Crystal Structure and Mechanism of the *Escherichia coli* ArnA (PmrI) Transformylase Domain. An Enzyme for Lipid A Modification with 4-Amino-4-deoxy-L-arabinose and Polymyxin Resistance[†,‡]

Petia Z. Gatzeva-Topalova, Andrew P. May,[§] and Marcelo C. Sousa*

Department of Chemistry and Biochemistry, University of Colorado at Boulder, Boulder, Colorado 80309

Received December 13, 2004; Revised Manuscript Received February 6, 2005

ABSTRACT: Gram-negative bacteria have evolved mechanisms to resist the bactericidal action of cationic antimicrobial peptides of the innate immune system and antibiotics such as polymyxin. The strategy involves the addition of the positively charged sugar 4-amino-4-deoxy-L-arabinose (Ara4N) to lipid A in their outer membrane. ArnA is a key enzyme in the Ara4N−lipid A modification pathway. It is a bifunctional enzyme catalyzing (1) the oxidative decarboxylation of UDP-glucuronic acid (UDP-GlcA) to the UDP-4″-ketopentose [UDP-*β*-(L-*threo*-pentapyranosyl-4″-ulose] and (2) the *N*-10-formyltetrahydrofolate-dependent formylation of UDP-Ara4N. Here we demonstrate that the transformylase activity of the *Escherichia coli* ArnA is contained in its 300 N-terminal residues. We designate it the ArnA transformylase domain and describe its crystal structure solved to 1.7 Å resolution. The enzyme adopts a bilobal structure with an N-terminal Rossmann fold domain containing the *N*-10-formyltetrahydrofolate binding site and a C-terminal subdomain resembling an OB fold. Sequence and structure conservation around the active site of ArnA transformylase and other *N*-10-formyltetrahydrofolate-utilizing enzymes suggests that the HxSLLPxxxG motif can be used to identify enzymes that belong to this family. Binding of an *N*-10-formyltetrahydrofolate analogue was modeled into the structure of ArnA based on its similarity with glycinamide ribonucleotide formyltransferase. We also propose a mechanism for the transformylation reaction catalyzed by ArnA involving residues N_{102}, H_{104}, and D_{140}. Supporting this hypothesis, point mutation of any of these residues abolishes activity.

Lipopolysaccharide (LPS)[1] is the major surface component of the outer membrane of Gram-negative bacteria. Its conserved element, lipid A, is the bioactive component predominantly responsible for many of the pathophysiological effects associated with Gram-negative bacterial infections (*1*, *2*). Phosphate groups in lipid A and LPS result in a net negative charge on the bacterial outer surface. This overall negative charge is exploited by many multicellular organisms that produce cationic antimicrobial peptides (CAMPs) to defend themselves from bacterial invasion. CAMPs are small, amphipathic, positively charged peptides that constitute a conserved branch of the innate immune system (*3*−*5*). In Gram-negative bacteria, CAMPs bind through electrostatic interactions to the negatively charged groups of lipid A, causing membrane permeabilization and bacterial death (*6*−*8*).

Gram-negative bacteria, including the human pathogens *Salmonella typhimurium* and *Pseudomonas aeruginosa*, can alter the structure of their cell wall to escape the host immune response (*9*−*11*). One mechanism that Gram-negative bacteria use to resist the bactericidal action of CAMPs is the addition of the positively charged sugar 4-amino-4-deoxy-L-arabinose (Ara4N) to lipid A (*9*, *11*−*13*). This modification results in a less negatively charged bacterial outer membrane, decreasing the electrostatic binding of CAMPs to the bacterial cell wall. It has been clearly shown that lipid A modification with Ara4N is responsible for bacterial resistance to CAMPs and clinical antimicrobials such as polymyxin (*14*−*16*). Importantly, *Ps. aeruginosa* isolated from the lungs of cystic fibrosis (CF) patients had lipid A modified by the addition of Ara4N (*17*). This modification facilitates the proliferation of bacteria in the CF lung (*18*). A clear understanding of the mechanism of Ara4N−lipid A biosynthesis is crucial for the development of combination therapies that combat bacterial resistance to CAMPs.

The enzymes responsible for lipid A modification with Ara4N are located in the *pmrHFIJKLM* operon and the *pmrE* gene (*pmr* = polymyxin resistance) (*19*). They are under

[†] This work was supported by a grant from the Cystic Fibrosis Foundation and an NIH grant (AI060841-01) to M.C.S. Support for P.Z.G.-T. was provided by an NIH training grant (GM65103). Structural biology research at the University of Colorado at Boulder is supported in part by a grant from the William M. Keck Foundation.

[‡] Coordinates and structure factors for ArnA transformylase have been deposited in the Protein Data Bank as entry 1YRW.

* To whom correspondence should be addressed. Phone: (303) 735-4341. Fax (303) 492-5894. E-mail: marcelo.sousa@colorado.edu.

[§] Current address: Fluidigm Corp., 7100 Shoreline Court, South San Francisco, CA 94080.

[1] Abbreviations: CAMPs, cationic antimicrobial peptides; LPS, lipopolysaccharide; FMT, methionyl-tRNA formyltransferase; GARF, glycinamide ribonucleotide formyltransferase; N_t-FDH, N-terminal, hydrolase domain of *N*-10-formyltetrahydrofolate dehydrogenase; ArnA_TF, ArnA transformylase domain; HEPES, *N*-(2-hydroxyethyl)-piperazine-*N*′-2-ethanesulfonic acid; UDP-GlcA, UDP-glucuronic acid; UDP-Ara4N, uridine 5′-diphospho-*β*-(4-amino-4-deoxy-L-arabinose); UDP-Ara4FN, uridine 5′-diphospho-*β*-(4-deoxy-4-formamido-L-arabinose); UDP-Ara4O, uridine 5′-diphospho-*β*-(L-*threo*-pentapyranosyl-4″-ulose); IPTG, isopropyl 1-thio-*β*-D-galactopyranoside.

10.1021/bi047384g CCC: $30.25
Published on Web 03/15/2005

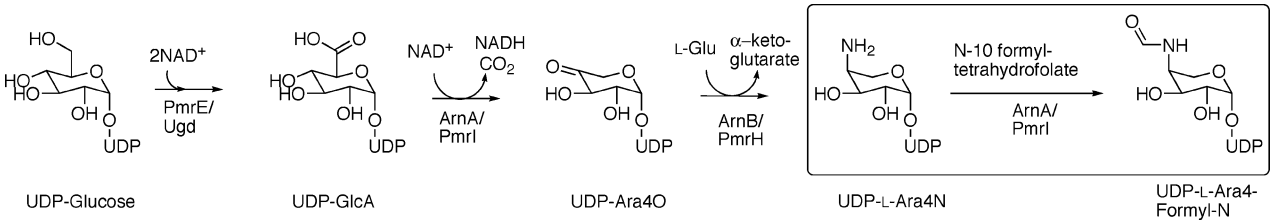

FIGURE 1: Part of the proposed pathway for the biosynthesis of Ara4N−lipid A (*24*, *25*). The pathway starts with UDP-Glc being oxidized by Ugd/PmrE to form UDP-glucuronic acid. UDP-GlcA is then oxidatively decarboxylated by the ArnA decarboxylase domain (*28*) to yield UDP-4-ketoarabinose (UDP-Ara4O). UDP-Ara4O is transaminated by ArnB, yielding the novel sugar nucleotide UDP-4-amino-4-deoxyarabinose (UDP-Ara4N). The N-terminal domain of ArnA can formylate UDP-Ara4N and has been proposed to help displace the reaction catalyzed by ArnB toward UDP-Ara4N synthesis and generate a transiently formylated product (*25*).

the direct control of the two-component system PmrA/PmrB (*13*, *20*), which also regulates some other modifications of LPS including the addition of phosphoethanolamine to lipid A (*21*, *22*). It has been shown that all of the gene products of the *pmrHFIJKLM* operon except *pmrM* are essential for the biosynthesis of Ara4N−lipid A and bacterial resistance to CAMPs and polymyxin (*12*). On the basis of sequence analyses, a pathway for the biosynthesis of Ara4N−lipid A has been proposed (*23*−*26*). In this pathway (Figure 1), UDP-Glc dehydrogenase (Ugd/PmrE) oxidizes UDP-glucose (UDP-Glc) to UDP-glucuronic acid (UDP-GlcA) in the presence of NAD^+. UDP-GlcA is then oxidatively decarboxylated by the C-terminal domain of ArnA/PmrI (ArnA decarboxylase domain) to yield uridine 5′-diphospho-β-(L-*threo*-pentapyranosyl-4″-ulose) (UDP-Ara4O). This intermediate is then transaminated by ArnB (PmrH) to produce UDP-4-aminoarabinose (UDP-Ara4N) in a thermodynamically unfavorable reaction ($K_{eq} = 0.1$) (*25*). Raetz and co-workers have shown that ArnA can formylate UDP-Ara4N using *N*-10-formyltetrahydrofolate as the formyl donor (*25*), to yield uridine 5′-diphospho-β-(4-deoxy-4-formamido-L-arabinose) (UDP-Ara4NF) (*27*). However, no formylated products have been found modifying lipid A. Therefore, it has been proposed that UDP-Ara4N is transiently formylated by ArnA to compensate for the unfavorable equilibrium in the conversion of UDP-Ara4O to UDP-Ara4N. Preliminary data also obtained by Raetz and co-workers show the next enzyme in the pathway, ArnC (catalyzing the transfer of Ara4NF to undecaprenyl phosphate, thus preparing the molecule for translocation through the inner membrane), to be selective for UDP-Ara4NF (*25*). Very recent work by Raetz and co-workers shows that the transient formylation of UDP-Ara4N catalyzed by the N-terminal domain of ArnA is essential for both lipid A modification with Ara4N and polymyxin resistance (*27*).

