

RESEARCH ARTICLE

The Biosynthesis of UDP-D-QuiNAc in *Bacillus cereus* ATCC 14579

Soyoun Hwang¹, Avi Aronov¹, Maor Bar-Peled^{1,2*}

1 Complex Carbohydrate Research Center, University of Georgia, Athens, Georgia, United States of America, **2** Departments of Plant Biology, University of Georgia, Athens, Georgia, United States of America

* peled@ccrc.uga.edu



OPEN ACCESS

Citation: Hwang S, Aronov A, Bar-Peled M (2015) The Biosynthesis of UDP-D-QuiNAc in *Bacillus cereus* ATCC 14579. PLoS ONE 10(7): e0133790. doi:10.1371/journal.pone.0133790

Editor: Raymond Schuch, ContraFect Corporation, UNITED STATES

Received: April 7, 2015

Accepted: July 1, 2015

Published: July 24, 2015

Copyright: © 2015 Hwang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by UG1021RR549475, US Department of Energy, (<http://energy.gov/>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

N-acetylquinovosamine (2-acetamido-2,6-di-deoxy-D-glucose, QuiNAc) is a relatively rare amino sugar residue found in glycans of few pathogenic gram-negative bacteria where it can play a role in infection. However, little is known about QuiNAc-related polysaccharides in gram-positive bacteria. In a routine screen for bacillus glycan grown at defined medium, it was surprising to identify a QuiNAc residue in polysaccharides isolated from this gram-positive bacterium. To gain insight into the biosynthesis of these glycans, we report the identification of an operon in *Bacillus cereus* ATCC 14579 that contains two genes encoding activities not previously described in gram-positive bacteria. One gene encodes a UDP-N-acetylglucosamine C4,6-dehydratase, (abbreviated Pdeg) that converts UDP-GlcNAc to UDP-4-keto-4,6-D-deoxy-GlcNAc (UDP-2-acetamido-2,6-dideoxy- α -D-xylo-4-hexulose); and the second encodes a UDP-4-reductase (abbr. Preq) that converts UDP-4-keto-4,6-D-deoxy-GlcNAc to UDP-N-acetyl-quinovosamine in the presence of NADPH. Biochemical studies established that the sequential Pdeg and Preq reaction product is UDP-D-QuiNAc as determined by mass spectrometry and one- and two-dimensional NMR experiments. Also, unambiguous evidence for the conversions of the dehydratase product, UDP- α -D-4-keto-4,6-deoxy-GlcNAc, to UDP- α -D-QuiNAc was obtained using real-time ¹H-NMR spectroscopy and mass spectrometry. The two genes overlap by 4 nucleotides and similar operon organization and identical gene sequences were also identified in a few other *Bacillus* species suggesting they may have similar roles in the lifecycle of this class of bacteria important to human health. Our results provide new information about the ability of *Bacilli* to form UDP-QuiNAc and will provide insight to evaluate their role in the biology of *Bacillus*.

Introduction

UDP-N-acetylglucosamine, UDP-GlcNAc, is an activated nucleotide sugar found in all organisms and is essential to life. In humans, UDP-GlcNAc is a precursor for synthesis of glycoprotein, surface glycans, and it is also found in extracellular matrices of cells which contain a variety of different sugar polymers; in fungi and arthropod, it is required to form chitin for

cell wall and exoskeleton structures, respectively, and in bacterium it is essential component for bacterial cell wall peptidoglycan assembly that is required to form the wall and protect the cells [1–4]. In most organisms, UDP-GlcNAc is not only a substrate for glycans but also an important precursor that is further metabolized to form many other nucleotide amino-sugars.

In the 1960's, the late glycobiologist Nathan Sharon, identified and characterized several 2-amino-sugars derivatives of GlcNAc in *Bacillus* sp. including a D-fucosamine (2-amino-2,6-dideoxy-D-galactose), D-galactosamine, and the di-amino-sugar bacillosamine [5]. The 4-epimer of D-fucosamine is D-quinovosamine (2-amino-2,6-dideoxy-D-glucose, abbr. quinovosamine, QuiN). Some of the amino-group of 2-amino-sugars can be found in glycans in acetylated or de-acetylated forms. The acetylated form of QuiN is QuiNAc (2-acetoamido-2,6-dideoxy-D-glucose). QuiNAc was identified in several important gram-negative human bacterial pathogens including the lipopolysaccharide (LPS) from *Brucella* [6] and *Legionella* [7]. QuiNAc is also an amino-sugar component of the LPS structure of plant fixing bacterium *Rhizobium* [8]. In *Rhizobium etli* CE3, a QuiNAc residue is located in the outer core of O-chain polysaccharide (OPS) linked to a 3-deoxy-2-octulosonic acid (Kdo) residue in the inner core of the LPS [9]. Mutant strains that lack the QuiNAc [10–12] fail to infect the host cells, and a specific *wreQ* mutant strain where a QuiNAc residue is replaced by its 4-keto derivatives [13] gives rise to bacterium that infect the host cells but very slowly. These studies suggest that QuiNAc contained in LPS is an important residue that participates in a symbiotic relationship between *Rhizobium etli* CE3 and its plant host. In other gram negative pathogenic bacterium like *Helicobacter pylori* and *Neisseria gonorrhoeae*, a modified QuiNAc sugar is found with additional acetamido group connected at C-4" forming a glycan with diNAcBac sugar residue (see [14]) for review of the biosynthetic route of UDP-diNAcBac).

