Structure and Function of Both Domains of ArnA, a Dual Function Decarboxylase and a Formyltransferase, Involved in 4-Amino-4-deoxy-L-arabinose Biosynthesis*

Received for publication, February 9, 2005, and in revised form, March 8, 2005
Published, JBC Papers in Press, April 4, 2005, DOI 10.1074/jbc.M501534200

Gareth J. Williams§§, Steven D. Breazeale¶¶, Christian R. H. Raetz¶¶, and James H. Naismith§§§§

From the §Centre for Biomolecular Science, The University of St. Andrews, St. Andrews KY16 9ST, United Kingdom and the ¶¶Department of Biochemistry, Duke University, Medical Center, Durham, North Carolina 27710

Modification of the lipid A moiety of lipopolysaccharide by the addition of the sugar 4-amino-4-deoxy-L-arabinose (L-Ara4N) is a strategy adopted by pathogenic Gram-negative bacteria to evade cationic antimicrobial peptides produced by the innate immune system. L-Ara4N biosynthesis is therefore a potential anti-infective target, because inhibiting its synthesis would render certain pathogens more sensitive to the immune system. The bifunctional enzyme ArnA, which is required for L-Ara4N biosynthesis, catalyzes the NAD⁺-dependent oxidative decarboxylation of UDP-glucuronic acid to generate a UDP-4'-keto-pentose sugar and also catalyzes transfer of a formyl group from N-10-formyltetrahydrofolate to the 4'-amino of UDP-L-Ara4N. We now report the crystal structure of the N-terminal formyltransferase domain in a complex with uridine monophosphate and N-5-formyltetrahydrofolate. Using this structure, we identify the active site of formyltransfer in ArnA, including the key catalytic residues Ser543, His644, and Asp1104. Additionally, we have shown that residues Ser433 and Glu434 of the decarboxylase domain are required for the oxidative decarboxylation of UDP-GlcUA. An E434Q mutant is inactive, suggesting that chemical rather than steric properties of this residue are crucial in the decarboxylation reaction. Our data suggest that the decarboxylase domain catalyzes both hydride abstraction (oxidation) from the C-4' position and the subsequent decarboxylation.

The outer leaflet of the outer membrane of Gram-negative bacteria predominantly consists of the lipopolysaccharide (LPS)1 molecule. This immunogenic glycolipid is anchored to the membrane by lipid A, which is the active moiety causing many of the pathophysiological effects associated with Gram-negative sepsis (1). The innate immune response activated by lipid A includes production of cationic antimicrobial peptides (CAMPs), cytokines, clotting factors, and other immunostimulatory molecules (2–5). Severe sepsis can lead to a shock syndrome, which is brought on by high levels of cytokines and procoagulants and in some cases causes organ failure and even death (6). CAMPs are small, positively charged peptides, and their antimicrobial properties arise through their interactions with the lipid A moiety (7–9). These peptides interact electrostatically with the negative groups of the LPS and then traverse the outer membrane, forming pores that result in permeabilization and cell death (10, 11). Many Gram-negative bacteria have evolved resistance to CAMPs by modifying the negative groups of LPS to reduce the net negative charge, thus reducing the attraction to CAMPs (12, 13). These modifications include acylation of the O antigen and the addition of phosphoethanolamine and/or 4-amino-4-deoxy-L-arabinose (L-Ara4N) to the 1'- and 4'-phosphates of lipid A (13–17).

