THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 277, No. 4, Issue of January 25, pp. 2886–2896, 2002

Printed in U.S.A.

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*

Received for publication, September 27, 2001
Published, JBC Papers in Press, November 8, 2001, DOI 10.1074/jbc.M109377200

Steven D. Breazeale‡, Anthony A. Ribeiro‡§, and Christian R. H. Raetz‡¶

From the ‡Department of Biochemistry and the §Duke NMR Spectroscopy Center and Department of Radiology, Duke University Medical Center, Durham, North Carolina 27710

Addition of the 4-amino-4-deoxy-L-arabinose (L-Ara4N) moiety to the phosphate groups of lipid A is implicated in bacterial resistance to polymyxin and cationic antimicrobial peptides of the innate immune system. The sequences of the products of the *Salmonella typhimurium pmrE* and *pmrF* loci, both of which are required for polymyxin resistance, recently led us to propose a pathway for L-Ara4N biosynthesis from UDP-glucuronic acid (Zhou, Z., Lin, S., Cotter, R. J., and Raetz, C. R. H. (1999) *J. Biol. Chem.* 274, 18503–18514). We now report that extracts of a polymyxin-resistant mutant of *Escherichia coli* catalyze the C-4″ oxidation and C-6″ decarboxylation of [α-^{32}P]UDP-glucuronic acid, followed by transamination to generate [α-^{32}P]UDP-L-Ara4N, when NAD and glutamate are added as co-substrates. In addition, the [α-^{32}P]UDP-L-Ara4N is formylated when *N*-10-formyltetrahydrofolate is included. These activities are consistent with the proposed functions of two of the gene products (PmrI and PmrH) of the *pmrF* operon. PmrI (renamed ArnA) was overexpressed using a T7 construct, and shown by itself to catalyze the unprecedented oxidative decarboxylation of UDP-glucuronic acid to form uridine 5′-(β-L-*threo*-pentapyranosyl-4″-ulose diphosphate). A 6-mg sample of the latter was purified, and its structure was validated by NMR studies as the hydrate of the 4″ ketone. ArnA resembles UDP-galactose epimerase, dTDP-glucose-4,6-dehydratase, and UDP-xylose synthase in oxidizing the C-4″ position of its substrate, but differs in that it releases the NADH product.

Lipopolysaccharide is an immunogenic glycolipid that is the major component of the outer leaflet of the outer membranes of Gram-negative bacteria (1–3). Lipid A, the hydrophobic membrane anchor of lipopolysaccharide, is the active (endotoxin) moiety that causes many of the pathophysiological effects associated with Gram-negative sepsis (1–3). Lipid A activates the innate immune system by interacting with Toll-like receptor 4 (4–7). The host response includes production of cationic antimicrobial peptides, cytokines, clotting factors, and additional immunostimulatory molecules (7–11). In severe sepsis, high levels of cytokines and procoagulant activity cause a shock syndrome characterized by increased vascular permeability, hypotension, multiple organ failure, and death (10, 12).

In *Escherichia coli* K-12, lipid A consists of a β-1′,6-linked disaccharide of glucosamine that is phosphorylated at the 1- and 4′-positions, and acylated with *R*-3-hydroxymyristate at the 2-, 3-, 2′-, and 3′-positions (Fig. 1) (1–3). The 2′ and 3′ *R*-3-hydroxymyristate chains are further acylated with laurate and myristate residues, respectively (1–3). Lipid A is glycosylated at position 6′ with two 3-deoxy-D-*manno*-octulosonic acid (Kdo) moieties (1–3). The Kdo-lipid A substructure of lipopolysaccharide is sufficient to support the growth of *E. coli* (1–3).

Although the structure of lipid A is relatively conserved among Gram-negative bacteria (1, 3), it may be modified in response to environmental stimuli (13–16) or in *pmrA* constitutive mutants of *E. coli* and *Salmonella typhimurium* (Fig. 1) (17–20). The latter are resistant to polymyxin, a cationic lipopeptide antibiotic (17–19), and to some of the cationic antimicrobial peptides of the innate immune system (21, 22). The covalent appendages that confer polymyxin resistance (Fig. 1) are not required for growth under laboratory conditions, but promote virulence when administered orally to animal hosts (23). Substituents attached to lipid A in various combinations include palmitate, 2-hydroxymyristate, phosphoethanolamine (pEtN),[1] and 4-amino-4-deoxy-L-arabinose (L-Ara4N) (Fig. 1) (14, 24, 25). PmrA-PmrB, a two-component regulatory system activated by low pH, PhoP/PhoQ, or certain *pmrA* mutations (15), is required for the modification of lipid A with pEtN and L-Ara4N moieties (20, 24, 25), resulting in less anionic lipid A species that do not bind cationic antimicrobial substances as strongly and prevent them from penetrating the outer membrane.

Mutations in the *pmrE/ugd* gene or in the *pmrF* operon render *S. typhimurium* incapable of making L-Ara4N-modified lipid A (20, 23), and abolish polymyxin resistance in *pmrA* constitutive strains. The genes in the *pmrF* operon, *pmrHFIJKLM*, are co-transcribed under PmrA control, and except for *pmrM*, are required for the maintenance of polymyxin resistance (23). The *pmrE* gene and the *pmrF* operon are also present in *E. coli* (20), and appear to be activated in polymyxin-resis-

* This work was supported in part by National Institutes of Health Grant GM-51310 (to C. R. H. R.), The Duke NMR Center is partially supported by National Institutes of Health NCI Grant P30-CA-14236 (to A. A. R.), NMR instrumentation in the Duke NMR Center was supported by the National Science Foundation, the National Institutes of Health, the North Carolina Biotechnology Center, and Duke University. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AY057445.

¶ To whom correspondence should be addressed. Tel.: 919-684-5326; Fax: 919-684-8885; E-mail: raetz@biochem.duke.edu.

[1] The abbreviations used are: pEtN, phosphoethanolamine; L-Ara4N, 4-amino-4-deoxy-L-arabinose; UDP-L-Ara4N, UDP-4-amino-4-deoxy-L-arabinose; UDP-L-Ara4O, uridine 5′-(β-L-*threo*-pentapyranosyl-4″-ulose diphosphate); UDP-GlcA, UDP-glucuronic acid; Kdo, 3-deoxy-D-*manno*-octulosonic acid; DTT, dithiothreitol; SW, spectral width; RD, relaxation delay.

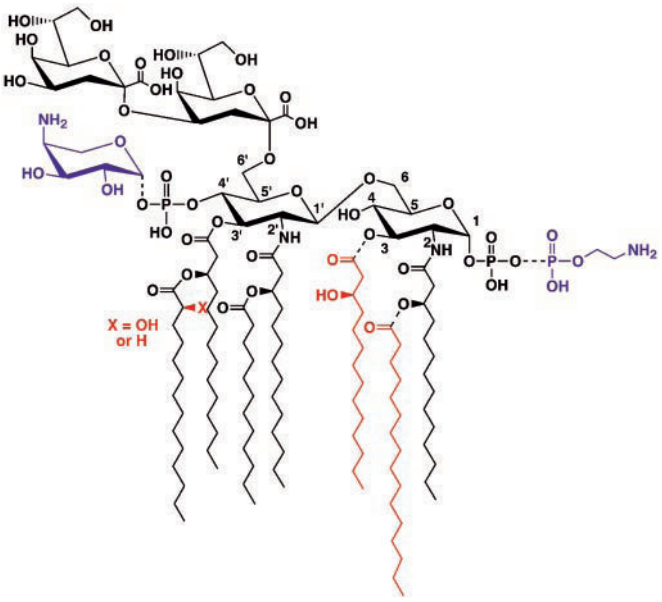

FIG. 1. **Regulated covalent modifications of lipid A.** The Kdo_2-lipid A substructure of lipopolysaccharide is sufficient to support the growth of *S. typhimurium* and *E. coli*. However, the phosphate residues and acyl chains of lipid A may be modified in a regulated fashion, as indicated by the *dashed* bonds. The phosphates may be substituted with L-Ara4N and/or pEtN groups, both of which are under PmrA control (*blue*) (20, 25). Minor lipid A species may be present in which the locations of the L-Ara4N and pEtN groups are reversed (not shown), or in which both phosphates are modified with the same group (25). Addition of the palmitoyl chain is catalyzed by the outer membrane enzyme PagP (59), removal of the hydroxymyristoyl chain at position 3 is carried out by PagL (60), and formation of 2-hydroxymyristate requires a novel hydroxylase homologue (61), encoded by *lpxO*. PagL and LpxO occur only in *S. typhimurium*, but can function when transferred to *E. coli*. Substituents that are thought to be under PhoP/PhoQ control (15) are shown in *red*.

tant mutants or in wild-type cells exposed to ammonium metavanadate (14). Based upon sequence similarity of the deduced proteins of the *pmrE* and *pmrF* loci to enzymes of known function, a putative biosynthetic pathway that starts with UDP-glucuronic acid (UDP-GlcA) was recently proposed by two independent laboratories as the source of the L-Ara4N substituent of lipid A (Fig. 2) (14, 26). Although the initial steps of the pathway have not yet been validated, we recently showed that *arnT* (previously *pmrK*) encodes an inner membrane enzyme that utilizes the novel glycolipid, undecaprenyl phosphate-L-Ara4N, to modify lipid A (27, 28). The existence of an undecaprenyl phosphate intermediate implies that L-Ara4N transfer to lipid A occurs on the periplasmic side of the inner membrane (Fig. 2).

We now show that extracts of a polymyxin-resistant *E. coli* mutant (but not wild type) can convert UDP-GlcA to the novel sugar nucleotide, uridine 5′-β-L-*threo*-pentapyranosyl-4″-ulose diphosphate (UDP-L-Ara4O) (Fig. 2). The latter is subsequently transaminated to make UDP-4-amino-4-deoxy-L-arabinose (UDP-L-Ara4N) (Fig. 2). Furthermore, UDP-L-Ara4N is formylated *in vitro* in the presence of *N*-10-formyltetrahydrofolate, but the biological significance of this reaction is unclear (Fig. 2). The UDP-GlcA C-4″-dehydrogenase activity of ArnA was demonstrated with overexpressed protein, made by cloning *arnA* behind the T7 promoter. ArnA (previously designated PmrI or Orf3) (20, 23) by itself catalyzed both the C-4″ oxidation and C-6″ decarboxylation of UDP-GlcA to generate UDP-L-Ara4O (Fig. 2). The latter was obtained in 6 mg quantities almost entirely in the form of its 4″-hydrate in aqueous solution, as judged by one-dimensional and two-dimensional NMR spectroscopy. The data support our proposed scheme for L-Ara4N biosynthesis and transfer to lipid A (Fig. 2), but are at variance with Baker *et al*. (26) and Gunn (29) who suggested that an additional enzyme (PmrJ) catalyzes the C-6″-decarboxylation.

ArnA is homologous to other enzymes that oxidize the C-4″ position of certain UDP sugars, such as UDP-galactose epimerase (30), dTDP-glucose-4,6-dehydratase (31), and the eucaryotic enzyme that converts UDP-GlcA to UDP-xylose (32, 33). ArnA is unique, however, in releasing the 4″ ketone and the NADH products without reutilizing the latter for a subsequent reduction step. ArnA, furthermore, contains a second catalytic domain that is highly homologous to methionyl-tRNA formyltransferase (34). Although the function of the formyltransferase domain of ArnA remains unclear, it is responsible for the UDP-L-Ara4N formylation observed *in vitro* (Fig. 2).

