THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 24, Issue of June 11, pp. 25400–25410, 2004
Printed in U.S.A.


# Oxidation and Transamination of the 3″-Position of UDP-*N*-Acetylglucosamine by Enzymes from *Acidithiobacillus ferrooxidans*

ROLE IN THE FORMATION OF LIPID A MOLECULES WITH FOUR AMIDE-LINKED ACYL CHAINS*[S]

Received for publication, January 20, 2004, and in revised form, March 1, 2004
Published, JBC Papers in Press, March 24, 2004, DOI 10.1074/jbc.M400596200

**Charles R. Sweet‡§, Anthony A. Ribeiro‡¶, and Christian R. H. Raetz‡‖**

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

**Lipid A, a major component of the outer membranes of *Escherichia coli* and other Gram-negative bacteria, is usually constructed around a β-1′,6-linked glucosamine disaccharide backbone. However, in organisms like *Acidithiobacillus ferrooxidans*, *Leptospira interrogans*, *Mesorhizobium loti*, and *Legionella pneumophila*, one or both glucosamine residues are replaced with the sugar 2,3-diamino-2,3-dideoxy-D-glucopyranose. We now report the identification of two proteins, designated GnnA and GnnB, involved in the formation of the 2,3-diamino-2,3-dideoxy-D-glucopyranose moiety. The genes encoding these proteins were recognized because of their location between *lpxA* and *lpxB* in *A. ferrooxidans*. Based upon their sequences, the 313-residue GnnA protein was proposed to catalyze the $NAD^+$-dependent oxidation of the glucosamine 3-OH of UDP-GlcNAc, and the 369-residue GnnB protein was proposed to catalyze the subsequent transamination to form UDP 2-acetamido-3-amino-2,3-dideoxy-α-D-glucopyranose (UDP-GlcNAc3N). Both *gnnA* and *gnnB* were cloned and expressed in *E. coli* using pET23c+. In the presence of L-glutamate and $NAD^+$, both proteins were required for the conversion of [α-$^{32}$P]UDP-GlcNAc to a novel, less negatively charged sugar nucleotide shown to be [α-$^{32}$P]UDP-GlcNAc3N. The latter contained a free amine, as judged by modification with acetic anhydride. Using recombinant GnnA and GnnB, ~0.4 mg of the presumptive UDP-GlcNAc3N was synthesized. The product was purified and subjected to NMR analysis to confirm the replacement of the GlcNAc 3-OH group with an equatorial $NH_2$. As shown in the accompanying papers (Sweet, C. R., Williams, A. H., Karbarz, M. J., Werts, C., Kalb, S. R., Cotter, R. J., and Raetz, C. R. H. (2004) *J. Biol. Chem.* 279, 25411–25419; Que-Gewirth, N. L. S., Ribeiro, A. A., Kalb, S. R., Cotter, R. J., Bulach, D. M., Adler, B., Saint Girons, I., Werts, C., and Raetz, C. R. H. (2004) *J. Biol. Chem.* 279, 25420–25429), UDP-GlcNAc3N is selectively acylated by LpxAs of *A. ferrooxidans*, *L. interrogans*, and *M. loti*. UDP-GlcNAc3N may be useful as a substrate analog for diverse enzymes that utilize UDP-GlcNAc.**

Many Gram-negative bacteria, including some human pathogens, synthesize lipid A molecules in which one or both glucosamine residues are replaced with the analog 2,3-diamino-2,3-dideoxy-D-glucopyranose (GlcN3N)[1] (Fig. 1) (1–7). *Acidithiobacillus ferrooxidans* (3) is a good model system for studying the origin of these lipid A variants, because genomic DNA sequences are available (www.tigr.org). *A. ferrooxidans* grows optimally at pH ~2 (8), and it is often found in the acidic effluents of mines (9).

Despite the prevalence of lipid A molecules containing GlcN3N units (10–12), the enzymes responsible for their formation have not been identified. Using the tBLASTn algorithm, one finds significant orthologs of the *Escherichia coli* lipid A biosynthetic enzymes LpxA, LpxC, LpxD, LpxH, LpxB, LpxK, WaaA (KdtA), and LpxL (Scheme 1) in *A. ferrooxidans* and other bacteria that make GlcN3N-containing lipid A, implying that the well characterized *E. coli* lipid A pathway (12) remains operative. These bioinformatic observations raise the possibility that lipid A molecules containing GlcN3N might be synthesized from the hypothetical sugar nucleotide UDP 2-acetamido-3-amino-2,3-dideoxy-α-D-glucopyranose (UDP-GlcNAc3N), an analog of UDP-GlcNAc in which the GlcNAc 3-OH group is replaced with $NH_2$ (Scheme 1). The selective utilization of this analog by LpxA variants present in certain bacteria could account for the formation of the observed GlcN3N-containing lipid A species (Fig. 1 and Scheme 1).

There is some precedent for the existence of sugar nucleotides resembling UDP-GlcNAc3N. For instance, Okuda *et al.* (13) have described the isolation and characterization of UDP-2,3-diacetamido-2,3-dideoxy-D-glucuronic acid from *Pseudomonas aeruginosa* P1-III, but its enzymatic synthesis was not studied. He and Liu (14) have described genes and enzymes for the biosynthesis of various 3-amino-3-deoxy-sugars. One of these enzymes, TylB from *Streptomyces fradiae* (15), transaminates TDP-3-keto-6-deoxy-α-D-glucose to form TDP-3-amino-

* This work was supported 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 Grant P30-CA-14236. NMR instrumentation in the Duke NMR Center was funded 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.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental figures and a supplemental table.

*The nucleotide sequence(s) reported in this paper has been submitted to the DDBJ/GenBank™/EBI Data Bank with accession number(s) AY541061 and AY541062.*

§ Supported by National Institutes of Health Training Grant GM08558 in Biological Chemistry to Duke University. Present address: Div. of Infectious Diseases, University of Massachusetts Medical Center, Worcester, MA 01655.

‖ To whom correspondence should be addressed: Dept. of Biochemistry, Duke University Medical Center, P.O. Box 3711, Durham, NC 27710. Tel.: 919-684-5326; Fax: 919-684-8885; E-mail: raetz@biochem.duke.edu.

[1] The abbreviations used are: GlcN3N, 2,3-diamino-2,3-dideoxy-D-glucopyranose; BisTris, 2,2-bis(hydroxymethyl)-2,2′,2″-nitrilotriethanol; Ni-NTA, nickel-nitrilotriacetic acid; UDP-GlcNAc3N, UDP 2-acetamido-3-amino-2,3-dideoxy-α-D-glucopyranose; PEI, polyethyleneimine.

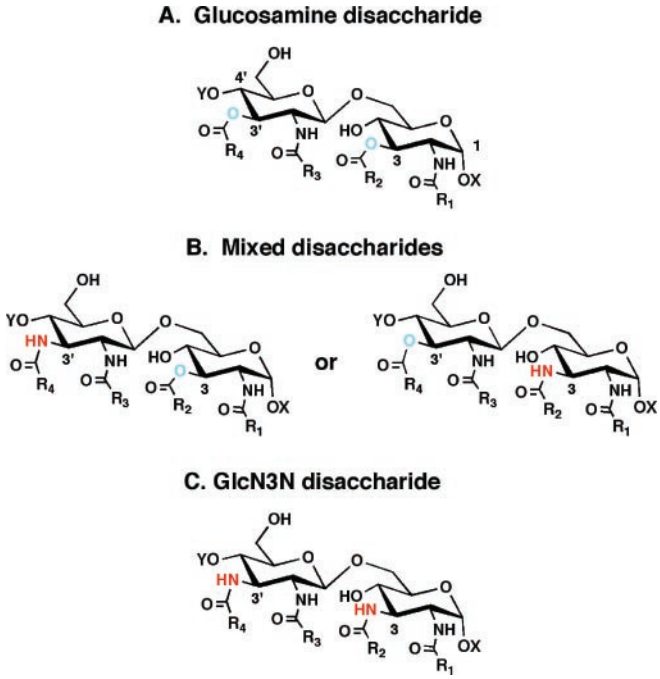


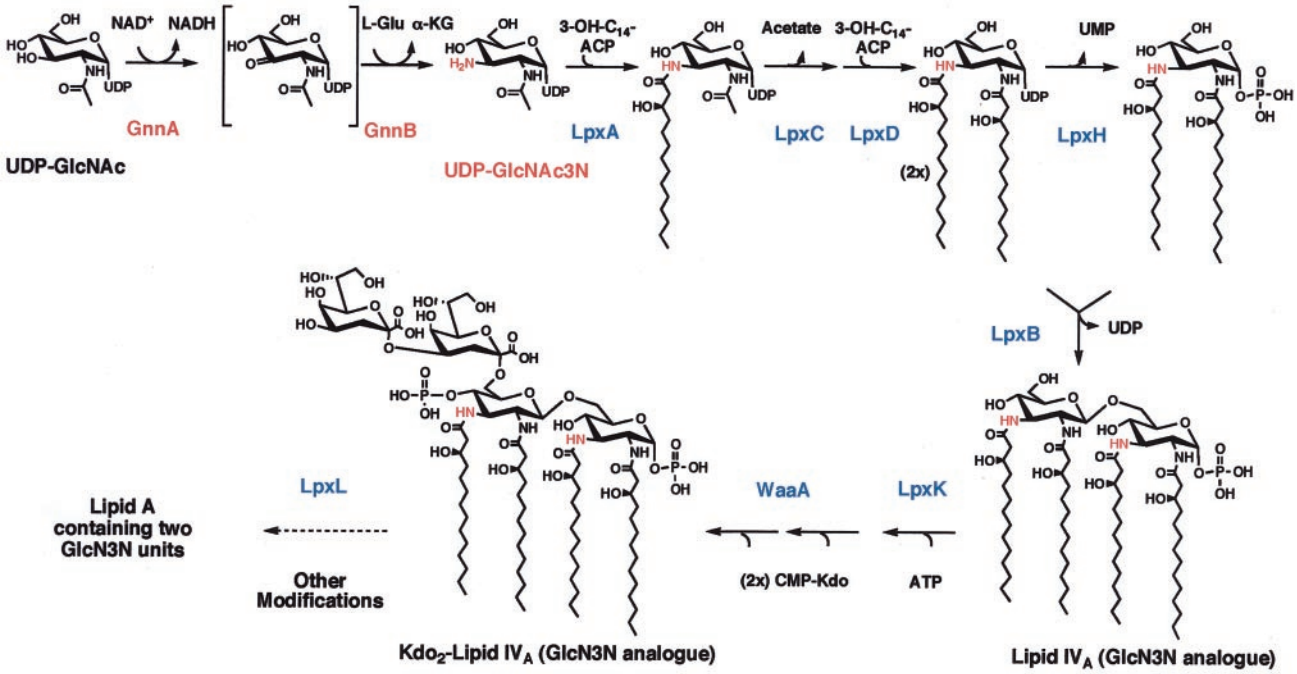

FIG. 1. **Lipid A disaccharides in which one or both glucosamine residues are replaced with GlcN3N units.** The nitrogen atom at the 3- or 3′-positions in the GlcN3N units is indicated in *red*. Oxygen, which occupies those positions in glucosamine, is shown in *cyan*. *A*, classical glucosamine-based lipid A disaccharide. The 1- and 4′-positions (*X* and *Y*, respectively) are typically substituted with phosphate or phosphate esters, as in *E. coli* but may also be H and/or galacturonic acid in other systems (11, 12). *B*, mixed lipid A disaccharides. *A. ferrooxidans* (3) or *Campylobacter jejuni* (47) make lipid A species containing both glucosamine and GlcN3N units. *C*, GlcN3N-based lipid A disaccharide. Organisms that make the GlcN3N disaccharide exclusively include *Leptospira interrogans* (7) and *Mesorhizobium loti* (4).