Modification of lipid A by L-Ara4N in Salmonella typhimurium can be abolished through mutations in the pmrE/ugd gene or the pmrF operon, rendering these cells sensitive to polymyxin B (a CAMP type antibiotic) (12, 18). The pmrF operon encodes seven genes, which are co-transcribed under PmrA control, that are required for polymyxin resistance, with the exception of pmrM (12, 18). The pmrF operon and pmrE gene are also conserved in many other pathogenic bacteria that are known to modify their lipid A with L-Ara4N, including Escherichia coli K-12, Yersinia pestis, and the major cystic fibrosis pathogen Pseudomonas aeruginosa (12, 13, 18, 19). The sequence similarity of the enzymes encoded by the pmrE and -F loci to enzymes of known function led to a proposed pathway for the biosynthesis of the l-Ara4N moiety and its attachment to lipid A (Fig. 1) (14, 20), which has since been largely validated by Raetz and co-workers (21–23). The first step is the conversion of UDP-Glc to UDP-glucuronic acid (UDP-GlcUA), catalyzed by the well characterized enzyme UDP-Glc dehydrogenase. UDP-GlcUA is then oxidatively decarboxylated by ArnA in the first unique step of the pathway, giving a UDP-4'-keto-pentose (21). This 4'-keto sugar is subsequently converted to a 4'-amino sugar, UDP-L-Ara4N, by the pyridoxal phosphate-dependent enzyme ArnB (22, 24). ArnA then catalyzes a second reaction, the transfer of a formyl group from N-10-formyltetrahydrofolate (N-10-THF) to the 4'-amino of UDP-L-Ara4N, yielding UDP-L-Ara4FN. The membrane-bound protein ArnC catalyzes the transfer of the L-Ara4FN moiety from UDP to undecaprenyl-phosphate, generating Und-P-L-Ara4FN (25). The ArnC enzymatic product is then deformylated by ArnD and...
transported to the outer surface of the inner membrane by a process that is not yet fully understood. The deformylated \( \text{L-Ara-m} \) moiety is then transferred from undecapeynyl-phospho-
phate to lipid A by ArnT (23) (Fig. 1).

Sequence analysis of ArnA reveals two discrete catalytic domains. The N-terminal formyltransferase domain is homologous to N-1-formyltetrahydrofolate, \( \text{L-methionyl-tRNA} \text{N-formyltransferase} \) (FMT), and glycaminde ribonucleotide transformylase (GART). The mechanism of transfer of the formyl group to a free amine group, catalyzed by the FMT and GART enzymes, is quite well understood (26). However, details of the active site structure of the ArnA formyltransferase domain are unknown. The C-terminal domain of ArnA is homologous to UDP-galactose-4-epimerase, dTDP-glucose-4,6-dehy-
dratase (RmlB), and other short chain dehydrogenase (SDR) enzymes that oxidize the C4’ position of sugar nucleotides. The structure of the decarboxylase domain has recently been re-
ported, and it confirms that the enzyme belongs to the SDR class of enzymes, although no biochemical characterization was reported (27). A key question about the ArnA C-terminal decar-
boxylase domain is the mechanism of decarboxylation. The formation of the keto sugar is well understood from studies of RmlB enzyme. However, whether the decarboxylation is spon-
taneous or catalyzed by ArnA is not known.

The enzymes involved in the biosynthesis of L-Ara-m are all potential therapeutic targets for the production of new anti-
fectives. ArnA is a good target because of its involvement in two steps of this pathway. The well known inhibitory effects of deaza analogues of folic acid on GART activity make the formyltransferase domain particularly interesting (26, 28). Therefore an understanding of the mechanism of action and the structure of these two domains will be useful for inhibitor design. Here we report the structure of the N-terminal formyl-
transferase domain as a complex with UMP and a substrate analog, N-5-formyltetrahydrofolate (N-5-THF), allowing iden-
tification of the active site. We have also determined the struc-
ture of the C-terminal decarboxylase domain and present bio-
chemical evidence for the involvement of SeCys\(_{4}\) and Glu\(_{4}\) in the decarboxylation reaction.

**EXPERIMENTAL PROCEDURES**

**Overexpression of Native and Selenomethionine Decarboxylase and Formyltransferase Domains of ArnA**—Full-length protein was cloned and expressed in E. coli, as described by the manufacturer. Briefly, plasmid T316M (450 bp) (22) was incubated with a cell-free extract overexpressing ArnA (750 ng) (21) at 4 °C for 90 min with shaking in 400 
ul of buffer containing 50 mM HEPES, pH 7.5, and 10% glycerol. Nontagged proteins were precipitated with 40 ul of Protein G-agarose (Sigma) before the 2 
ul of penta-His antibody was added. Fresh Protein G-agarose was then added to precipitate ArnB-N6H. The precipitated Protein G-ag-
araose was washed twice with 100 ul of 50 mM HEPES, pH 7.5. Protein samples from each step were resolved by 10% SDS-PAGE and analyzed by silver staining.

