Biosynthesis of a new UDP-sugar, UDP-2-Acetamido-2-deoxy-xylose, in the human pathogen *Bacillus cereus* subsp. cytotoxis NVH 391-98

Xiaogang Gu[1], John Glushka[1], Sung G Lee[1] and Maor Bar-Peled[1,2,3]

[1]Complex Carbohydrate Research Center (CCRC); [2]Dept. of Plant Biology; University of Georgia, Athens, GA 30602

[3]Address correspondences to: Maor Bar-Peled, CCRC 315 Riverbend Rd., Athens, GA 30602 Fax: 706-542-4412; Email: peled@ccrc.uga.edu

We have identified an operon and characterized the functions of two genes from the severe food poisoning bacterium, *Bacillus cereus* subsp. cytotoxis NVH 391-98 that are involved in the synthesis of a unique UDP-sugar, UDP-2-Acetamido-2-deoxy-xylose (UDP-*N*-acetyl-xylosamine, UDP-XylNAc). *UGlcNAcDH* encodes a UDP-*N*-acetyl-glucosamine 6-dehydrogenase converting UDP-*N*-acetyl-glucosamine (UDP-GlcNAc) to UDP-*N*-acetyl-glucosaminuronic acid (UDP-GlcNAcA). The second gene in the operon, *UXNAcS,* encodes a distinct decarboxylase not previously described in the literature, which catalyzes the formation of UDP-XylNAc from UDP-GlcNAcA in the presence of exogenous NAD$^+$. *UXNAcS* is specific and cannot utilize UDP-glucuronic acid and UDP-galacturonic acid as substrates. UXNAcS is active as a dimer with catalytic efficiency of 7 mM^{-1}s^{-1}. The activity of UXNAcS is completely abolished by NADH but unaffected by UDP-xylose. Real time NMR based assay showed unambiguously the dual enzymatic conversions of UDP-GlcNAc to UDP-GlcNAcA and subsequently to UDP-XylNAc. From the analyses of all publicly available sequenced genomes, it appears that UXNAcS is restricted to pathogenic Bacillus species including *B. anthracis* and *B. thuringiensis*. The identification of UXNAcS provides insight into the formation of UDP-XylNAc. Understanding the metabolic pathways involved in the utilization of this amino-sugar may allow the development of drugs to combat and eradicate the disease.

A food-poisoning outbreak in France in 1998 led to the isolation of a new strain of *Bacillus cereus* subsp. cytotoxis NVH 391-98 (1). This rod-shape Gram-positive bacterium causes a disease that initially produces emetic (nausea and vomiting) like symptoms, and/or in a more severe case produces a diarrheal form that causes abdominal cramps and diarrhea (2). Similar to its close relatives, the notorious human pathogen *Bacillus anthracis* and the insecticidal *Bacillus thuringiensis*, the *Bacillus cereus* strain can form spores. Due to its cell surface, a spore can survive harsh conditions (e.g. soil, air); and when the environment becomes appropriate, will germinate, resulting in a vegetative cell that can produce emetic toxin and different enterotoxins (3). The cell surface of many pathogenic bacteria is composed of diverse and complex carbohydrate structures, some of which are known virulence factors. Indeed, different *Bacillus cereus* peptidoglycans and glycoproteins were isolated and some were reported to play a role in spore formation and infection (3-5). It is also clear, however, that among those different types of Bacillus glycans, many are not fully characterized yet. Regardless, compared with the limited knowledge of these glycan structures, the pathways leading to their biosyntheses are still elusive. To identify such metabolic pathways we initially decided to look for putative genes encoding enzymes involved in the synthesis of glycan precursors (i.e. nucleotide-sugars).

Different UDP-GlcA decarboxylases with distinct functions exist in both eukaryotes and prokaryotes (Fig. 1A). UDP-xylose synthase (abbr. Uxs) from animals (6), plants (7), fungi (8) and bacteria (Gu and Bar-Peled, unpublished), for example, decarboxylates UDP-GlcA into UDP-xylose via an enzyme-bound NAD$^+$. A different decarboxylase in plants, Uaxs, converts UDP-GlcA in the presence of NAD$^+$ to both UDP-apiose and UDP-xylose (9). A bi-functional decarboxylase UDP-4-keto-pentose/UDP-xylose synthase (abbr. U4kpxs) from *Ralstonia solanacearum* str. GMI1000 converts UDP-GlcA

and NAD$^+$ to UDP-4-keto-pentose and then in the presence of NADH, turns UDP-4-keto-pentose to UDP-xylose (10). ArnA is also a decarboxylase, identified in *E. coli*, and it converts UDP-GlcA and NAD$^+$ to UDP-4-keto-pentose and NADH (11).

The above enzymes supply sugar precursors for the biosynthesis of various glycans, including proteoglycans in human and animals (6,12,13), glucuronoxylomannan in the pathogenic fungus *Cryprtococcus neofromans* (14-16), diverse plant polysaccharides such as xylan and xyloglucan (17,18), and different types of LPS in bacteria (11). We are interested in studying the evolution of nucleotide-sugar biosynthetic pathways across all species, as a way to understand the richness of diverse glycans and to evaluate how this diversity provides the specific organism an advantage for ecological adaptation. Here, we report the first identification and characterization of two genes (*UGlcNAcDH* and *UXNAcS*) involved in the biosynthesis of UDP-XylNAc (Fig. 1B) from *Bacillus cereus* subsp. cytotoxis NVH 391-98. The identification of these enzymes provides insight related to the formation of a new UDP-amino sugar and means to explore their roles within the lifecycle of this human pathogen. Hopefully this may lead to the eradication of the disease.

EXPERIMENTAL PROCEDURES

Cloning of UGlcNAcDH and UXNAcS from Bacillus cereus subsp. cytotoxis NVH 391-98- Genomic DNA isolated from *Bacillus cereus* subsp. cytotoxis NVH 391-98 was used as templates to clone the coding sequences of two genes in an operon predicted to be involved in nucleotide-sugars syntheses. The genes herein named UDP-GlcNAc 6-dehydrogenase (abbr. *UGlcNAcDH*) and UDP-XylNAc synthase (abbr. *UXNAcS*) were amplified by PCR using 1 unit of proof-reading Platinum Taq DNA polymerase high-fidelity (Invitrogen), 200 μM dNTPs and 0.2 μM of following primers: 5'-CCATGGAAAAAGAGAAAGGAGAAG-3' and 5'-GGATCCAAGCTTTGCACTCACCTTCTTTAG-3' for *UGlcNAcDH*; 5'-CCATGGCTAAGAAACGTTGTTTGATTACAGGTG-3' and 5'-

AAGCTTATCATTCTCTATTTCACGAAACCAC-3' for *UXNAcS*. The PCR products were isolated from agarose gel; cloned to generate plasmids pCR4:BcUGlcNAcDH#1 and pGEM-T:BcUXNAcS#1 respectively. The corresponding DNA sequences were deposited in GenBank with the following accession number: GU784842, (*UGlcNAcDH*) and GU784843 (*UXNAcS*). The Nco1-HindIII fragments of *UGlcNAcDH* (1273 bp) and *UXNAcS* (970 bp) were cloned into an *E. coli* expression vector to form pET28b:BcUGlcNAcDH#1 and pET28b:BcUXNAcS#1, respectively. The recombinant enzymes were designed to have a six histidine extension at their C-terminus to facilitate affinity-purification.