ArnA (formerly PmrI) is a bifunctional 74 kDa protein (*24*, *25*, *28*). We have previously shown that the C-terminal region of *Escherichia coli* ArnA (amino acids 306−660) constitutes a separable domain fully responsible for the C-4″ oxidation of UDP-GlcA to UDP-Ara4O (*28*). The N-terminal domain of ArnA (amino acids 1−300) displays similarity to enzymes involved in formyl transfer including glycinamide ribonucleotide formyltransferase, methionyl-tRNA formyltransferase, and the hydrolase domain of *N*-10-formyltetrahydrofolate dehydrogenase (*23*, *25*, *26*). Here we show that the N-terminal domain of *E. coli* ArnA is, indeed, a separable domain responsible for the transformylation reaction. We designate it the ArnA transformylase domain and analyze its crystal structure refined to 1.7 Å resolution.

MATERIALS AND METHODS

Cloning of the ArnA Transformylase Domain. The *E. coli* ArnA N-terminus was amplified by polymerase chain reaction (PCR) from genomic *E. coli* DNA using the following primers: sense primer, 5′-GGA AAA TGA CAT ATG AAA ACC GTC GTT TTT GCC containing an *Nde*I restriction site, and antisense primer, 5′-GCT ATT CCC GGG TGA ACC TTG CAC CAG GCC containing a *Xma*I restriction site. The PCR amplification was performed with Pfu Turbo polymerase (Invitrogen) according to manufacturer's instructions. The PCR product was purified with the QIAquick PCR purification kit (Qiagen) followed by digestion with *Nde*I and *Xma*I (New England BioLabs) overnight at 37 °C. The digested gene was purified on a 1% agarose gel using the QIAquick gel purification kit (Qiagen). The purified gene was ligated into the pTYB2 vector (New England BioLabs), resulting in a C-terminal intein−chitin binding domain (CBD) fusion. The resulting plasmid, pMS195, was sequenced to confirm that no mutations had been introduced in the ArnA N-terminal sequence.

Protein Expression and Purification. The plasmid pMS195 was transformed into *E. coli* Rosetta (DE3) pLysS cells (Novagen) and plated on LB media supplemented with 100 μg/mL ampicillin. An overnight culture (100 mL) from a single colony containing 100 μg/mL ampicillin and 34 μg/mL chloramphenicol was used to inoculate 6 × 1 L of LB medium supplemented with the above antibiotics. Cultures were grown at 37 °C to an OD_{600} of 0.6 and cooled on ice to approximately 4 °C. Expression was induced with 0.4 mM isopropyl β-D-thiogalactopyranoside (IPTG; Gold Bio Technology Inc.), and cultures were allowed to grow overnight at room temperature. Cells were harvested by centrifugation at 6000 rpm for 15 min at 4 °C, and the cell pellet was resuspended in lysis buffer containing 20 mM Tris-HCl, pH 8.0, 1 mM EDTA, and 0.1% Triton X-100. Cells were frozen in liquid nitrogen, thawed, and then sonicated to achieve lysis. Following lysis, NaCl was added to a final concentration of 500 mM and cell debris removed by centrifugation at 16000 rpm for 30 min at 4 °C. The supernatant was applied to a 20 mL chitin bead column (New England BioLabs) preequilibrated with column buffer containing 20 mM Tris-HCl, pH 8.0, 0.1 mM EDTA, 500 mM NaCl, and 0.1% Triton X-100. The column was washed with 10 column volumes of the above buffer, followed by a quick wash with 3 column volumes of column buffer containing 30 mM freshly diluted DTT (cleavage buffer). The flow was stopped and the column left to incubate with the cleavage buffer overnight at 4 °C. The protein was eluted with 3 column volumes of the column buffer. Fractions containing the

protein were loaded on a size exclusion (HiLoad 26/60 Superdex 200, Amersham Pharmacia Biotech) column pre-equilibrated with 25 mM Tris-HCl, pH 8.0, 200 mM KCl, 10% glycerol, and 10 mM 2-mercaptoethanol and eluted in the same buffer. Elution was monitored by measuring the absorption at 280 nm. The ArnA N-terminus eluted as a monomer from the column. The fractions containing protein were combined, and the protein was concentrated to approximately 9.5 mg/mL (Bio-Rad protein assay; Bio-Rad Laboratories). This protein stock was used for crystallization experiments.

Full-length *E. coli* ArnA and the ArnA decarboxylase domain were overexpressed and purified as described previously (*28*).

Protein Crystallization and Data Collection. Initial crystallization screening with the Hampton Research sparse matrices (crystal screen I and II) yielded small, thin rhombohedral plates from an unbuffered solution of 2.0 M NaCl and 10% PEG 6000 as precipitant. Larger, single crystals were obtained refining the condition to a precipitant solution of 12% PEG 6000, 0.5 M NaCl, and 100 mM HEPES, pH 7.25, using the hanging drop method at 16 °C and a 1:1 protein:precipitant ratio. Single, well-formed crystals with a maximum size of 0.5 × 0.3 × 0.03 mm were grown in 1 week. Prior to X-ray data collection the crystals were transferred to a cryoprotecting solution composed of mother liquor containing 20% glycerol, soaked for 10 min, and flash cooled in a nitrogen stream. Initial data to 2.0 Å were collected with a rotating anode generator using Cu Kα radiation and a Rigaku RAXIS IV^{2+} detector. A data set to higher resolution of 1.7 Å was collected on beam line 5.0.3 at the Advanced Light Source in Berkeley. Data were indexed and integrated with DENZO and scaled with SCALEPACK (*29*). X-ray data collection statistics are shown in Table 1.

Structure Determination and Refinement. The structure of the ArnA N-terminus was solved by molecular replacement with the 2.0 Å data set. A sequence similarity search with the ArnA N-terminus against the PDB database returned a few structures as possible models for molecular replacement. Methionyl-tRNA (f)Met formyltransferase from *E. coli* had the highest similarity score (29% identity, *E* value 1×10^{-33}), and its 2.0 Å refined structure (PDB_ID 1FMT) (*30*) was selected as the search model. All non-glycine side chains of the model were set to Ala. Rotation/translation searches, performed with the program AMoRe (*31*) and data between 15 and 4 Å, yielded a solution clearly above the noise level. Inspection of the crystal packing revealed no unfavorable molecular contacts. Using CNS (*32*), 10% of the data were removed for cross-validation, and the model was subjected to rounds of simulated annealing with torsion angle dynamics (*33*, *34*) alternated with manual rebuilding using the program O (*35*). Model phases were improved by solvent flipping as implemented in CNS (solvent content of 62%). A new map calculated with the improved phases showed unambiguous density for most side chains and connectivity for most of the molecule. The model was subjected to a round of simulated annealing with Cartesian dynamics followed by rounds of positional and *B*-factor refinement with data to 2.0 Å until no further improvement of *R*-factors was achieved. At this point a higher resolution data set became available. After transferring the test set of reflections to the higher resolution data set, the structure was further refined to 1.7 Å resolution. During this refinement the electron density maps showed clear density for several water molecules, and these were included in the model as well. The final model has good stereochemistry as determined using the program PROCHECK (*36*) with all amino acids lying in the most favorable or allowed regions on the Ramachandran plot. Refinement statistics and model stereochemistry are summarized in Table 1.

Table 1: Data Collection and Refinement Statistics[a]

Data Collection Statistics	
wavelength (Å)	1.00
space group	$C222_1$
cell parameters (Å)	$a = 59.1$, $b = 113.3$, $c = 125.0$; $\alpha = \beta = \gamma = 90°$
resolution (Å)	30.0−1.70 (1.76−1.70)
measured reflections	131346 (13027)
unique reflections	46108 (4557)
I/σ	18.1 (7.2)
redundancy	2.8 (2.9)
data completeness (%)	99.0 (98.8)
R_{merge} (%)	2.5 (15.6)
Refinement Statistics	
no. of reflections used in refinement	45262
% of reflections used in R_{free}	10.1
no. of protein atoms	2209
no. of water molecules	274
R_{work}	19.8 (21.5)
R_{free}	22.5 (25.0)
mean *B*-value, overall (Å^2)	21.8
mean *B*-value, protein (Å^2)	20.7
mean *B*-value, solvent (Å^2)	29.9
B-factor deviation bonds (Å^2)	1.21
B-factor deviation angles (Å^2)	1.81
RMS deviation from ideal values	
bond lengths (Å)	0.011
bond angles (deg)	1.6
Ramachandran: residues in	
most favored region (%)	93.1
allowed regions (%)	6.9

[a] $R_{work} = \sum|F_o - F_c|/\sum F_o$, where F_o = the observed structure factor amplitude and F_c = the structure factor calculated from the model. R_{free} is computed in the same manner as R_{work}, using the test set of reflections.