3,6-dideoxy-α-D-glucose, a key precursor of the mycaminose unit in the antibiotic tylosin.

Based on what was known about mycaminose, it seemed plausible that the enzymatic formation of UDP-GlcNAc3N might involve the $NAD^+$-dependent oxidation of the GlcNAc 3-OH moiety of UDP-GlcNAc, followed by transamination (Scheme 1). While inspecting the *lpxA* region of the *A. ferrooxidans* genome, we noticed that two genes, one encoding an oxidoreductase and the other encoding a transaminase, are inserted between *lpxA* and *lpxB* (Scheme 1). We reasoned that these genes might be involved in GlcN3N formation, given that *lpxA* and *lpxB* are contiguous in *E. coli* (16–18) and other bacteria. We have now cloned and purified the two proteins encoded by these *A. ferrooxidans* genes and demonstrate that they can convert UDP-GlcNAc to UDP-GlcNAc3N in the presence of $NAD^+$ and glutamate. The recombinant oxidoreductase (GnnA) and the transaminase (GnnB) can be used to prepare milligram quantities of UDP-GlcNAc3N, facilitating analysis by NMR and characterization of its role as a substrate for diverse LpxA orthologs. As shown in the accompanying manuscripts (7, 19), the LpxAs of bacteria that contain GlcN3N in their lipid A are highly selective for UDP-GlcNAc3N. A preliminary communication of our results has appeared in abstract form (20).

## EXPERIMENTAL PROCEDURES

*Materials*—All of the growth media were purchased from Difco, Oxoid, or Mallinckrodt Chemical Works through VWR Scientific. Ampicillin, $NAD^+$, $FAD^+$, $NADP^+$, pyridoxal phosphate, pyridoxine, α-ketoglutarate, amino acids, and fine chemicals were purchased from Sigma-Aldrich. PerkinElmer Life Sciences was the source for radiolabeled [$\alpha$-$^{32}$P]UTP. PEI-cellulose plates were purchased from VWR Scientific. T4 ligase and preparative grade phenol/chloroform/isoamyl alcohol (25:24:1, v/v/v) were from Invitrogen. The *Pfu* polymerase was from Strat-




SCHEME 1. **Proposed pathway for biosynthesis of lipid A molecules containing two GlcN3N units.** The *A. ferrooxidans* enzymes GnnA and GnnB leading from UDP-GlcNAc to the novel sugar nucleotide UDP-GlcNAc3N are highlighted in *red*. The putative ketone intermediate (in *brackets*), UDP 2-acetamido-2-deoxy-α-D-*ribo*-hexopyranos-3-ulose, has not yet been isolated and characterized. The distal conserved enzymes of the constitutive lipid A pathway (12), which are encoded in the genomes of all Gram-negative bacteria (including the ones that make GlcN3N-containing lipid A), are indicated in *blue*. Lipid A molecules usually have at least one acyloxyacyl moiety incorporated by LpxL (11, 12) and may contain longer or shorter *R*-3-hydroxylacyl chains in place of the indicated *R*-3-hydroxymyristate groups. Those bacteria that make mixed lipid A disaccharides (Fig. 1B), which include *A. ferrooxidans* (3), have an ortholog of LpxA that is not entirely selective for UDP-GlcNAc3N over UDP-GlcNAc (19). *L. interrogans* and *M. loti* LpxAs display absolute selectivity for UDP-GlcNAc3N over UDP-GlcNAc (19) and synthesize lipid A containing only GlcN3N (4, 7).

TABLE I
*Strains and plasmids*

| | Description | Source |
|---|---|---|
| Bacterial strains | | |
| *A. ferrooxidans* 23270 | | ATCC |
| *E. coli* BL21(DE3)/pLysS | DE3 lysogen with pLysS plasmid | Stratagene |
| *E. coli* BL21(DE3)/pRP | Codon plus plasmid | Stratagene |
| *E. coli* XL1-Blue | | Stratagene |
| Plasmids | | |
| pET23c+ | T7 expression vector, $Amp^r$ | Novagen |
| pCS321 | pET23c+ containing *A. ferrooxidans gnnA* | This work |
| pCS331 | pET23c+ containing *A. ferrooxidans gnnB* | This work |
| pCS355 | pET23c+ containing *A. ferrooxidans gnnA* and *gnnB* | This work |
| pCS462 | pET23c+ containing $His_6$-tagged *A. ferrooxidans gnnA* | This work |
| pCS484 | pET23c+ containing $His_6$-tagged *A. ferrooxidans gnnB* | This work |
| pCS591 | pET23c+ containing *A. ferrooxidans gnnA* and *gnnB* K190A | This work |
| pCS601 | pET23c+ containing $His_6$-tagged *A. ferrooxidans gnnB* K190A | This work |

agene. Restriction endonucleases, other modification enzymes, and deoxy-ribonucleotide triphosphates were from New England Biolabs.

*Bacterial Strains and Plasmids*—All of the constructs utilized in this study are summarized in Table I. *A. ferrooxidans* 23270 was grown on modified 9K medium (21). *E. coli* strains were grown on LB agar or in LB broth adjusted to pH 7.4 (22). Bacteria bearing plasmid DNA were selected using ampicillin at 100 $\mu$g/ml.

*Recombinant DNA Techniques*—Preparation of growth media, competent cells, transformation, nucleic acid purification, and electrophoresis were performed according to published procedures (23, 24). The plasmids were purified using the Qiaprep miniprep spin column kit (Qiagen), and DNA was extracted from agarose gels using the Qiaquick gel extraction spin column kit (Qiagen). *E. coli* cells were made competent for transformation by the $CaCl_2$ method (23, 24).

*PCR of gnnA and gnnB from A. ferrooxidans Genomic DNA*—*A. ferrooxidans* genomic DNA was prepared from cells grown on modified 9K medium, as described (48). The *gnnA* and *gnnB* genes were amplified with *Pfu* polymerase. PCR primers were made based on available DNA sequences for *A. ferrooxidans* (see www.tigr.org). The N-terminal oligonucleotide primer for *gnnA* was 5′-GGG TTT ATT AAT GAT CCA TAT GAG AAC CGG GGT CAT CG-3′, and the C-terminal primer was 5′-CGG GAT CCC GTC AAA ACC CGC ACC GGC ATA CGC-3′. The AseI and BamHI sites are underlined. AseI makes sticky ends matching those generated by NdeI. The plasmid containing the desired *gnnA* insert was designated pCS321. The N-terminal primer used to clone *gnnB* was 5′-GAA ATT CCA TAT GCA TGA CTC AAA ATA CAG CAA TTC CC-3′, and the C-terminal *gnnB* primer was 5′-CCC AAA AAG CTT CGA GAT TTT CGC CGG AAC GCT C-3′. The NdeI and HindIII restriction sites are underlined. The desired plasmid containing the *gnnB* gene ligated into NdeI/HindIII-cut pET23c+ was designated pCS331.

The contiguous *gnnA* and *gnnB* genes were also cloned in tandem as a single PCR fragment into the same expression vector. To create this construct, PCR was performed using the above *gnnA* N-terminal primer and *gnnB* C-terminal primer. The successful ligation of this PCR product into pET23c+ cut with NdeI and BamHI generated the construct pCS355. All of the inserts were confirmed by DNA sequencing.

To construct plasmids expressing GnnA or GnnB with a C-terminal hexahistidine tag, two additional primers were required. The *gnnA* His tag C-terminal primer was 5′-CCC GCC TCG AGA CGC TGC AGA AAA GCC TCC ACG GC-3′, and the *gnnB* His tag C-terminal primer was 5′-CCC GCC TCG AGG CCG TGC AGG GTG CGA CGG ATC AC-3′. The XhoI site is underlined in both primers. These two primers were used in conjunction with the *gnnA* and *gnnB* N-terminal primers listed above to create the His-tagged constructs. The desired ligations of these PCR products into pET23c+ (cut with NdeI and XhoI) were designated pCS462 and pCS484, respectively.

The predicted amino acid sequence of GnnA is MIHMRTGVIG VGHLGRFHAQ KYAAISQLAG VFDENAERAA EVAAELRCRA FPSVDALLAE VDAVSIVTPT STHFAVAEVA MQAGVHCLIE KPFTLDTEEA DALIGMAQER HLVLAIGHIK RVHPAIQYLR QAGFGAPRYL EAERLAPFKP RSLDIDVIMD LMIHDLDLTL LLTGAEPVDV RAVGVAAVTD KADMATAWMT LNNGTVANLA ASRVVREPAR RMRIFWQDRY ASVDFLNNTL HIYHRGAGTV PGIPGVRDEA VDLAKRDALA AEIEDFLNAI AAHRPVFCDG VAGRRVLAAA LQVRVAVEAF LQR. The predicted amino acid sequence of GnnB is MTQNTAIPM VDLRAHFAPL RDEILTGIGK ILDDASFILG NQGRALEAEV AGLSGVAHGV GCASGTDALM LALRALEIGP GDEVIVPTFT FIATAEAVLY VGATPVFVDV DDRFYAMTIA GIEAAITPRT KAIIPVHLYG LPADMPGIMA LAQKHGLRVI EDCAQAIGAQ INGQGVGSFG DIGCFSFFPS KNLGAAGDGG MVVTADAELE RKLRGLRNHG SWQTYHHDVL GYNSRLDEMQ AVILRAEFPH LAAYNDGRRR AAGWYAEHLV GLDLQLPEAP AGYHHVFHQF TIQLNARDAV KTALHAEGIA SAIYYPIPGH QQKMFAHQAQ THCPVAEHLA ERVLSLPMFP ELREEQIARI ATVIRRTLHG.