Protein expression and purification - E. coli cells containing pET28b:BcUGlcNAcDH#1, pET28b:BcUXNAcS#1 or an empty vector control (pET28b) were cultured for 16 h at 37 ℃ in 20 ml LB medium supplemented with kanamycin (50 μg/ml) and chloramphenicol (34 μg/ml). A portion (7 ml) of the cultured cells was transferred into fresh LB liquid medium (250 ml) supplemented with the same antibiotics, and the cells were then grown at 37 ℃ at 250 rpm until the cell density reached A_{600} = 0.8. The cultures were then transferred to 30 ℃ and gene expression was induced by the addition of isopropyl β-D-thiogalactoside to a final concentration of 0.5 mM. After 4 h growth while shaking (250 rpm), the cells were harvested by centrifugation (6,000 × g for 10 min at 4 ℃), re-suspended in lysis buffer (20 ml 50 mM sodium-phosphate, pH 7.5, containing 10% (v/v) glycerol, 1 mM EDTA, 1 mM DTT and 0.5 mM phenylmethylsulfonyl fluoride). Ammonium sulfate (50 mM) was also included for cells harboring *UGlcNAcDH* or empty vector. Cells were lysed in an ice bath by 24 sonication cycles (10-sec pulse; 20-sec rest) using a Misonix S-4000 (Misonix incorporated, Farmingdale, New York) equipped with 1/8'' microtip probe. The lysed cells were centrifuged at 4 ℃ for 10 min at 6,000 × g and the supernatant was supplemented with 1 mM DTT, and centrifuged again (30 min at 20,000 × g). The resulting supernatant (termed S20) was recovered and kept at -20 ℃.

His-tagged proteins were purified over a Ni-Sepharose fast flow column (2 ml resin (GE, NJ) packed in 10 mm id × 150 mm long column). The column was pre-equilibrated with 50 mM sodium-phosphate (pH 7.5), 0.1 M NaCl. The bound His-tagged protein was eluted with the same buffer containing increasing concentrations of imidazole. The fractions containing enzyme activity were pooled; supplemented with 1 mM DTT and 10% (v/v) glycerol, and dialyzed using 6,000-8,000 MWCO (Spectrum Laboratories, Inc, CA) tubing at 4^oC for 3X 30 min each with 800 ml dialysis buffer: 50 mM sodium-phosphate buffer (pH 7.5), 0.1 M NaCl, 10% (v/v) glycerol, 1 mM DTT. The purification and dialysis of UGlcNAcDH included 50 mM ammonia sulfate in all the buffers. The purified and dialyzed UXNAcS and UGlcNAcDH were flash frozen in liquid nitrogen and stored in aliquots at -80 ℃. Proteins extracted from *E. coli* cells expressing empty vector, were passed via the same Ni-column; and fractions eluted with imidazole were collected and served as controls in enzyme assays and SDS-PAGE analyses. The concentration of protein was determined using the Bradford reagent with bovine serum albumin (BSA) as standard. The estimated molecular weight of the active UXNAcS was determined by gel-filtration chromatography using Superdex-200 column (5).

UGlcNAcDH and UXNAcS enzyme assays and HPLC analysis– The UGlcNAcDH assays were conducted in a total volume of 50 μl consisting of 50 mM sodium phosphate (pH 7.6), 1 mM NAD^+, 1 mM UDP-GlcNAc (50 nmol), 50 mM ammonia sulfate, and 2 μl S20 fraction or purified recombinant UGlcNAcDH (1.5 μg).

The UXNAcS standard reaction assay, unless otherwise indicated, consisted of 50 mM sodium phosphate, pH 8.2, 0.4 mM NAD^+, 0.4 mM UDP-GlcNAcA (20 nmol) and 1.5 μg recombinant UXNAcS in final volume of 50 µl. Following incubation at 37 ℃ for up to 30 min for UGlcNAcDH and 15 min for UXNAcS, assays were terminated by heating (45 seconds at 80-100^oC). Chloroform (50 µl) was added and after vortexing (30 seconds) and centrifugation (12,000 rpm for 5 min, at room temperature), the entire upper aqueous phase was collected and subjected to chromatography on a Q15 anion-exchange column (4.6 mm id × 250 mm long, Amersham) using an Agilent Series 1100 HPLC system

equipped with an autosampler, diode-array detector and ChemStation software Ver. B.04.02. After sample injection, the Q15 column was washed at 1 ml min^{-1} for 5 min with 20 mM ammonium formate and subsequently nucleotides were eluted with a linear 20 to 600 mM ammonium formate gradient formed over 25 min. Nucleotides were detected by their UV absorbance. The maximum absorbance for UDP-sugars, NAD^+ and NADH were 261, 259, and 259/340 nm respectively in ammonium formate. The peak area of analytes was determined based on standard calibration curves. The peak corresponding to UDP-GlcNAcA (eluted at 21.4 min) or to UDP-XylNAc (eluted at 15.1 min) was collected, lyophilized, resuspended in D_2O, and further analyzed by NMR.

MALDI-MS analysis - Negative-ion MALDI mass spectra were recorded using a Microflex LT mass spectrometer (Bruker Daltonik, Bremen, Germany). Aqueous samples (1 µl of 1 mM UDP-GlcNAcA or UDP-XylNAc) were mixed with an equal volume of matrix solution (1 μg/µl 2,5-dihydroxybenzoic acid in 50% methanol) and dried on the plate. Spectra from 500 laser (N_2, 337nm) shots were summed to generate a mass spectrum.

^{1}H-NMR analyses –1D proton NMR spectrum of the HPLC-purified UDP-GlcNAcA peak (lyophilized, and resuspended in 250 µl D_2O buffered with 50 mM sodium phosphate pH/pD 7.6) was collected at 37^oC using a Varian DirectDrive^{TM} 600 MHz spectrometer equipped with a cryogenic probe.

The 1D proton and 2D HSQC NMR spectra of the HPLC-purified UDP-XylNAc peak (lyophilized, and resuspended in 250 µl D_2O buffered with 50 mM sodium phosphate pH/pD 8.2) were collected at 25°C using a Varian DirectDrive^{TM} 900 MHz spectrometer. The HSQC was collected with a carbon spectral width of 80 ppm, centered at 80.1 ppm, and with 64 points of 48 transients each.