ArnB Cloning and Purification. ArnB was amplified by PCR from genomic *E. coli* DNA using the following primers: sense primer, 5′-AGC AAA AAC ATA TGG ACA GGT ATT CAA TGG CGG containing an *Nde*I site, and antisense primer, 5′-TTC AAC GAG CTC TTA TTG TCC TGC GAG TTG CTG containing a *Sac*I site. The PCR product was purified with the QIAquick PCR purification kit (Qiagen) followed by digestion with *Nde*I and *Sac*I (New England BioLabs) overnight at 37 °C. The digested gene was purified on a 1% agarose gel using the QIAquick gel purification kit (Qiagen). The purified gene was ligated into pMS122 [an engineered variant of the pET28 vector that generates an N-terminal His-tag fusion that can be efficiently and specifically cleaved with the tobacco etch virus (TEV) protease]. The resulting plasmid was sequenced to confirm that no mutations had been introduced in the ArnB sequence. pMS122-ArnB was transformed into *E. coli* Rosetta (DE3) cells (Novagen) and plated on LB media supplemented with 50 μg/mL kanamycin. Cells were grown from a single colony in LB media supplemented with 50 μg/mL kanamycin to an OD_{600} of 0.6 and induced with 0.4 mM IPTG for 4 h at 37 °C. Cells were harvested by centrifugation at 6000 rpm for 10 min at 4 °C, and the cell pellet was resuspended in lysis

buffer containing 25 mM Tris-HCl, pH 8.0, 5 mM 2-mercaptoethanol, and Complete EDTA-free protease inhibitor cocktail (Roche). Cells were lysed on ice by sonication. Following lysis, KCl was added to a final concentration of 300 mM and cell debris removed by centrifugation at 16000 rpm for 30 min at 4 °C. The supernatant was applied to a Ni-NTA column and washed with 5 column volumes of the above lysis buffer, followed by 5 column volumes of wash buffer (25 mM HEPES/KOH, pH 7.5, 300 mM KCl, 10% glycerol, 10 mM 2-mercaptoethanol, 25 mM imidazole). Protein was eluted in 10 column volumes of elution buffer (wash buffer containing 300 mM imidazole). The fractions containing protein were combined, and the 6× His tag was removed by overnight incubation at room temperature with TEV protease. The protein was then dialyzed against 25 mM Tris-HCl, pH 8.0, and 10 mM 2-mercaptoethanol, loaded on a preequilibrated (with 25 mM Tris-HCl, pH 8.0, 150 mM KCl, 10% glycerol, 1 mM EDTA, 10 mM 2-mercaptoethanol) size exclusion column (HiLoad 26/20 Superdex 200; Amersham Pharmacia Biotech), and eluted in the same buffer. Elution was monitored by measuring the absorption at 280 nm.

Site-Directed Mutagenesis. Point mutations in the ArnA transformylase domain cloned in pTYB2 were introduced using the QuickChange site-directed mutagenesis kit (Stratagene) according to the manufacturer's instructions. During the mutagenesis procedure a unique restriction site was also introduced (by silent mutation) next to the site of the point mutation to facilitate the selection of clones containing the mutant. After the introduction of the mutation the vector was transformed into *E. coli* XL-10 cells (Stratagene) and plated on LB ampicillin plates. Plasmids were isolated by using the QIAprep spin miniprep kit (Qiagen). Clones carrying the mutation were identified by restriction digest, taking advantage of the introduced unique restriction site, and sequenced to ensure the absence of random mutations. The resulting mutants were overexpressed and purified as already described for the ArnA transformylase domain.

Synthesis of N-10-Formyltetrahydrofolate. N-10-Formyltetrahydrofolate was obtained from the commercially available 5-formyltetrahydrofolate (Sigma) as described previously (*37*, *38*). In short, 7 mg of 5-formyltetrahydrofolate was dissolved in 1.5 mL of deionized water and the pH lowered to 1.9 by the addition of 0.1 N HCl. The volume was brought to 2.2 mL with deionized water and the reaction allowed to proceed at room temperature until completion. The formation of 5,-10-methenyltetrahydrofolate was followed by measuring the increase in absorbance at 355 nm. The obtained compound was stored at −20 °C. Conversion of *N*-5,10-methenyltetrahydrofolate to *N*-10-formyltetrahydrofolate was achieved before use in assays by performing a 10 min preincubation of 5,10-methenyltetrahydrofolate in 50 mM HEPES, pH 7.5, in 1:1 (v/v) ratio. This allowed us to obtain an approximately 3 mM *N*-10-formyltetrahydrofolate stock solution.

Thin-Layer Chromatography Assays. A solution containing 25 mM Tris-HCl, pH 8.0, 5 mM 2-mercaptoethanol, 0.1 mg/mL purified ArnA C-terminal domain, 16.5 μM UDP-GlcA [glucuronyl-^{14}C (U); Perkin-Elmer], and 3 mM NAD^+ (Roche) was incubated at room temperature for 90 min to obtain the UDP-4″-ketopentose. A portion of the above mixture was diluted two times with L-Glu (final concentration of 1 mM) and ArnB (final concentration of 0.2 mg/mL) and incubated for 90 min at room temperature to obtain UDP-4-aminoarabinose (UDP-Ara4N). Conversion of UDP-Ara4N to the N-formylated product (UDP-Ara4FN) was obtained as follows. *N*-10-Formyltetrahydrofolate and enzyme (ArnA full length or the ArnA transformylase domain) were added to 10 μL of the UDP-Ara4N reaction mixture to bring the final volume to 20 μL. The reaction was set up to yield final concentrations of 1 mM *N*-10-formyltetrahydrofolate and 0.1 mg/mL enzyme. The reaction was allowed to proceed for 60 min at room temperature. Three microliters of each reaction mixture was then spotted on a polyethylenimine (PEI)−cellulose plate (Merck) prewashed in methanol. The plate was developed in a solvent system containing 0.25 M acetic acid and 0.4 M LiCl (*24*, *25*). Radioactivity in the plate was visualized with a PhosphorImager.

Spectrophotometric Assay of Transformylase Activity. The conversion of *N*-10-formyltetrahydrofolate to tetrahydrofolate can be monitored by the increase in absorbance at 300 nm and can be used to measure transformylation reactions (*39*). Since UDP-Ara4N, the formyl acceptor for the reaction catalyzed by ArnA, is not commercially available, it was generated enzymatically as follows. A reaction mixture containing 400 μM UDP-GlcA, 400 μM NAD^+, and 750 nM ArnA decarboxylase in a buffer containing 25 mM Tris-HCl, pH 8.0, 10% glycerol, 100 mM KCl, and 5 mM 2-mercaptoethanol was incubated at room temperature until complete conversion of UDP-GlcA to UDP-Ara4O was achieved. Conversion was followed by detecting the increase of absorbance at 340 nm due to the release of NADH. Purified ArnB and glutamate were added to the above reaction mixture to achieve a 2-fold increase of volume. The final concentrations were 800 μM glutamate and 0.02 mg/mL ArnB. The reaction was allowed to proceed for 2 h at room temperature. This assured the conversion of UDP-Ara4O to UDP-Ara4N.

To measure ArnA transformylase activity, 400 μL of the above reaction cocktail was mixed with 75 μM *N*-10-formyltetrahydrofolate (produced by 10 min preincubation of 5,10-methenyltetrahydrofolate in 50 mM HEPES, pH 7.5) and wild-type ArnA_TF at a concentration of 0.02 mg/mL or mutant enzyme at a concentration of 0.1 mg/mL to obtain a final volume of 800 μL. The conversion of *N*-10-formyltetrahydrofolate to tetrahydrofolate (THF) was monitored by measuring the increase of absorbance at 300 nm. The rate of THF production was calculated from the blank-corrected absorbance rate using an extinction coefficient of $19.5 \times 10^3\ M^{-1}\ cm^{-1}$ (*39*).

RESULTS AND DISCUSSION

The N-Terminal Domain of ArnA Contains the Transformylase Activity. As already mentioned, ArnA is a bifunctional enzyme with its 300 N-terminal residues showing sequence similarity to transformylases that utilize *N*-10-formyltetrahydrofolate as a formyl donor. This suggested that the N-terminal domain of ArnA may represent a separable domain responsible for the transformylation reaction catalyzed by ArnA. Therefore, we prepared a fragment of *E. coli* ArnA containing residues 1−302 as a fusion with the self-cleavable intein−chitin binding domain tag. Using this construct, the ArnA N-terminal domain was purified to homogeneity as described in Materials and Methods. The

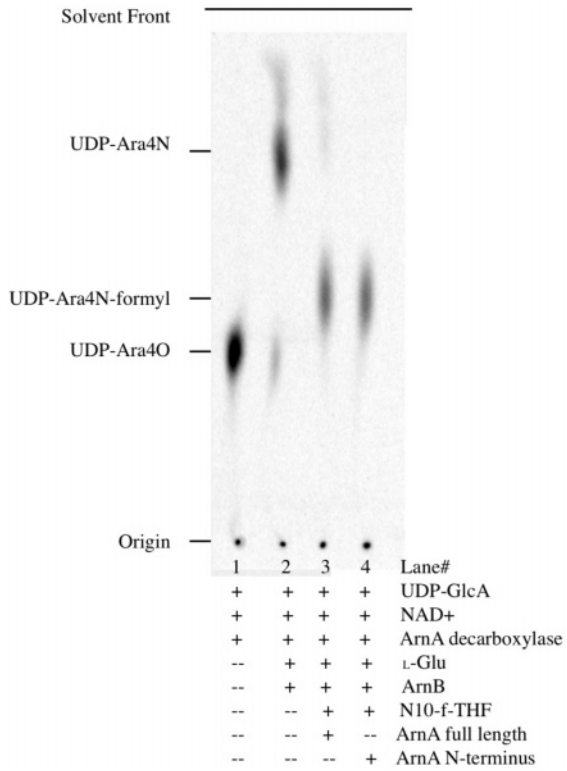

FIGURE 2: The N-terminal domain of ArnA is a separable domain capable of performing the transformylation reaction. Lane 1: Formation of UDP-Ara4O from UDP-GlcA and NAD^+ by the ArnA decarboxylase domain. Lane 2: Formation of UDP-Ara4N in the presence of L-glutamate by ArnB. Formation of UDP-Ara4-formyl-N from UDP-Ara4N and *N*-10-formyltetrahydrofolate by ArnA full length (lane 3) and the ArnA N-terminal domain (lane 4).