**Crystal Structure of ArnA**

**Crystallization and Data Collection of Native and Se-Met Formy-
transferase Protein**—Protein was incubated overnight at 4 °C in the presence of 5 mM N-5-THF and 5 mM UDP (Sigma) before crystal screens were set up. Crystallization of the native protein was tested against commercially available screens (screens 1 and 2, index, pegfon, and Wizard 1 and 2; Hampton), using the sitting drop vapor diffusion method at 20 °C. Drops contained 1 ul of protein with 1 ul of precipi-
tant, against a well of 100 ul of precipitant. Good quality crystals were obtained in 0.1 M ammonium acetate, 0.1 M BisTris, pH 5.5, 17%PEG
10000, using a drop containing 4.5 ul of protein and 4.5 ul of precipi-
tant against a reservoir of 100 ul of precipitant. Se-Met protein was screened against sparse matrix screens from Hampton at 20 and 4 °C. Good quality crystals formed at 4 °C in 0.22 M ammonium fluoride and 22% PEG 3350, with drops containing 3 ul of protein and 3 ul of precipi-
tant against a well of 100 ul of mother liquor. Crystals were cryocoled to 100 K using a cryoprotectant of 20% glycerol. Data to 1.2 
Å were collected on native crystals at European Synchrotron Radiation Facility beamline ID14–4. Data were indexed and integrated in MOSFLM (30) and scaled in SCALA (31). Merging of the data and analysis of the systematic absences identified the space group as P2\(_1\)2\(_1\)2\(_1\), with cell dimensions \( a = 67.2 \, \text{Å}, \, b = 90.0 \, \text{Å}, \, c = 97.9 \, \text{Å}, \, \alpha = \beta = \gamma = 90° \).

**Crystallization and Data Collection of Native and Se-Met Decar-
boxylase Protein**—Good quality crystals were obtained with a precipitant solution of 3.2 M NaCl, 0.1 M BisTris, pH 5.2, using a drop containing 4 
ul of protein and 4 ul of precipitant equilibrated against a reservoir of 100 ul of precipitant. Se-Met protein crystallized in similar conditions. The crystals were cryo-cooled to 100 K using a cryoprotectant of 4 M sodium formate. Data to 2.3 Å were collected on native crystals on beamline ID14–2 at the European Synchrotron Radiation Facility using an exposure time of 2 s and crystal to detector distance of 91 mm. A Se-Met crystal was used to collect MAD data to 1.6 
Å on a MAR-CCD detector, at three wavelengths, on the European Synchrotron Radiation Facility beamline ID14–4. Data were indexed and integrated in MOSFLM (30) and scaled in SCALA (31). Merging of the data and analysis of the systematic absences identified the space group as P2\(_2\)_1\(_2\)_1, with cell dimensions \( a = 67.2 \, \text{Å}, \, b = 90.0 \, \text{Å}, \, c = 149.4 \, \text{Å}, \, \alpha = \beta = \gamma = 90° \).

**Structure Determination and Refinement of the Decarboxylase and Formyltransferase Proteins**—SOLVE (32) located four selenomethionines for the decarboxylase data and eight selenomethionines for the
ArnA catalyzes the oxidative decarboxylation of UDP-

lipid A. As shown in Fig. 1, the C-terminal decarboxylase domain of ArnA catalyzes the oxidative decarboxylation of UDP-GlcUA to yield UDP-Ara4O, which is then converted to UDP-

l-Ara4N by a separate enzyme, ArnB. To synthesize UDP-

l-Ara4F, the N-terminal formyltransferase domain of ArnA then catalyzes the transfer of a formyl group from N-10-fTHF to the 4′-amine of UDP-l-Ara4N. Although fusion proteins are common in biology, it is unusual for a bifunctional enzyme to require an intervening enzymatic step. We therefore investigated whether the purified discrete domains of ArnA form a complex by themselves or together with ArnB. To this end, the ArnA domains were combined and passed over an S200 gel filtration column in the presence and absence of cofactors. Only the individual domains were observed, heterodimer formation was not seen. We conservatively estimate that the lower level of detection was about 0.06 mg/ml (1 μM), correlating to a lower limit of the $K_{D}$ as 5 mM. A similar analysis was performed by incubating purified ArnB with the N-terminal formyltransferase domain and the C-terminal decarboxylase domain of ArnA or the full-length ArnA protein. Again, formation of a complex between ArnB and ArnA (discrete domains or full-length protein) was not observed, suggesting that under the conditions the $K_{D}$ for a complex would be in the millimolar range if not considerably higher. Furthermore, we were not able to detect formation of a complex by co-precipitation experiments with His-tagged ArnB and ArnA, using the penta-His antibody in combination with Protein G-agarose.