The 250 µl real-time ^{1}H-NMR based enzymatic reactions were performed at 37^oC in a final mixture of D_2O:H_2O (1:9 v/v). The UGlcNAcDH assay consists of 50 mM sodium phosphate, pH/pD 7.6, 1 mM NAD^+, 1 mM UDP-GlcNAc and 10 µl S20 UGlcNAcDH. The UXNAcS reaction contains 50 mM sodium phosphate, pH/pD 8.2, 1 mM NAD^+, 1 mM UDP-

3

GlcNAcA and 7.5 µg recombinant UXNAcS. The dual-enzyme assay was performed in 50 mM sodium phosphate, pH/pD 7.6, supplemented with 2 mM NAD$^+$, 1 mM UDP-GlcNAc, 50 mM ammonia sulfate, 20 µl S20 UGlcNAcDH and 3 µg recombinant UXNAcS. Real-time ^{1}H-NMR spectra were obtained using the Varian DirectDrive^{TM} 600 MHz spectrometer. Data acquisition was started approximately 3 minutes after the addition of enzyme in order to optimize the spectrometer. Sequential 1D proton spectra with pre-saturation of the water resonance were acquired over the course of the enzymatic reaction. All NMR spectra were referenced to the resonance of 2,2-dimethyl-2-silapentane-5-sulphonate (DSS).

Synthesis of UDP-GlcNAcA - UDP-GlcNAcA was produced in a total volume of 5 ml consisting of 50 mM sodium phosphate, pH 7.6, 1 mM NAD$^+$, 1.5 mM UDP-GlcNAc and 200 µl S20 fraction of UGlcNAcDH. Reactions were kept at 37 ºC for up to 2 hours and terminated by heat (2 min at 80-100°C), and after chloroform extraction, was injected into Q15-HPLC column. The peak corresponding to UDP-GlcNAcA was collected, lyophilized, resuspended in H$_2$O, and an aliquot was re-analyzed by HPLC and 600MHz ^{1}H-NMR to check its identity and purity. The amount of UDP-GlcNAcA produced was calculated by comparing the HPLC chromatogram peak area with calibration curves of standards (UDP-Glc and UDP-GlcA).

Characterization and kinetic analyses of recombinant UXNAcS - The enzyme activity was tested in a variety of buffers, at different temperatures, or with different potential inhibitors. For optimal pH experiments, 1.5 µg recombinant enzyme was first mixed with 50 mM of each individual buffer (Tris-HCl or sodium phosphate). NAD$^+$ (0.4 mM) and UDP-GlcNAcA (0.4 mM) were then added and after 15 min incubation at 37°C, the amount of UDP-XylNAc formed was determined by HPLC. Inhibition assays were performed by first mixing the enzyme and sodium phosphate buffer with various additives (e.g. nucleotides) on ice for 10 min. UDP-GlcNAcA and NAD$^+$ (0.4 mM each) were then added. After 15 min at 37°C, the amount of UDP-XylNAc formed was calculated from the HPLC chromatogram. For the optimal temperature experiments, assays were performed under standard conditions except that reactions were incubated at different temperature for 15 min. Subsequently, the activities were terminated and the amount of UDP-XylNAc produced was measured by HPLC.

The catalytic activity of UXNAcS was determined at 37 ºC for 8 min in 50 mM sodium-phosphate pH 8.2, with variable concentrations of UDP-GlcNAcA (0.08 to 1 mM), 1 mM NAD$^+$ and 0.8 µg protein. The reciprocal initial velocity was plotted against the reciprocal UDP-GlcNAcA concentration according to Lineweaver and Burk to calculate K_m value. Enzyme kinetic was linear with respect to the reaction time and to the protein amount.

RESULTS

Identification of UXNAcS and UGlcNAcDH in Bacillus cereus subsp. cytotoxis NVH 391-98 – Glycan diversity can be generated by the ability of an organism to: 1) synthesize new and different types of NDP-sugars; 2) attach the sugar to different molecules in a specific linkage and configuration; 3) modify the glycan with different acetyl, methyl or amino-residues; and 4) link the oligosaccharides to specific lipid, protein molecules, or to large capsule of polysaccharides. A major obstacle in studying the metabolic pathways and diversity of glycans across species, especially with pathogenic microbes, is the need to work with live infectious organisms. The available genomes of various species thus, provide an opportunity to explore novel genes involved in glycan metabolic pathways as a way to study glycan diversity.

The characteristic signatures of proteins belonging to the large family of short chain dehydrogenase / reductase (SDR), such as UDP-GlcA 4-epimerase (19), TDP-glucose 4,6-dehydratse (20,21), and UDP-GlcA decarboxylase (8,10) are the GxxGxxG motif for cofactor NAD$^+$ binding (22), and the catalytic triad S/Y/K amino acids commonly found in the SE and YxxxK motifs at the catalytic pocket (23). To identify new enzyme activities belonging to the large SDR family, we first did a BLAST search of known proteins against various protein databases. Selected bacterial candidate proteins were further re-examined by 1) their sequence identities to

other proteins, and 2) analyses of the genes in the operon in which they were found.

BLAST analysis using AtUxs3 (AF387789) led to the identification of a gene in the *Bacillus cereus* subsp. cytotoxis NVH 391-98 chromosome, encoding a protein (we later named it UXNAcS) that shared relatively low amino acid sequence identity (Fig. S1 in the supplemental section) to known decarboxylases: 32% identity to Sinorhizobium SmUxs1 (GU062741), and 29% identity to the Ralstonia RsU4kpxs (GQ369438). The operon containing *UXNAcS* also consists of a gene encoding a putative UDP/GDP-sugar dehydrogenase (that we later named UGlcNAcDH).

To validate our approach and to identify the functions of both enzymes, their genes were cloned, expressed in *E. coli* and the function of the individual recombinant enzyme was studied.

UGlcNAcDH encodes active UDP-GlcNAc 6-dehydrogenase- Compared with control, a highly expressed recombinant protein band (47 kDa) was isolated and purified from *E. coli* cells expressing UGlcNAcDH (Fig. 2A, lane 1, 3). Preliminary HPLC-based assays revealed the enzyme couldn't use UDP-glucose as a substrate. However, the recombinant *B. cereus* enzyme did convert NAD^+ and UDP-GlcNAc to NADH (Fig. 2B, panel 5, marked by asterisk) and a new UDP-sugar that eluted from the column at 21.4 min (Fig. 2B, panel 5, marked by arrow). The identity of NADH was confirmed by its dual UV absorbance (259 and 340 nm) and its HPLC retention time when compared to the standard. The peak (21.4 min) was collected; analyzed by MALDI-MS, and had a mass of 621.2 (Fig. S2 panel 1 in the supplemental section) that would be expected for a UDP-hexo-*N*-acetylamino-uronic acid. Further analysis of the eluted peak using proton-NMR (Fig. 3) provided chemical shifts that are consistent with UDP-GlcNAcA (Table S1 in the supplemental section). We thus proposed that this Bacillus gene is active as UDP-GlcNAc 6-dehydrogenase, hence its name *UGlcNAcDH*. The Bacillus UGlcNAcDH is a very specific dehydrogenase and in our analyses, no activity was observed with UDP-glucose, UDP-galactose and UDP-*N*-acetyl-galactosamine as substrates (data not shown). HPLC peak integration indicates that the dehydrogenase requires around 2 mol of

NAD^+ to convert 1 mol of UDP-GlcNAc into 1 mol of UDP-GlcNAcA and 2 mol of NADH.