EXPERIMENTAL PROCEDURES

Buffers and Reagents—UDP-glucose pyrophosphorylase (yeast), inorganic pyrophosphatase (yeast), UDP-glucose dehydrogenase (bovine, type VI), UDP-glucose, UDP-glucuronic acid, NAD, NADP, glucose 1-phosphate, triethylammonium bicarbonate buffer (pH 8.5), L-glutamate, polymyxin B sulfate, kanamycin, 5-formyltetrahydrofolate (citrovorum factor), and HEPES were purchased from Sigma. [α-^{32}P]UTP was purchased from PerkinElmer Life Sciences. Restriction enzymes (*Nde*I and *Xho*I) were obtained from New England Biolabs. T4 DNA ligase, oligonucleotides, and isopropyl-1-thio-β-D-galactopyranoside were purchased from Invitrogen. The *Pfu* polymerase was purchased from Stratagene. The polyethyleneimine-cellulose TLC plates were purchased from E. Merck, Darmstadt, Germany. DEAE-Sepharose™ Cl-6B was purchased from Amersham Biosciences AB. Construction of the polymyxin-resistant mutant WD101 derived from *E. coli* W3110 was described previously (28).

Synthesis of [α-^{32}P]UDP-glucose—The synthesis of [α-^{32}P]UDP-glucose was analogous to that of [α-^{32}P]UDP-GlcNAc (35), and accomplished in a reaction containing the following: [α-^{32}P]UTP (250 μCi, 800 Ci/mmol), 0.5 mM glucose-1-phosphate, 1 mM $MgCl_2$, 5 mM DTT, 100 mM Tris-HCl (pH 8.0), UDP-glucose pyrophosphorylase (3 units), and inorganic pyrophosphatase (12.5 units) in a total volume of 65 μl. The reaction was allowed to proceed at 30 °C for 1 h, diluted to 1.5 ml with deionized water, and loaded onto a 2-ml column of DEAE-Sepharose™ CL-6B (pre-equilibrated with 10 mM triethylammonium bicarbonate, pH 8.5). The column was washed with 5 ml of deionized water followed by elution of the product with a 0–100 mM triethylammonium bicarbonate (pH 8.5) gradient (40 ml total volume). The eluted material was concentrated using a Speed-Vac centrifuge, and was re-dissolved in 200 μl of deionized water. The yield was generally greater than 75% based on recovery of radioactivity.

Synthesis of [α-^{32}P]UDP-glucuronic acid—The conversion of [α-^{32}P]UDP-glucose to [α-^{32}P]UDP-glucuronic acid proceeded in the following reaction: 50 mM Tris (pH 8.7), 2 mM DTT, 6 mM NAD, [α-^{32}P]UDP-glucose (100 μCi), 3 mM $MgCl_2$, and 75 μg of UDP-glucose dehydrogenase in 150 μl total volume for 60 min at 30 °C. The reaction was diluted to 1.5 ml with deionized water and loaded onto a 2-ml column of DEAE-Sepharose™ CL-6B (pre-equilibrated with 10 mM triethylammonium bicarbonate, pH 8.5). The column was washed with 5 ml of deionized water and 5 ml of 50 mM triethylammonium bicarbonate (pH 8.5). The product was eluted with 30 ml of 100 mM triethylammonium bicarbonate (pH 8.5). The eluted material was concentrated by lyophilization, and re-dissolved in deionized water. The yield was generally greater than 75% based on recovery of radioactivity.

Preparation of Cell Extracts—Typically, bacteria were grown in LB medium (36) at 37 °C to late log phase, and harvested at an A_{600} = 1.0 by centrifugation at 6500 × *g* for 15 min. Cells from a 100-ml culture were resuspended in 5 ml of 50 mM HEPES (pH 7.5), 2 mM DTT, and 10% glycerol, passed through a French pressure cell at 18,000 psi, and then clarified by centrifugation at 16,500 × *g* for 30 min at 4 °C. These cell-free extracts were stored at −80 °C. Protein concentrations were determined by bicinchoninic acid using the Pierce assay reagent with bovine serum albumin as the standard (37).

In Vitro Assays for Detecting [α-^{32}P]UDP-glucuronic Acid Conversion to Novel α-^{32}P-Labeled Sugar Nucleotides—Assays contained 50 mM HEPES (pH 7.5), 2 mM DTT, 10 μM [α-^{32}P]UDP-glucuronic acid (8×10^5 cpm/nmol), 0.5 mg/ml protein, and when appropriate, 3 mM NAD, 3 mM NADP, 1 mM L-glutamate, and/or 1 mM 5,10-methenyltetrahydrofolate in a total volume of 50 μl, as indicated. Reactions were incubated at 30 °C, and at the desired time points 1 μl was withdrawn and spotted

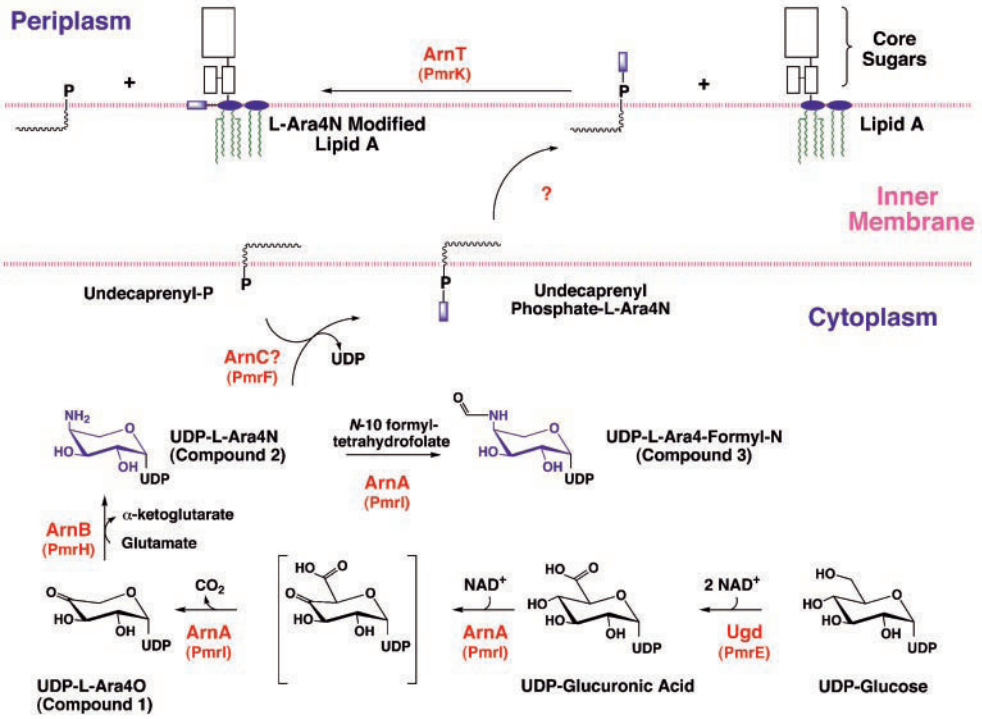

FIG. 2. **Biosynthesis of UDP-L-Ara4N from UDP-GlcA in polymyxin-resistant *E. coli* and *S. typhimurium*.** The pathway starts with the conversion of UDP-glucose to UDP-GlcA by the well characterized dehydrogenase, Ugd (PmrE). ArnA (previously Orf3 or PmrI) (20, 23) catalyzes the oxidative decarboxylation of UDP-glucuronic acid to form the novel sugar nucleotide, UDP-L-Ara4O. A separate enzyme to catalyze the decarboxylation is not needed. ArnB (previously Orf1 or PmrH) (20, 23) then catalyzes the further transamination to make UDP-L-Ara4N. Based upon its homology to dolichoyl phosphate-mannose synthase of yeast, we suggest that ArnC (PmrF) transfers the L-Ara4N moiety to undecaprenyl phosphate (14), forming the novel compound undecaprenyl phosphate-α-L-Ara4N, which has recently been isolated from polymyxin-resistant mutants (28). After translocation to the outer surface of the inner membrane, ArnT (Orf5, PmrK, or YfbI) transfers the L-Ara4N unit to lipid A (27). The ArnA protein has a second catalytic domain with strong similarity to methionyl-tRNA formyltransferase (34). This may account for the efficient transfer of a formyl group from *N*-10-formyltetrahydrofolate to UDP-L-Ara4N in our *in vitro* system (see below), but the significance of this modification is unclear. It may be that the true substrate for ArnC is the formylated derivative of UDP-L-Ara4N, but an assay for ArnC is not yet available.

onto a polyethyleneimine-cellulose TLC plate. The spots were allowed to dry, and the plate was washed in methanol for 5 min. After drying, the plate was developed using 0.25 M acetic acid with 0.4 M LiCl. The plate was dried and exposed to a PhosphorImager screen. Radioactive enzymatic products were identified and quantified with a PhosphorImager (ImageQuant software, Molecular Dynamics).

Synthesis of N-5,N-10-Methenyltetrahydrofolate—The conversion of *N*-5-formyltetrahydrofolate to *N*-5,*N*-10 methenyltetrahydrofolate was carried out as described (38). In 1.5 ml of deionized water containing 1% (v/v) 2-mercaptoethanol, 7.5 mg of *N*-5-formyltetrahydrofolate was dissolved, and the pH was adjusted to 1.9 with 0.1 N HCl. The material was diluted to 2.2 ml with deionized water, and the reaction was allowed to proceed at room temperature. Generation of *N*-5,*N*-10 methenyltetrahydrofolate was followed by measuring the increase in absorbance at 355 nm. Upon completion, the reaction mixture was stored at −20 °C.

Conversion of N-5,N-10-Methenyltetrahydrofolate to N-10-Formyltetrahydrofolate—The *N*-5,*N*-10-methenyltetrahydrofolate was converted to *N*-10-formyltetrahydrofolate at pH 7.5 (38) by performing a 10-min preincubation of this substrate in the enzymatic assay buffer, which contains 50 mM HEPES (pH 7.5).

Recombinant DNA Techniques—Preparation of competent cells, transformation by the $CaCl_2$ method, genomic DNA purification, and electrophoresis were performed according to published procedures (39). Plasmids were purified using the QIAprep spin miniprep kit (Qiagen). Restriction enzymes, *Pfu* DNA polymerase, and T4 DNA ligase were used as recommended by the manufacturer. DNA sequencing was performed on an ABIprism 377 instrument at the Duke University DNA Analysis Facility.

Construction of pETArnA—The predicted coding region for ArnA (PmrI) in the *pmrF* operon (23) was amplified by polymerase chain reaction from *E. coli* W3310 genomic DNA (40) with primers to the *arnA* 5′ end (5′-CGGGATCCATATGAAAACCGTCGTTTTTGCCT-3′) and to the *arnA* 3′ end (5′-CGGGATCCTCGAGTCATGATGGTTTATCCGTAAG-3′), containing engineered *Nde*I and *Xho*I restriction sites, respectively. The Opti-Prime™ PCR Optimization kit (Stratagene) was used as described by the manufacturer. Amplification in a 50-μl reaction mixture with 10 × Opti-Prime Buffer 9 (100 mM Tris-HCl, pH 9.2, 35 mM $MgCl_2$, 250 mM KCl), 0.2 mM of each dNTP, 250 ng of genomic DNA, 0.5 μM of each primer, and 2.5 units of cloned *Pfu* DNA polymerase for 30 cycles (94 °C for 1 min, 52 °C for 1 min, and 72 °C for 2.5 min) followed by one cycle of 94 °C for 1 min, 52 °C for 1 min, and 72 °C for 15 min resulted in a single product of the expected size. The PCR product was isolated with the QIAquick PCR Purification kit (Qiagen), digested with *Xho*I and *Nde*I, and ligated into a T7 expression vector, pET24b (Novagen), previously digested with the same enzymes. The fidelity of the *Pfu* DNA polymerase was confirmed by DNA sequencing of the cloned *arnA* gene in pETArnA.