**Overall Structure of the Formyltransferase Domain**—The crystal structure of the formyltransferase domain of ArnA has been solved to 1.2 Å in complex with N-5-fTHF and UMP. Crystals were grown in the presence of UDP, but only UMP was observed in the active site. The formyltransferase domain of ArnA crystallizes as a dimer, and the overall structure, shown in Fig. 2c, is similar to that of FMT. The ArnA formyltransferase domain structure consists of two subdomains: an N-terminal subdomain (residues Met1–Lys181) and a C-terminal subdomain (residues Phe 207–Asn305), which are common in biology, it is unusual for a bifunctional enzyme to require an intervening enzymatic step. We therefore investigated whether the purified discrete domains of ArnA form a complex by themselves or together with ArnB. To this end, the ArnA domains were combined and passed over an S200 gel filtration column in the presence and absence of cofactors. Only the individual domains were observed, heterodimer formation was not seen. We conservatively estimate that the lower level of detection was about 0.06 mg/ml (1 μM), correlating to a lower limit of the $K_{D}$ as 5 mM. A similar analysis was performed by incubating purified ArnB with the N-terminal formyltransferase domain and the C-terminal decarboxylase domain of ArnA or the full-length ArnA protein. Again, formation of a complex between ArnB and ArnA (discrete domains or full-length protein) was not observed, suggesting that under the conditions the $K_{D}$ for a complex would be in the millimolar range if not considerably higher. Furthermore, we were not able to detect formation of a complex by co-precipitation experiments with His-tagged ArnB and ArnA, using the penta-His antibody in combination with Protein G-agarose.

**RESULTS**

**Association of Domains with Each Other and with ArnB**—The full-length ArnA protein contains two discrete catalytic domains that are both required for the l-Ara4N modification of lipid A. As shown in Fig. 1, the C-terminal decarboxylase domain of ArnA catalyzes the oxidative decarboxylation of UDP-

lipid A. As shown in Fig. 1, the C-terminal decarboxylase domain of ArnA catalyzes the oxidative decarboxylation of UDP-

lipid A. Results for these assays are shown in Table II. The mutants

| ArnA decarboxylase domain | ArnA formyltransferase domain |
|---------------------------|------------------------------|
| Peak                      | Inflection                   | Remote | Native |
| Peak                      | Inflection                   | Remote | Native |
| Data collection           |                              |        |        |
| Resolution (Å)            | 61 to 3.35                   | 61 to 3.35 | 61 to 3.35 |
| Wavelength (Å)            | 3.43 to 3.35                 | 3.43 to 3.35 | 3.43 to 3.35 |
| Unique reflections        | 15066                        | 15065  | 15047  |
| Multiplicity              | 65 (67)                      | 21 (21) | 21 (21) |
| Completeness (%)          | 100 (100)                    | 100 (100) | 100 (100) |
| $R_{merge}$ (%)           | 14.3 (31.1)                  | 12.4 (27.1) | 13.4 (29.7) |
| $I_a$ (Å)                 | 4.4 (2.2)                    | 4.7 (1.5) | 5.3 (2.5) |
| Refinement                |                              |        |        |
| r.m.s.d. bond length (Å)  | 0.019/1.63                   | 0.019/1.93 |
| $R_{merge}$ (%)           | 18.6/23.0                    | 13.5/15.7 |
| Residues in most favored regions (%) | 90 | 92 |
| Protein Data Bank code    | 2bln                         | 2b11   |

$^{a}$ Numbers in parentheses correspond to highest resolution shell.

$^{b}$ r.m.s.d., root mean square deviation.

$^{c}$ These refer to the Ramachandran plot and are defined by PROCHECK (49).
classic Rossmann fold. This subdomain is similar in structure to GART (26) and the N-terminal domain of FMT (41). The C-terminal subdomain of the formyltransferase domain is similar in structure to the C-terminal domain of FMT and is folded into two large $\beta$-sheets ($\beta_8$ and $\beta_9$) and two small $\beta$-sheets ($\beta_{10}$ and $\beta_{11}$) with two $\alpha$-helices ($\alpha_7$ and $\alpha_8$). The $\beta$-sheets fold to form a small $\beta$-barrel, which is flanked on either side by one of the helices, with $\alpha_7$ located parallel to $\alpha_5$ of the N-terminal subdomain (Fig. 2a).