*Construction of the GnnB Mutant K190A*—The Stratagene QuikChange double-stranded PCR mutagenesis kit was used to introduce the lysine to alanine substitution at position 190 of GnnB. This mutagenesis was carried out using pCS355 or pCS484 as template, and positive clones were designated pCS601 and pCS591, respectively.

*Preparation of A. ferrooxidans Cell-free Extracts*—To prepare extracts of *A. ferrooxidans* 23270, a culture was grown in 1 liter of modified 9K medium (21) at 30 °C with 215 rpm rotary shaking for 90 h. The culture was harvested by centrifugation at 4000 × *g* at 25 °C for 15 min, as for the genomic DNA isolation (48). The bacteria in 2 ml of 100 mM HEPES, pH 8.0, were broken by two passages through a French pressure cell at 16,000 p.s.i. The extract was centrifuged at 10,000 × *g* to remove large debris and residual inorganic precipitate. The protein concentration was determined using bicinchoninic acid (25) with bovine serum albumin as the standard. The sample was stored at −80 °C.

*Preparation of E. coli Cell-free Extracts*—To prepare cell extracts of the overexpressing constructs for assays, cultures were grown at 30 °C in 50 ml of LB broth containing 80 $\mu$g/ml ampicillin, as described previously (26). When $A_{600}$ had reached 0.5, the cultures were shifted to 18 °C for 10 min, induced with 1 mM isopropyl-1-thio-$\beta$-D-galactopyranoside, and grown overnight. The cultures expressing GnnB were supplemented with pyridoxine.

*Ni-NTA Purification of the Hexahistidine-tagged GnnA and GnnB Proteins*—GnnA and GnnB were purified from extracts of 500-ml BL21(DE3)/pLysS/pCS462 and BL21(DE3)/pRP(CodonPlus)/pCS484 cultures, respectively, using Ni-NTA affinity chromatography (Qiagen). A 2-ml column of Ni-NTA resin was poured using 4 ml of the commercial 50% resin, 30% ethanol slurry and was washed with 20 ml of 100 mM HEPES, pH 7.5, containing 10% glycerol. The extract then was loaded onto the column at 1 ml/min, and the flow through was collected as a single fraction. The column was washed with 20 ml of 100 mM HEPES, pH 7.5, containing 10% glycerol, before stepwise elution with 10 ml each of 25, 50, 100, 200, and 400 mM imidazole in 100 mM HEPES, pH 7.5, containing 10% glycerol. The effluent was collected in 2-ml fractions, which were analyzed by 12% SDS-PAGE and by activity assays.

*Substrate Preparation and in Vitro Coupled Assay for GnnA/GnnB*—The [$\alpha$-$^{32}$P]UDP-GlcNAc substrate was prepared as described (26). The assay for GnnA/GnnB monitors conversion of [$\alpha$-$^{32}$P]UDP-GlcNAc to [$\alpha$-$^{32}$P]UDP-GlcNAc3N based upon the more rapid mobility of the latter during TLC on PEI-cellulose. Each 10-$\mu$l GnnA/GnnB reaction contained 50 mM HEPES, pH 8.0, 1 mM $NAD^+$, 100 mM L-glutamic acid (for extracts), or 10 mM L-glutamic acid (for pure protein), 100 $\mu$M UDP-GlcNAc (for extracts), or 10 $\mu$M UDP-GlcNAc (for pure protein), [$\alpha$-$^{32}$P]UDP-GlcNAc ($\approx 2 \times 10^5$ dpm/reaction), 0.25 mg/ml bovine serum albumin (for pure protein), and an appropriate amount of enzyme.

The assays were carried out at 30 °C, and product formation at various times was analyzed by spotting 1-$\mu$l portions of each reaction mixture onto a PEI-cellulose plate. The spots were allowed to air dry. The plate was then soaked in anhydrous methanol for 10 min and dried again. The plate was then developed using 0.2 M guanidine HCl and was analyzed with a PhosphorImager (Molecular Dynamics Storm 840 system) equipped with Molecular Dynamics ImageQuant software (26).

Chemical modification of the [$\alpha$-$^{32}$P]UDP-GlcNAc3N product with acetic anhydride was accomplished by combining 26 $\mu$l of an assay

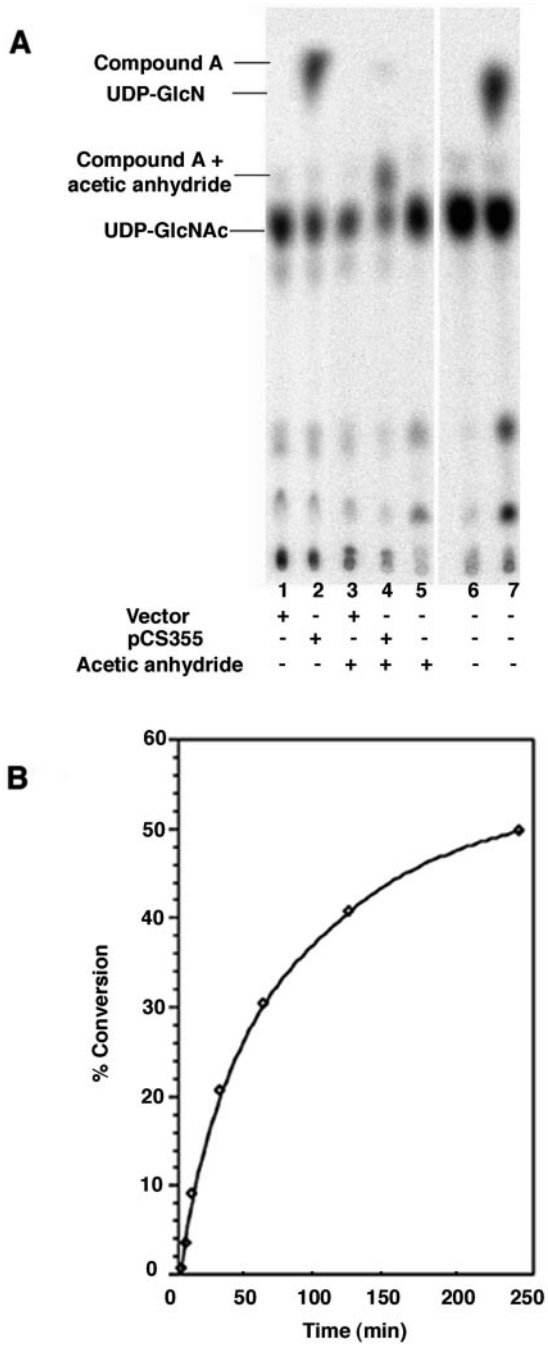



FIG. 2. **Conversion of UDP-GlcNAc to Compound A by *E. coli* extracts expressing *A. ferrooxidans* GnnA and GnnB.** Enzymatic conversion of 100 $\mu$M [$\alpha$-$^{32}$P]UDP-GlcNAc to Compound A after 4 h at 30 °C was monitored by PEI-cellulose TLC. The plate was developed using 0.2 M guanidine HCl and analyzed with a PhosphorImager. *A*, *lanes 1* and *3* are from a reaction containing 0.5 mg/ml of a vector control extract (pET23c+). *Lanes 2* and *4* are from a reaction containing 0.5 mg/ml of a pCS355 extract, which expresses both GnnA and GnnB. The reaction mixtures in *lanes 3* and *4* were treated with acetic anhydride for 3 min at room temperature prior to TLC. *Lane 5* demonstrates that acetylation of a mixture of [$\alpha$-$^{32}$P]UDP-GlcN and [$\alpha$-$^{32}$P]UDP-GlcNAc standards with acetic anhydride generates only one spot migrating with [$\alpha$-$^{32}$P]UDP-GlcNAc. *Lane 6* is untreated [$\alpha$-$^{32}$P]UDP-GlcNAc control, and *lane 7* is a mixture of untreated [$\alpha$-$^{32}$P]UDP-GlcN and [$\alpha$-$^{32}$P]UDP-GlcNAc standards. *B*, time course of the conversion of 100 $\mu$M [$\alpha$-$^{32}$P]UDP-GlcNAc to Compound A by the pCS355 extract at 0.5 mg/ml.

mixture with 30 $\mu$l of methanol, 24 $\mu$l of $H_2O$, 1 $\mu$l of saturated $NaHCO_3$, and 1 $\mu$l of acetic anhydride. Acetylation of the amine group was complete after 3 min at room temperature.

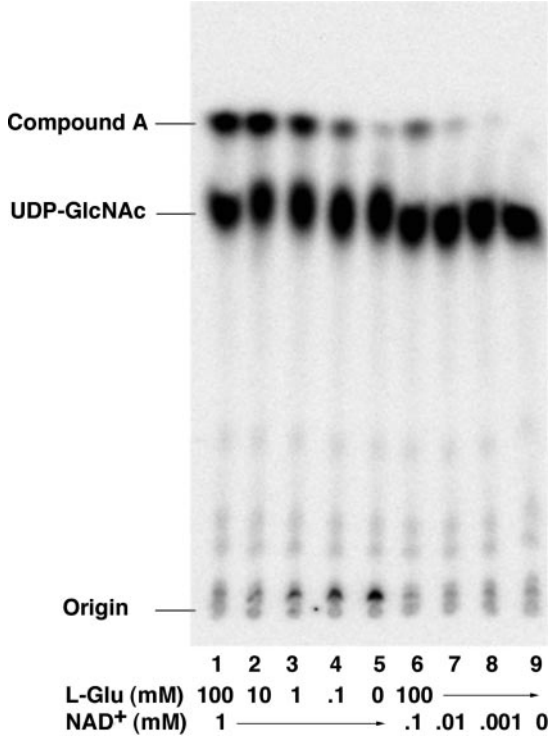



FIG. 3. **Formation of Compound A is dependent on NAD$^+$ and L-glutamate.** Purified, His-tagged GnnA and GnnB (10 $\mu$g/ml each) were assayed for 10 min as indicated with 10 $\mu$M [$\alpha$-$^{32}$P]UDP-GlcNAc as the sugar nucleotide probe.