Our initial enzymatic characterization of the Bacillus UGlcNAcDH indicates that the enzyme has identical activity with a previously described UGlcNAcDH (WpbA*)* isolated from *Pseudomonas aeruginosa* (24). Therefore, we decided to concentrate our effort in the characterization of the other Bacillus gene product (UXNAcS) in that operon which has no homologous protein in Pseudomonas.

Identification of a new enzyme activity, UDP-XylNAc synthase- As shown in Fig. 4A, the recombinant Bacillus UXNAcS (37 kDa) was expressed in *E. coli* and subsequently column-purified. Initial analyses in the presence of exogenous NAD^+ with different commercially available UDP-, TDP- or GDP-sugars revealed no activity (data not shown). The enzyme assays were then carried out using UDP-GlcNAcA, the product of UGlcNAcDH, as an alternative substrate. Crude or purified UXNAcS readily converted UDP-GlcNAcA to another UDP-sugar with the retention time of 15.1 min (Fig. 4B panel 2 and 3, marked by asterisk) and such conversion requires exogenous NAD^+ (Fig. 4B panel 4).

MALDI-MS analysis of the new product indicates that it has a mass of 577.1 that would be expected for a UDP-pento-*N*-acetylamino sugar (Fig. S2 panel 2). To elucidate the product identity, the HPLC peak was collected and analyzed by NMR. Both 1D proton (Fig. 5A) and 2D HSQC (Fig. 5B) NMR spectroscopy provide unambiguous evidence that the enzymatic product is UDP-2-Acetimido-2-deoxy-α-D-xylopyranoside (UDP-XylNAc).

The chemical shifts assignments for UDP-XylNAc are summarized in Table S1. The diagnostic $J_{1",2"}$ value of 3.3 Hz and $J_{2",3"}$, $J_{3",4"}$, $J_{5b",4"}$, $J_{5a",5b"}$ values of 9.6, 10, 10.2, 11.2 Hz respectively, indicates an α-xylo configuration. The linkage of the anomeric XylNAc residue with the phosphate is given by the coupling constant values of 6.8 and 3.0 Hz for $J_{1",P}$ and $J_{2",P}$ respectively, and is also supported by the distinct chemical shift of H"1 (5.454 ppm). In addition, the NAc-C=O resonance was assigned 2.064 ppm and the specific H"2 chemical shift (3.962 ppm) is consistent for a C"-2-acetamido moiety of UDP-2-acetamido-2-deoxy-α-D-xylo*pyranose* (UDP-XylNac).

5

Collectively, we thus concluded that *UXNAcS* encodes a newly identified specific UDP-GlcNAcA decarboxylase that forms UDP-XylNAc, hence the name *UXNAcS*.

Real-Time NMR analyses of UDP-GlcNAc dehydrogenase and UDP-XylNAc synthase activities - The real-time NMR assays were performed with each individual enzyme and its corresponding substrate as well as a combined assay that comprises both enzymes (the dehydrogenase and the decarboxylase). The latter assay was conducted to determine if any intermediates could be observed. The UGlcNAcDH activity is shown in Fig. 6 panel 1. As time progresses, the peak corresponding to the UDP-GlcNAc anomeric proton (peak I) is decreased while peaks related with NADH (H1) and UDP-GlcNAcA (peak II) are increased, confirming the enzymatic conversion of UDP-GlcNAc and NAD$^+$ into UDP-GlcNAcA and NADH.

The UXNAcS assay is shown in Fig. 6 panel 2. As enzymatic reaction progresses, the diagnostic anomeric proton of UDP-GlcNAcA (peak II) is decreased and concomitantly an increase of UDP-XylNAc (peak III) is observed. It should be noted that no variation was observed for the peaks belonging to NAD$^+$ protons (N1, N2) (Fig. 6 panel 2) in the UXNAcS assay, suggesting that the enzyme-bound NAD$^+$ is not released as NADH during the C-4 oxidation (NAD$^+$→NADH)/ reduction (NADH→NAD$^+$) cycle.

The dual-enzyme assay including both Bacillus UGlcNAcDH and UXNAcS is shown in Fig. S3 in the supplemental section. Over the reaction time, as the dehydrogenase converts UDP-GlcNAc to UDP-GlcNAcA, the peaks corresponding to the anomeric proton of UDP-GlcNAc (peak I in Fig. S3 panel 2) were decreased, and those of UDP-GlcNAcA increase (Peak II); and simultaneously, an increase in peaks related to NADH proton (H1 in Fig. S3 panel 1) and the two diagnostic protons attached to C4 of the nicotinamide ring, H-4a,4b (Fig. S3 panel 3), were observed. As reaction progresses, the increase of UDP-XylNAc anomeric proton (peak III) is also observed (Fig. S3 panel 2), indicating the formation of UDP-XylNAc by UXNAcS. No other intermediate compounds were observed.

Characterization and properties of UXNAcS- The optimal pH and temperature for the activity of Bacillus UDP-XylNAc synthase is around 8.2 and 37°C (Fig. S4 in the supplemental section). The enzyme shows substantial activity at high pH as well, but no activity was observed below pH 4 (Fig. S4). The Bacillus UXNAcS was completely inhibited by NADH but not by its analog, NADPH (Table 1), suggesting the strict recognition of the enzyme for the C2-OH ribosyl group of the adenine moiety in NADH. UDP-xylose and UDP-GlcA had no adverse effects on UXNAcS activity. This might suggest that the UXNAcS recognition of the C"2-acetamido moiety of the UDP-XylNAc is critical. Other nucleotides, aside from UDP, or nucleotide-sugars (Table 1) had no significant effects on the decarboxylase activity. The enzyme is specific and did not react when NADP$^+$ was substituted for NAD$^+$, or when UDP-GlcNAcA was substituted with other UDP-uronates, such as UDP-GlcA and UDP-GalA (data not shown). Size-exclusion chromatography (Fig. S5 in the supplemental section) analyses suggest that the decarboxylase is active as a dimer since it elutes from the Superdex-200 gel filtration column with a mass corresponding to 70,200. Kinetics analyses of the UDP-XylNAc synthase activity are summarized in Table 2. The apparent K_m value was 67 μM, V_{max} value (μM s^{-1}) was 0.2, and k_{cat}/K_m value (s^{-1}mM^{-1}) was 7. These kinetic values are comparable with other bacterial decarboxylase enzyme such as RsU4kpxs (Table 2).

DISCUSSION

We have described the biosynthesis of a new UDP-amino sugar, UDP-XylNAc, and the cloning and characterization of a previously unknown decarboxylase activity. UXNAcS shares approximately 30% amino acid sequence identity with known UDP-GlcA decarboxylases; despite that low similarity, they appeared to maintain a conserved fold and most likely, similar catalytic mechanism throughout evolution. A putative structural model was created using human Uxs1 with bound NAD$^+$ and UDP as a template (PDB ID: 2b69.pdb) by the aid of the PHYRE (25) structural prediction program. These two enzymes share a highly conserved structural fold (See Fig. S6 in supplemental section) consists of a central

6

domain with structure resembling the Rossmann-fold. In addition to the conserved structure, the catalytic motifs and the NAD^+-binding domain, the predicted 3D-model of UXNAcS is almost indistinguishable from the human Uxs. However, the predicted UXNAcS model has some differences, specifically at region around amino acid 266-282 when compared with the human Uxs1 (Fig. S6). We speculate that this region is likely involved in the interaction with the C"-2 acetamido moiety. Based on the predicted UXNAcS model, another region appears different near amino acid 47-53. Obviously, pointing out the exact UXNAcS amino acids that are involved in ligands binding (the nature of NAD^+ association and the specificity for the acetamido group binding) and catalysis will require site directed mutagenesis and a real crystal structure.