Overexpression of the ArnA Gene Product—The plasmid pETArnA was transformed into *E. coli* NovaBlue(DE3) cells (Novagen). Cultures were grown overnight at 37 °C from single colonies in LB broth supplemented with 30 μg/ml kanamycin. The overnight culture was then used to inoculate 100 ml of the same LB broth, and cells were grown at 37 °C to A_{600} of 0.6. Expression was induced by the addition of isopropyl-β-D-thiogalactopyranoside to a final concentration of 1 mM, and growth was continued at 30 °C for 3.5 h. Cells were harvested by centrifugation at 6,500 × *g* for 15 min at 4 °C. The cell pellet was resuspended in 3 ml of buffer containing 50 mM HEPES (pH 7.5), 10% glycerol, and 2 mM DTT, and lysis was achieved by passage through a French pressure cell at 18,000 psi. Cell debris was removed by centrifugation at 16,500 × *g* for 15 min at 4 °C to give crude cell-free extracts. Expression was monitored by SDS-polyacrylamide gel electrophoresis.

Enzymatic Synthesis and Isolation of Uridine 5′-(β-L-Threo-pentapyranosyl-4″-ulose diphosphate) (Compound 1)—A 15-ml reaction mixture containing 0.1 mg/ml of the recombinant ArnA crude cell-free extract, 1 mM UDP-glucuronic acid, and 6 mM NAD, in buffer consisting of 50 mM HEPES (pH 7.5) and 2 mM DTT, was allowed to proceed for 8 h at 30 °C. Protein was removed by passing the reaction mixture through a Centricon™ (YM-10) membrane. The enzymatic product was loaded onto a 10-ml column of DEAE-Sepharose™ CL-6B (pre-equilibrated with 10 mM triethylammonium bicarbonate, pH 8.5). The column was

washed with 40 ml of deionized water followed by elution of the product with a 0–100 mM triethylammonium bicarbonate (pH 8.5) gradient (200 ml total volume). Chromatography was monitored by measuring the absorbance at 254 nm, and fractions containing the desired product, the putative UDP-L-Ara4O, were pooled and concentrated by lyophilization to give the triethylammonium salt of the sugar nucleotide. Conversion to the sodium salt was accomplished by passage over a 1-ml column of AG™ 50W-X8 resin (sodium form), followed by lyophilization to give the pure compound (6 mg, 68% yield).

Nuclear Magnetic Resonance Spectroscopy—The putative UDP-L-Ara4O (Na^+ salt, total of 6 mg) and the UDP-glucuronic acid (34 mg) were each dissolved in 0.6 ml of 99% D_2O in 5-mm NMR tubes. Measurements with a 3-mm pH electrode revealed pD values of 6.51 and 5.44 for the ArnA product and the UDP-GlcA samples, respectively. All ^1H and ^{13}C NMR chemical shifts were referenced to 2,2-dimethylsilapentane-5-sulfonic acid at 0.00 ppm. The ^{31}P NMR spectra were referenced to 85% phosphoric acid at 0.00 ppm.

^1H and ^{13}C NMR spectra were obtained on a Varian Inova 600 spectrometer equipped with a Sun Ultra 5 computer and a 5-mm Varian triple resonance probe. One-dimensional ^1H spectra were obtained with a spectral width (SW) of 6.5 kHz, a 72° pulse flip angle (6 μs), a 6.3-s acquisition time, and a 1.2-s relaxation delay (RD), and digitized using 82000 points to obtain a digital resolution of 0.159 Hz/pt. ^1H-decoupled one-dimensional ^{13}C spectra were recorded with a 35 kHz SW, a 54° pulse flip angle (6 μs), and a 3.1-s repeat time, and digitized into 91000 points to give a digital resolution of 0.777 Hz/pt.

Two-dimensional COSY data were recorded in absolute value mode with 2048 points, 1-s RD, and 32 scans per increment; 512 time increments were collected and zero-filled to 2048 points with sine-bell weighting in both dimensions before Fourier transformation, followed by symmetrization of the two-dimensional matrix.

The two-dimensional NOESY data were recorded in hypercomplex phase-sensitive mode with two sets of 256 time incremented spectra, 32 scans per increment, a 1-s delay and a mixing time of 500 ms. The final two-dimensional matrix was 2048 × 2048 with Gaussian weighting in both dimensions, and was not symmetrized.

Proton-detected single bond ^1H-^{13}C two-dimensional chemical shift correlation spectra were recorded using the HMQC method (41) with ^{13}C decoupling during acquisition. Two sets of 200 time increments were obtained in hypercomplex phase-sensitive mode with 2048 points in t_2; 64 scans were recorded per time increment, and the RD was 1.2 s. The two-dimensional data were processed using Gaussian functions and zero-filled to a final size of 2048 × 2048.

The two-dimensional HMBC multiple bond correlation spectra (41, 42) were recorded in hypercomplex phase-sensitive mode without ^{13}C decoupling during acquisition. Two sets of 200 time increments were recorded with 2048 points in t_2. The RD was 1.2 s, the filter delay was set to an average $^1J_{CH}$ of 140 Hz, and 96 transients were obtained per increment. The long-range ^1H-^{13}C couplings were allowed to evolve for a delay of 83 ms (6 Hz optimization). The data were processed using a shifted Gaussian weighting function along f_2 and a cosine weighting function along f_1, and zero-filled to a final size of 2048 × 2048.

The one-dimensional ^1H and ^{31}P NMR spectra were obtained on a Varian Unity 500 spectrometer equipped with a Sun Ultra 5 computer and a 5-mm Varian inverse probe. ^1H spectra at 500 MHz were obtained with a 4.5 kHz SW, 67° pulse flip angle (6 μs), 9.4-s acquisition time, and 2-s RD, and digitized using 84000 points to obtain a digital resolution of 0.107 Hz/pt. ^{31}P spectra at 202 MHz were obtained with a 12 kHz SW, 59° pulse flip angle (13 μs), 2.4-s acquisition time, and 1.6-s RD, and digitized using 25000 points to obtain a digital resolution of 0.97 Hz/pt.

Selective inverse decoupling difference spectroscopy (43, 44) was implemented by detecting the ^1H NMR spectrum at high digitization while selectively exciting a phosphorus signal with 1 s of low-power (estimated 0.2 mW, ~4 db on the Unity 500 spectrometer) continuous wave on- and off-resonance irradiation during the acquisition period. To improve subtraction in the generation of one-dimensional difference ^{31}P-decoupled ^1H NMR spectra, the samples were thermally equilibrated in the magnet for over 1 h before recording data, and the ^1H difference spectra were obtained in an interleaved manner with 4 scans accumulated for each free induction decay and looping around n times to achieve a good signal-to-noise ratio with 4n scans per free induction decay.

RESULTS

Enzymatic Transformations of [α-^{32}P]UDP-GlcA in Cell Extracts of Polymyxin-resistant E. coli—Given that *S. typhimurium* mutants defective in UDP-glucose dehydrogenase (Fig. 2) are unable to make L-Ara4N-modified lipid A (20), we used 10 μM [α-^{32}P]UDP-GlcA as a probe for novel enzymes that might generate [α-^{32}P]UDP-L-Ara4N *in vitro*. Cell extracts of wild-type and polymyxin-resistant *E. coli* (28) were tested for their ability to catalyze the oxidative decarboxylation and transamination of [α-^{32}P]UDP-GlcA (Fig. 2), using a combination of TLC and PhosphorImager analysis to detect the products. In the presence of exogenous NAD, extracts of the polymyxin-resistant *E. coli* mutant WD101 (28) rapidly converted [α-^{32}P]UDP-GlcA to a faster migrating species designated compound **1** (Fig. 3, *lane 5*). The relatively slow migration of the substrate *versus* the product suggests that **1** is less negative than [α-^{32}P]UDP-GlcA. Formation of **1** was absolutely dependent upon the presence of added NAD, but did occur with NADP at a slow rate (Fig. 3, *lane 6*). An extract of the polymyxin-sensitive wild-type strain W3110 (the parent of WD101) did not metabolize [α-^{32}P]UDP-GlcA in the presence of NAD or NADP (Fig. 3, *lanes 1–3*). We propose that the biosynthesis of **1** is catalyzed by ArnA(PmrI/Orf3) of the PmrF operon (Fig. 2), given the distant sequence similarity of ArnA to dTDP-glucose-4,6-dehydratase and UDP-galactose epimerase (14).

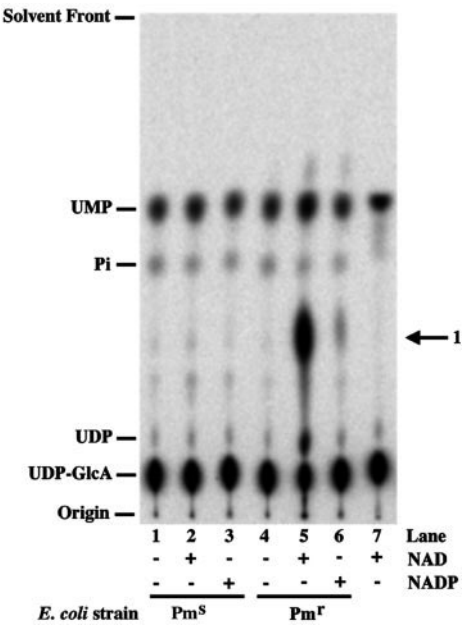

FIG. 3. **NAD dependent conversion of [α-^{32}P]UDP-GlcA to compound 1 in extracts of a polymyxin-resistant mutant.** Cell-free extracts (0.5 mg/ml) of wild-type *E. coli* W3110 (Pm^s) or mutant WD101 (Pm^r) were assayed at 30 °C for 40 min with 10 μM [α-^{32}P]UDP-GlcA (8×10^5 cpm/nmol). A portion of each reaction mixture (1 μl) was then spotted onto a polyethyleneimine-cellulose plate, which was developed in 0.4 M LiCl containing 0.25 M acetic acid. Products generated from [α-^{32}P]UDP-GlcA were detected with a PhosphorImager. *Lanes 1–3,* wild-type *E. coli* (Pm^s) with no cofactor, with 3 mM NAD, or with 3 mM NADP, respectively. *Lanes 4–6,* WD101 (Pm^r) with no cofactor, with 3 mM NAD, or with 3 mM NADP, respectively. *Lane 7,* no enzyme control with 3 mM NAD. The positions of UDP-GlcA, compound **1,** and minor impurities are shown.

The formation of the 4″-keto intermediate (Fig. 2) should facilitate 6″-decarboxylation, which might be spontaneous or enzyme catalyzed. In either case, compound **1** provides the substrate for an aminotransferase (proposed to be ArnB/PmrH/Orf1) (14) that generates [α-^{32}P]UDP-L-Ara4N (Fig. 2). In accordance with this idea, crude extracts of mutant WD101 efficiently converted [α-^{32}P]UDP-GlcA to an even faster migrating species, designated compound **2** (Fig. 4, *lane 3*), when L-glutamate was present in addition to NAD. The R_F of **2** in this solvent system was about the same as that of UDP-glucosa-

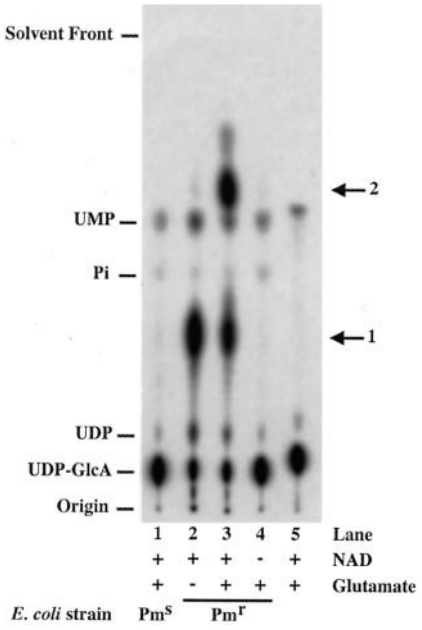

FIG. 4. **Conversion of [α-^{32}P]UDP-GlcA to compound 2 in the presence of NAD and glutamate.** Cell-free extracts (0.5 mg/ml) of wild-type *E. coli* (Pm^s) or WD101 (Pm^r) were assayed at 30 °C for 90 min with 10 μM [α-^{32}P]UDP-GlcA. Next, a 1-μl portion of each reaction mixture was spotted onto a polyethyleneimine-cellulose plate, which was developed and analyzed as described in the legend to Fig. 3. *Lane 1,* wild-type *E. coli* (Pm^s) with 3 mM NAD and 1 mM L-glutamate. *Lanes 2–4,* WD101 (Pm^r) with 3 mM NAD, with 3 mM NAD plus 1 mM L-glutamate, or 1 mM L-glutamate, respectively. *Lane 5,* no enzyme control with 3 mM NAD and 1 mM L-glutamate.