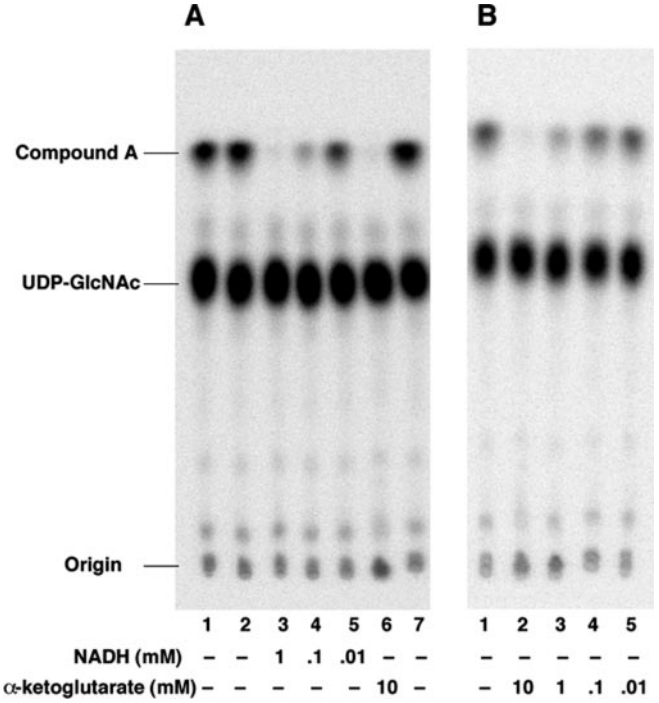



FIG. 4. **Inhibition of Compound A synthesis by NADH or $\alpha$-ketoglutarate.** *A*, reaction mixtures contained 14 $\mu$g/ml of each purified protein and were incubated for 30 min with 10 $\mu$M [$\alpha$-$^{32}$P]UDP-GlcNAc as the sugar nucleotide substrate. *Lanes 1* and *2* represent the standard assay conditions, but the enzyme samples used in *lane 2* were subjected to four cycles of freezing and thawing. The reaction mixtures in *lanes 3–5* contained NADH, as indicated. *Lane 6* contained 10 mM $\alpha$-ketoglutarate. The reaction in *lane 7* was supplemented with 1 milliunit of purified NADH-oxidase (Sigma). *B*, *lane 1* is the standard assay condition. The reaction mixtures in *lanes 2–5* were supplemented with $\alpha$-ketoglutarate, as indicated.

*Large Scale Purification of the Product Synthesized in Vitro by GnnA and GnnB*—The product of the *in vitro* GnnA/GnnB reaction was purified on a milligram scale by three ion exchange chromatography steps from a 20-ml reaction mixture containing 50 mM HEPES, pH 8.0, 200 mM L-glutamate, 200 $\mu$M UDP-GlcNAc, 1 mM $NAD^+$, and 0.5 mg/ml extract of induced cells harboring pCS355. The reaction was held at 30 °C for 5 h, chilled on ice for 10 min, and then diluted 3-fold with cold ethanol. The diluted reaction mixture was chilled for 10 min on ice and centrifuged at 5000 × *g* at 4 °C to remove protein. The supernatant was diluted another 4-fold with $H_2O$ and loaded at 1.5 ml/min onto a 20-ml DEAE-cellulose column (Whatman DE52), equilibrated with 10 mM BisTris, pH 6.0. The column was washed with 80 ml of $H_2O$, and 60 ml of 10 mM BisTris, pH 6.0, and 60 ml of 20 mM BisTris, pH 6.0. The column was then eluted at 2 ml/min with 64 ml of 30 mM BisTris, pH 6.0, 136 ml of 40 mM BisTris, pH 6.0, and 128 ml of 80 mM BisTris, pH 6.0. Fractions of 8 ml were collected. The absorbance of the fractions was monitored by $A_{254}$, and the fractions were analyzed by capillary electrophoresis (Beckman P/ACE 5010 system) using a bare silica capillary and 25 mM sodium tetraborate, pH 9.4, as the running buffer (27, 28). The samples were analyzed at 22 kV and were detected by absorbance at 254 nm. The capillary electrophoresis revealed that $NAD^+$ elutes in the 20 to 40 mM BisTris fractions, the putative UDP-GlcNAc3N elutes in the 40 mM fractions, and UDP-GlcNAc elutes in the 80 mM fractions.

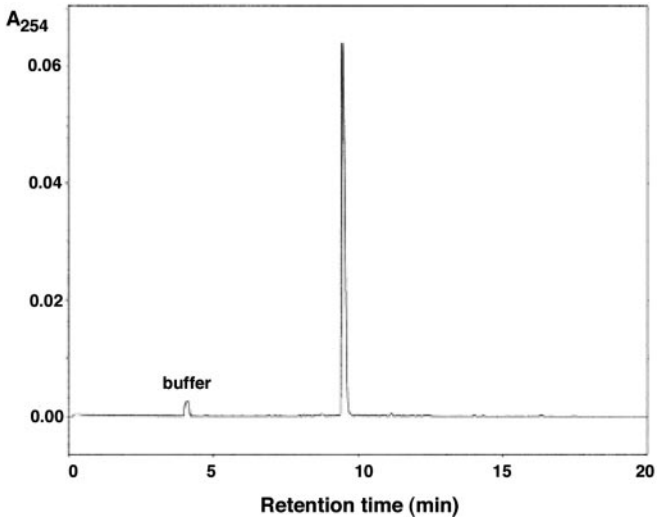



FIG. 5. **Capillary electrophoresis of purified Compound A.** The sodium salt of the purified product was subjected to capillary electrophoresis using a bare silica capillary at 22 kV in 25 mM sodium tetraborate buffer, pH 9.4 (27, 28).

To purify the UDP-GlcNAc3N further, a second DEAE column was run at pH 8.5. The UDP-GlcNAc3N fractions from the first column were pooled, diluted 4-fold with water, and loaded onto a 10-ml DEAE column (Whatman DE52) equilibrated with 10 mM triethylammonium bicarbonate at pH 8.5. The column washed with 20 ml of $H_2O$ and eluted with triethylammonium bicarbonate at pH 8.5 as follows: 16 ml at 10 mM, 20 ml at 20 mM, 36 ml at 30 mM, 16 ml at 40 mM, 16 ml at 50 mM, 16 ml at 60 mM, 16 ml at 70 mM, 24 ml at 80 mM, and 8 ml at 1 M. The elution was monitored by $A_{254}$, and the fractions were analyzed by capillary electrophoresis (27, 28). The 70 mM fractions contained the desired product. They were pooled and lyophilized to yield the triethylammonium salt of UDP-GlcNAc3N. This material was redissolved in 1 ml of $H_2O$. The concentration was calculated using the UDP extinction coefficient of $9.9 \times 10^3$ $M^{-1}$ $cm^{-1}$ at 262 nm (29).

Before NMR analysis could be performed on the above UDP-GlcNAc3N sample, the excess triethylamine was removed with a Dowex AG-50W 8X column. The final yield after purification was ~0.36 mg of the sodium salt of UDP-GlcNAc3N, about 15% of the 2.4 mg of UDP-GlcNAc used as the starting material.

*NMR Analysis of the Purified Sugar Nucleotide*—NMR spectroscopy was performed at the Duke University NMR Spectroscopy Center. The sodium salt of the putative UDP-GlcNAc3N (~1 mM) and the substrate UDP-GlcNAc (~10 mM) were each dissolved in 0.6 ml of 99% $D_2O$ in 5 mm NMR tubes. Chemical shifts were referenced to 2,2-dimethylsilapentane-5-sulfonic acid at 0.00 ppm.

The UDP-GlcNAc3N was studied initially at pD of 7.8, and the substrate UDP-GlcNAc was evaluated at pD 7.8, as measured with a 3 mm pH electrode. Because this pD is in the range expected for the p$K_a$ of a primary amine on a pyranose ring, the putative UDP-GlcNAc3N was later acidified with 1 $\mu$l of 1 M perdeuterated acetic acid in $D_2O$, and the NMR spectroscopy was repeated. The spectra were recorded on a Varian Inova 600 or 800 MHz spectrometer, as indicated, equipped with Sun Ultra 10 computers and 5-mm Varian triple resonance probes (30).

TABLE II
*$^1$H NMR assignments for UDP-GlcNAc and the putative UDP-GlcNAc3N*

The values shown are for $^1$H chemical shifts at 25 °C in $D_2O$ relative to internal 2,2-dimethylsilapentane-5-sulfonic acid.

| Moiety | UDP-GlcNAc (pD 7.8) | | | UDP-GlcNAc3N | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | pD 7.8 | | | pD 4.0 | | |
| | δH [mult] | *J* | [Hz] | δH [mult] | *J* | [Hz] | δH [mult] | *J* | [Hz] |
| Uracil | | | | | | | | | |
| H-5 | 5.972 [d] | $J_{5,6}$ | 8.13 | 5.968 [d] | $J_{5,6}$ | 8.13 | 5.967 [d] | $J_{5,6}$ | 8.13 |
| H-6 | 7.962 [d] | | | 7.958 [d] | | | 7.956 [d] | | |
| Ribose | | | | | | | | | |
| H-1′ | 5.988 [d] | $J_{1',2}$ | 4.46 | 5.986 [d] | $J_{1',2}$ | 4.36 | 5.989 [d] | $J_{1',2}$ | 4.47 |
| H-2′ | ~4.37 | | | ~4.37 | | | ~4.37 | | |
| H-3′ | ~4.37 | | | ~4.37 | | | ~4.37 | | |
| H-4′ | ~4.292 [m] | | | 4.296 [m] | | | 4.300 [m] | | |
| H-5′a | 4.193 [ddd] | $J_{5'a,P}$ | 5.6 | 4.200 [ddd] | $J_{5'a,P}$ | 5.7 | 4.203 [ddd] | $J_{5'a,P}$ | 5.5 |
| | | $J_{5'a,4}$ | 3.1 | | $J_{5'a,4}$ | 3.2 | | $J_{5'a,4}$ | 3.3 |
| | | $J_{5'a,5'b}$ | 11.8 | | $J_{5'a,5'b}$ | 11.8 | | $J_{5'a,5'b}$ | 11.7 |
| H-5′b | 4.248 [ddd] | $J_{5'b,P}$ | 4.5 | 4.261 [ddd] | $J_{5'b,P}$ | 4.5 | 4.268 [ddd] | $J_{5'b,P}$ | 4.6 |
| | | $J_{5'b,4}$ | 2.6 | | $J_{5'b,4}$ | 2.6 | | $J_{5'b,4}$ | 2.7 |
| Pyranose | | | | | | | | | |
| H-1″ | 5.517 [dd] | $J_{1'',2'}$ | 3.2 | 5.541 [dd] | $J_{1'',2''}$ | 3.3 | 5.577 [dd] | $J_{1'',2''}$ | 3.3 |
| | | $J_{1'',P}$ | 7.2 | | $J_{1'',P}$ | 6.9 | | $J_{1'',P}$ | 7.0 |
| H-2″ | 3.994 [dt] | $J_{2'',3''}$ | 10.3 | ~4.21 [dt][a] | $J_{2'',3''}$ | 9.8 | 4.343 [dt] | $J_{2'',3''}$ | 11.5 |
| | | $J_{2'',P}$ | 3.0 | | | | | $J_{2'',P}$ | 3.0 |
| H-3″ | 3.810 [dd] | $J_{3'',4''}$ | 9.3 | 3.394 [m] | | | 3.577 [dd] | $J_{3'',4''}$ | 10.0 |
| H-4″ | 3.504 [dd] | $J_{4'',5''}$ | 10.1 | 3.674 [m] | $J_{4'',5''}$ | 10.0 | 3.820 [dd] | $J_{4'',5''}$ | 10.1 |
| H-5″ | 3.934 [ddd] | $J_{5'',6''a}$ | 4.9 | 3.968 [ddd] | $J_{5'',6''a}$ | 4.2 | 4.001 [ddd] | $J_{5'',6''a}$ | 3.9 |
| | | $J_{5'',6''b}$ | 2.3 | | $J_{5'',6''b}$ | 2.2 | | $J_{5'',6''b}$ | 2.3 |
| H-6″a | 3.80 [dd] | $J_{6''a,6''b}$ | 12.5 | 3.820 [dd] | $J_{6''a,6''b}$ | 12.6 | 3.836 [dd] | $J_{6''a,6''b}$ | 12.6 |
| H-6″b | 3.874 [dd] | | | 3.882 [dd] | | | 3.894 [dd] | | |
| Acetyl | | | | | | | | | |
| $CH_3$ | 2.082 [s] | | | 2.097 [s] | | | 2.105 [s] | | |

[a] Overlaps ribose H-5′a. Coupling constant estimated from the resolved H-3″ signal.