N-terminal fragment obtained after cleavage of the purification tag contained amino acids 1−302 from ArnA plus two additional amino acids (PG) at the C-terminus.

Breazeale et al. showed that the formation of the ArnA formylated product, UDP-β-(4-deoxy-4-formamido-L-arabinose) (UDP-Ara4FN), could be detected by thin-layer chromatography, where UDP-Ara4FN migrates slower than the substrate UDP-Ara4N (*24*, *25*, *27*). We used this observation to test if our N-terminal construct retained the transformylation activity of the full-length ArnA. As shown in Figure 2, the ArnA N-terminal domain is also able to produce the formylated species. We therefore conclude that the ArnA N-terminal domain is a separable, functional domain responsible for the transfer of a formyl group from *N*-10-formyltetrahydrofolate to UDP-Ara4N. We shall refer to it as the ArnA transformylase domain (ArnA_TF).

The two-domain organization of ArnA is reminiscent of *N*-10-formyltetrahydrofolate dehydrogenase (FDH), an enzyme that converts *N*-10-formyltetrahydrofolate into tetrahydrofolate (THF) (*40*). In vitro, FDH can catalyze the production of THF and CO_2 in a $NADP^+$-dependent dehydrogenase reaction or of THF and formate in a $NADP^+$-independent, hydrolase reaction (*41*). The 310 amino acid, N-terminal domain of FDH (N_t-FDH) is similar to other *N*-10-formyltetrahydrofolate binding enzymes and shares 26% sequence identity with the ArnA_TF domain. The larger, 592 amino acid C-terminal fragment of FDH contains a domain similar to aldehyde dehydrogenases (*40*, *42*). However, the dehydrogenase domain of FDH does not appear to have significant sequence similarity with the 350 amino acid C-terminal domain of ArnA (ArnA decarboxylase).

When expressed separately, the N-terminus of FDH retains in vitro hydrolase activity (*43*). However, this activity requires the presence of 2-mercaptoethanol (*41*, *43*). The dehydrogenase activity cannot be catalyzed by the isolated FDH C-terminus. This reaction requires full-length FDH (*42*). This is in sharp contrast with ArnA, where both the N-terminal and C-terminal domains can be expressed separately, retaining their enzymatic activities (*28*).

The C-terminal domain of ArnA, ArnA decarboxylase, catalyzes the second reaction in the pathway for lipid A−Ara4N biosynthesis (Figure 1), while its N-terminal domain, ArnA_TF, catalyzes the fourth reaction. The intervening reaction is catalyzed by a separate enzyme (ArnB). It is tempting to speculate that these enzymes may form a complex where the products are channeled from one active site to the next. However, in preliminary experiments we have not been able to detect any interaction between ArnA and ArnB. It is also possible that additional enzymes in the pathway, like ArnC, are needed to form such multienzyme complex.

Crystal Structure of the ArnA Transformylase Domain. Crystals of ArnA_TF were obtained by the hanging drop method of vapor diffusion from a precipitant containing 100 mM HEPES, pH 7.25, 0.5 M NaCl, and 12% PEG 6000. The crystals belonged to space group $C222_1$, with typical unit cell dimensions of $a = 59.1$ Å, $b = 113.3$ Å, $c = 125.0$ Å, $\alpha = \beta = \gamma = 90°$, and one molecule per asymmetric unit. The structure was solved by molecular replacement and refined to 1.7 Å resolution as detailed in Materials and Methods. The final model comprises residues 1−300 and 274 water molecules with a crystallographic *R*-factor = 19.8 ($R_{free} = 22.5$). There are two small breaks in the chain (N_{35}−A_{40} and H_{250}−S_{252}) for which no clear density was visible, presumably due to conformational flexibility. The model displays good stereochemistry with all non-glycine residues in the allowed regions of the Ramachandran plot and 93.1% of them lying in the most favorable regions. Complete data collection and refinement statistics are shown in Table 1.

The structure of ArnA_TF comprises two subdomains (Figure 3A): a bigger N-terminal subdomain containing residues M_1−G_{183} (blue) and a smaller C-terminal subdomain consisting of residues P_{213}−Q_{300} (green). A long interdomain connector, containing residues G_{183}−P_{213} (magenta), links the two subdomains together. The N-terminal subdomain is identifiable as a Rossmann fold. It contains a seven-stranded mixed-type β-sheet ($\beta1$−$\beta5$, $\beta7$, $\beta8$) sandwiched by α-helices on both sides. The $\beta6$ strand does not participate in this fold but instead forms a small two-stranded parallel β-sheet with $\beta9$, which is part of the interdomain connector. The C-terminal subdomain contains a five-stranded antiparallel β-sheet flanked by one α-helix on each side. This arrangement forms an open β-barrel that resembles an OB fold as previously described for methionyl-tRNA formyltransferase (FMT) (*30*).

Comparison of ArnA Transformylase with Other N-10-Formyltetrahydrofolate Binding Enzymes. The structure of ArnA_TF reveals that the enzyme is closely related to other enzymes that use *N*-10-formyltetrahydrofolate as a substrate, such as methionyl-tRNA formyltransferase (*30*) and the