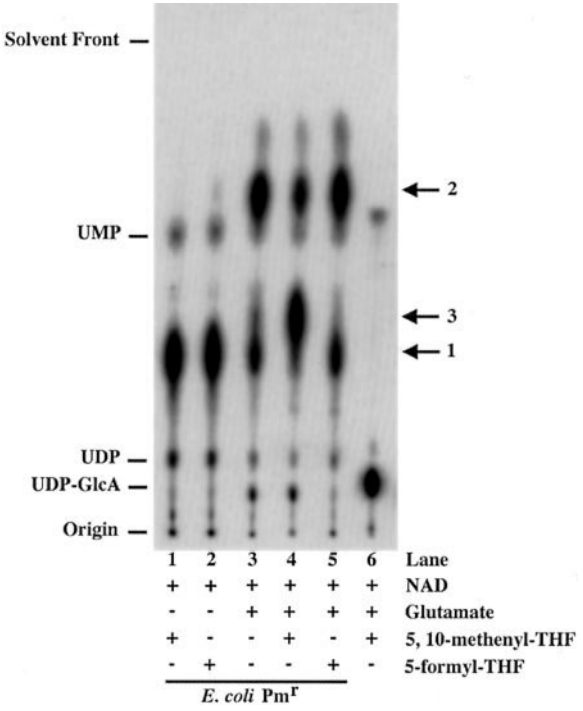

FIG. 5. **Conversion of [α-^{32}P]UDP-GlcA to compound 3 in the presence of NAD, glutamate, and *N*-10-formyltetrahydrofolate.** Cell-free extracts (0.5 mg/ml) of *E. coli* mutant WD101 (Pm^r) were assayed at 30 °C for 180 min in the presence of 10 μM [α-^{32}P]UDP-GlcA. Next, a 1-μl portion of each reaction mixture was spotted onto a polyethyleneimine-cellulose plate, which was developed and analyzed as described in the legend to Fig. 3. As indicated for *lanes 1–5*, 3 mM NAD, 1 mM L-glutamate, and/or various formyl donors (each at 1 mM) were also included in the assay. The *N*-5,*N*-10-methenyltetrahydrofolate is converted *in situ* to *N*-10-formyltetrahydrofolate by means of a 10-min preincubation (38). Formation of compound **3** from **2** is absolutely dependent upon the presence of *N*-10-formyltetrahydrofolate. *Lane 6* is the no enzyme control.

mine, which has the same net charge as [α-^{32}P]UDP-L-Ara4N. Presumably, the NAD is first used to convert [α-^{32}P]UDP-GlcA to compound **1,** given that [α-^{32}P]UDP-GlcA by itself is not converted directly to **2** in the presence of L-glutamate alone (Fig. 4, *lane 4*). A matched extract of a polymyxin-sensitive wild-type *E. coli* strain was completely inactive in the formation of **2** (Fig. 4, *lane 1*).

L-Glutamate likely functions as the co-substrate for a pyridoxal phosphate-dependent aminotransferase (Fig. 2). L-Glutamine also supports this conversion, but less efficiently than L-glutamate. L-Alanine and L-aspartate are inactive (data not shown). The natural cosubstrate of the putative aminotransferase is probably L-glutamate, which would be converted to α-ketoglutarate during the formation of UDP-L-Ara4N (Fig. 2).

Formylation of [α-^{32}P]UDP-L-Ara4N in Extracts of WD101—ArnA, which is predicted based upon its sequence to catalyze the NAD-dependent conversion of UDP-GlcA to **1,** possesses an additional N-terminal domain of unknown function that displays significant similarity to methionyl-tRNA formyltransferases (34). Given this peculiar observation, we investigated the possibility of a formyltransferase that acts on [α-^{32}P]UDP-L-Ara4N in extracts of WD101 by including exogenous *N*-5,*N*-10-methenyltetrahydrofolate in addition to NAD and glutamate in the assay system (Fig. 5). The role of a formylated-nucleotide intermediate in the biosynthesis of L-Ara4N-modified lipid A is unclear *a priori*, because formylated L-Ara4N moieties have not been reported in association with lipid A, or elsewhere (14, 25, 44–47). Nevertheless, when exogenous *N*-5,*N*-10-methenyltetrahydrofolate was included in our system, an additional more slowly migrating substance, designated compound **3,** was formed in significant quantities (Fig. 5, *lane 4*). The low R_F of **3** (Fig. 5) suggested formylation of the free amino group in **2** (Fig. 2), but *O*-formylation could not be excluded. The formylation reaction was highly specific for *N*-5,*N*-10-methenyltetrahydrofolate, which at pH 7.5 rapidly rearranges to *N*-10-formyltetrahydrofolate (38). The latter is the actual formyl donor (Fig. 2). The commercially available analog, 5-formyltetrahydrofolate, was inactive (Fig. 5, *lane 5*). The substrate requirements for the formation of **3** indicated that UDP-GlcA and **1** (Fig. 5, *lanes 1 and 2*) were not utilized directly by the formyltransferase. As in the case of the other transformations described above, extracts of wild-type cells did not catalyze formyl transfer to **2** (data not shown).

Cloning and Overexpression of arnA—Although the metabolism of [α-^{32}P]UDP-GlcA observed in extracts of *E. coli* WD101 (Figs. 3–5) supported the proposed pathway for L-Ara4N biosynthesis and transfer to lipid A (Fig. 2), the actual enzymes catalyzing these transformations had not been identified, and data establishing the structures of the proposed biosynthetic intermediates (14) were lacking. To prove that *arnA* encodes the NAD-dependent dehydrogenase that converts UDP-GlcA to compound **1** (Fig. 3), the full-length *arnA* gene was cloned and expressed in the pET24b vector behind a T7 RNA polymerase promoter, inducible with isopropyl-β-D-thiogalactopyranoside (48). Expression was carried out in *E. coli* NovaBlue(DE3), a polymyxin-sensitive strain that does not modify its own lipid A with the L-Ara4N moiety (27). A protein of the predicted molecular mass of ArnA, 73 kDa, was observed by SDS-gel electrophoresis in cell-free extracts of induced cells, but was absent in uninduced controls (Fig. 6*A*).

The induced cell-free extracts derived from cells containing pETArnA catalyzed very rapid NAD-dependent conversion of UDP-GlcA to **1**; this activity was not present in matched extracts derived from cells harboring the vector pET24b (data not

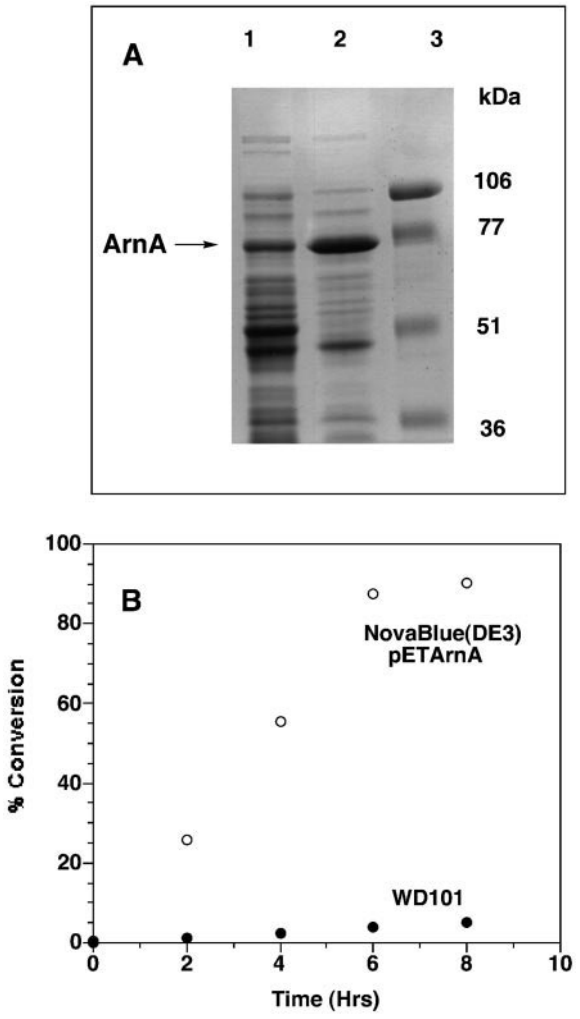

FIG. 6. **Overexpression of ArnA on pETArnA in *E. coli* NovaBlue(DE3).** *Panel A* shows the overexpression of the cloned ArnA protein using an 8% SDS-polyacrylamide gel. *Lane 1*, uninduced pETArnA. *Lane 2*, pETArnA induced for 3 h at 30 °C with 1 mM isopropyl-1-thio-β-D-galactopyranoside. *Lane 3*, protein molecular weight standards: phosphorylase *b* (106 kDa), bovine serum albumin (77 kDa), ovalbumin (51 kDa), and carbonic anhydrase (36 kDa). *Panel B* shows the rate of conversion of [α-^{32}P]UDP-GlcA to compound **1**. Cell-free extracts of mutant WD101 (Pm^r) at 0.5 mg/ml or of induced *E. coli* NovaBlue(DE3)/pETArnA at 0.1 mg/ml were assayed at 30 °C with 1 mM [α-^{32}P]UDP-GlcA (8×10^3 cpm/nmol). At various times, 1-μl portions were spotted onto a polyethyleneimine-cellulose plate that was developed in 0.4 M LiCl containing 0.25 M acetic acid. Compound **1** was detected and quantified with a PhosphorImager.

shown). The extracts of cells containing the induced pETArnA plasmid showed ~100-fold higher ArnA specific activity than comparable preparations of WD101 (Fig. 6*B*).

Given these results, the predicted genes of the *pmrF* operon (14, 20) were renamed to emphasize the enzymatic activities they encode (27), in accordance with the proposal of Reeves *et al.* (49). Although these genes are necessary for maintenance of polymyxin resistance, they are in fact specifically involved in L-Ara4N biosynthesis and transfer of lipid A (hence "Arn"), with ArnA catalyzing the first reaction of the pathway (Fig. 2).

Evaluation of the Covalent Structure of Compound 1 by One-dimensional ^{1}H NMR Spectroscopy—Compound **1** was isolated on a 6-mg scale using the overexpressed recombinant ArnA protein, as described under "Experimental Procedures." The full ^{1}H NMR spectra of UDP-glucuronic acid and **1** are shown in Fig. 7, *A* and *B*, respectively. The spectra reveal the presence of a uracil and ribose moiety in both compounds with virtually identical chemical shifts and coupling constants (Table I and Fig. 7). The remaining ^{1}H NMR signals, which are well resolved and arise from the pyranose, differ in the two compounds (Table I and Fig. 7). The spectrum of **1** is simplified in the 3.6–4.0 ppm region, where three distinct doublets are now seen (Fig. 7*B*) when compared with the more complex multiplets of the UDP-GlcA (Fig. 7*A*). The ^{1}H NMR spectra thus demonstrate that the enyzmatic reaction catalyzed by ArnA altered the glucuronic acid moiety of UDP-GlcA but left the uridine portion of the molecule unchanged.