The active site of the ArnA-formyltransferase domain is located in the same position as in GART and FMT, consistent with all three enzymes utilizing $N$-5-fTHF as the formyl donor. In the structure presented here (Fig. 2a) $N$-5-fTHF binds adjacent to $\beta_3$ and $\beta_4$, and is held in place by hydrogen bonds between the bicyclic ring and two loops, one between $\beta_3$ and $\alpha_4$ (residues Ile$^{88}$ and His$^{90}$) and the other between $\beta_3$ and $\beta_7$ (residues Val$^{136}$ and Asp$^{146}$). Hydrophobic interactions between $N$-5-fTHF and residues Leu$^{87}$ (on the loop between $\beta_3$ and $\alpha_4$), Asn$^{102}$ (on $\beta_2$), Met$^{135}$, and Ala$^{139}$ (on the loop between $\beta_6$ and $\beta_7$) also assist in the binding of the cofactor analogue. Fig. 3a shows these interactions in more detail. Analysis of the thermal factors of the atoms of $N$-5-fTHF indicates that the bicyclic ring is very well ordered, whereas the benzoyl ring is less ordered, and the glutamate displays the least order. This is in agreement with the proposal of Almassy et al. that the bicyclic ring is more tightly bound than the benzoyl or glutamate moieties in GART (26, 41).

UMP is bound mainly by the N-terminal domain and is adjacent to $N$-5-fTHF (Figs. 2a and 3a). Tyr$^{42}$ is hydrogen-bonded to O2P and Arg$^{301}$ to O1P of the UMP phosphate, with a further hydrogen bond between O4 of uracil and Arg$^{301}$ (Fig. 3a). Hydrophobic interactions occur between uracil and Asn$^{118}$, Thr$^{202}$, and Trp$^{228}$. Arg$^{201}$ and Thr$^{202}$ are part of the loop connecting the N- and C-terminal subdomains, and Trp$^{228}$ is part of $\alpha_7$ of the C-terminal subdomain (Fig. 3a). These are the only interactions that occur outside of the N-terminal subdo-

### Table II

| Enzyme                        | Specific activity $\mu$mol min$^{-1}$ mg$^{-1}$ | $V_{\text{max}}$ $\mu$mol min$^{-1}$ | Apparent $K_m$ mM | $V_{\text{max}}$ $\mu$mol min$^{-1}$ | Apparent $K_m$ mM | Protein |
|-------------------------------|-----------------------------------------------|--------------------------------------|-------------------|--------------------------------------|-------------------|---------|
| ArnA decarboxylase domain     | 350                                           | 6.5                                  | 0.7               | 6.8                                  | 1.3               | 0.5     |
| S433A decarboxylase mutant    | 8.8                                           | 0.7                                  | 0.2               | 0.8                                  | 1.4               | 2.5     |
| E434A decarboxylase mutant    | 3.3                                           | 0.3                                  | 0.4               | 0.3                                  | 1.6               | 2.5     |
| E434Q decarboxylase mutant    | NA$^*$                                         | NA                                   | NA                | NA                                   | NA                | NA      |

$^*$ No activity.
main between the formyltransferase domain and its cofactor or substrate.

Fig. 4 shows the sequence alignment between the formyltransferase domain of ArnA, GART, and FMT. Highlighted are Asn102, His104, and Asp140 (in ArnA), the three conserved residues thought to be involved in catalysis. Their conservation suggests a similar mechanism of formyl transfer as described by Almassy et al. (26), with the major differences between the proteins being how they bind their respective substrates UDP-L-Ara4N (ArnA), β-GAR (GART), and M-tRNA (FMT). Although we have been unable to obtain a true substrate co-crystal, the location of UMP and the spatial conservation of the key catalytic residues allow identification of the ArnA formyltransferase active site (Fig. 3a).

The Decarboxylase Domain of ArnA—We solved the apo structure of the decarboxylase domain of ArnA to a resolution of 2.3 Å (Fig. 2b) and based our biochemical studies on this structure prior to the publication by Gatzea-Topalova et al. (27) describing the structure of the decarboxylase domain to 2.4 Å. The structures are essentially identical. The decarboxylase domain is similar in overall structure to RmlB (42) and UDP-galactose-4-epimerase (43). The structure can be split into two domains: a C-terminal sugar nucleotide binding domain and an N-terminal NAD$^+$ binding domain. A cleft is formed between the two functional domains, allowing NAD$^+$ and UDP-GlcUA access to the active site.