By contrast, little is known about QuiNAc in gram-positive bacteria. In 1993, Ito et al [15] identified QuiN that was isolated from the cell walls of the alkaliphilic *Bacillus* sp. Y-25. QuiNAc, however, was not found in other *Bacillus* or any other gram-positive bacteria. *Bacillus cereus* is a food borne, spore-forming, and pathogenic bacterium, which is capable of motion by flagella. The bacterium is present in soil, dust, water, and plants (15–16). This *Bacillus* can also inhabit the intestinal tract of insects and mammals [16]. The bacterium is a facultative anaerobe and likely transiently present in insect cadavers and in decaying organic matter [17]. *Bacillus* is thus an attractive model to study the role of different polysaccharide structures that are made in response to different environments. Surprisingly, during routine analyses of polysaccharides derived from *Bacillus* sp, we detected an unusual amino-sugar and further analyses revealed this to be 2-acetamide-2,6-dideoxy-glucose, QuiNAc. However, little was known about the biochemical pathways and the corresponding genes involved in the formation of QuiN or its acetylated form QuiNAc in *Bacillus*. This prompted us to identify genes involved in the synthesis of QuiNAc-containing glycans. Here, we report the identification of a QuiNAc operon and the functional characterization of two enzymes that sequentially convert UDP-GlcNAc to UDP-WhoNAc (see Fig 1A) in *Bacillus cereus* ATCC 14579. Two *Bacillus* enzymes encode UDP-GlcNAc-4,6-dehydratase and 4-reductase, which we named Pdeg and Preq. We used combined instrumentations with NMR spectroscopy and mass spectrometry to show that Pdeg converts UDP-D-GlcNAc to UDP-4-keto-6-deoxy-D-GlcNAc, and Preq immediately converts the 4-keto sugar to UDP-WhoNAc. Such enzyme activities have not previously been described in *Bacillus*, and thus our data provides the basis for understanding the formation of QuiNAc-containing glycans by *Bacillus* and their roles.