^{31}P NMR Spectroscopy of Compound 1—^{31}P NMR analysis (not shown) confirmed the presence of an unaltered diphosphodiester linkage in compound **1** *versus* UDP-GlcA from the observation of two distinct ^{31}P NMR signals near −11.9 and −10.2 ppm with ~20.7 Hz ^{31}P-^{31}P coupling. Subtraction of two ^{1}H NMR spectra obtained with on- and off-resonance selective decoupling of the −10.2 ppm ^{31}P signal in compound **1** revealed decoupling changes only at ribose H-5′ and H-4′ (Fig. 7*C*), clearly assigning the −10.2 ppm ^{31}P signal as the phosphate group linked to ribose C-5′. The ^{1}H NMR difference spectrum obtained upon selective decoupling of the −11.9 ppm ^{31}P resonance (Fig. 7*D*) reveals decoupling changes only at the anomeric H-1″ and the H-2″ of compound **1,** assigning the −11.9 ppm signal to the phosphate group linked to C-1″.

Analysis of Compound 1 by Two-dimensional ^{1}H NMR Spectroscopy—The two-dimensional COSY analysis (Fig. 8) confirms the identity of the ribose and uracil rings in the two compounds. The cross-peak connectivity from H-1′ to H-2′, the degeneracy of H-2′ and H-3′ (~4.40 ppm), the connectivity from H-3′ to H-4′ (~4.31 ppm), and the connectivities of H-4′ to H-5a′ and H-5b′ are the same in both compounds. The uracil H-6 and H-5 cross-peaks are likewise the same in both (not shown).

The assignments of the pyranose sugars were easily derived from the COSY analysis (Fig. 8). The double-doublet at 5.65 ppm ($J_{1''2''}$ = 3.3, $J_{H1''P}$ = 7.5 Hz) in the UDP-GlcA spectrum (Figs. 8*A*) arises from the glucuronic acid anomeric H-1″ proton. The cross-peak from H-1″ locates the H-2″ multiplet at 3.61 ppm (dt, $J_{1''2''}$ = 3.3, $J_{2''3''}$ = 9.8, $J_{H2''P}$ = 3.2 Hz). The second COSY cross-peak from H-2″ connects to H-3″ at 3.81 ppm (dd, $J_{2''3''}$ = 9.8, $J_{3''4''}$ = 9.4 Hz). Further tracing of the COSY cross-peak connectivities locate H-4″ (3.53 ppm; dd, $J_{3''4''}$ = 9.4, $J_{4''5''}$ = 10.1 Hz) and H-5″ (4.17 ppm; d, $J_{4''5''}$ = 10.1 Hz).

The COSY analysis of **1** (Fig. 8*B*) reveals cross-peak connectivities from H-1″ (5.61 ppm, dd, $J_{1''2''}$ = 3.3, $J_{H1''P}$ = 7.1 Hz) to H-2″ (3.69 ppm, dt, $J_{1''2''}$ = 3.3, $J_{2''3''}$ = 9.8, $J_{H2''P}$ = 3.3 Hz), and from H-2″ to H-3″ (3.84 ppm, d, $J_{2''3''}$ = 9.8 Hz). However, the connectivity pattern is broken at position 4″ (Fig. 8*B*), consistent with the proposed structure (Fig. 7*B*). The remaining cross-peak connectivity reflects the geminal coupling between H-5b″ (3.96 ppm, d, $J_{5b''5a''}$ = 12.0 Hz) and H-5a″ (3.58 ppm, d, $J_{5b''5a''}$ = 12.0 Hz) (Fig. 8*B*). The COSY data demonstrate conclusively that no proton is attached at C-4″ and fits with the assignment of the three simple doublets to H-3″, H-5a″, and H-5b″. The presence of two protons at position 5″ (*versus* a single H-5″ in UDP-GlcA) shows that 6″-decarboxylation occurred during the formation of **1**.

The small $J_{1''2''}$ couplings (3.3 Hz) and the large $J_{2''3''}$ coupling (9.8 Hz) of the pyranose moiety in both the substrate and the product (Table I) demonstrate that these sugars have an equatorially disposed H-1″, and contain axially disposed H-2″ and H-3″ protons (50, 51). Additional two-dimensional NOESY experiments with **1** (not shown) reveal strong intramolecular NOE cross-peaks between H-1″ and H-2″, further supporting the idea of an equatorial H-1″ and an axial H-2″ in an α-glucopyranosyl configuration (50). The 9.8 Hz $J_{2''3''}$ coupling implies that H-3″ is axial and trans-relative to H-2″, as in UDP-glucuronic acid.

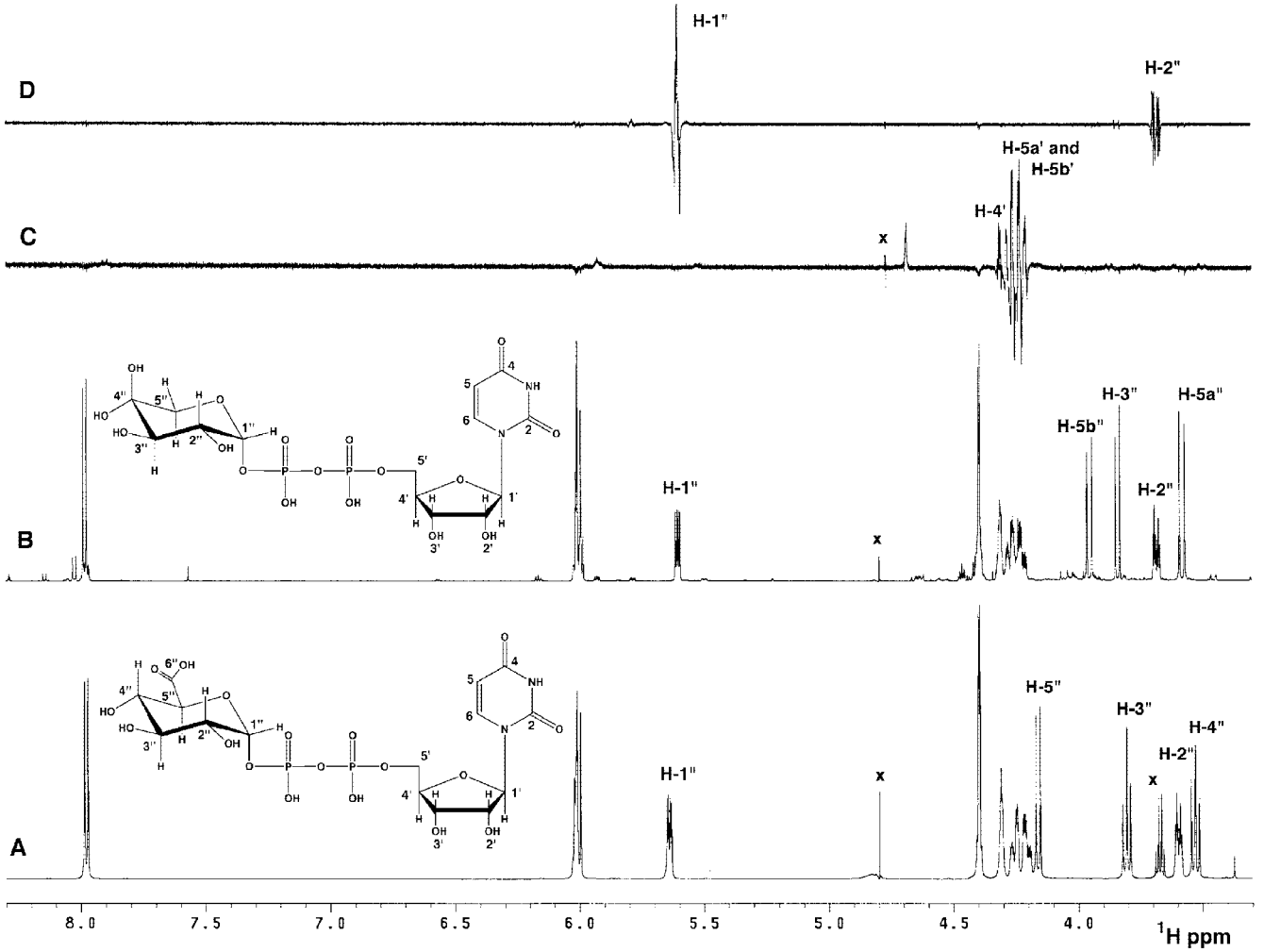

FIG. 7. **Partial ^{1}H-NMR spectra of UDP-GlcA and compound 1 at 600 MHz.** *Panel A*, 34 mg of UDP-GlcA in D_2O at 25 °C; *Panel B*, 6 mg of compound **1**; *Panel C*, selective inverse ^{31}P decoupled ^{1}H difference spectrum at 500 MHz of compound **1** with selective irradiation of the ^{31}P signal at −10.2 ppm; *Panel D*, selective inverse ^{31}P decoupled ^{1}H difference spectrum at 500 MHz of compound **1** with selective irradiation of the ^{31}P signal at −11.9 ppm. *X* indicates an impurity or residual HOD resonances.

The H-5a″ (3.58 ppm) and H-5b″ (3.96 ppm) signals of compound **1** (Figs. 7*B* and 8*B*) were assigned to the axial and equatorial positions, respectively, based upon the consideration that a proton in an axial position on a given pyranose carbon usually resonates ~0.5 ppm higher than the corresponding equatorial proton (51). However, attempts to validate the assignments of the two H-5″ protons of compound **1** independently with a two-dimensional NOESY experiment were unsuccessful, as no NOE cross-peak was observed between the axially disposed H-3″ and either H-5a″ or H5b″.

Because the ketone group of a hexulose may be hydrated (52), it is possible that both the ketone form and its corresponding hydrate co-exist in solution. The ^{1}H NMR spectrum of CDP-3,6-dideoxy-D-glycero-4-hexulose in D_2O (52) does in fact show two sets of hexulose sugar resonances, whereas our product shows only one. Furthermore, compound **1** is very stable, given that identical ^{1}H NMR spectra were recorded over the course of three months in D_2O solution, and repeated integrations showed no evidence for deuterium exchange at the axial H-3″. The ^{1}H NMR data indicate only one predominant form of compound **1** in D_2O, most likely the hydrate, given the negligible rate of deuterium exchange at H-3″.

Evaluation of Compound 1 by ^{13}C NMR Spectroscopy—To distinguish between the ketone and/or hydrate forms of compound **1,** and to validate its structure, natural abundance ^{13}C NMR spectra were recorded for both UDP-GlcA and **1** (Figs. 9, *A* and *B,* respectively). Similar to the ^{1}H NMR results, the ^{13}C spectra show that the uracil and ribose moieties are the same in both compounds (Fig. 9 and Table I). The ^{13}C spectra differ in the locations of the pyranose sugar resonances between 67 and 77 ppm, and in the appearance of a new carbon signal at 95.1 ppm in compound **1** (Fig. 9*B*). The 95.1 ppm signal is in the anomeric region, indicating an additional carbon substituted with two oxygens, as would be expected for the hydrated form of the ketone. Furthermore, the glucuronic acid carbonyl of the substrate at 178.8 ppm (Fig. 9*A*) is entirely absent in **1** (Fig. 9*B*), confirming that 6″-decarboxylation has occurred.