RESULTS

*Identification of the gnnA and gnnB Genes in A. ferrooxidans*—The *gnnA* and *gnnB* genes are situated between *lpxA* and *lpxB* (Supplementary Table I) in *A. ferrooxidans*. In *E. coli*, the *lpxA* and *lpxB* genes are contiguous, and their transcription is coupled (17, 31). The fact that the reading frames of *lpxA* and *gnnA* overlap by 4 base pairs in *A. ferrooxidans* suggests that *gnnA* and *gnnB* may be coordinately expressed with *lpxA* and *lpxB* in this system and may encode novel enzymes of lipid A biosynthesis.

The predicted GnnA protein is homologous to $NAD^+$-dependent dehydrogenases of the Gfo/Idh/MocA family (32). It might therefore function to oxidize the glucosamine 3-OH group of UDP-GlcNAc. Close orthologs of GnnA have a relatively limited distribution among diverse bacteria. GnnB is homologous to TylB, which is a pyridoxal phosphate-dependent 3″-keto-pyranose transaminase involved in mycaminose biosynthesis (15). Accordingly, GnnB may transaminate the putative 3″-keto-sugar generated by GnnA (Scheme 1). A good ortholog ($E^{-48}$ or better) of GnnB is present in all bacteria known to synthesize lipid A with three or four *N*-linked hydroxyacyl chains (Supplementary Table I). Furthermore, GnnB is present in every organism that possesses a full-length GnnA. The *gnnA* and *gnnB* genes are often adjacent to each other and to *lpxA* and *lpxB*, as in *A. ferrooxidans* (Supplementary Table I).

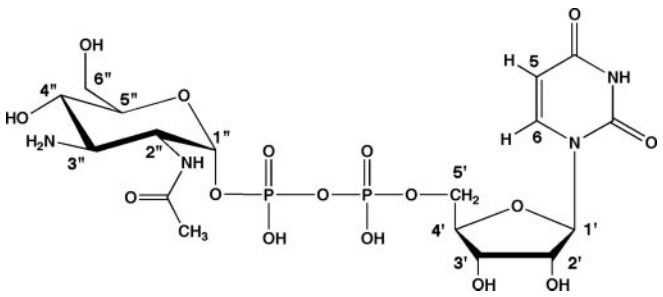



FIG. 6. **Proposed covalent structure and numbering scheme for UDP-GlcNAc3N (Compound A).**

*Recombinant GnnA and GnnB Convert UDP-GlcNAc to a Less Negatively Charged Sugar Nucleotide*—Cell-free extracts of *E. coli* BL21(DE3)/pLysS/pCS355, which expresses both *gnnA* and *gnnB*, were assayed alongside vector controls for their ability to convert [$\alpha$-$^{32}$P]UDP-GlcNAc to a less negatively charged compound in the presence of $NAD^+$ and L-glutamate. Following incubation with enzyme, the substrate and product were resolved on a PEI-cellulose plate, which was analyzed with a PhosphorImager. In the presence of both GnnA and GnnB, [$\alpha$-$^{32}$P]UDP-GlcNAc is converted to a more quickly migrating (*i.e.* less negatively charged) derivative (Fig. 2*A*, *lane 2*, *Compound A*) with an $R_f$ close to that of [$\alpha$-$^{32}$P]UDP-GlcN (Fig. 2*A*, *lane 7*). Compound A appears to contain a free amine group, as judged by its susceptibility to modification with acetic anhydride (Fig. 2*A*, *lane 4*). However, Compound A is not [$\alpha$-$^{32}$P]UDP-GlcN, because it is converted by acetic anhydride to substance that migrates slightly more quickly than [$\alpha$-$^{32}$P]UDP-GlcNAc (Fig. 2*A*, *lane 4*). Formation of Compound A is dependent upon expression of both *gnnA* and *gnnB* (see below). Product formation is efficient in cell extracts of *E. coli* BL21(DE3)/pLysS/pCS355 and is linear with time below 10% conversion of substrate (Fig. 2*B*). Small amounts of Compound A are formed in extracts of wild-type *A. ferrooxidans* under similar conditions (not shown).

*Purification and Assay of GnnA and GnnB*—To characterize the above reactions in more detail, C-terminal hexahistidine-tagged versions of GnnA and GnnB were constructed, and following expression in *E. coli*, these proteins were purified by Ni-NTA affinity chromatography. Both proteins, together with both $NAD^+$ and L-glutamate (Fig. 3), are required to support the formation of Compound A. $FAD^+$ and $NADP^+$ do not substitute for $NAD^+$ (data not shown). NADH and $\alpha$-keto-glutarate inhibit the reaction (Fig. 4). Other amine donors, such as alanine and glutamine, can substitute for glutamate (data not

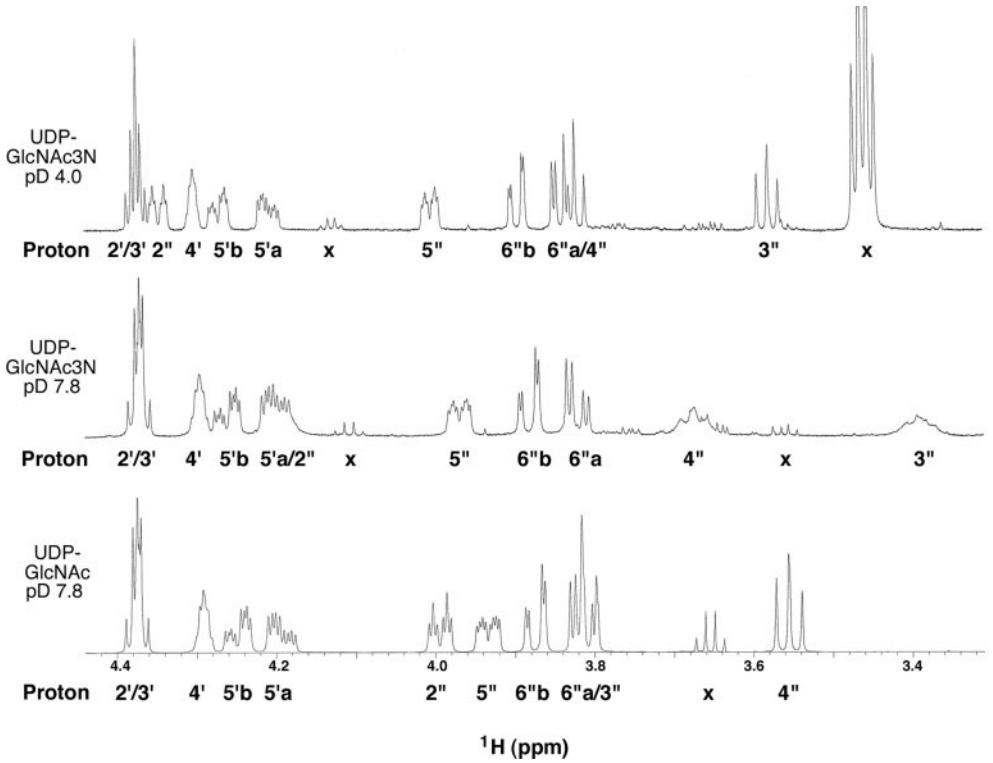



FIG. 7. **Partial one-dimensional $^1$H NMR spectra of UDP-GlcNAc and UDP-GlcNAc3N.** The protons are labeled according to the scheme shown in Fig. 6. The UDP-GlcNAc concentration is 10 mM. *x* indicates impurities, most likely residual triethylamine in the case of UDP-GlcNAc3N.