The genomes of *Bacillus cereus* subsp. cytotoxis NVH 391-98, its closest pathogenic relatives *Bacillus anthracis,* and the insecticidal *Bacillus thuringiensis* all contain the same operon for UDP-XylNAc synthesis. Interestingly, the non-pathogen *Bacillus subtilis* does not carry this operon and based on the current genome sequences, it is likely that no other organism carries this UXNAcS gene. But, we cannot exclude the possibility that an organism may have UXNAcS homolog with a low sequence similarity to the Bacillus UXNAcS. Our work also illustrates that functional, biochemical analysis is essential; and that homology is an insufficient criterion to infer functional specificity.

In addition, several questions regarding the Bacillus UXNAcS can be raised: is there any specific advantage for those Bacillus species that harbor UXNAcS? Did other microbes originally have UXNAcS and then lose it? Was the UXNAcS gene in these Bacillus species evolved from a rapidly evolving Uxs-like gene? From an evolutionary point of view, the UDP-sugar decarboxylase protein family is currently divided into three separate clades: the Uaxs, the ArnA and U4kpxs, and the Uxs, all proposed to be derived from the same ancestral Uxs/Uaxs-like enzyme (10). Phylogeny analyses as shown in Fig S1B, places the Bacillus UXNAcS as a distinct clade from the Uxs. This suggests that the bacteria Uxs is unlikely the ancestral enzyme to the UXNAcS. How UXNAcS originated and what was led to the selective mutations resulting in the utilization of

the UDP-GlcNAcA as substrate, however, is still ambiguous.

Nature produced diverse glycans decorated with different amino sugars, such as 2-deoxy-hexosamines and 6-deoxy-hexosamines, that are common in both Gram-negative and Gram-positive bacteria species. Peptidoglycan for example, a critical bacterial cell wall polymer, is composed of alternating residues of β-1,4-linked N-acetylmuramic acid and GlcNAc (26,27), and the aminoglycoside-antibiotics produced by certain microbes contains different types of amino sugars as well (28). Amino sugar modified glycans are also common in eukaryotes. GalNAc-residue, for example, is a predominant glycosyl-residue in the core structure of O-linked glycans in human and certain animals (29-32). The metabolism leading to the synthesis of such diverse glycans requires specific NDP-amino sugar precursors and as such, malfunction in the activated-sugar biosynthesis pathways might be lethal. Some UDP-amino sugar may also serve as intermediates for the synthesis of other nucleotide-amino sugars, like UDP-GlcNAcA for the synthesis of UDP-ManNAc(3Ac)A, a proposed sugar-precursor for the synthesis of O-antigen in *Pseudomonas aeruginosa* (33). However, the role for UDP-XylNAc, specifically in this food-poisoning Bacillus strain, remains elusive and is currently under investigation using a non-human pathogen strain as a model.

XylNAc residue has not yet been observed in polysaccharides. The XylNac residue however, was used as a chemical substitution for GlcNAc residue in peptidoglycan for the study of the polysaccharide hydrolytic enzyme, lysozyme (34). Interestingly, the *Bacillus thuringiensis* peptidoglycan cell wall is recalcitrant to hydrolysis by lysozyme (35). Whether this is in part due to XylNAc residues in the Bacillus wall polysaccharide remained to be determined. On the other hand, there are limited reports of xylosamines in glycosides. For example, some diamino xylosides have been found in certain antibiotics (36,37), one of which was isolated from members of Nocardiopsis species (36). Whether the pathogenic Bacillus species also synthesize antimicrobial xylosamine-glycosides, however, is still unclear. Interestingly, analyses of *Nocardiopsis dassonvillei* subsp. dassonvillei DSM 43111 genome, with the Bacillus UXNAcS

7

and UGlcNAcDH proteins point to an operon containing several genes and among them are NdasDRAFT_3140 (ZP_04334040, we named "DH") and NdasDRAFT_3141 (ZP_04334041, we named "DC"). The Nocardiopsis "DH" and "DC" shares 27% amino acid sequence identity to the UGlcNAcDH and 38% identity to the UXNAcS described in this report, respectively. Despite the low sequence identities, it is possible that these Nocardiopsis enzymes are involved in the formation of xylosamine antibiotics such as, glycocinnamoylspermidines or macrolactam glycosides-like incednine. In addition, the organization of the *N. dassonvillei* operon is found also in genomes of several other members of the Nocardiopsis species. Within this Nocardiopsis operon and its surrounding gene clusters, there are several addition genes encoding potential glycosyltransferases, acyl-carrier protein, and other NDP-sugar biosynthetic enzymes. Whether these Nocardiopsis proteins are involved in the formation of these powerful antibiotics against Gram-negative bacteria, is currently under investigation. Taken together, the genes described in this report along with other surrounding Bacillus genes could be involved in production of xylosamine antimicrobial agents in this pathogenic bacterium.

In conclusion, this unique UXNAcS enzyme in pathogenic Bacillus species has not yet been found in any other organism and could be a potential drug-target to eradicate the devastating Anthrax and food poisoning diseases both in human and livestock.

FOOTNOTES

ABBREVIATIONS: 1D, one-dimensional; 2D, two-dimensional; UDP-GlcA, UDP-glucuronic acid, UDP-GlcNAc, UDP-*N*-acetyl-glucosamine; UDP-GlcNAcA, UDP-*N*-acetyl-glucosaminuronic acid; UDP-XylNAc, UDP-*N*-acetyl-xylosamine, UDP-α-D-2-acetamido-2-deoxy-xylose; UGlcNAcDH, *Bacillus cereus* UDP-GlcNAc 6-dehydrogenase; UXNAcS, *Bacillus cereus* UDP-XylNAc synthase; HPLC, high pressure liquid chromatography; DTT, Dithiothreitol; HSQC, heteronuclear single quantum coherence.

FIGURE LEGENDS

FIGURE 1. **NDP-uronate sugar decarboxylases in eukaryote and prokaryote.**
A. In animals, plants, some fungi and bacteria, UDP-GlcA is converted to UDP-xylose by UDP-xylose synthase (abbr., Uxs). Plants also have a bifunctional UDP-apiose /UDP-xylose synthase (abbr., Uaxs) that can inter-convert UDP-GlcA to UDP-apiose and UDP-xylose. In some bacteria, UDP-GlcA can be converted into UDP-4-keto-pentose in the presence of ArnA. Other bacteria may have a bifunctional UDP-4-keto-pentose/UDP-xylose synthase (abbr., U4kpxs) that first converts UDP-GlcA to UDP-4-keto-pentose and subsequently form UDP-xylose from UDP-4-keto-pentose, albeit at lower rate compared with the first activity.
B. The biosynthesis of UDP-XylNAc in *Bacillus cereus* subsp. cytotoxis NVH 391-98. In the presence of UGlcNAcDH, UDP-GlcNAc was first converted into UDP-GlcNAcA, which is further decarboxylated into UDP-XylNAc with the activity of UXNAcS. Similar operon organization is found in other Bacillus species such as *Bacillus anthracis* and *Bacillus thuringiensis*.