The 95.1 ppm carbon signal of compound **1** is further shown to arise from a quaternary ketal carbon by the two-dimensional HMQC ^{1}H-^{13}C correlation map (Fig. 10*A*). There are only three direct ^{1}H-^{13}C single-bond correlations in the anomeric region (90–105 ppm). The anomeric H-1″ (~5.6 ppm) correlates to the carbon resonance of the pyranose near 98 ppm (C-1″) (Fig. 10*A*), as in UDP-GlcA (not shown), consistent with an axially disposed oxygen atom and an equatorially disposed H-1″ (50). The anomeric ribose H-1′ proton near 6.01 ppm correlates to the carbon resonance at 90.9 ppm (C-1′), while the uracil H-5 near 6.00 ppm correlates to the carbon signal near 105.2 ppm. The fourth carbon signal in the anomeric region near 95.1 ppm has no proton partner in the two-dimensional spectrum (Fig. 10*A*), consistent with the presence of a quaternary ketal carbon.

Examination of the two-dimensional HMQC ^{1}H-^{13}C data for

TABLE I
NMR spectroscopy of UDP-GlcA and UDP-L-Ara4O

^{31}P chemical shifts of UDP-glucuronic acid are −10.1 (α-P) and −11.8 ppm (β-P), respectively. ^{31}P-^{31}P coupling = 20.6 Hz. ^{31}P chemical shifts of UDP-L-Ara4O are −10.2 (α-P) and −11.9 ppm (β-P), respectively. ^{31}P-^{31}P coupling = 20.7 Hz.

Position	^{13}C NMR signals[a]		^{1}H NMR signals[a]					
	UDP-GlcA	UDP-L-Ara4O	UDP-GlcA			UDP-L-Ara4O		
	δ (ppm)		δ (ppm)	Hz		δ (ppm)	Hz	
C-2	154.4	154.4						
C-4	168.8	168.8						
C-5	105.3	105.2	6.01(d)	$J_{5,6}$	8.1	6.00(d)	$J_{5,6}$	8.1
C-6	144.2	144.2	7.98(d)			7.99(d)		
C-1′	90.9	90.9	6.02(d)			6.02(d)		
C-2′	76.4	76.3	~4.4			~4.4		
C-3′	72.2	72.2	~4.4			~4.4		
C-4′	85.9	85.8	~4.3			~4.3		
C-5′	67.5	67.4	4.27(ddd)	$^{3}J_{5'a,4'}$	2.8	4.27(ddd)	$^{3}J_{5'a,4'}$	2.8
				$J_{5'a,P}$	5.4		$J_{5'a,P}$	5.4
				$J_{5'a,5'b}$	11.9		$J_{5'a,5'b}$	11.8
			4.21(ddd)	$^{3}J_{5'b,4'}$	2.5	4.22	$^{3}J_{5'b,4'}$	2.5
				$J_{5'b,P}$	4.4		$J_{5'b,P}$	4.4
C-1″	97.9	98.4	5.65(dd)	$J_{1'',2''}$	3.3	5.61(dd)	$J_{1'',2''}$	3.3
				$^{3}J_{1'',P}$	7.5		$^{3}J_{1'',P}$	7.1
C-2″	74.0	73.1	3.61(dt)	$J_{2'',3''}$	9.8	3.69(dt)	$J_{2'',3''}$	9.8
				$^{4}J_{2'',P}$	3.2		$^{4}J_{2'',P}$	3.3
C-3″	75.2	75.1	3.81(dd)	$J_{3'',4''}$	9.4	3.84(d)		
C-4″	74.4	95.1	3.53(dd)	$J_{4'',5''}$	10.1			
C-5″	75.6	67.8	4.17(d)			3.96(d)	$J_{5''a,5''b}$	12.0
						3.58(d)		
C-6″	178.9							

[a] ^{1}H and ^{13}C chemical shifts were determined at 25 °C and are relative to internal 2,2-dimethylsilapentane-5-sulfonic acid at 0.00 ppm.
[b] s, singlet; d, doublet; t, triplet; dd, double doublet; m, multiplet.

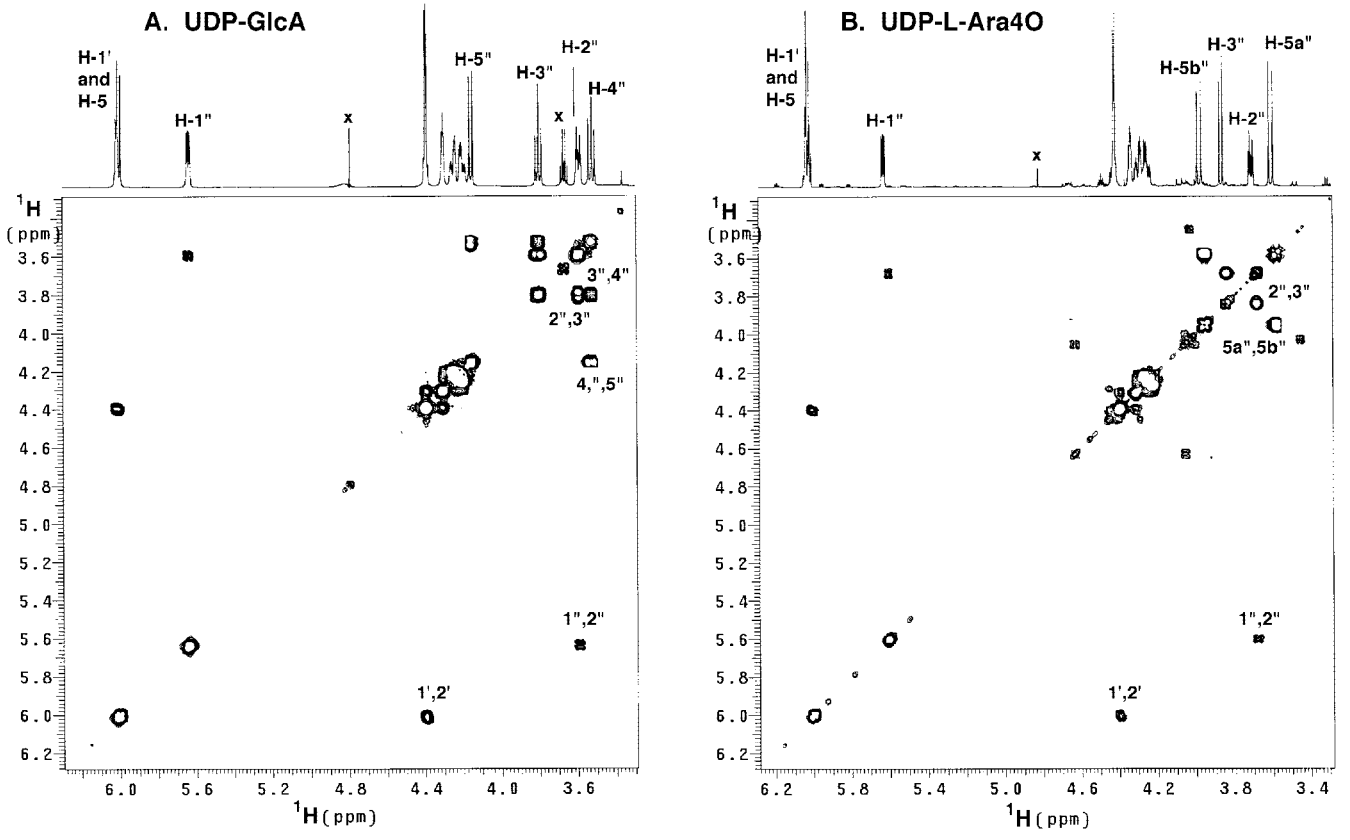

FIG. 8. **Partial two-dimensional ^{1}H-^{1}H COSY of UDP-GlcA and compound 1 at 600 MHz.** *Panel A*, 34 mg of UDP-GlcA in D_2O at 25 °C; *Panel B*, 6 mg of compound **1**. The ribose coupling connectivities are the same in both compounds, whereas the pyranose coupling connectivities differ. *X* indicates an impurity or residual HOD resonances.

the other sugar resonances of **1** (Fig. 10*A*) reveals that the H-5a″ and H-5b″ doublets correlate to a carbon at 68.0 ppm (C-5″), while the H-2″ multiplet and the H-3″ doublet, respectively, correlate to carbon resonances at 73.0 and 75.1 ppm. The other protonated carbon cross-peaks arise from the ribose ring (Table I), the assignments of which agree with literature values (50).

The multibond HMBC map in the sugar region of **1** (Fig. 10*B*) further verifies the above assignments. Multibond correlations are clearly observed from the H-3″ and the H-5a″ doublets to the 95.1 ppm carbon resonance, unequivocally demonstrating

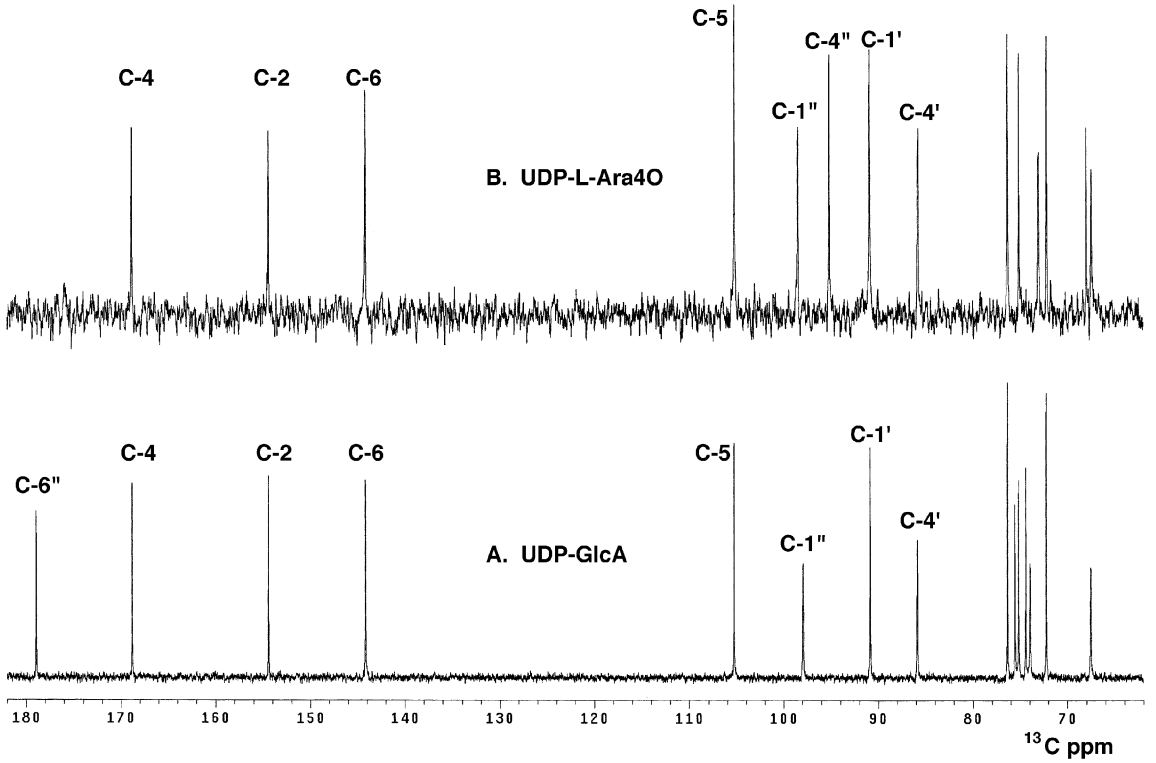

FIG. 9. **Partial ^{13}C-NMR spectra of UDP-GlcA and compound 1 at 150 MHz.** *Panel A*, 34 mg of UDP-GlcA in D_2O at 25 °C; *Panel B*, 6 mg of compound **1**. The uracil CO, uracil CH, and ribose CH resonances are the same in both compounds. The UDP-GlcA C-6″ carboxyl signal at 178 ppm is absent in compound **1,** concomitant with the appearance of a new carbon signal in the latter in the anomeric region at 95.1 ppm.

that the latter must arise from C-4″. Taken together, the NMR data are fully consistent with the identification of **1** as the hydrated form of the novel sugar nucleotide, uridine 5′-β-L-*threo*-pentapyranosyl-4″-ulose diphosphate, designated UDP-L-Ara4O (Fig. 2).