Co-crystallization and soaking experiments failed to produce a complex of the ArnA decarboxylase domain with NAD$^+$ and/or UDP-GlcUA or substrate analogues (UDP-Glc and UDP-galactose). ArnA is a member of the SDR superfamily, which is characterized by a catalytic triad of Thr/Ser (Thr in ArnA and RmlB), Tyr, and Lys. These three residues and NAD(P)$^+$ are required for hydride transfer between the substrate and the nicotinamide cofactor, the defining chemical reaction catalyzed by all SDR enzymes. Reduction of NAD$^+$ is consistent with the first step in the ArnA decarboxylase mechanism, the oxidation of UDP-GlcUA to 4'-keto-UDP-GlcUA. This oxidation is very similar to the first step of the RmlB mechanism, the oxidation of the dTDP-Glc to 4'-keto-dTDP-Glc. The spatial requirements of hydride transfer from the substrate to NAD$^+$ and the positioning of the catalytic triad have been established for RmlB and are quite stringent (42, 44). We have superimposed the RmlB structure containing NAD$^+$ and dTDP-Glc onto ArnA and used this to generate a conceptual model of UDP-GlcUA at the active site of ArnA (Fig. 3b). In addition to hydride transfer, the SDR superfamily catalyzes a number of other chemical transformations; these are usually dependent upon the presence of oxidized (activated) product. In RmlB, Glu135 abstracts a proton from C5' and Asp134 protonates 6'-OH, resulting in the elimination of water (42). Our superposition of RmlB and ArnA shows that the ArnA residues Ser$^{433}$ and Glu$^{434}$, which are
conserved throughout all ArnA decarboxylase domains, superimpose with these two key RmlB residues Asp\(^{134}\) and Glu\(^{135}\). Gatzeva-Topalova et al. (27) hypothesized, but did not test, a role for the conserved Ser\(^{433}\) in the reaction as well as a role for substrate recognition by Arg\(^{619}\). They suggested that Glu\(^{344}\) served to block rotation of the oxidized 4'-keto intermediate. To probe the mechanism further, we mutated Ser\(^{433}\) and Glu\(^{434}\) to determine whether this would affect the decarboxylation reaction. Since neither residue is conserved across the SDR superfamily, we did not expect these mutations to affect the oxidation step in the reaction (Fig. 1).

**Biochemical Studies**—The ArnA decarboxylase domain and the S433A, E434A, and E434Q mutant decarboxylase enzymes were assayed for their ability to reduce NAD\(^+\) to NADH in the presence of UDP-GlcUA, the results of which are shown in Table II. The E434Q mutant has no activity, and the specific activities of the S433A and E434A mutants are 40 and 100 times lower than that of the native enzyme, respectively. These mutations do not significantly perturb the \(K_m\) for the substrate. The assay only measures the first step of the reaction, oxidation of UDP-GlcUA to 4'-keto-UDP-GlcUA, which we did not expect to be affected by these mutations. Furthermore, in wild-type, S433A, and E434A reactions, only the decarboxylated product, UDP-Ara4O, and the substrate UDP-GlcUA are observed by HPLC. The putative intermediate, 4'-keto-UDP-GlcUA, was not seen in any reaction. Instead, it is stoichiometrically decarboxylated by the mutant and native enzymes to form the 4'-keto-pentose, UDP-Ara4O.

**DISCUSSION**

The ArnA protein has two recognizable domains, an N-terminal formyltransferase and a C-terminal decarboxylase, which are both required for the \(\iota\)-Ara4N modification of lipid A and polymyxin resistance. These domains were expressed, purified, crystallized, and analyzed as discrete proteins. We have previously shown that these specific formyltransferase and decarboxylase domains are catalytically functional and together (expressed in trans) can substitute for the ArnA protein (expressed as a single polypeptide) in the biosynthesis of \(\iota\)-Ara4N-modified lipid A (25). Thus, we do not expect any artifacts from studying the domains separately. All bacteria known to modify their lipid A molecules with \(\iota\)-Ara4N have full-length ArnA orthologues. Although their lipid As have not been characterized, the *Rasbtsonia* spp., *solanarearum* and *metalilidurans* genomes contain all of the genes required for the \(\iota\)-Ara4N modification of lipid A and have the two domains of ArnA annotated as distinct proteins. Thus, separate ArnA formyltransferase and ArnA decarboxylase domains or enzymes may occur in vivo in some bacteria. Under the conditions we have employed, if an association between the discrete ArnA
domains exists, it is very weak. Our data also suggest that ArnA does not interact strongly with ArnB. The ArnA formyltransferase structure has two domains, both of which are structurally similar to the two domains of FMT, but only the N-terminal domain bears similarity to GART. The key catalytic residues, Asn102, His104, and Asp140 in ArnA, are found in the N-terminal domain and are conserved through all three proteins. This suggests that the mechanism of formyl transfer described by Almassy et al. (26) also occurs in ArnA. For ArnA, this transfer is likely to involve direct nucleophilic attack on the formyl carbon of N-10-fTHF by the 4'-H of UDP-L-Ara4N, producing a tetrahedral intermediate. Similarly, Asp 140 could be the ultimate base accepting the proton from the 4'-H of UDP-L-Ara4N, although the proton may transfer first to His 104. As in GART, Asn 102 is positioned close to the formyl binding site and could stabilize the negatively charged oxygen of the intermediate (26) (Fig. 3a).