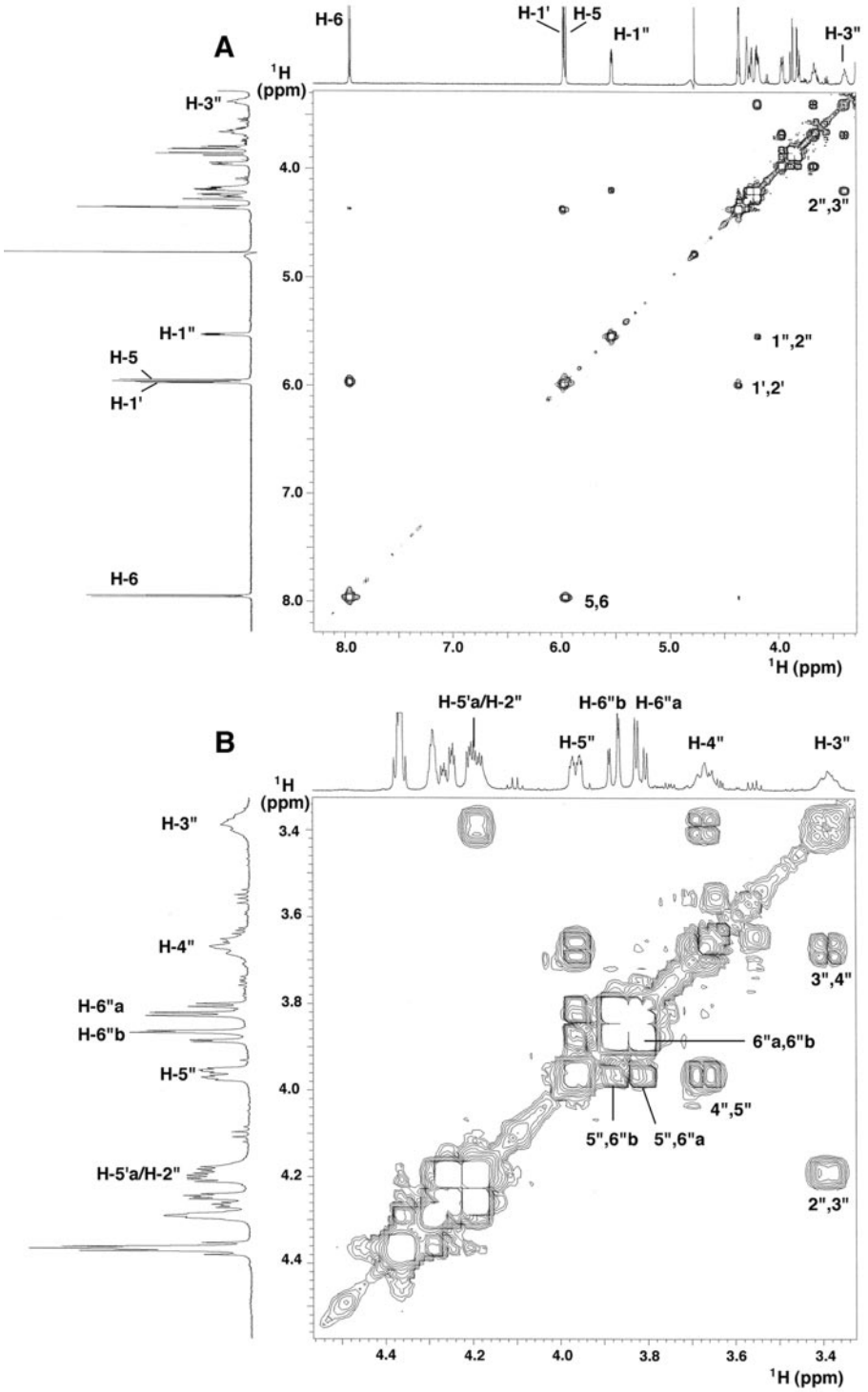



FIG. 8. **$^1$H-$^1$H COSY NMR analysis of UDP-GlcNAc3N.** This two-dimensional spectrum was recorded at 600 MHz on a ~1 mM sample at pD 7.8. *A* shows the sugar and nucleotide regions. *B* is a partial expanded spectrum highlighting the sugar resonances. Relevant cross-peaks are labeled as in Fig. 6.

shown), but product formation is slower. The rate of formation of Compound A by purified GnnA and GnnB is relatively constant between pH 6.0 and pH 10.0.

Transamination reactions generally require pyridoxal phosphate, but the coupled GnnA/GnnB reaction observed in crude extracts (Fig. 2) does not demonstrate any stimulation by added cofactor. When the purified GnnB protein is dialyzed in the absence of glycerol, however, pyridoxal phosphate must be added back to recover activity (data not shown).

*The GnnB Mutant K190A Is Inactive*—The predicted N-terminal cofactor-binding domain of GnnB is homologous to the γ-subfamily of the aspartate aminotransferase class of pyridoxal phosphate-dependent enzymes (33), which includes certain cysteine metabolizing enzymes. X-ray crystallography has shown that the cofactor-binding residue at the active site of cystathionine β-lyase is lysine 210 (34), which corresponds to lysine 190 in GnnB, the only absolutely conserved lysine in all GnnB orthologs. The K190A mutant of GnnB is completely inactive (data not shown).

*Purification of Compound A*—Several ion exchange chromatography steps were necessary to purify mg quantities of Compound A for NMR characterization. The first DEAE-cellulose column, run at pH 6, separated Compound A from the substrate UDP-GlcNAc, but not from residual $NAD^+$. Another DEAE-cellulose column was therefore run at pH 8.5, above the pK of the free amine of Compound A. Under these conditions,

TABLE III
*$^{13}C$ NMR assignments for UDP-GlcNAc and the putative UDP-GlcNAc3N*

The values shown are for $^{13}C$ chemical shifts at 25 °C in $D_2O$ relative to internal 2,2-dimethylsilapentane-5-sulfonic acid.

| Moiety | δC | | |
|---|---|---|---|
| | UDP-GlcNAc[a] (pD 7.8) | UDP-GlcNAc3N[b] | |
| | | pD 7.8 | pD 4.0 |
| Uracil | | | |
| C-2 | 154.451 [s] | ~154.4 | ~154.4 |
| C-4 | 168.902 [s] | ~168.9 | ~168.9 |
| C-5 | 105.203 [d] | ~105.2 | ~105.2 |
| C-6 | 144.161 [d] | ~144.2 | ~144.2 |
| Ribose | | | |
| C-1′ | 90.989 [d] | ~91.0 | ~92.0 |
| C-2′ | 76.319 [d] | ~76.3 | ~76.3 |
| C-3′ | 72.038 [d] | ~72.2 | ~72.2 |
| C-4′ | 85.757 [dd,9.08][c] | ~85.7 | ~85.8 |
| C-5′ | 67.490 [dt,5.88][c] | ~67.7 | ~67.8 |
| Pyranose | | | |
| C-1″ | 97.041 [dd,6.41][c] | ~96.4 | ~97.0 |
| C-2″ | 56.205 [dd,8.54][c] | ~54.0 | ~53.0 |
| C-3″ | 73.493 [d] | ~56.4 | ~57.0 |
| C-4″ | 72.187 [d] | ~69.8 | ~68.3 |
| C-5″ | 75.536 [d] | ~75.4 | ~75.4 |
| C-6″ | 62.856 [t] | ~62.8 | ~62.8 |
| Acetyl | | | |
| $CH_3$ | 24.610 [q] | ~24.5 | ~24.5 |
| C=O | 177.315 [s] | ~177.3 | ~177.8 |

[a] [(mult), *J*(Hz)]. Shift and coupling data obtained from natural abundance one-dimensional $^{13}C$ spectrum at high digitization.

[b] Shift data estimated from heteronuclear multi-quantum correlation two-dimensional map at coarse digitization.

[c] Carbon-phosphorus coupling constant obtained from one-dimensional spectrum at high digitization.

Compound A was well separated from $NAD^+$ and eluted at about the same salt concentration as UDP-GlcNAc, which had already removed by the first column. This two-step process generated a highly purified triethylammonium salt of Compound A. Excess triethylamine was removed using a Dowex AG-50W 8X column in the sodium form. Capillary electrophoresis was used to validate the purity (Fig. 5).

*Evaluation of the Structure of Compound A by One-dimensional and Two-dimensional $^1H$ NMR Spectroscopy*—Full one-dimensional $^1H$ NMR spectra of the substrate, UDP-GlcNAc at pD 7.8, and of the putative UDP-GlcNAc3N (*Compound A*) at pD 7.8 and 4.0, are shown in Supplementary Figs. 1–3. Key chemical shifts and coupling constants are summarized in Table II. The numbering scheme is shown in Fig. 6. The results demonstrate the presence of the same uracil and ribose moieties in both compounds. The spectra of both compounds also contain the characteristic methyl singlet resonance near 2.1 ppm from the 2″ *N*-acetyl group (Table II). Peaks from residual triethylamine are evident in the UDP-GlcNAc3N sample, indicated with *x* (Fig. 7, *top panel*, and Supplementary Figs. 2 and 3).

The $^1H$ NMR signals that arise from the pyranose sugar differ in UDP-GlcNAc and the putative UDP-GlcNAc3N (Figs. 7 and 8 and Table II). In UDP-GlcNAc at pD 7.8 (Supplementary Fig. 1; Fig. 7, *bottom panel*; and Table II), the anomeric H-1″ double-doublet (arising from the H-1″-H-2″ and the H-1″-P couplings) and the double-triplet from H-2″ resonate near 5.52 and 3.99 ppm, respectively. In the putative UDP-GlcNAc3N at pD 7.8 (Supplementary Fig. 2; Fig. 7, *middle panel*; and Table II), both of these signals appear slightly downfield at 5.54 and 4.21 ppm, respectively. When the pD is lowered to 4.0, both H-1″ and H-2″ of UDP-GlcNAc3N resonate even further downfield (Supplementary Fig. 3; Fig. 7, *top panel*; and Table II). These proton assignments, deduced both from standards and two-dimensional COSY analysis (see below), also confirm that H-4″ resonates *upfield* from H-3″ in the spectrum of UDP-GlcNAc (Fig. 7, *bottom panel*, and Table II) but *downfield* of H-3″ in UDP-GlcNAc3N (Fig. 7, *middle* and *top panels*; Fig. 8; and Table II).

Although the spectra for UDP-GlcNAc3N at pD 4.0 (Fig. 7, *top panel*) and UDP-GlcNAc at pD 7.8 (Fig. 7, *bottom panel*) display sharp H-3″ and H-4″ signals, the spectrum for UDP-GlcNAc3N at pD 7.8 (Fig. 7, *middle panel*) shows well resolved but relatively broad H-3″ and H-4″ signals. The H-2″ signal also appears to be broadened, but it overlaps with the ribose H-5′a signal. At both pD 7.8 and 4.0, however, the other pyranose protons of UDP-GlcNAc3N all yield sharp multiplets with similar line shapes. At pD 7.8, the UDP-GlcNAc3N would be near the expected $pK_a$ of a primary amine group on a pyranose (35). Consequently, the broad H-3″ and H-4″ signals at pD 7.8 may reflect the presence of both protonated and unprotonated forms of the 3-amino group of UDP-GlcNAc3N and exchange between the protonated and deprotonated forms. At pD 4.0 the UDP-GlcNAc3N primary amine would be fully protonated, and H-2″, H-3″ and H-4″ should yield sharp, resolved multiplets (Fig. 7, *top panel*).