DISCUSSION

The covalent modification of the phosphate groups of lipid A with L-Ara4N moieties confers resistance to polymyxin and cationic anti-microbial peptides in *E. coli* and *S. typhimurium* (17, 18, 20, 25, 27, 28). The L-Ara4N unit is thought to arise by a novel pathway from UDP-GlcA. The proposed scheme (Fig. 2) was initially developed based upon bioinformatic analyses (14, 26) of the gene products encoded by the *pmrE* and *pmrF* loci, which are required for L-Ara4N biosynthesis and polymyxin resistance.

We now present the first *in vitro* evidence for the enzymatic conversion of UDP-GlcA to UDP-L-Ara4N (Fig. 2). Using [α-^{32}P]UDP-GlcA as the probe, three distinct enzymatic activities were discovered in cell-free extracts of WD101, a polymyxin-resistant mutant of *E. coli* harboring a constitutively active *pmrA* gene (28). These enzymes, which were detected based on mobility shifts of [α-^{32}P]UDP-GlcA, initially convert [α-^{32}P]UDP-GlcA to [α-^{32}P]UDP-L-Ara4O (compound **1**) in a reaction requiring NAD as the co-substrate (Fig. 3). This step involves the C-4″ oxidation and C-6″ decarboxylation of the glucuronic acid moiety of UDP-GlcA (Fig. 2). The [α-^{32}P]UDP-L-Ara4O is less negatively charged than [α-^{32}P]UDP-GlcA and migrates faster in the TLC system shown in Fig. 3. In the presence of L-glutamate, [α-^{32}P]UDP-L-Ara4O is further metabolized to [α-^{32}P]UDP-L-Ara4N (compound **2**), which migrates even faster (Fig. 4) with about the same R_F as UDP-glucosamine. These data are consistent with a direct transamination between glutamic acid and the 4″-ketone of UDP-L-Ara4O (Fig. 2). Lastly, [α-^{32}P]UDP-L-Ara4N is converted to compound **3** in the presence of *N*-10-formyltetrahydrofolate (Fig. 5). The significantly slower migration of compound **3** *versus* [α-^{32}P]UDP-L-Ara4N (compound **2,** Fig. 5) suggests that the charge of the nitrogen atom of UDP-L-Ara4N is neutralized by formylation (Fig. 2).

All three enzymes are absent in extracts of W3110, the polymyxin-sensitive, wild-type strain of *E. coli* from which the *pmrA* constitutive mutant WD101 is derived (28). These findings are consistent with the fact that PmrA activates the transcription of the L-Ara4N biosynthesis/polymyxin-resistance genes of the *pmrE* and *pmrF* loci (20, 23).

The *E. coli arnA* gene was cloned and overexpressed in *E. coli* NovaBlue(DE3) behind an inducible T7 promoter (Fig. 6). The massive overproduction of the NAD-dependent conversion of UDP-GlcA to compound **1** in association with the synthesis of the expected protein of 74 kDa (Fig. 6) demonstrates that *arnA* is the structural gene for the enzyme. Preparation of 6 mg of compound **1** with the ArnA overproducing extracts greatly facilitated detailed structural analysis by NMR spectroscopy (Figs. 7–10). Our results demonstrate unequivocally that compound **1** is the proposed UDP-L-Ara4O intermediate (Fig. 2), which exists almost entirely as the 4″-hydrate in aqueous solution (Fig. 7). In the case of CDP-3,6-dideoxy-D-glycero-D-glycero-4-hexulose (a precursor of CDP-paratose) (52), distinct ^{13}C NMR signals were observed at 210 ppm (the 4″-keto form) and at 90 ppm (the 4″-hydrate), whereas the ketone could not be detected by NMR spectroscopy in our product (Fig. 9). The purified UDP-L-Ara4O is stable in water, however, and can be converted to compounds **2** and **3** in milligram quantities in the presence of cloned, purified ArnB and ArnA, respectively, facilitating the validation of the proposed structures shown in Fig. 2.[2]

The NMR analysis (Figs. 7–10) demonstrates that ArnA is the only enzyme required for the conversion of UDP-GlcA to UDP-L-Ara4O (Fig. 2). Although it is possible that the C-6″

[2] S. Breazeale and C. R. H. Raetz, manuscript in preparation.

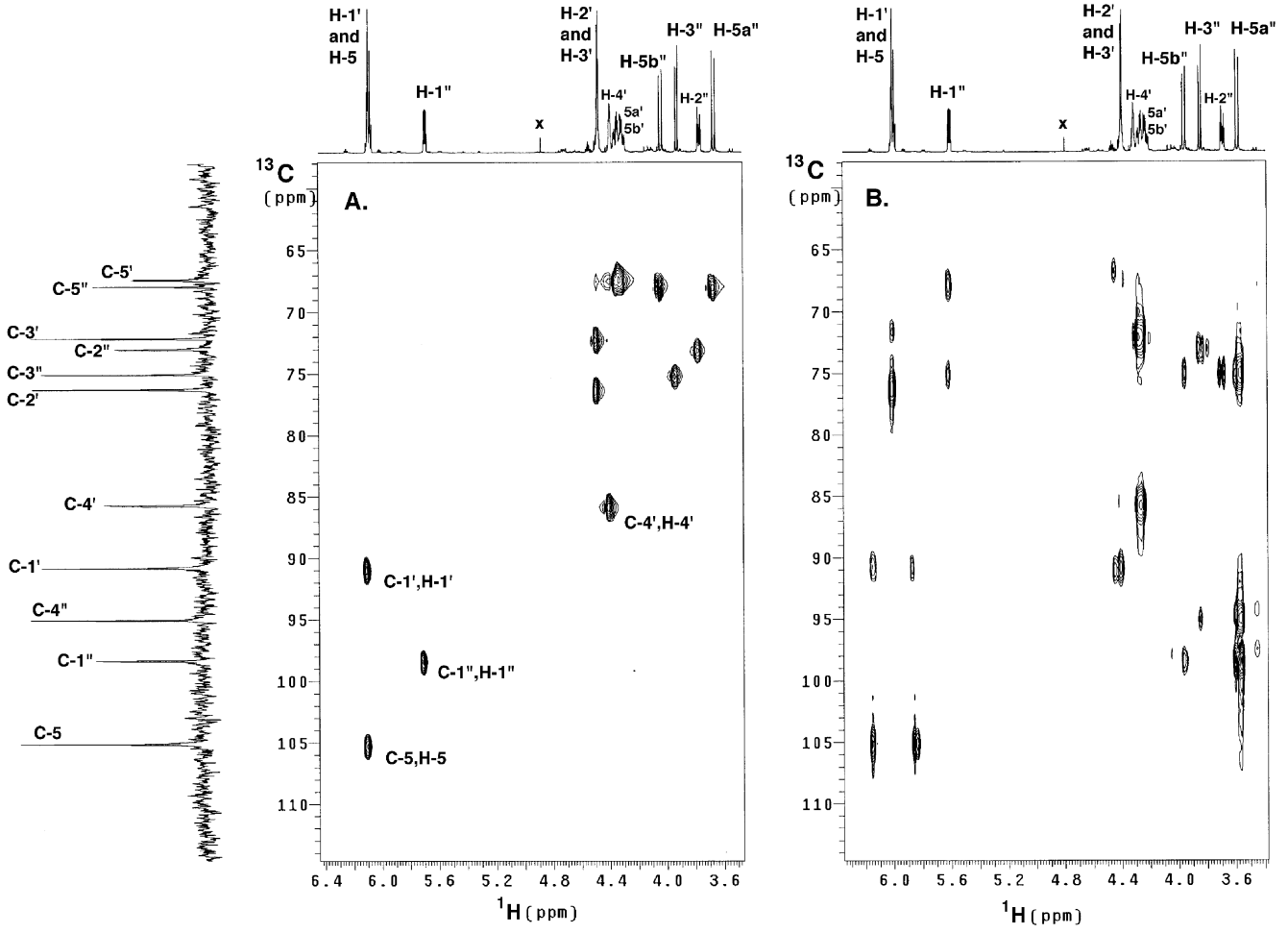

FIG. 10. **Partial two-dimensional HMQC and HMBC spectra of compound 1.** *Panel A* shows the HMQC analysis of 6 mg of compound **1** in D_2O at 25 °C, demonstrating the single bond ^{1}H-^{13}C correlations from the uracil H-5, ribose, and pyranose protons to their directly bonded carbons. The carbon signal at 95.1 ppm is consistent with a quaternary ketal carbon, as it shows no connectivity to a proton. *Panel B* is a partial two-dimensional HMBC spectrum of **1** under the same conditions, demonstrating the multibond ^{1}H-^{13}C correlations from the uracil H-5, ribose, and pyranose protons. Of special interest are the multibond correlations from H-3″ and H-5a″ to the 95.1 ppm carbon signal, confirming its assignment as C-4″. *X* indicates an impurity or residual HOD resonances.

decarboxylation is nonenzymatic following C-4″ oxidation, we favor the view that C-6″ oxidation is also catalyzed by ArnA. The fact that ArnA alone is sufficient to catalyze both the C-4″ oxidation and the C-6″ decarboxylation of UDP-GlcA is at variance with the biosynthetic scheme proposed by Baker *et al*. (26) and Gunn (29), who suggest that PmrJ is the decarboxylase. Since PmrJ is not homologous to other proteins of known function, we cannot yet assign a role for PmrJ in our scheme (Fig. 2). However, the genetic evidence does clearly demonstrate that PmrJ is needed for polymyxin resistance (23).

ArnA represents the first reported enzyme that catalyzes the oxidative decarboxylation of UDP-GlcA with the concomitant release of NADH, which can be detected spectrophotometrically when purified ArnA is incubated with UDP-GlcA and NAD (data not shown). UDP-xylose synthase, the structural gene of which was recently identified and cloned based on its intriguing functional similarity to ArnA,[3] likewise catalyzes the C-4″ oxidation and the C-6″ decarboxylation of UDP-GlcA (32, 33), but the ketone intermediate is immediately reduced by NADH to generate UDP-xylose (32, 33). Several other types of enzymes catalyze the oxidation at C-4″ of related sugar nucleotides, and display limited sequence similarity to the dehydrogenase domain of ArnA. The best characterized of these is UDP-galactose epimerase (30, 53), which interconverts UDP-galactose and UDP-glucose. In this case, the presumed 4″-ketopyranose intermediate is rotated within the active site and immediately reduced by the adjacent NADH molecule (30), inverting the configuration at C-4″. Lastly, there is a large family of sugar nucleotide-4,6-dehydratases that catalyze the oxidation of the pyranose C-4″ to activate the C-5″ proton for elimination of water between C-5″ and C-6″ (31, 54). Like the UDP-galactose epimerases, these dehydratases reutilize the NADH that is generated during the initial C-4″ oxidation to reduce C-6″ to a methyl group following elimination of water (31). It may be interesting to study UDP-L-Ara4O as an inhibitor or as an alternative substrate for some of these enzymes.