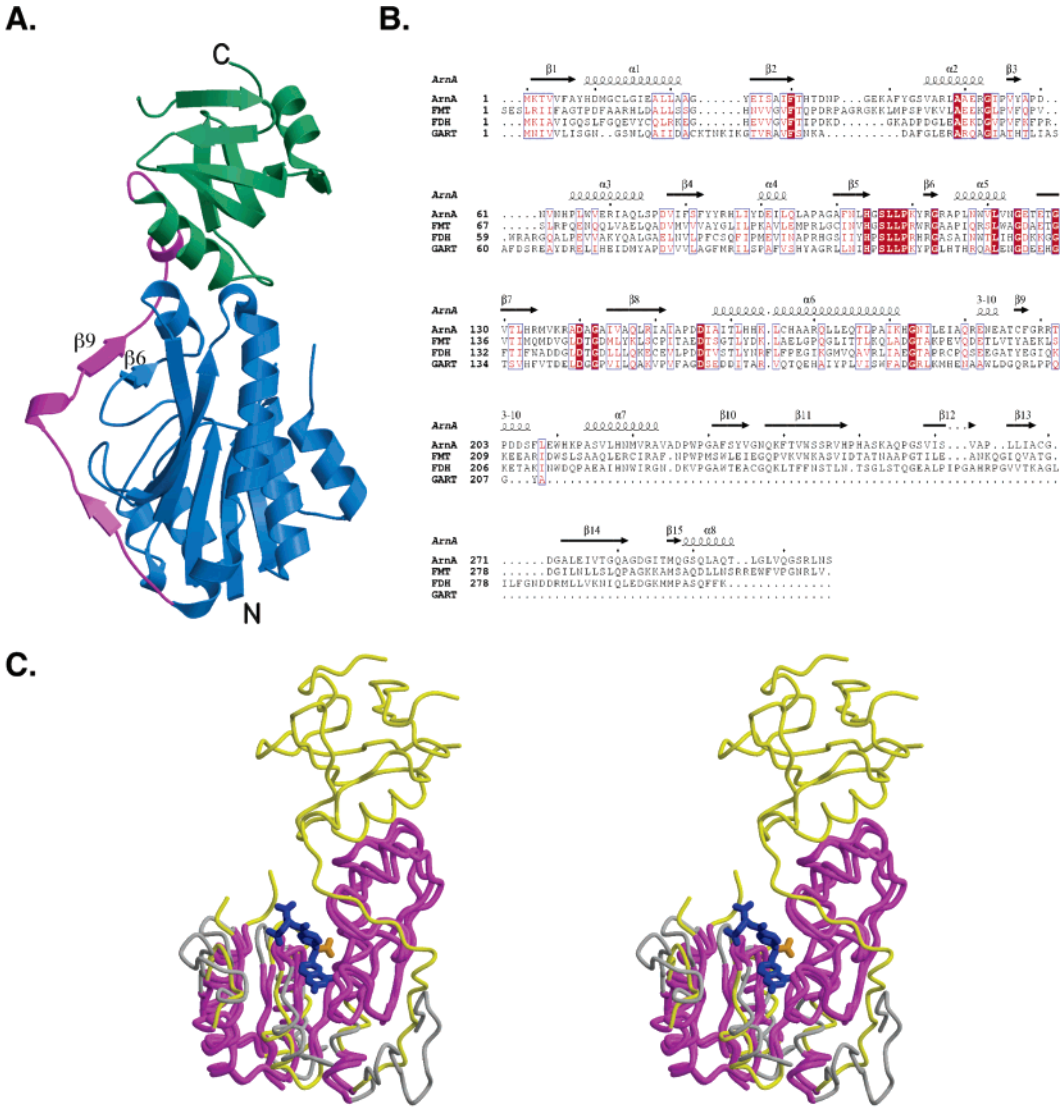

FIGURE 3: (A) Overall structure of ArnA transformylase. The N-terminal domain (in blue) is formed by residues M_1 to G_{183}, and the C-terminal domain (in green) is formed by residues P_{213}−Q_{300}. Both domains are connected by a long stretch of polypeptide (in magenta) containing residues G_{183}−P_{213}. β6 does not participate in the Rossmann fold formed by the N-terminal residues but instead forms a small parallel β-sheet with β9 from the interdomain connector. This figure was prepared using MOLSCRIPT (*57*) and Raster 3D (*58*). (B) Structure-based sequence alignment of ArnA transformylase with other *N*-10-formyltetrahydrofolate utilizing enzymes with known structure. The proteins are the *E. coli* ArnA transformylase domain (ArnA), *E. coli* methionyl-tRNA formyltransferase (FMT), *R. norvegicus N*-10-formyltetrahydrofolate dehydrogenase, N-terminal hydrolase domain (N_t-FDH), and *E. coli* glycinamide ribonucleotide transformylase (GART). The structural alignment was performed with 3Dcoffee (*59*). The secondary structure assignment for ArnA transformylase is shown with arrows representing β-strands and coils representing α-helices. This figure was prepared with the program ESPript (*60*). (C) Stereoview showing superposition between ArnA transformylase (in yellow) and GART (in gray); α carbons of both proteins used to calculate the RMS deviation (RMSD of 1.32 Å over 148 Cα) are colored in magenta. The substrate analogue 10-formyl-5,8,10-trideazafolic acid (NHS; PDB_ID 1C2T) cocrystallized with GART is colored in blue. NHS is bound as the hydrated (10*S*) diastereomer (*47*), and the hydrated formyl of NHS is colored in orange.

hydrolase domain of *N*-10-formyltetrahydrofolate dehydrogenase (N_t-FDH) (*41*).

While the sequence identity between ArnA_TF and FMT is relatively low (29%), the two structures superimpose with an RMS deviation of 1.23 Å for 173 Cα in structurally conserved regions. Similarly, ArnA_TF and N_t-FDH share 26% sequence identity, but the two structures are homologous and superimpose with an RMS deviation of 1.36 Å for 172 Cα in the conserved core. The similarity between ArnA_TF and FMT and N_t-FDH extends to both the N-terminal Rossmann fold subdomain and the C-terminal OB fold subdomain.

The N-terminal subdomain of ArnA_TF (residues 1−203) also shows 26% sequence identity to glycinamide ribonucleotide formyltransferase (GARF), a 10-formyltetrahydrofolate binding enzyme extensively studied both mechanistically and structurally (*44*−*51*). The GARF structure superimposes with the N-terminal subdomain of ArnA_TF with an RMS deviation of 1.32 Å for 148 conserved Cα atoms, but it lacks the C-terminal, OB fold-like domain present in ArnA_TF (Figure 3C).

Sequence conservation among the four enzymes mentioned above is low (Figure 3B). However, an asparagine−histidine−aspartate triad identified as important for catalysis in FMT (*52*) and GARF (*50*) is conserved in ArnA_TF in terms of both sequence (N_{102}, H_{104}, D_{140}) and conformation (Figure 4A). All four enzymes also retain the SLLP motif common to many 10-formyltetrahydrofolate binding enzymes (*30*).

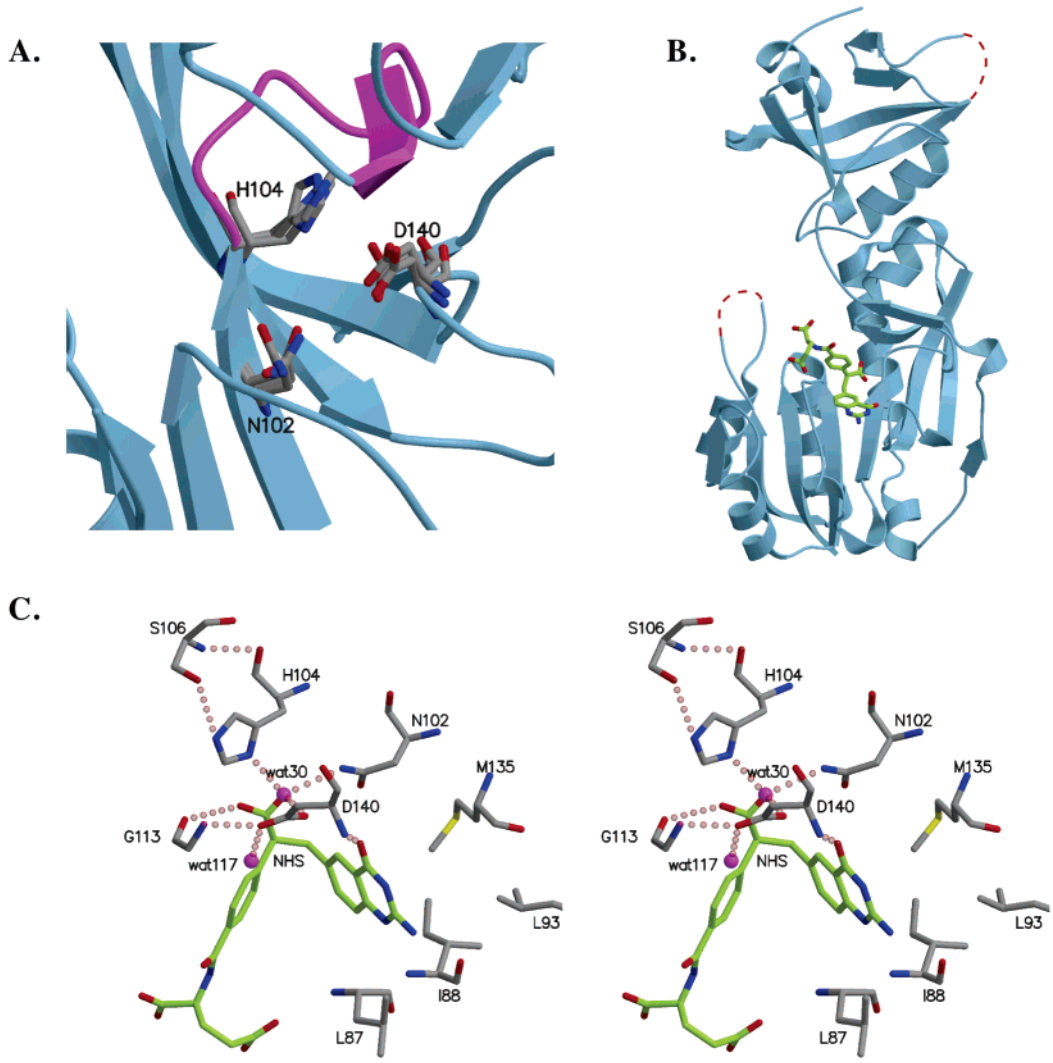

FIGURE 4: (A) Superposition of catalytic residues (N_{102}, H_{104}, D_{140}) in ArnA transformylase, FMT, and GART showing that the side chains of these residues adopt similar conformations in all three enzymes. The HxSLLPxxxG motif in ArnA transformylase is colored in magenta. (B) ArnA transformylase with NHS modeled in the active site. The resulting model shows no steric conflicts between ArnA transformylase and the *N*-10-formyltetrahydrofolate analogue. The two flexible loops ($N_{35}-A_{40}$ and $H_{250}-S_{252}$) for which no electron density is observed in our structure are shown as dashed lines colored in red. (C) Stereoview of the NHS binding pocket. In our crystal structure there is a water molecule (Wat30, in magenta) sitting where the second hydroxyl of the hydrated aldehyde is modeled. The water molecule (Wat117) proposed to play a role of a proton shuttle in the transformylation mechanism is also shown in magenta. Dotted lines represent the hydrogen-bonding network in the vicinity of the hydrated formyl of NHS. The residues (L_{87}, I_{88}, L_{93}, and M_{135}) forming a hydrophobic pocket in which the bicyclic ring of NHS sits are also shown (A_{139} is not shown for clarity of the figure). Carbon atoms of NHS are colored in green, and carbon atoms from side chains of residues belonging to ArnA transformylase are colored in gray.