*Further Analysis of the Proposed UDP-GlcNAc3N by Two-dimensional $^1H$ NMR Spectroscopy*—Although the H-3″ and H-4″ multiplets of the presumed UDP-GlcNAc3N are broader at pD 7.8 than at pD 4.0 (Fig. 7), they are sufficiently narrow to yield clear two-dimensional correlations (Fig. 8). Thus, the COSY of the UDP-GlcNAc3N sample at pD 7.8 (Fig. 8*A* and Table II) reveals a distinct correlation from H-1″ (5.54 ppm; dd, $J_{1'',2''} = 3.3$, $J_{H1'',P} = 6.9$ Hz) to H-2″ at 4.21 ppm (dt, $J_{1'',2''} = 3.2$, $J_{2'',3''} = 9.8$, $J_{H2'',P}$ not measured because of spectral overlap). At this shift position, H-2″ overlaps the ribose H-5′a multiplet (Fig. 8*B*) and is shifted downfield by ≈0.2 ppm compared with H-2″ in the spectrum of UDP-GlcNAc (Table II). The second cross-peak from H-2″ locates H-3″ at 3.39 ppm (broadened multiplet, estimated $J_{2'',3''} = 9.8$ Hz) (Fig. 8*B*). This position is 0.5 ppm upfield relative to H-3″ in the spectrum of UDP-GlcNAc (Table II). H-3″ of the putative UDP-GlcNAc3N in turn connects to H-4″ at 3.67 ppm (broadened multiplet) (Fig. 8*B*), which is about 0.2 ppm downfield relative to H-4″ in UDP-GlcNAc. Further tracing of the COSY cross-peaks locates H-5″ (3.97 ppm; ddd, $J_{4'',5''} = 10.0$, $J_{5'',6''b} = 2.2$, $J_{5'',6''a} = 4.2$, Hz), H-6″b (3.88 ppm; dd, $J_{5'',6''b} = 2.2$, $J_{6''b,6''a} = 12.6$ Hz), and H-6″a (3.82 ppm; dd, $J_{5'',6''a} = 4.2$, $J_{6''b,6''a} = 12.6$ Hz). The chemical shifts of H-5″, H-6″b, and H-6″a of the putative UDP-GlcNAc3N are almost identical to those of UDP-GlcNAc (Table II).

COSY analysis of the UDP-GlcNAc3N sample at pD 4.0 (Supplementary Fig. 4) gave similar results and allowed better estimates of the coupling constants from the sharp and well resolved H-2″, H-3″, and H-4″ multiplets (Fig. 7 and Table II). Comparison of the pD 4.0 and 7.8 data reveals that protonation of the 3-amino group shifted all the glucosamine protons downfield relative to their positions at pD 7.8, whereas the ribose and uracil proton shifts were unchanged (Table II). The small $J_{1'',2''}$ couplings (3.3 Hz) and the large $J_{2'',3''}$ couplings (9–11 Hz) indicate that the glucosamine rings of UDP-GlcNAc3N and UDP-GlcNAc are both in the α-anomeric configuration with axially disposed H-2″ and H-3″ protons. The large values of the $J_{3'',4''}$ and $J_{4'',5''}$ couplings (9–10 Hz) further indicate that H-4″ and H-5″ are also in axial positions.

*Evaluation of the Carbon Structure of UDP-GlcNAc3N*—The $^{13}C$ data of the presumed UDP-GlcNAc3N at ~1 mM (Table III) were obtained indirectly through $^1H$ -detected HMQC and heteronuclear multi-bond correlation two-dimensional NMR experiments. The $^{13}C$ data for UDP-GlcNAc at 10 mM (Table III) were obtained both directly by recording the one-dimensional $^1H$ -decoupled $^{13}C$ NMR spectrum and indirectly by the $^1H$-detected two-dimensional NMR analysis (data not shown). As

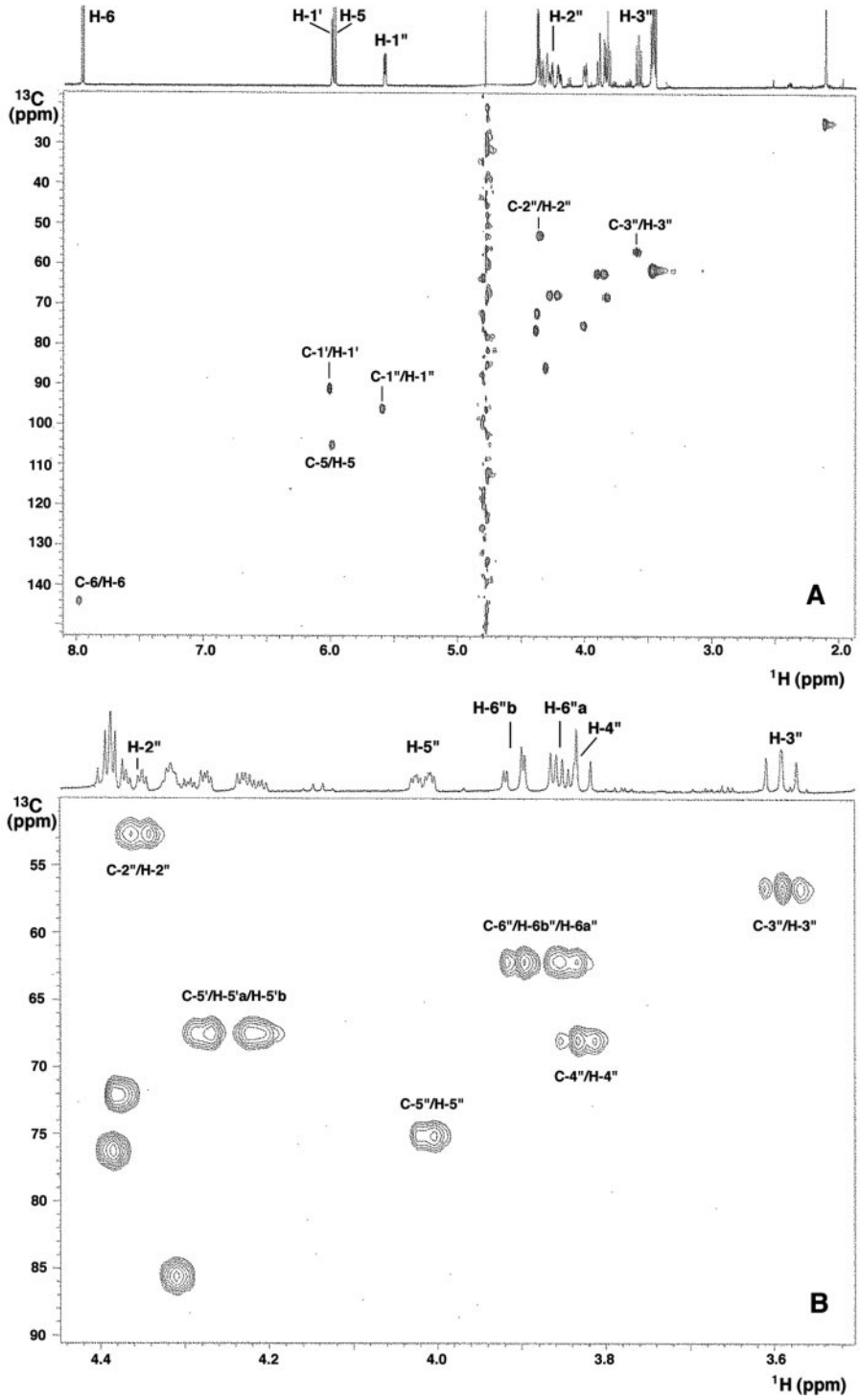



FIG. 9. **Two-dimensional HMQC NMR analysis of UDP-GlcNAc3N.** This spectrum was recorded at 600 MHz on a ~1 mM sample at pD 4.0. *A* is the complete heteronuclear multi-quantum correlation. *B* is a partial expanded spectrum highlighting the sugar region. Coupling of the 2″ and 3″ protons of UDP-GlcNAc3N to inferred carbon signals near 53 and 57 ppm, respectively, is diagnostic of nitrogen-linked carbon atoms. Relevant cross-peaks are labeled as in Fig. 6.

shown for UDP-GlcNAc3N at pD 4.0 (Fig. 9*A*), the ribose H-1′ proton signal (≈5.99 ppm) correlates to the anomeric carbon signal near 92 ppm (C-1′). The anomeric H-1″ of the pyranose correlates to the carbon resonance near 97 ppm (C-1″), consistent with an axially disposed oxygen atom and an equatorially disposed H-1″ (36). The uracil H-5 (≈5.97 ppm) correlates to the 105 ppm carbon signal. In addition, the uracil H-6 (7.96 ppm) correlates to a 144.2-ppm carbon signal, and the 2″ *N*-acetyl methyl protons correlate to a carbon signal near 24.5 ppm (Fig. 9*A* and Table III).

Scrutiny of the heteronuclear multi-quantum correlation data in the sugar region for the presumptive UDP-GlcNAc3N at pD 4.0 (Fig. 9*B*) shows that the H-6b″, H-6a″, and H-4″ protons correlate to oxygen-substituted carbon positions near 62.8 and 69.3 ppm (C-6″ and C-4″); however, in addition to the correlation of H-2″ to a nitrogen-substituted carbon near 53 ppm, H-3″ in the UDP-GlcNAc3N now correlates to a nitrogen-substituted carbon resonating near 57 ppm. This chemical shift is further diagnostic evidence for C-3″ as the site of the amino group substitution (36, 37). The remaining sugar cross-peaks of the UDP-GlcNAc3N arise from the ribose ring and are virtually identical to those for UDP-GlcNAc (Table III) (36, 38).

The heteronuclear multi-bond correlation data in the sugar region further validate the C-3″ and C-2″ assignments of UDP-GlcNAc3N at pD 4.0 (Supplementary Fig. 5). A multi-bond correlation is clearly observed from H-3″ to a 53-ppm carbon

resonance, identified as C-2″. Conversely, a multi-bond correlation from H-4″ in UDP-GlcNAc3N locates C-3″ at 57 ppm (Supplementary Fig. 5).

## DISCUSSION

A significant subset of evolutionarily diverse Gram-negative bacteria synthesize lipid A molecules in which one or both glucosamine residues are replaced with GlcN3N units (10–12). This substitution results in lipid A molecules containing three or four *N*-linked hydroxyacyl chains (Fig. 1), possibly providing stability during exposure of cells to elevated temperatures or extremes of pH. Prior to the present work, the enzymes required for the biosynthesis of GlcN3N substituted lipid A species were unknown. We have now demonstrated that the *A. ferrooxidans* proteins GnnA and GnnB, which are encoded by contiguous genes located between *lpxA* and *lpxB* in this organism (Supplementary Table I), catalyze the formation of the novel sugar nucleotide UDP-GlcNAc3N from UDP-GlcNAc (Scheme 1). The glucosamine 3-OH of UDP-GlcNAc is replaced with an amine in UDP-GlcNAc3N (Fig. 6). LpxA acyltransferases from bacteria that make lipid A molecules containing only GlcN3N units are highly selective for UDP-GlcNAc3N over UDP-GlcNAc (7, 19). Given that orthologs of the constitutive *E. coli* enzymes of lipid A biosynthesis (Scheme 1) are present in all of the bacteria that make GlcN3N substituted lipid A (Fig. 1), we propose that 3-*N*-acylated UDP-GlcNAc3N is converted to lipid A by the same reactions as in *E. coli* (Scheme 1). In this regard, *E. coli* LpxB efficiently utilizes synthetic lipid X analogs containing GlcN3N in place of glucosamine *in vitro* (39, 40). Wild-type *E. coli* cells do not have the *gnnA* and *gnnB* genes and therefore cannot make UDP-GlcNAc3N.