The function of the formylated derivative of UDP-L-Ara4N (compound **3,** Fig. 5) is unclear. Formylated L-Ara4N moieties have not been found in association with lipid A or other natural products (14, 25, 44–47). Undecaprenyl phosphate-L-Ara4N, the donor substrate in the ArnT catalyzed modification of lipid A (Fig. 2), has recently been purified from WD101 (28). Extensive characterization by NMR spectroscopy and mass spectrometry did not reveal the presence of a formyl group in this compound (28). However, the existence of such a modification might have been overlooked, given the scheme that was used to purify undecaprenyl phosphate-L-Ara4N (28). If the formylated derivative of UDP-L-Ara4N that can be generated *in vitro* (Fig.

[3] T. Doering, personal communication.

5) is an obligatory intermediate in the biosynthesis of L-Ara4N-modified lipid A, a novel deformylase must also be required at some later stage. One possibility is that compound **3** (rather than UDP-L-Ara4N) is the actual physiological substrate of ArnC (Fig. 2). An *in vitro* assay for ArnC has not yet been developed to test this possibility, and furthermore, a candidate gene encoding a putative deformylase that would have to function prior to the formation of undecaprenyl phosphate-L-Ara4N has not been identified (14, 26). However, we have recently shown that the formylation of UDP-L-Ara4N to make compound **3** is catalyzed by purified ArnA, which is a bifunctional protein with a dehydrogenase domain that synthesizes UDP-L-Ara4O (compound **1**) and a methionyl-tRNA formyltransferase-like domain that generates compound **3**.[2]

The sequence similarity of the dehydrogenase domain of ArnA to the epimerases and the dehydratases is not sufficient to permit homology modeling based on available x-ray structures (53, 54). On the other hand, the homology of the formyltransferase domain of ArnA to the methionyl-tRNA formyltransferases (34) is much greater, and may be amenable to modeling. Many procaryotic ArnA sequences are in fact currently annotated as methionyl-tRNA formyltransferases. It will be important to determine the x-ray structure of ArnA to evaluate its mechanism of action in relationship to the UDP-galactose epimerases and the sugar nucleotide 4,6-dehydratases, and to assess the interaction of the dehydrogenase and formyltransferase domains.

Orthologs of *E. coli* and *S. typhimurium* ArnA are present in the genomes of all pathogenic bacteria known to modify their lipid A with the L-Ara4N moiety (14, 45, 55–57), including all types of *Salmonella*, *Yersinia pestis, Pseudomonas aeruginosa*, and *Burkholderia cepacia*. The existence of full-length ArnA homologues in these organisms indicates that the Ara4N pathway for lipid A modification and polymyxin resistance is likely to be conserved (Fig. 2). The best eucaryotic homologues of ArnA are present in *Arabidopsis thaliana*, but since they correspond only to the dehydrogenase domain of ArnA, they are likely to encode plant UDP-xylose synthases (32, 58).

REFERENCES

1. Raetz, C. R. H. (1990) *Annu. Rev. Biochem.* **59,** 129–170
2. Raetz, C. R. H. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology* (Neidhardt, F. C., ed) Vol. 1, Second Ed., pp. 1035–1063, American Society for Microbiology, Washington, D. C.
3. Brade, H., Opal, S. M., Vogel, S. N., and Morrison, D. C. (eds) (1999) *Endotoxin in Health and Disease*, Marcel Dekker, Inc., New York
4. Poltorak, A., He, X., Smirnova, I., Liu, M. Y., Huffel, C. V., Du, X., Birdwell, D., Alejos, E., Silva, M., Galanos, C., Freudenberg, M., Ricciardi-Castagnoli, P., Layton, B., and Beutler, B. (1998) *Science* **282,** 2085–2088
5. Hoshino, K., Takeuchi, O., Kawai, T., Sanjo, H., Ogawa, T., Takeda, Y., Takeda, K., and Akira, S. (1999) *J. Immunol.* **162,** 3749–3752
6. Aderem, A., and Ulevitch, R. J. (2000) *Nature* **406,** 782–787
7. Medzhitov, R., and Janeway, C., Jr. (2000) *N. Engl. J. Med.* **343,** 338–344
8. Dinarello, C. A. (1991) *Blood* **77,** 1627–1652
9. Beutler, B., and Cerami, A. (1988) *Annu. Rev. Biochem.* **57,** 505–518
10. Bernard, G. R., Vincent, J. L., Laterre, P. F., LaRosa, S. P., Dhainaut, J. F., Lopez-Rodriguez, A., Steingrub, J. S., Garber, G. E., Helterbrand, J. D., Ely, E. W., and Fisher, C. J., Jr. (2001) *N. Engl. J. Med.* **344,** 699–709
11. Ohl, M. E., and Miller, S. I. (2001) *Annu. Rev. Med.* **52,** 259–274
12. Parillo, J. E. (1993) *N. Engl. J. Med.* **328,** 1471–1477
13. Groisman, E. A., Kayser, J., and Soncini, F. C. (1997) *J. Bacteriol.* **179,** 7040–7045
14. Zhou, Z., Lin, S., Cotter, R. J., and Raetz, C. R. H. (1999) *J. Biol. Chem.* **274,** 18503–18514
15. Groisman, E. A. (2001) *J. Bacteriol.* **183,** 1835–1842
16. Wosten, M. M., Kox, L. F., Chamnongpol, S., Soncini, F. C., and Groisman, E. A. (2000) *Cell* **103,** 113–125
17. Nummila, K., Kilpeläinen, I., Zähringer, U., Vaara, M., and Helander, I. M. (1995) *Mol. Microbiol.* **16,** 271–278
18. Helander, I. M., Kilpeläinen, I., and Vaara, M. (1994) *Mol. Microbiol.* **11,** 481–487
19. Roland, K. L., Martin, L. E., Esther, C. R., and Spitznagel, J. K. (1993) *J. Bacteriol.* **175,** 4154–4164
20. Gunn, J. S., Lim, K. B., Krueger, J., Kim, K., Guo, L., Hackett, M., and Miller, S. I. (1998) *Mol. Microbiol.* **27,** 1171–1182
21. Hancock, R. E., Falla, T., and Brown, M. (1995) *Adv. Microb. Physiol.* **37,** 135–175
22. Guo, L., Lim, K. B., Poduje, C. M., Daniel, M., Gunn, J. S., Hackett, M., and Miller, S. I. (1998) *Cell* **95,** 189–198
23. Gunn, J. S., Ryan, S. S., Van Velkinburgh, J. C., Ernst, R. K., and Miller, S. I. (2000) *Infect. Immun.* **68,** 6139–6146
24. Guo, L., Lim, K. B., Gunn, J. S., Bainbridge, B., Darveau, R. P., Hackett, M., and Miller, S. I. (1997) *Science* **276,** 250–253
25. Zhou, Z., Ribeiro, A. A., Lin, S., Cotter, R. J., Miller, S. I., and Raetz, C. R. H. (2001) *J. Biol. Chem.* **276,** 43111–43121
26. Baker, S. J., Gunn, J. S., and Morona, R. (1999) *Microbiology* **145,** 367–378
27. Trent, M. S., Ribeiro, A. A., Lin, S., Cotter, R. J., and Raetz, C. R. H. (2001) *J. Biol. Chem.* **276,** 43122–43131
28. Trent, M. S., Ribeiro, A. A., Doerrler, W. T., Lin, S., Cotter, R. J., and Raetz, C. R. H. (2001) *J. Biol. Chem.* **276,** 43132–43144
29. Gunn, J. S. (2001) *J. Endotoxin Res.* **7,** 57–62
30. Frey, P. A. (1996) *FASEB J.* **10,** 461–470
31. Hegeman, A. D., Gross, J. W., and Frey, P. A. (2001) *Biochemistry* **40,** 6598–6610
32. Gebb, C., Baron, D., and Grisebach, H. (1975) *Eur. J. Biochem.* **54,** 493–498
33. Schutzbach, J. S., and Feingold, D. S. (1970) *J. Biol. Chem.* **245,** 2476–2482
34. Schmitt, E., Panvert, M., Blanquet, S., and Mechulam, Y. (1998) *EMBO J.* **17,** 6819–6826
35. Kelly, T. M., Stachula, S. A., Raetz, C. R. H., and Anderson, M. S. (1993) *J. Biol. Chem.* **268,** 19866–19874
36. Miller, J. R. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
37. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., and Klenk, D. C. (1985) *Anal. Biochem.* **150,** 76–85
38. Blanquet, S., Dessen, P., and Kahn, D. (1984) *Methods Enzymol.* **106,** 141–152
39. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (eds) (1989) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York
40. Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997) *Science* **277,** 1453–1474
41. Summers, M. F., Marzili, L. G., and Bax, A. (1986) *J. Am. Chem. Soc.* **108,** 4285–4294
42. Bax, A., and Summers, M. F. (1986) *J. Am. Chem. Soc.* **108,** 2093–2094
43. Ribeiro, A. A., Zhou, Z., and Raetz, C. R. H. (1999) *Magn. Res. Chem.* **37,** 620–630
44. Zhou, Z., Ribeiro, A. A., and Raetz, C. R. H. (2000) *J. Biol. Chem.* **275,** 13542–13551
45. Volk, W. A., Galanos, C., and Lüderitz, O. (1970) *Eur. J. Biochem.* **17,** 223–229
46. Raetz, C. R. H., Purcell, S., Meyer, M. V., Quershi, N., and Takayama, K. (1985) *J. Biol. Chem.* **260,** 16080–16088
47. Strain, S. M., Armitage, I. M., Anderson, L., Takayama, K., Quershi, N., and Raetz, C. R. H. (1985) *J. Biol. Chem.* **260,** 16089–16098
48. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorf, J. W. (1990) *Methods Enzymol.* **185,** 60–89
49. Reeves, P. R., Hobbs, M., Valvano, M. A., Skurnik, M., Whitfield, C., Coplin, D., Kido, N., Klena, J., Maskell, D., Raetz, C. R. H., and Rick, P. D. (1996) *Trends Microbiol.* **4,** 495–503
50. Agrawal, P. K. (1992) *Phytochemistry* **31,** 3307–3330
51. van Halbeek, H. (1996) in *Encyclopedia of NMR* (Grant, D. M., and Harris, R. K., eds) Vol. 2, pp. 1107–1137, Wiley, Chichester
52. Hallis, T. M., Lei, Y., Que, N. L., and Liu, H. (1998) *Biochemistry* **37,** 4935–4945
53. Thoden, J. B., Frey, P. A., and Holden, H. M. (1996) *Biochemistry* **35,** 5137–5144
54. Allard, S. T., Giraud, M. F., Whitfield, C., Graninger, M., Messner, P., and Naismith, J. H. (2001) *J. Mol. Biol.* **307,** 283–295
55. Lebbar, S., Karibian, D., Deprun, C., and Caroff, M. (1994) *J. Biol. Chem.* **269,** 31881–31884
56. Ernst, R. K., Yi, E. C., Guo, L., Lim, K. B., Burns, J. L., Hackett, M., and Miller, S. I. (1999) *Science* **286,** 1561–1565
57. Burtnick, M. N., and Woods, D. E. (1999) *Antimicrob. Agents Chemother.* **43,** 2648–2656
58. Kyossev, Z. N., Drake, R. R., Kyosseva, S. V., and Elbein, A. D. (1995) *Eur. J. Biochem.* **228,** 109–112
59. Bishop, R. E., Gibbons, H. S., Guina, T., Trent, M. S., Miller, S. I., and Raetz, C. R. H. (2000) *EMBO J.* **19,** 5071–5080
60. Trent, M. S., Pabich, W., Raetz, C. R. H., and Miller, S. I. (2001) *J. Biol. Chem.* **276,** 9083–9092
61. Gibbons, H. S., Lin, S., Cotter, R. J., and Raetz, C. R. H. (2000) *J. Biol. Chem.* **275,** 32940–32949