It is interesting that the ArnA formyltransferase C-terminal domain is present in FMT, but absent in GART. The role of this domain in FMT is to bind a nucleotide (adenosine) in the tRNA substrate to correctly orient this substrate at the active site (41). A similar role appears to be played in ArnA; this domain interacts with the nucleotide UMP (Fig. 3a) presumably to appropriately orient the 4'-amine of the l-Ara4N moiety. The GART substrate, glycaminide ribonucleotide, is much smaller than the FMT substrate, l-methionyl-tRNA and closer in size to the sugar-nucleotide substrate for ArnA. Interestingly, the N-terminal domain of GART is sufficient to bind both substrate and cofactor. The C-terminal domain of the ArnA formyltransferase may also be important in the full-length ArnA, acting to spatially separate the formyltransferase and decarboxylase domains to avoid interference with each other.

The overall structure of the ArnA decarboxylase domain confirms that it belongs to the SDR class of enzymes, consistent with the first step in the reaction being hydride abstraction. Mutations remote from the catalytic triad and NAD+ would not be expected to perturb hydride transfer. Yet mutation of either Ser433 or Glu434 (both of which are conserved) significantly impact catalysis but do not significantly affect UDP-GlcUA or NAD+ binding (Table II). Furthermore, the S433A and E434A mutants still produce the decarboxylated product, UDP-Ara4O, and do not accumulate the 4'-keto-UDP-GlcUA intermediate. From the biochemical data, we draw two important conclusions.

First, the decarboxylation of 4'-keto-UDP-GlcUA is not spontaneous but is indeed enzyme catalyzed. The inactivity of E434Q and the over 30-fold reduced activity of S433A and E434A suggest that the initial NAD+ dependent oxidation of the 4'-OH in the substrate UDP-GlcUA has not been decoupled from the decarboxylation leading to the UDP-Ara4O product. Gatzeva-Topalova et al. (27) postulated that Glu434 serves as a steric blocking group preventing rotation of the substrate and the 4'-keto intermediate in the active site. Implicit in this mechanism is that the rates of decarboxylation and hydride transfer are significantly slower than rotation of the carboxylate at the active site. We felt that Glu434 may instead be involved in catalysis and decided to probe its role by making the isosteric E434Q mutant. Although E434Q should be able to block rotation as efficiently as the native enzyme, this mutant is inactive. Further, the E434A mutant, which should have sufficient volume in the active site to allow ring rotation, retains activity. Together these data suggest that Glu434 serves a different purpose than blocking rotation the glucuronic acid moiety. Glu434 could act as a base ensuring deprotonation of the carboxylic acid, thus promoting decarboxylation upon formation of 4'-keto sugar (Fig. 5). The observed activity of the
E434A mutant may be accounted for by water molecules filling the newly created void to provide a bridge between the carboxylate of UDP-GlcUA and a second unidentified base in the protein, permitting the deprotonation of UDP-GlcUA. In E434Q, the water bridge to the second base would not be present, and therefore decarboxylation does not occur. We suggest that Ser433, by hydrogen bonding to the carboxylate group, may aid the decarboxylation. The S434A mutant shows that this serine does not have a significant role in substrate recognition (3-fold change in apparent $K_m$).