FIGURE 2. **Expression and characterization of recombinant UGlcNAcDH.**
A. SDS-PAGE of total soluble protein isolated from *E. coli* cells expressing UGlcNAcDH (lane 1) or control empty vector (lane 2), and of Nickel-column purified fractions (lane 3, UGlcNAcDH; lane 4 is control).
B. HPLC chromatogram of UGlcNAcDH enzyme reaction. Purified recombinant UGlcNAcDH was incubated with UDP-GlcNAc in the presence (panel 5), or absence (panel 7) of exogenous NAD^+. As a control, the corresponding column-purified protein isolated from cells expressing control empty vector was incubated with UDP-GlcNAc and NAD^+ (panel 6). Activity of total

protein isolated from cell expressing recombinant UGlcNAcDH or vector control is shown in panel 2 and 3, respectively. The reaction products were separated on Q15 column and the UDP-sugar peak (marked by arrow, in panel 2 and 5) were collected and analyzed by MALDI-MS and NMR.

FIGURE 3. **^{1}H-NMR analysis of the UGlcNAcDH enzymatic product, UDP-GlcNAcA**
The enzymatic product (Fig. 2B, panel 5, marked by arrow) was collected and analyzed by NMR at 37°C. 1D 600-MHz NMR spectrum of the product is shown. Location for each proton residue on the spectrum is indicated: the position of protons on the uracil (Ura) ring are indicated by H, the ribose (Rib) protons by H', the GlcNAcA protons by H", and the C"-2-linked acetamido by NAc-H. The inset shows an expansion of the 1D NMR spectrum from 3.5 to 4.5 ppm.

FIGURE 4. **Expression and characterization of recombinant UXNAcS**
A. SDS-PAGE of total soluble protein isolated from *E. coli* cells expressing UXNAcS (lane 2) or control empty vector (lane 3), and of Nickel-column purified fractions (lane 4, UXNAcS; lane 5 is control).
B. HPLC chromatogram of UXNAcS enzyme reaction. Purified recombinant UXNAcS was incubated with UDP-GlcNAcA in the presence (panel 3), or absence (panel 4) of exogenous NAD$^+$. As a control, the corresponding column-purified protein isolated from cells expressing control empty vector was incubated with UDP-GlcNAcA and NAD$^+$ (panel 6). Activity of total protein isolated from cell expressing recombinant UXNAcS or vector control is shown in panel 2 and 5, respectively. The reaction products were separated on Q15 column and the UDP-sugar peak (marked by asterisk, in panel 2 and 3) were collected separately and analyzed by MALDI-MS and NMR.

FIGURE 5. **1D and 2D ^{1}H-NMR analyses of the UXNAcS enzymatic product, UDP-XylNAc**
A. 1D 900 MHz NMR spectrum of the enzymatic product collected from HPLC column (see Fig. 4B, panel 2 and 3, marked by asterisk). The chemical shift of each proton residue on the spectrum is indicated: the position of protons on the uracil (U) ring are indicated by H, the ribose (R) protons by H', the XylNAc (X) protons by H", and the C"-2 linked acetamido protons by NAc-H. Peak marked as HDO is the residual water signal.
B. 2D 900 MHz NMR spectrum of UDP-XylNAc. Selected region of a carbon-proton HSQC shows the ring protons and carbons of XylNAc (X) and Ribose (R), with the corresponding region of the proton spectrum above. Peaks corresponding to column impurities are indicated by ***.

FIGURE 6. **Monitoring individual UGlcNAcDH and UXNAcS assay by Real-time 1D NMR.**
The single-enzyme assay of UGlcNAcDH is shown in panel 1; of UXNAcS in panel 2. All reactions were performed at 37°C. The selected regions for the diagnostic peaks for NAD$^+$, NADH, and the anomeric protons of the NDP-sugars, are shown between 5.4 and 6.2 ppm at 5 min or 60 min after each reaction was initiated. The labeling for NAD$^+$ and NADH protons referred to the corresponding compounds shown above. Peaks labeled as I, II or III indicate the anomeric protons of UDP-GlcNAc, UDP-GlcNAcA or UDP-XylNAc, respectively. Peaks marked by * (from 5.9 to 6.0 ppm) in panel 1 are the mixture of signals for Rib-H1' and Ura-H5 protons of UDP-GlcNAc and UDP-GlcNAcA. Peaks marked by ** (from 5.9 to 6.0 ppm) in panel 2 are the mixture of signals for Rib-H1' and Ura-H5 protons of UDP-GlcNAcA and UDP-XylNAc.

Table 1. *The effects of nucleotides and nucleotide sugars on UXNAcS activity.*

Additives	Relative activity of UXNAcS (%)[a]
Water	100 ± 1.4
UDP-glucose	99 ± 3.8
UDP-galactose	105 ± 2.4
UDP-xylose	101 ± 2.0
UDP-arabinose	95 ± 3.3
UDP-GlcNAc	103 ± 1.2
UDP-GalNAc	97 ± 0.9
UDP-GlcA	70 ± 2.0
UDP	48 ± 2.4
UMP	93 ± 2.1
NADP	97 ± 5.3
NADPH	92 ± 6.8
NADH	0

[a] Recombinant enzyme was incubated with 0.5 mM NAD^+ in the absence or presence of different 0.5 mM nucleotides and nucleotide sugars for 15 min prior the addition of 0.5 mM UDP-GlcNAcA. 100% activity corresponds to 15 nmol of UDP-XylNAc produced. The data presented are the average from three experiments.

Table 2. *Enzymatic kinetics of recombinant UXNAcS*

	$K_m{}^{a}$ (μM)	V_{max} (μM s^{-1})	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1}s^{-1})
UXNAcS	67 ± 6.0	0.2 ± 0.01	0.47 ± 0.02	7 ± 0.4
RsU4kpxs[b]	22 ± 1.2	0.07 ± 0.005	0.29 ± 0.02	13 ± 0.3

[a] UXNAcS activity was measured with varied concentrations of UDP-GlcNAcA (0.08–1.0 mM) and 1mM NAD$^+$, after 8 min at standard conditions. The reciprocal initial velocity was plotted against the reciprocal UDP-GlcNAcA concentration according to Lineweaver and Burk to calculate the corresponding K_m values. The data presented are the average K_m values from three experiments.

[b] The kinetic data for the decarboxlyase activity of RsU4kpxs were taken from (10).

