by Gly8 to Leu12. This region, however, does not classify as a canonical "P-loop", which requires the consensus sequence of Gly-Xaa-Xaa-Xaa-Xaa-Gly-Lys-Thr/Ser. Finally, in the model presented in Figure 5b, Lys314 lies within hydrogen-bonding distance of N5 of AIR which, in turn, is situated at ~7 Å from the γ-phosphorus. As discussed below, this residue may play an important role in catalysis. Asp127, contributed by the B-domain, lies within hydrogen-bonding distance of Lys314.

A Possible Catalytic Mechanism. Members of the ATP-grasp superfamily catalyze the same basic reaction, namely, a nucleophilic attack on an ATP-activated electrophile. The specifics of the nucleophile and electrophile are different for each reaction, but usually the nucleophile is an amine and the group to be activated is a carboxylic acid. Many members of the family, including N^5-CAIR synthetase, biotin carboxylase, and carbamoyl phosphate synthetase, utilize bicarbonate as the carboxylic acid component of the reaction. In all of these enzymes, research has demonstrated that one oxygen atom from bicarbonate is transferred to the inorganic

Scheme 2



phosphate product during the course of the reaction, a result consistent with the generation of carboxyphosphate (*3, 22, 23*). The structure of the ADP/P_i complex, along with the proposed AIR binding site, provides us with an opportunity to offer a model of the ADP/AIR/carboxyphosphate complex (Figure 5c). This complex, in turn, suggests a catalytic mechanism for N^5-CAIR synthetase (Scheme 2). The first step is the attack of bicarbonate on the γ-phosphorus atom of ATP. Here we propose that the attack takes place via a pseudochair transition state in which the proton on bicarbonate forms an intramolecular hydrogen bond to one of the oxygen atoms of the γ-phosphate. Herschlag and Jencks previously demonstrated that hydrogen bond formation in phosphoryl transfer reactions to bicarbonate serves to remove unfavorable interactions in the transition state (*24*). Once phosphoryl transfer occurs, quantum mechanical calculations indicate that carboxyphosphate is approximately 13 kcal/mol more stable in the chair conformation than in the extended conformation (S. Firestine, personal communication) suggesting that it remains in the pseudochair conformation.

We superimposed the phosphate group of the pseudochair conformation onto the experimentally observed inorganic phosphate in the ADP/P_i complex. According to our model (Figure 5c), the carboxyl group of carboxyphosphate is ~4 Å away from N5 of AIR, a distance somewhat longer than would be expected for a direct attack on the carbonyl group. Furthermore, the angle of attack of the amine on the carboxyl group is not ideal. These observations suggested to us that carboxyphosphate decomposes to CO_2, which migrates to a position in the active site that is favorable for attack by AIR.

What triggers the decarboxylation of carboxyphosphate? The pseudochair conformation of carboxyphosphate allows for the phosphate group to act as an intramolecular base to remove the proton from the acid resulting in the generation of CO_2 and HPO_4^{-2}. Such a reaction has precedent in the decomposition of carbamoyl phosphate (*25*) and in product inhibition studies on biotin carboxylase, which have shown that HPO_4^{-2} is the product released from the enzyme (*26*).

Once CO_2 is positioned near AIR, N5 of AIR needs to be deprotonated. Both Lys314 and Asp127 are close to N5, and thus either one or both in conjunction could serve as the base. Thus Asp127, once proposed to be a ligand to the magnesium ion, more likely plays a role in activating AIR

for attack. Both Asp127 and Lys314 are strictly conserved in 100 aligned N^5-CAIR synthetase amino acid sequences deposited in data banks to date. Site-directed mutagenesis experiments are presently underway to examine the role of these amino acids in catalysis.

In summary, the structures described here provide a more complete description of the N^5-CAIR synthetase active site and allow for the proposal of a catalytic mechanism that can be tested by both site-directed mutagenesis experiments and kinetic measurements. Understanding the subtleties of the N^5-CAIR synthetase active site will have important ramifications for the development of new antibacterial and antifungal agents.

ACKNOWLEDGMENT

We thank Norma E. C. Duke, Ph.D., for assistance at the beam line during X-ray data collection. The insightful conversations of Dr. W. W. Cleland are gratefully acknowledged.