The second conclusion is more subtle; in the stepwise mechanism of decarboxylation, why should decarboxylation affect the preceding step of hydride transfer? The experimental evidence shows that these mutations do not perturb the structure; nor do they significantly affect NAD$^+$ or UDP-GlcUA binding (where activity can be measured). One might therefore expect an accumulation of NADH as the enzyme generates the 4'-keto intermediate. That NADH production does not occur suggests that hydride transfer has an unfavorable free energy. Apparently, the equilibrium for the oxidative step lies farther to the

---

**Fig. 4.** Sequence alignment of formyl transfer domains, numbering corresponds to full-length ArnA. The key conserved catalytic residues are marked with an asterisk as in Fig. 3a; other conserved residues have a plus sign below. ArnA residues that recognize N-5-THF and UMP are marked with a double dagger and number symbol, respectively.

**Fig. 5.** A plausible mechanism for the catalysis of the decarboxylation reaction. The carboxylic acid group of substrate is shown protonated, although in solution we would have expected it to be deprotonated. However, the E434Q mutant is inactive, and we suggest that it is required to ensure deprotonation of the keto intermediate. The carboxylate form of the keto intermediate is expected to spontaneously decompose without the need for any further catalysis.
substrate (UDP-GlcUA) side rather than the intermediate (4'-keto-UDP-GlcUA). It is the decarboxylation step, which is effectively irreversible, that serves to pull the substrate to product. When the irreversible step is either abolished or slowed significantly, the hydride remains principally on the substrate, consistent with a slower rate of NAD" consumption.

Reexamining other carbohydrate utilizing SDR enzymes suggests that this may be a common theme. In RmlB, an irreversible step, the elimination of water, occurs after hydride transfer from substrate. Prior to that, the reaction is reversible, and both substrate and keto sugar have been shown to exist (45, 46). UDP-galactose-4-epimerase oxidizes and reduces its substrates without releasing the intermediate keto sugar (47, 48). Thus, in each case, the equilibrium between oxidized substrate and NAD" is finely balanced. High resolution crystallographic studies and electronic structure calculations have shown that SDR enzymes fine tune the redox potential of NAD(P)H by distortion of the nicotinamide ring (44). ArnA, UDP-galactose-4-epimerase, and RmlB operate on quite similar carbohydrates (Glc and GlcUA), and their active sites have similar volumes. By posing the equilibrium of hydride transfer toward substrate such that only a second irreversible step will create flux through the pathway, a subtle substrate specificity can be created by the second irreversible step. ArnA is unable to reduce NAD" in the presence of UDP-Glc (very similar to UDP-GlcUA) and therefore does not oxidize the C4' position. The structure of ArnA does not reveal any reason why UDP-Glc should not bind; it is slightly smaller than UDP-GlcUA. The presence of Glu^{435} and Ser^{495} in ArnA would prevent catalysis of the dehydration of UDP-Glc, since the proton source required for water elimination is absent. By ensuring in ArnA that the equilibrium for the first oxidation reaction favors substrate, the accumulation of reactive and undesired 4'-keto sugars closely related in structure to the true substrate is prevented. This may well be a general feature of these very similar SDR enzymes.

REFERENCES

1. Raetz, C. R. H. (1990) Annu. Rev. Biochem. 59, 129–170
2. Medzhitov, R., and Janeway, C., Jr. (2000) N. Engl. J. Med. 343, 338–344
3. Dinarello, C. A. (1991) Blood 77, 1627–1652
4. Beutler, B., and Cerami, A. (1986) Annu. Rev. Biochem. 57, 505–518
5. Ohl, M. E., and Miller, S. I. (2000) Annu. Rev. Med. 52, 259–274
6. Parrillo, J. E. (1993) N. Engl. J. Med. 328, 1471–1477
7. Zasloff, M. (2002) Nature 415, 389–395
8. Hoffmann, J. A., Kafatos, F. C., Janeway, C. A., and Esekowitz, R. A. (1999) Science 284, 1313–1318
9. Scott, M. G., and Hancock, R. E. (2000) Crit. Rev. Immunol. 20, 407–431
10. Matsuzaki, K. (1999) Biochim. Biophys. Acta 1462, 1–10
11. Shiau, Y. (1999) Biochim. Biophys. Acta 1462, 55–70
12. Gunn, J. S., Ryan, S. S., Van Velkinburgh, J. C., Ernst, R. K., and Miller, S. I. (2000) Infect. Immun. 68, 6139–6146
13. Gunn, J. S. (2001) J. Endotoxin Res. 7, 57–62