The polypeptide containing the SLLP motif in ArnA_TF ($S_{106}LLP_{109}$) adopts a type VIa turn conformation in which the proline is involved in a cis peptide bond. This turn is followed by a short β-strand (β6) ending on a glycine residue (G_{113}) that is also conserved (Figures 3B and 4A). The structure of this turn−strand motif is stabilized by electrostatic interactions with the β9 strand of the interdomain connector (Figures 3A and 4A), and a similar conformation is observed in FMT (*30*) and N_t-FDH (*41*). Since GARF does not have a C-terminal subdomain nor an interdomain connector, this turn−strand motif is disordered in the unliganded structure of GARF (*44*) but becomes ordered upon binding of tetrahydrofolate analogues and adopts a conformation similar to that observed in ArnA_TF (*44*, *48*). All four enzymes (ArnA_TF, FMT, N_t-FDH, and liganded-GARF) superimpose in this region with an RMS deviation of 0.23 Å or less.

On the basis of this sequence and structural conservation, we propose that the extended sequence motif HxSLLPxxxG (in ArnA_TF this is represented by residues **H**$_{104}$G**SLLP**KYR**G**$_{113}$) can be used to identify *N*-10-formyltetrahydrofolate binding enzymes with Rossmann folds such as ArnA_TF. We used the program CoSMoS (www.biology.lsa.umich.edu/~jakoblab/cosmos/) to search for SLLP or HxSLLPxxxG as conserved sequence motifs. While SLLP retrieved formyltransferases together with many other families of proteins, the HxSLLPxxxG motif retrieved only proteins annotated as formyltransferases (242 sequences). In all of these sequences the histidine, proline, and glycine are strictly conserved while some S to T and L to hydrophobic substitutions were observed. The cases of strict conservation can be explained by the histidine being crucial for catalysis (see below), the *cis*-proline being the defining feature of the type VIa turns observed in all crystal structures of these

enzymes, and the glycine adopting ϕ and ψ angles (ϕ = 93.2° and ψ = −153.8° for G_{113} in ArnA_TF) that fall in the disallowed region of the Ramachandran plot.

Substrate Binding Model. The N−H−D triad of residues identified to be important for catalysis in GARF and FMT (*51*, *52*) is represented by residues N_{102}, H_{104}, and D_{140} in ArnA_TF. The structure superposition shows that the side chain conformations of these residues are also very well conserved (Figure 4A). The position of the catalytic residues is stabilized by an extensive network of hydrogen-bonding interactions (Figure 4C). The side chain of the conserved H_{104} is stabilized by two hydrogen bonds: one between the histidine N2 and the side chain oxygen of S_{106} and another between the histidine N1 and a well-ordered water molecule (Wat30) that brings together all three catalytic residues.

The structure of GARF has been solved with various *N*-10-formyltetrahydrofolate analogues bound to the active site including the formylated analogue 10-formyl-5,8,10-trideazafolic acid (NHS) (*47*). On the basis of the sequence and structure similarities between GARF and ArnA_TF described above, we modeled NHS in the active site of ArnA_TF using as a guide the position of NHS in GARF [PDB_ID 1C2T (*47*)]. The resulting model showed no steric clashes between side chains of ArnA_TF and NHS except for the NH1 atom in the side chain of R_{114} making a very close contact with the carbonyl group at C17 of NHS. Given the large conformational flexibility of the arginine side chain, we assume that it could adopt a different conformation in the liganded structure. Otherwise, the inhibitor fits very well in the active site of ArnA_TF (Figure 4B).

As shown in Figure 4C, the three catalytic residues, N_{102}, H_{104}, and D_{140}, are in proximity to the C10 of NHS (the NHS analogue has a carbon atom at position 10 instead of N and is therefore unable to donate the formyl group). Moreover, the O4 of the NHS bicyclic ring is within hydrogen-bonding distance of the $O_{\delta}1$ and the backbone N atom of D_{140}. Finally, the bicyclic ring of NHS sits in a hydrophobic pocket formed by residues L_{87}, I_{88}, L_{93}, M_{135}, and A_{139} (Figure 4C).

The electron density observed in the NHS-GARF crystal structure as well as NMR studies suggested that, at neutral pH, the formyl group of NHS is actually bound as the hydrated (10*S*) diastereomer (*gem*-diol) (*47*). Remarkably, a water molecule in our crystal structure (Wat30 in Figure 4C) sits where the formyl oxygen of NHS is modeled. This water molecule makes hydrogen bonds with $O_{\delta}1$ of D_{140} (2.9 Å), $N_{\delta}2$ of N_{102} (2.9 Å), and $N_{\delta}1$ of H_{104} (2.7 Å), the three residues implicated in the formyl transfer reaction (Figure 4C). The main chain atoms of G_{113} hydrogen bond with both the hydroxyl of the hydrated aldehyde of NHS [that represents where the primary amino group from the second substrate, UDP-Ara4N, would be (*47*)] and the side chain oxygen of the catalytically important D_{140} bringing them together. Therefore, G_{113} positions the amino group of UDP-Ara4N and the catalytic D_{140} in a favorable conformation for catalysis. The ϕ and ψ angles required for these interactions are best accommodated by a glycine, which likely explains its conservation.

The structures of both GARF and FMT have also been solved with the formyl acceptor substrate bound to the enzymes (*44*, *53*). However, it was not possible to model the binding of UDP-Ara4N to ArnA_TF based on these structures. GARF lacks the C-terminal, OB fold domain present in ArnA_TF and binds its substrate (GAR) in a way incompatible with ArnA_TF. FMT does have the C-terminal OB fold domain, which is important for binding of the tRNA portion of the methionyl-tRNA substrate, but again this large substrate is bound in a way incompatible with ArnA_TF. It is, however, likely that the C-terminal domain of ArnA_TF is important for UDP-Ara4N binding because a truncated fragment of ArnA_TF (with removed C-terminal subdomain) was unable to catalyze formyl transfer to UDP-Ara4N (data not shown). Further experimentation is required to test this hypothesis.

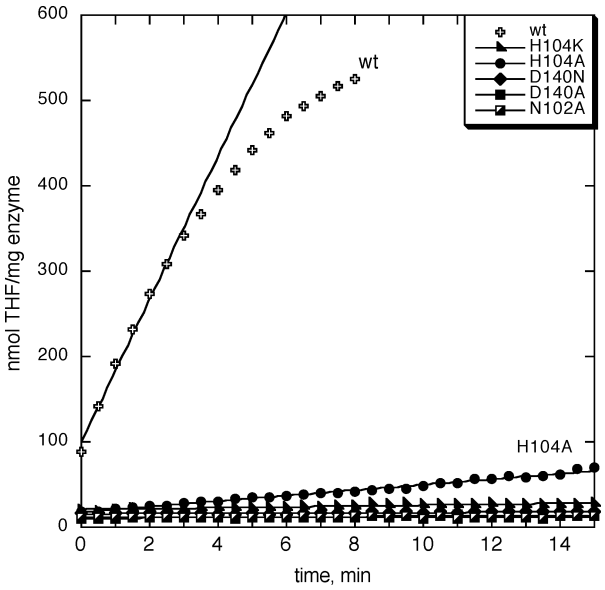

FIGURE 5: Transformylase activity of ArnA_TF and mutants. Wild-type ArnA_TF and mutants of interest were assayed spectrophotometrically for conversion of UDP-Ara4N to UDP-Ara4FN as described in Materials and Methods. The following enzymes were used: wild-type ArnA_TF (crosses), H104K (triangles), H104A (circles), D140N (diamonds), D140A (squares), and N102A (semi-filled squares). All of the mutants except H104A had activity at least 150 times lower than the wild-type enzyme.

Table 2: Characteristics of Wild-Type ArnA_TF and Mutants

enzyme	rate, nmol/(min·mg)
wild-type ArnA_TF	84 ± 3
H104A	3.24 ± 0.06
H104K	0.51 ± 0.04
D140N	0.16 ± 0.02
D140A	0.25 ± 0.03
N102A	0.00 ± 0.03

Proposed Catalytic Mechanism for the Transformylation Reaction. The conservation of the catalytic triad N_{102}, H_{104}, and D_{140} suggested the involvement of these residues in formyl transfer by ArnA_TF. We confirmed this hypothesis by generating five point mutants (H104A, H104K, D140A, D140N, and N102A) of ArnA_TF and using a spectrophotometric assay to test their transformylase activity (*39*). As shown in Figure 5 and Table 2, the activity of the H104K, D140A, D140N, and N102A mutants was at least 150-fold lower than that of the wild-type Arn_TF, whereas a residual activity 25-fold lower than wild type was observed for the H104A mutant. These results support the involvement of the N_{102}, H_{104}, and D_{140} triad in catalysis. The residual activity of the H104A mutant is in line with results obtained by Benkovic and co-workers for GARF (*51*). Using a large panel