REFERENCE

1. Lapidus, A., Goltsman, E., Auger, S., Galleron, N., Segurens, B., Dossat, C., Land, M. L., Broussolle, V., Brillard, J., Guinebretiere, M. H., Sanchis, V., Nguen-The, C., Lereclus, D., Richardson, P., Wincker, P., Weissenbach, J., Ehrlich, S. D., and Sorokin, A. (2008) *Chem Biol Interact* **171**, 236-249
2. Stenfors Arnesen, L. P., Fagerlund, A., and Granum, P. E. (2008) *FEMS Microbiol Rev* **32**, 579-606
3. Henriques, A. O., and Moran, C. P., Jr. (2007) *Annu Rev Microbiol* **61**, 555-588
4. Fox, A., Stewart, G. C., Waller, L. N., Fox, K. F., Harley, W. M., and Price, R. L. (2003) *J Microbiol Methods* **54**, 143-152
5. Kotiranta, A., Lounatmaa, K., and Haapasalo, M. (2000) *Microbes Infect* **2**, 189-198
6. Hwang, H. Y., and Horvitz, H. R. (2002) *Proc Natl Acad Sci U S A* **99**, 14218-14223
7. Harper, A. D., and Bar-Peled, M. (2002) *Plant Physiol* **130**, 2188-2198
8. Bar-Peled, M., Griffith, C. L., and Doering, T. L. (2001) *Proc Natl Acad Sci U S A* **98**, 12003-12008
9. Guyett, P., Glushka, J., Gu, X., and Bar-Peled, M. (2009) *Carbohydr Res* **344**, 1072-1078
10. Gu, X., Glushka, J., Yin, Y., Xu, Y., Denny, T., Smith, J. A., Jiang, Y., and Bar-Peled, M. *J Biol Chem*
11. Breazeale, S. D., Ribeiro, A. A., and Raetz, C. R. (2002) *J Biol Chem* **277**, 2886-2896
12. Gotting, C., Kuhn, J., Zahn, R., Brinkmann, T., and Kleesiek, K. (2000) *J Mol Biol* **304**, 517-528
13. Kuhn, J., Gotting, C., Schnolzer, M., Kempf, T., Brinkmann, T., and Kleesiek, K. (2001) *J Biol Chem* **276**, 4940-4947
14. Cherniak, R., and Sundstrom, J. B. (1994) *Infect Immun* **62**, 1507-1512
15. Vaishnav, V. V., Bacon, B. E., O'Neill, M., and Cherniak, R. (1998) *Carbohydr Res* **306**, 315-330
16. Klutts, J. S., and Doering, T. L. (2008) *J Biol Chem* **283**, 14327-14334
17. Mohnen, D. (2008) *Curr Opin Plant Biol* **11**, 266-277
18. Pena, M. J., Zhong, R., Zhou, G. K., Richardson, E. A., O'Neill, M. A., Darvill, A. G., York, W. S., and Ye, Z. H. (2007) *Plant Cell* **19**, 549-563
19. Gu, X., Wages, C. J., Davis, K. E., Guyett, P. J., and Bar-Peled, M. (2009) *J Biochem* **146**, 527-534
20. Allard, S. T., Giraud, M. F., Whitfield, C., Graninger, M., Messner, P., and Naismith, J. H. (2001) *J Mol Biol* **307**, 283-295
21. Gerratana, B., Cleland, W. W., and Frey, P. A. (2001) *Biochemistry* **40**, 9187-9195
22. Rao, S. T., and Rossmann, M. G. (1973) *J Mol Biol* **76**, 241-256
23. Weirenga, R. K., Terpstra, P., and Hol, W. G. J. (1986) *J Mol Biol* **187**, 7
24. Miller, W. L., Wenzel, C. Q., Daniels, C., Larocque, S., Brisson, J. R., and Lam, J. S. (2004) *J Biol Chem* **279**, 37551-37558
25. Kelley, L. A., and Sternberg, M. J. (2009) *Nat Protoc* **4**, 363-371
26. Glaser, L. (1973) *Annu Rev Biochem* **42**, 91-112

27. Vollmer, W. (2008) *FEMS Microbiol Rev* **32**, 287-306
28. Kudo, F., and Eguchi, T. (2009) *J Antibiot (Tokyo)* **62**, 471-481
29. Larsson, J. M., Karlsson, H., Sjovall, H., and Hansson, G. C. (2009) *Glycobiology* **19**, 756-766
30. Hanover, J. A. (2001) *FASEB J* **15**, 1865-1876
31. Ten Hagen, K. G., Fritz, T. A., and Tabak, L. A. (2003) *Glycobiology* **13**, 1R-16R
32. Kubota, T., Shiba, T., Sugioka, S., Furukawa, S., Sawaki, H., Kato, R., Wakatsuki, S., and Narimatsu, H. (2006) *J Mol Biol* **359**, 708-727
33. Larkin, A., and Imperiali, B. (2009) *Biochemistry* **48**, 5446-5455
34. Ballardie F, C. B., Cuthbert M and Dearie W. (1977) *Bioorganic Chemistry* **6**, 483-509
35. Temeyer, K. B. (1987) *Journal of General Microbiology* **133**, 503-506
36. Martin, J. H., Kunstmann, M. P., Barbatschi, F., Hertz, M., Ellestad, G. A., Dann, M., Redin, G. S., Dornbush, A. C., and Kuck, N. A. (1978) *J Antibiot (Tokyo)* **31**, 398-404
37. Takaishi, M., Kudo, F., and Eguchi, T. (2008) *Tetrahedron* **64**, 6651-6656
38. Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007) *Mol Biol Evol* **24**, 1596-1599

Figure 1

A.

UDP-xylose

UDP-xylose

Uxs

UDP-apiose ← UDP-GlcA → *U4kpxs* → UDP-4-keto-pentose → *U4kpxs* → UDP-xylose

Uaxs

ArnA

UDP-4-keto-pentose

B.