REFERENCES

1. Hartman, S. C., and Buchanan, J. M. (1959) Nucleic acids, purines, pyrimidines (nucleotide synthesis). *Annu. Rev. Biochem. 28*, 365–410.
2. Meyer, E., Leonard, N. J., Bhat, B., Stubbe, J., and Smith, J. M. (1992) Purification and characterization of the purE, purK, and purC gene products: Identification of a previously unrecognized energy requirement in the purine biosynthetic pathway. *Biochemistry 31*, 5022–5032.
3. Mueller, E. J., Meyer, E., Rudolph, J., Davisson, V. J., and Stubbe, J. (1994) N^5-Carboxyaminoimidazole ribonucleotide: Evidence for a new intermediate and two new enzymatic activities in the *de novo* purine biosynthetic pathway of *Escherichia coli*. *Biochemistry 33*, 2269–2278.
4. Meyer, E., Kappock, T. J., Osuji, C., and Stubbe, J. (1999) Evidence for the direct transfer of the carboxylate of *N*5-carboxyaminoimidazole ribonucleotide (N5-CAIR) to generate 4-carboxy-5-aminoimidazole ribonucleotide catalyzed by *Escherichia coli* PurE, an N5-CAIR mutase. *Biochemistry 38*, 3012–3018.
5. Firestine, S. M., and Davisson, V. J. (1994) Carboxylases in *de novo* purine biosynthesis. Characterization of the *Gallus gallus* bifunctional enzyme. *Biochemistry 33*, 11917–11926.
6. Ivanovics, G., Marjai, E., and Dobozy, A. (1968) The growth of purine mutants of *Bacillus anthracis* in the body of the mouse. *J. Gen. Microbiol. 53*, 147–162.
7. Mahan, M. J., Slauch, J. M., and Mekalanos, J. J. (1993) Selection of bacterial virulence genes that are specifically induced in host tissues. *Science 259*, 686–688.
8. Polissi, A., Pontiggia, A., Feger, G., Altieri, M., Mottl, H., Ferrari, L., and Simon, D. (1998) Large-scale identification of virulence genes from *Streptococcus pneumoniae*. *Infect. Immun. 66*, 5620–5629.
9. Samant, S., Lee, H., Ghassemi, M., Chen, J., Cook, J. L., Mankin, A. S., and Neyfakh, A. A. (2008) Nucleotide biosynthesis is critical for growth of bacteria in human blood. *PLoS Pathog. 4*, e37.
10. Kirsch, D. R., and Whitney, R. R. (1991) Pathogenicity of *Candida albicans* auxotrophic mutants in experimental infections. *Infect. Immun. 59*, 3297–3300.
11. Perfect, J. R., Toffaletti, D. L., and Rude, T. H. (1993) The gene encoding phosphoribosylaminoimidazole carboxylase (ADE2) is essential for growth of *Cryptococcus neoformans* in cerebrospinal fluid. *Infect. Immun. 61*, 4446–4451.
12. Donovan, M., Schumuke, J. J., Fonzi, W. A., Bonar, S. L., Gheesling-Mullis, K., Jacob, G. S., Davisson, V. J., and Dotson, S. B. (2001) Virulence of a phosphoribosylaminoimidazole carboxylase-deficient *Candida albicans* strain in an immunosuppressed murine model of systemic candidiasis. *Infect. Immun. 69*, 2542–2548.
13. Thoden, J. B., Kappock, T. J., Stubbe, J., and Holden, H. M. (1999) Three-dimensional structure of *N*5-carboxyaminoimidazole ribonucleotide synthetase: a member of the ATP grasp protein superfamily. *Biochemistry 38*, 15480–15492.
14. Murzin, A. G. (1996) Structural classification of proteins: New superfamilies. *Curr. Opin. Struct. Biol. 6*, 386–394.
15. Galperin, M. Y., and Koonin, E. V. (1997) A diverse superfamily of enzymes with ATP-dependent carboxylate-amine/thiol ligase activity. *Protein Sci. 6*, 2639–2643.
16. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307–326.
17. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr. 40*, 658–674.
18. Tronrud, D. E., Ten Eyck, L. F., and Matthews, B. W. (1987) An efficient general-purpose least-squares refinement program for macromolecular structures. *Acta Crystallogr. A43*, 489–501.
19. Thoden, J. B., Firestine, S., Nixon, A., Benkovic, S. J., and Holden, H. M. (2000) Molecular structure of *Escherichia coli* PurT-encoded glycinamide ribonucleotide transformylase. *Biochemistry 39*, 8791–8802.
20. Thoden, J. B., Raushel, F. M., Benning, M. M., Rayment, I., and Holden, H. M. (1999) The structure of carbamoyl phosphate synthetase determined to 2.1 Å resolution. *Acta Crystallogr. D55*, 8–24.
21. Yount, R. G., Babcock, D., Ballantyne, W., and Ojala, D. (1971) Adenylyl imidodiphosphate, an adenosine triphosphate analog containing a P−N−P linkage. *Biochemistry 10*, 2484–2489.
22. Jones, M. E., and Spector, L. (1960) The pathway of carbonate in the biosynthesis of carbamyl phosphate. *J. Biol. Chem. 235*, 2897–2901.
23. Ogita, T., and Knowles, J. R. (1988) On the intermediacy of carboxyphosphate in biotin-dependent carboxylations. *Biochemistry 27*, 8028–8033.
24. Herschlag, D., and Jencks, W. P. (1990) The effects of Mg^{+2}, hydrogen bonding, and steric factors on rate and equilibrium constants for phosphoryl transfer between carboxylate ions and pyridines. *J. Am. Chem. Soc. 112*, 1942–1950.
25. Waldrop, G. L., Braxton, B. F., Urbauer, J. L., Cleland, W. W., and Kiick, D. M. (1994) Secondary ^{18}O and primary ^{13}C isotope effects as a probe of transition-state structure for enzymatic decarboxylation of oxalacetate. *Biochemistry 33*, 5262–5267.
26. Tipton, P. A., and Cleland, W. W. (1988) Catalytic mechanism of biotin carboxylase: steady-state kinetic investigations. *Biochemistry 27*, 4317–4325.

BI801734Z