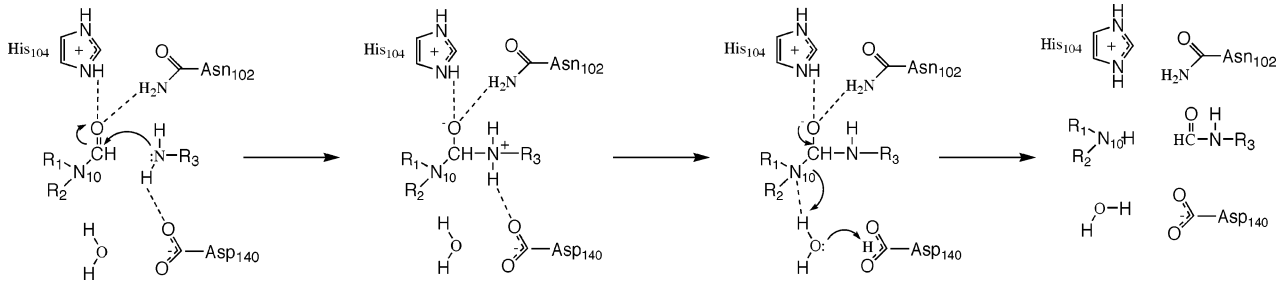

FIGURE 6: Proposed catalytic mechanism for the transformylation reaction. In this scheme, the formyl group is hydrogen bonded to the protonated imidazole group of H_{104} and to N_{102}. The amino group of UDP-Ara4N (which is in a free base form) nucleophilically attacks the activated carbonyl carbon of the formyl group. A tetrahedral intermediate is formed, and the generated oxyanion is stabilized by hydrogen bonding to H_{104} and N_{102}. Proton transfer from UDP-Ara4N to the N10 of the folate, mediated by D_{140} and a water molecule (Wat117 in Figure 4), is followed by breakdown of the intermediate and release of the reaction products.

of point mutants of the catalytic triad, they showed that a subset of mutations yielded enzymes with residual activity (*51*). This supports a mechanism of transformylation in which all three residues are important, but none is indispensable, for catalysis (*51*). On the basis of the similarities of ArnA with GARF and the 10f-THF binding model discussed above, we propose a GARF-like mechanism (*51*, *54*, *55*) for the transformylation reaction catalyzed by ArnA_TF as shown in Figure 6.

In this proposed mechanism (Figure 6), H_{104} and N_{102} activate the carbonyl carbon of the formyl group, which undergoes nucleophilic attack by the primary amine of UDP-Ara4N. The oxyanion of the putative tetrahedral intermediate is stabilized by H_{104} and N_{102}. A water molecule, properly positioned in the active site by hydrogen bonding with the side chain of D_{140}, works as a proton shuttle, mediating proton transfer from UDP-Ara4N to the folate N10 (*56*). In our crystal structure of ArnA_TF, a well-ordered water molecule (Wat117 in Figure 4C) is well situated to serve as a proton shuttle, with its oxygen within hydrogen-bonding distance to both D_{140} (2.76 Å) and the C10 of the modeled NHS (2.80 Å). The proton transfer is followed by decomposition of the tetrahedral intermediate and release of the products UDP-Ara4FN and tetrahydrofolate.

The involvement of both H_{104} and N_{102} in formyl activation and oxyanion stabilization may help to explain the residual activity we observed in the H104A mutant as the N_{102} could partially fulfill this role. Crystallographic data for GARF suggested that, in that enzyme, the catalytic histidine (H_{108}) imidazole resides in two different conformations (*56*). One of those conformations is equivalent to the one we observe in ArnA_TF (Figure 4A) and serves the proposed role in catalysis (we do not see any evidence for alternate conformations of H_{104} in ArnA_TF). The second conformation of H_{108} in GARF puts the imidazole ring far from the formyl group, preventing its participation in catalysis (*51*, *56*). The residual activity observed in the GARF H108P and H108L mutants also suggests that some catalysis can occur in the absence of the histidine (*51*). On the other hand, we were not able to detect any residual activity in the ArnA_TF N102A mutant. Saturation mutagenesis experiments of all three catalytic residues in ArnA_TF would be needed to fully assess the relative importance of these residues in catalysis.

The transformylase activity of ArnA is required for lipid A modification with Ara4N and bacterial resistance to CAMPs and related antibiotics such as polymyxin. A clear understanding of ArnA_TF structure and mechanism is important for the design and evaluation of selective inhibitors. The crystal structure of the ArnA transformylase domain and the accompanying hypothesis presented here provide a platform for designing and testing of such selective inhibitors that can overcome resistance to CAMPs and polymyxin-like antibiotics.

ACKNOWLEDGMENT

We thank Sandra Metzner for excellent technical assistance and Katherine Kenerson for the expression and purification of ArnB.

REFERENCES

1. Raetz, C. R., Ulevitch, R. J., Wright, S. D., Sibley, C. H., Ding, A., and Nathan, C. F. (1991) Gram-negative endotoxin: an extraordinary lipid with profound effects on eukaryotic signal transduction, *FASEB J. 5*, 2652−2660.
2. Raetz, C. R., and Whitfield, C. (2002) Lipopolysaccharide endotoxins, *Annu. Rev. Biochem. 71*, 635−700.
3. Hoffmann, J. A., Kafatos, F. C., Janeway, C. A., and Ezekowitz, R. A. (1999) Phylogenetic perspectives in innate immunity, *Science 284*, 1313−1318.
4. Scott, M. G., and Hancock, R. E. (2000) Cationic antimicrobial peptides and their multifunctional role in the immune system, *Crit. Rev. Immunol. 20*, 407−431.
5. Zasloff, M. (2002) Antimicrobial peptides of multicellular organisms, *Nature 415*, 389−395.
6. Matsuzaki, K. (1999) Why and how are peptide-lipid interactions utilized for self-defense? Magainins and tachyplesins as archetypes, *Biochim. Biophys. Acta 1462*, 1−10.
7. Yang, L., Weiss, T. M., Lehrer, R. I., and Huang, H. W. (2000) Crystallization of antimicrobial pores in membranes: magainin and protegrin, *Biophys. J. 79*, 2002−2009.
8. Shai, Y. (1999) Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by alpha-helical antimicrobial and cell nonselective membrane-lytic peptides, *Biochim. Biophys. Acta 1462*, 55−70.
9. Gunn, J. S. (2001) Bacterial modification of LPS and resistance to antimicrobial peptides, *J. Endotoxin Res. 7*, 57−62.
10. Gunn, J. S., Ernst, R. K., McCoy, A. J., and Miller, S. I. (2000) Constitutive mutations of the *Salmonella enterica* serovar *typhimurium* transcriptional virulence regulator phoP, *Infect. Immun. 68*, 3758−3762.
11. Guo, L., Lim, K. B., Gunn, J. S., Bainbridge, B., Darveau, R. P., Hackett, M., and Miller, S. I. (1997) Regulation of lipid A modifications by *Salmonella typhimurium* virulence genes phoP-phoQ, *Science 276*, 250−253.
12. Gunn, J. S., Ryan, S. S., Van Velkinburgh, J. C., Ernst, R. K., and Miller, S. I. (2000) Genetic and functional analysis of a PmrA-PmrB-regulated locus necessary for lipopolysaccharide modification, antimicrobial peptide resistance, and oral virulence of Salmonella enterica serovar typhimurium, *Infect. Immun. 68*, 6139−6146.
13. Gunn, J. S., and Miller, S. I. (1996) PhoP-PhoQ activates transcription of pmrAB, encoding a two-component regulatory