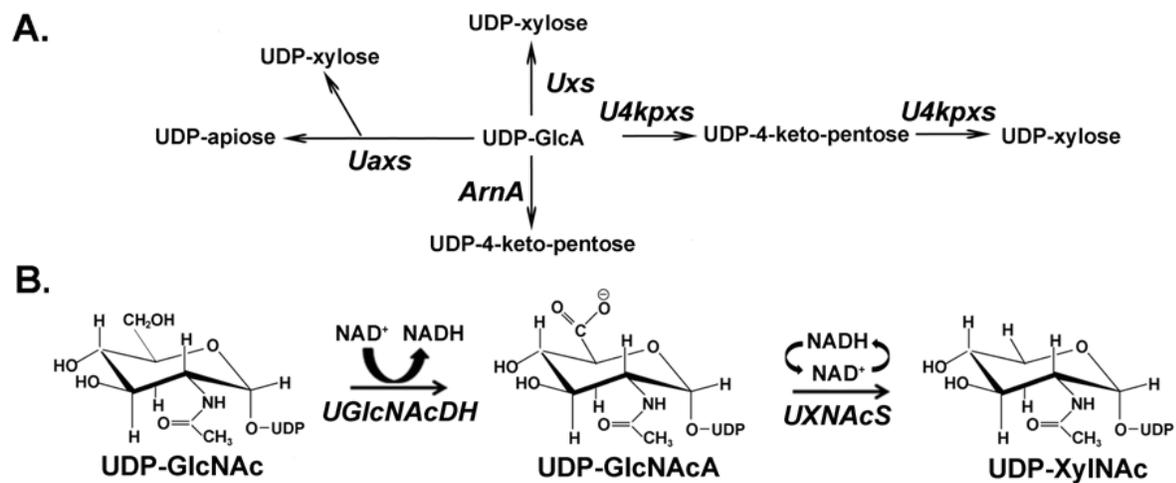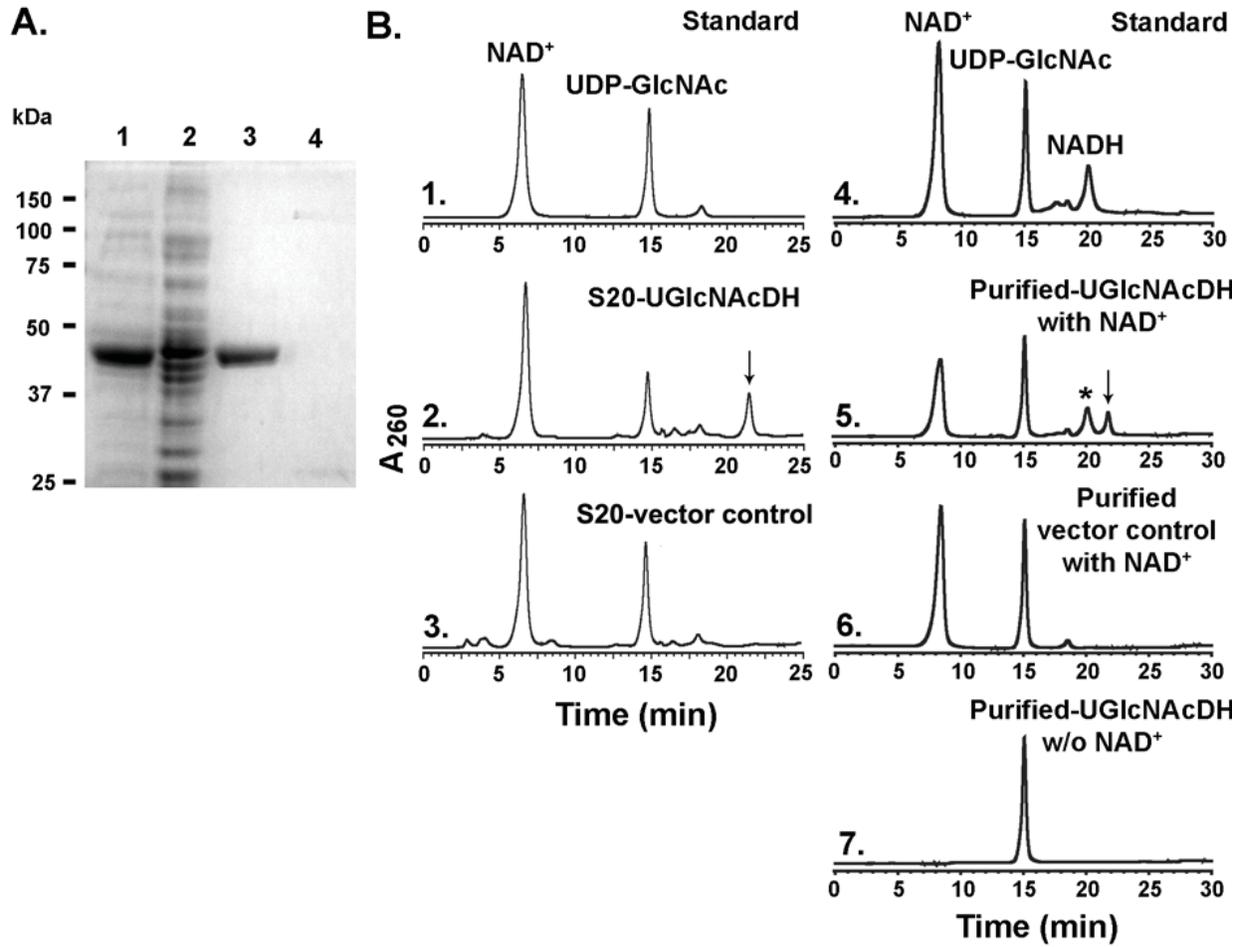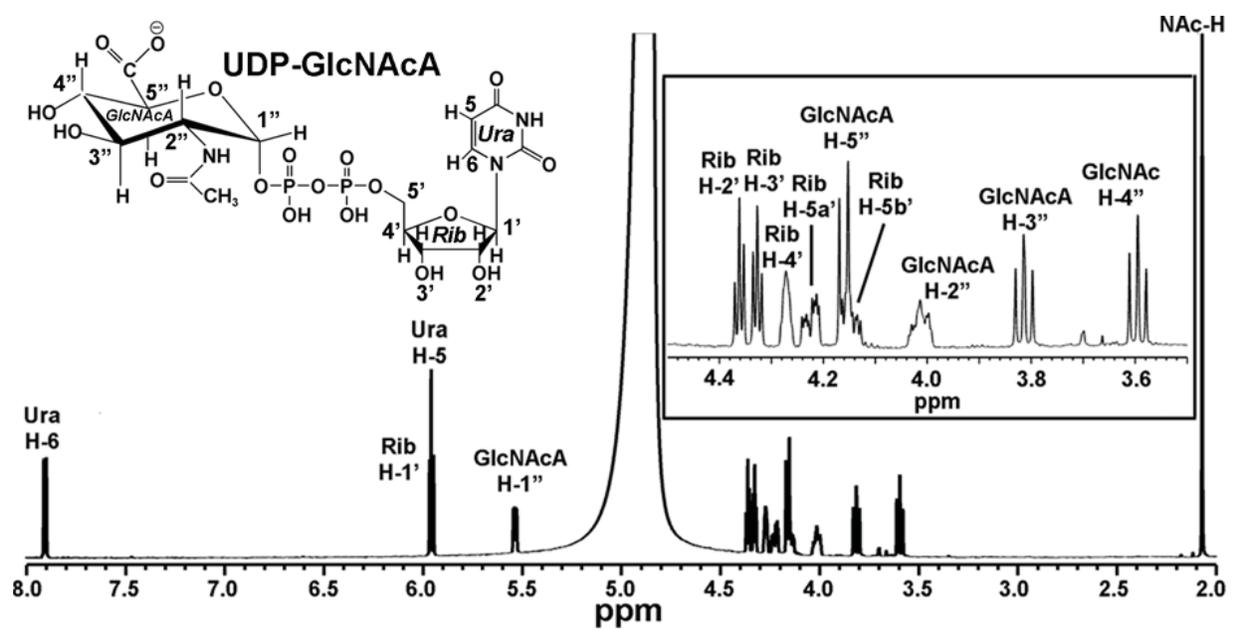UDP-GlcNAc → NAD⁺ NADH / *UGlcNAcDH* → UDP-GlcNAcA → NADH / NAD⁺ / *UXNAcS* → UDP-XylNAc

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6