Purified GnnA, GnnB, L-glutamate, and $NAD^+$ are sufficient to support the conversion of UDP-GlcNAc to UDP-GlcNAc3N (Fig. 3). The NMR studies in Figs. 7–9 and in Supplementary Figs. 1–5 demonstrated unequivocally that a nitrogen atom is present at the pyranose 3-position of UDP-GlcNAc3N and that it is equatorial (Fig. 6). This means that UDP-GlcNAc3N is essentially isosteric with UDP-GlcNAc and may function as a substrate or inhibitor of many enzymes that utilize UDP-GlcNAc. Introduction of the GlcN3N unit in place of GlcNAc at selected sites in complex glycoconjugates could prove useful for cell surface modification or affinity labeling (41).

The homology of GnnA and GnnB to proteins of known function suggested that GnnA might be a $NAD^+$-dependent dehydrogenase and GnnB a pyridoxal phosphate-dependent aminotransferase. The simplest scenario is that GnnA oxidizes the glucosamine 3-position of UDP-GlcNAc to a ketone moiety, which is subsequently transaminated by GnnB in the presence of excess L-glutamate to generate UDP-GlcNAc3N (Scheme 1). The striking inhibition of the coupled GnnA/GnnB system by NADH or $\alpha$-ketoglutararate (Fig. 4) suggests that oxidation of the glucosamine 3-position of UDP-GlcNAc might be thermodynamically unfavorable, which would be consistent with the reduction potentials of $NAD^+$ and typical ketones (42). The proposed keto-sugar intermediate (Fig. 6) was not detected by TLC or capillary electrophoresis (27, 28) under our conditions, and purified GnnA was inactive in the absence of purified GnnB and excess L-glutamate, as judged by spectrophotometric detection of NADH formation at 340 nm (data not shown). To demonstrate the GnnA-dependent formation of the proposed 3″-ketone intermediate in the absence of added glutamate and GnnB (Scheme 1), it may be necessary to deplete the GnnA reaction system of NADH.

The lack of significant homology of GnnA to TylA2 is consistent with the idea that GnnA carries out a direct dehydrogenation of the 3″-position of UDP-GlcNAc. TylA2 is necessary for the modification of the pyranose 3-position of TDP-glucose, but it does so by first oxidizing the pyranose 4″-position (43, 44) and catalyzing the reductive elimination of the 6″-OH group. This step is followed by the action of the TylM3 isomerase, which converts the 4″-ketone to the 3″-ketone (15, 44). No ortholog of TylM3 is required for UDP-GlcNAc3N formation, given that purified GnnA and GnnB alone are sufficient (Fig. 3). We consider unlikely but cannot yet exclude the possibility that GnnA first catalyzes the oxidation of the glucosamine 4″-position, which is then isomerized (perhaps by GnnA itself) to the 3″-keto-sugar prior to transamination by GnnB.

All organisms known for certain to make lipid A substituted with GlcN3N units contain at least one close GnnB ortholog and a full-length GnnA ortholog. However, highly significant orthologs of GnnB display a much wider distribution among diverse bacterial genomes than do full-length GnnA orthologs. This reflects the existence of diverse sugar nucleotide transaminases involved in the biosynthesis of important bacterial amino sugars, such as mycaminose, perosamine, or L-4-aminoarabinose. For instance, the ArnB transaminase (45) that generates UDP-L-4-aminoarabinose in *E. coli* displays 36% identity and 53% similarity over 312 residues with an *E* value of $4 \times 10^{-44}$ when compared pairwise with GnnB (46). Consequently, a full-length ortholog of GnnA may be a better indicator of the presence of the UDP-GlcNAc3N pathway in an uncharacterized organism than an ortholog of GnnB. However, *Chloroflexus aurantiacus*, *Pyrococcus abyssi*, *Pyrococcus furiosus*, *Methanosarcina barkeri*, *Methanothermobacter thermautotrophicus*, and several other organisms that do not produce any lipid A nevertheless do possess full-length GnnA and GnnB (Supplementary Table I), suggesting that UDP-GlcNAc3N (or a related sugar nucleotide) may be utilized for the biosynthesis of alternative glycoconjugates, such as exopolysaccharides or capsules.

*Acknowledgment*—We thank Laurens Anderson for advice on nomenclature.

## REFERENCES

1. Roppel, J., and Mayer, H. (1975) *Carbohydr. Res.* **40,** 31–40
2. Keilich, G. (1976) *Carbohydr. Res.* **51,** 129–134
3. Yokota, A., Rodriguez, M., Yamada, Y., Imai, K., Borowiak, D., and Mayer, H. (1987) *Arch. Microbiol.* **149,** 106–111
4. Russa, R., Urbanik-Sypniewska, T., Lindström, K., and Mayer, H. (1995) *Arch. Microbiol.* **163,** 345–351
5. Zähringer, U., Knirel, Y. A., Lindner, B., Helbig, J. H., Sonesson, A., Marre, R., and Rietschel, E. T. (1995) *Prog. Clin. Biol. Res.* **392,** 113–139
6. Plötz, B. M., Lindner, B., Stetter, K. O., and Holst, O. (2000) *J. Biol. Chem.* **275,** 11222–11228
7. Que-Gewirth, N. L. S., Ribeiro, A. A., Kalb, S. R., Cotter, R. J., Bulach, D. M., Adler, B., Saint Girons, I., Werts, C., and Raetz, C. R. H. (2004) *J. Biol. Chem.* **279,** 25420–25429
8. Rawlings, D. E., Tributsch, H., and Hansford, G. S. (1999) *Microbiol.* **145,** 5–13
9. Ingledew, W. J. (1982) *Biochim. Biophys. Acta* **683,** 89–117
10. Weckesser, J., and Mayer, H. (1988) *FEMS Microbiol. Rev.* **4,** 143–153
11. Zähringer, U., Lindner, B., and Rietschel, E. T. (1999) in *Endotoxin in Health and Disease* (Brade, H., Opal, S. M., Vogel, S. N., and Morrison, D. C., eds) pp. 93–114, Marcel Dekker, Inc., New York
12. Raetz, C. R. H., and Whitfield, C. (2002) *Annu. Rev. Biochem.* **71,** 635–700
13. Okuda, S., Murata, S., and Suzuki, N. (1986) *Biochem. J.* **239,** 733–738
14. He, X. M., and Liu, H. W. (2002) *Annu. Rev. Biochem.* **71,** 701–754
15. Chen, H., Yeung, S., Que, N. L. S., Müller, T., Schmidt, R. R., and Liu, H. W. (1999) *J. Am. Chem. Soc.* **121,** 7166–7167
16. Crowell, D. N., Anderson, M. S., and Raetz, C. R. H. (1986) *J. Bacteriol.* **168,** 152–159
17. Coleman, J., and Raetz, C. R. H. (1988) *J. Bacteriol.* **170,** 1268–1274
18. 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
19. Sweet, C. R., Williams, A. H., Karbarz, M. J., Werts, C., Kalb, S. R., Cotter, R. J., and Raetz, C. R. H. (2004) *J. Biol. Chem.* **279,** 25411–25419
20. Sweet, C. R., and Raetz, C. R. H. (2001) *FASEB J.* **15,** A194
21. Silverman, M. P., and Lundgren, D. G. (1959) *J. Bacteriol.* **77,** 642–647
22. Miller, J. R. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
23. 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

24. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor, Cold Spring Harbor, NY
25. 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
26. Sweet, C. R., Lin, S., Cotter, R. J., and Raetz, C. R. H. (2001) *J. Biol. Chem.* **276,** 19565–19574
27. Creuzenet, C., Schur, M. J., Li, J., Wakarchuk, W. W., and Lam, J. S. (2000) *J. Biol. Chem.* **275,** 34873–34880
28. Creuzenet, C., Bélanger, M., Wakarchuk, W. W., and Lam, J. S. (2000) *J. Biol. Chem.* **275,** 19060–19067
29. Basu, S. S., Dotson, G. D., and Raetz, C. R. H. (2000) *Anal. Biochem.* **280,** 173–177
30. Breazeale, S. D., Ribeiro, A. A., and Raetz, C. R. H. (2002) *J. Biol. Chem.* **277,** 2886–2896
31. Tomasiewicz, H. G., and McHenry, C. S. (1987) *J. Bacteriol.* **169,** 5735–5744
32. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402
33. Jansonius, J. N. (1998) *Curr. Opin. Struct. Biol.* **8,** 759–769
34. Clausen, T., Huber, R., Laber, B., Pohlenz, H., and Messerschmidt, A. (1996) *J. Mol. Biol.* **262,** 202–224
35. Park, J. W., Choi, K.-H., and Park, K. K. (1983) *Bull. Korean Chem. Soc.* **4,** 68–72
36. Agrawal, P. K. (1992) *Phytochemistry* **31,** 3307–3330
37. van Halbeek, H. (1996) in *Encyclopedia of NMR* (Grant, D. M., and Harris, R. K., eds) Vol. 2, pp. 1107–1137, Wiley, Chichester, UK
38. Rosenthal, S. N., and Fendler, J. H. (1976) *Adv. Phys. Org. Chem.* **13,** 280–424
39. Haselberger, A., Hildebrandt, J., Lam, C., Liehl, E., Loibner, H., Macher, I., Rosenwirth, B., Schütze, E., Vyplel, H., and Unger, F. M. (1987) *Triangle* **26,** 33–49
40. Garrett, T. A., Kadrmas, J. L., and Raetz, C. R. H. (1997) *J. Biol. Chem.* **272,** 21855–21864
41. Mahal, L. K., Yarema, K. J., and Bertozzi, C. R. (1997) *Science* **276,** 1125–1128
42. Berg, J. M., Tymoczko, J. L., and Stryer, L. (2002) *Biochemistry*, 5th Edition, p. 495, W. H. Freeman and Co., New York
43. Hallis, T. M., and Liu, H.-W. (1998) *Acc. Chem. Res.* **32,** 579–588
44. Trefzer, A., Salas, J. A., and Bechthold, A. (1999) *Nat. Prod. Rep.* **16,** 283–299
45. Breazeale, S. D., Ribeiro, A. A., and Raetz, C. R. H. (2003) *J. Biol. Chem.* **279,** 24731–24739
46. Tatusova, T. A., and Madden, T. L. (1999) *FEMS Microbiol. Lett.* **174,** 247–250
47. Moran, A. P., Zähringer, U., Seydel, U., Scholz, D., Stütz, P. l., and Rietschel, E. T. (1991) *Eur. J. Biochem.* **198,** 459–469
48. Sweet, C. R. (2002) *Novel Substrates of Sugar Nucleotide Acyltransferases that Initiate Lipid A Biosynthesis*. Ph.D. thesis, Duke University, Durham, NC