system involved in *Salmonella typhimurium* antimicrobial peptide resistance, *J. Bacteriol. 178*, 6857−6864.
14. Roland, K. L., Esther, C. R., and Spitznagel, J. K. (1994) Isolation and characterization of a gene, pmrD, from *Salmonella typhimurium* that confers resistance to polymyxin when expressed in multiple copies, *J. Bacteriol. 176*, 3589−3597.
15. Roland, K. L., Martin, L. E., Esther, C. R., and Spitznagel, J. K. (1993) Spontaneous pmrA mutants of *Salmonella typhimurium* LT2 define a new two-component regulatory system with a possible role in virulence, *J. Bacteriol. 175*, 4154−4164.
16. Shafer, W. M., Casey, S. G., and Spitznagel, J. K. (1984) Lipid A and resistance of *Salmonella typhimurium* to antimicrobial granule proteins of human neutrophil granulocytes, *Infect. Immun. 43*, 834−838.
17. Ernst, R. K., Yi, E. C., Guo, L., Lim, K. B., Burns, J. L., Hackett, M., and Miller, S. I. (1999) Specific lipopolysaccharide found in cystic fibrosis airway *Pseudomonas aeruginosa*, *Science 286*, 1561−1565.
18. Peschel, A. (2002) How do bacteria resist human antimicrobial peptides?, *Trends Microbiol. 10*, 179−186.
19. Gunn, J. S., Lim, K. B., Krueger, J., Kim, K., Guo, L., Hackett, M., and Miller, S. I. (1998) PmrA-PmrB-regulated genes necessary for 4-aminoarabinose lipid A modification and polymyxin resistance, *Mol. Microbiol. 27*, 1171−1182.
20. Soncini, F. C., and Groisman, E. A. (1996) Two-component regulatory systems can interact to process multiple environmental signals, *J. Bacteriol. 178*, 6796−6801.
21. Zhou, Z., Ribeiro, A. A., Lin, S., Cotter, R. J., Miller, S. I., and Raetz, C. R. (2001) Lipid A modifications in polymyxin-resistant *Salmonella typhimurium*: PMRA-dependent 4-amino-4-deoxy-L-arabinose, and phosphoethanolamine incorporation, *J. Biol. Chem. 276*, 43111−43121.
22. Guo, L., Lim, K. B., Poduje, C. M., Daniel, M., Gunn, J. S., Hackett, M., and Miller, S. I. (1998) Lipid A acylation and bacterial resistance against vertebrate antimicrobial peptides, *Cell 95*, 189−198.
23. Zhou, Z., Lin, S., Cotter, R. J., and Raetz, C. R. (1999) Lipid A modifications characteristic of *Salmonella typhimurium* are induced by NH4VO3 in *Escherichia coli* K12. Detection of 4-amino-4-deoxy-L-arabinose, phosphoethanolamine and palmitate, *J. Biol. Chem. 274*, 18503−18514.
24. Breazeale, S. D., Ribeiro, A. A., and Raetz, C. R. (2002) Oxidative decarboxylation of UDP-glucuronic acid in extracts of polymyxin-resistant *Escherichia coli*. Origin of lipid a species modified with 4-amino-4-deoxy-L-arabinose, *J. Biol. Chem. 277*, 2886−2896.
25. Breazeale, S. D., Ribeiro, A. A., and Raetz, C. R. (2003) Origin of lipid A species modified with 4-amino-4-deoxy-L-arabinose in polymyxin-resistant mutants of *Escherichia coli*. An aminotransferase (ArnB) that generates UDP-4-deoxyl-L-arabinose, *J. Biol. Chem. 278*, 24731−24739.
26. Baker, S. J., Gunn, J. S., and Morona, R. (1999) The *Salmonella typhi* melittin resistance gene pqaB affects intracellular growth in PMA-differentiated U937 cells, polymyxin B resistance and lipopolysaccharide, *Microbiology 145* (Part 2), 367−378.
27. Breazeale, S. D., Ribeiro, A. A., Raetz, C. R., and McClerren, A. L. (2005) A formyltransferase required for polymyxin resistance in *Escherichia coli* and the modification of lipid A with 4-amino-4-deoxy-L-arabinose. Identification and function of UDP-4-deoxy-4-formamido-L-arabinose, *J. Biol. Chem*. (in press).
28. Gatzeva-Topalova, P. Z., May, A. P., and Sousa, M. C. (2004) Crystal structure of *Escherichia coli* ArnA (PmrI) decarboxylase domain. A key enzyme for lipid A modification with 4-amino-4-deoxy-L-arabinose and polymyxin resistance, *Biochemistry 43*, 13370−13379.
29. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
30. Schmitt, E., Blanquet, S., and Mechulam, Y. (1996) Structure of crystalline *Escherichia coli* methionyl-tRNA(f)Met formyltransferase: comparison with glycinamide ribonucleotide formyltransferase, *EMBO J. 15*, 4749−4758.
31. Navaza, J. (2001) Implementation of molecular replacement in AMoRe, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 57*, 1367−1372.
32. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 54*, 905−921.
33. Brunger, A. T., Adams, P. D., and Rice, L. M. (1997) New applications of simulated annealing in X-ray crystallography and solution NMR, *Structure 5*, 325−336.
34. Brünger, A. T., Krukowski, A., and Erickson, J. (1990) Slow-cooling protocols for crystallographic refinement by simulated annealing, *Acta Crystallogr. A46*, 585−593.
35. Jones, A. (1978) A graphics model building and refinement system for macromolecules, *J. Appl. Crystallogr. 11*, 268−272.
36. Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) Procheck—a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
37. Blanquet, S., Dessen, P., and Kahn, D. (1984) Properties and specificity of methionyl-tRNAfMet formyltransferase from *Escherichia coli*, *Methods Enzymol. 106*, 141−152.
38. Rabinowitz, J. C. (1960) in *The Enzymes* (Boyer, P. B., Ed.) p 185, Academic Press, New York.
39. Case, G. L., Kaisaki, P. J., and Steele, R. D. (1988) Resolution of rat liver 10-formyltetrahydrofolate dehydrogenase/hydrolase activities, *J. Biol. Chem. 263*, 10204−10207.
40. Cook, R. J., Lloyd, R. S., and Wagner, C. (1991) Isolation and characterization of cDNA clones for rat liver 10-formyltetrahydrofolate dehydrogenase, *J. Biol. Chem. 266*, 4965−4973.
41. Chumanevich, A. A., Krupenko, S. A., and Davies, C. (2004) The crystal structure of the hydrolase domain of 10-formyltetrahydrofolate dehydrogenase: mechanism of hydrolysis and its interplay with the dehydrogenase domain, *J. Biol. Chem. 279*, 14355−14364.
42. Krupenko, S. A., Wagner, C., and Cook, R. J. (1997) Expression, purification, and properties of the aldehyde dehydrogenase homologous carboxyl-terminal domain of rat 10-formyltetrahydrofolate dehydrogenase, *J. Biol. Chem. 272*, 10266−10272.
43. Krupenko, S. A., Wagner, C., and Cook, R. J. (1997) Domain structure of rat 10-formyltetrahydrofolate dehydrogenase. Resolution of the amino-terminal domain as 10-formyltetrahydrofolate hydrolase, *J. Biol. Chem. 272*, 10273−10278.
44. Almassy, R. J., Janson, C. A., Kan, C. C., and Hostomska, Z. (1992) Structures of apo and complexed *Escherichia coli* glycinamide ribonucleotide transformylase, *Proc. Natl. Acad. Sci. U.S.A. 89*, 6114−6118.
45. Desharnais, J., Hwang, I., Zhang, Y., Tavassoli, A., Baboval, J., Benkovic, S. J., Wilson, I. A., and Boger, D. L. (2003) Design, synthesis and biological evaluation of 10-CF3CO-DDACTHF analogues and derivatives as inhibitors of GAR Tfase and the de novo purine biosynthetic pathway, *Bioorg. Med. Chem. 11*, 4511−4521.
46. Greasley, S. E., Marsilje, T. H., Cai, H., Baker, S., Benkovic, S. J., Boger, D. L., and Wilson, I. A. (2001) Unexpected formation of an epoxide-derived multisubstrate adduct inhibitor on the active site of GAR transformylase, *Biochemistry 40*, 13538−13547.
47. Greasley, S. E., Yamashita, M. M., Cai, H., Benkovic, S. J., Boger, D. L., and Wilson, I. A. (1999) New insights into inhibitor design from the crystal structure and NMR studies of *Escherichia coli* GAR transformylase in complex with beta-GAR and 10-formyl-5,8,10-trideazafolic acid, *Biochemistry 38*, 16783−16793.
48. 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) Rational design, synthesis, evaluation, and crystal structure of a potent inhibitor of human GAR Tfase: 10-(trifluoroacetyl)-5,10-dideazaacyclic-5,6,7,8-tetrahydrofolic acid, *Biochemistry 42*, 6043−6056.
49. Shim, J. H., and Benkovic, S. J. (1998) Evaluation of the kinetic mechanism of *Escherichia coli* glycinamide ribonucleotide transformylase, *Biochemistry 37*, 8776−8782.
50. Shim, J. H., and Benkovic, S. J. (1999) Catalytic mechanism of *Escherichia coli* glycinamide ribonucleotide transformylase probed by site-directed mutagenesis and pH-dependent studies, *Biochemistry 38*, 10024−10031.
51. Warren, M. S., Marolewski, A. E., and Benkovic, S. J. (1996) A rapid screen of active site mutants in glycinamide ribonucleotide transformylase, *Biochemistry 35*, 8855−8862.
52. Newton, D. T., and Mangroo, D. (1999) Mapping the active site of the *Haemophilus influenzae* methionyl-tRNA formyltransferase: residues important for catalysis and tRNA binding, *Biochem. J. 339* (Part 1), 63−69.

53. Schmitt, E., Panvert, M., Blanquet, S., and Mechulam, Y. (1998) Crystal structure of methionyl-tRNAfMet transformylase complexed with the initiator formyl-methionyl-tRNAfMet, *EMBO J. 17*, 6819−6826.
54. Su, Y., Yamashita, M. M., Greasley, S. E., Mullen, C. A., Shim, J. H., Jennings, P. A., Benkovic, S. J., and Wilson, I. A. (1998) A pH-dependent stabilization of an active site loop observed from low and high pH crystal structures of mutant monomeric glycinamide ribonucleotide transformylase at 1.8 to 1.9 Å, *J. Mol. Biol. 281*, 485−499.
55. Caperelli, C. A. (1987) N10-substituted 5,8-dideazafolate inhibitors of glycinamide ribonucleotide transformylase, *J. Med. Chem. 30*, 1254−1256.
56. Klein, C., Chen, P., Arevalo, J. H., Stura, E. A., Marolewski, A., Warren, M. S., Benkovic, S. J., and Wilson, I. A. (1995) Towards structure-based drug design: crystal structure of a multisubstrate adduct complex of glycinamide ribonucleotide transformylase at 1.96 Å resolution, *J. Mol. Biol. 249*, 153−175.
57. Kraulis, P. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr. 24*, 946−950.
58. Merritt, E. A., and Bacon, D. J. (1997) Raster3D: Photorealistic molecular graphics, *Methods Enzymol. 277*, 505−524.
59. Poirot, O., Suhre, K., Abergel, C., O'Toole, E., and Notredame, C. (2004) 3DCoffee@igs: a web server for combining sequences and structures into a multiple sequence alignment, *Nucleic Acids Res. 32*, W37−W40.
60. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript, *Bioinformatics 15*, 305−308.

BI047384G