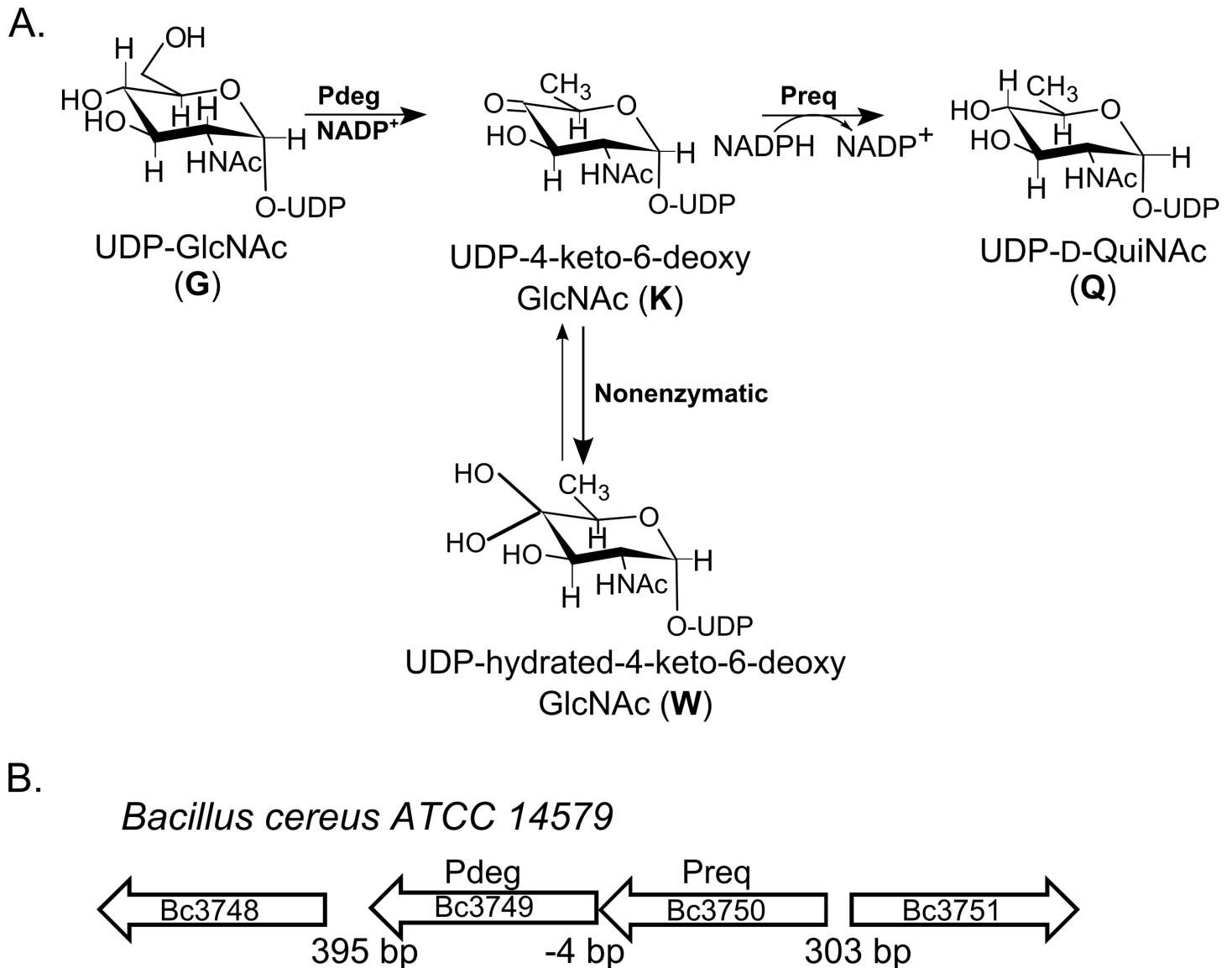


Fig 1. A. A proposed pathway for the formation of UDP-QuiNAc in *Bacillus cereus* ATCC 14579. The enzyme encoded by Bc3750, UDP-GlcNAc C4,6-dehydratase (Abbr. Pdeg), converts UDP-GlcNAc to UDP-4-keto-6-deoxy-GlcNAc. At steady state, the UDP-4-keto-sugar form (K) is converted non-enzymatically to a hydrated form W. The enzyme encoded by Bc3749 (Abbr. Preq) is a UDP-4-keto-sugar C4"-reductase and UDP-D-QuiNAc. **B. Organization of the two-genes operon and flanking regions in *B. cereus* ATCC 14579.**

doi:10.1371/journal.pone.0133790.g001

Materials and Method

Bacterial strain and growth

Stock of wild type *Bacillus cereus* ATCC 14579 was stored in 30% glycerol at -80°C , streaked onto agar plate, and grown for 18 hours at 30°C . The medium (agar or liquid) used was Luria Bertani (LB per liter: 10 g tryptone, 5 g yeast extract, 10 g NaCl). Stock of *E. coli* strain DH10B (LifeTechnologies) was used for cloning, and strain Rosetta2(De3)pLysS (Novagen), was used to generate recombinant proteins.

Cloning of Pdeg and Preq

A single colony of *Bacillus cereus* ATCC 14579 grown on LB-agar was suspended in 50 μ l sterile water. The suspension was heat-treated (5 min, 96°C), centrifuged (13,000 \times g, 2 min), and a 5 μ l portion of the supernatant was used as a source of genomic DNA for PCR amplification. The PCR primer sets used to amplify the coding region were designed to include at their 5' a 15-nucleotide extension with sequence homology to the cloning site of the pET28b-Tev plasmid. The primers used for Preq were SY120: 5' -CAGGGCGCCATGTCCatgaaaaaaaaatgcgagccttttaataaac and SY121: 5' -CTCGAGTGC GGCCGctcatgcatgcagatgtcactacacttgcg; for Pdeg SY122: 5' -CAGGGCGCCATGTCCatgttaataaaaaataattttaattactgg, and SY123: 5' -CTCGAGTGC GGCCGctcatgcaaaaaccctccttttaataag. Individual genes (Preq or Pdeg) were PCR-amplified in a 20 μ l reaction volume that included buffer, dNTP's (0.4 μ l of 10 mM), *Bacillus cereus* genomic DNA (5 μ l), PCR primer sets (1 μ l each of 10 μ M), and high fidelity *Pyrococcus* DNA polymerase (0.4U Phusion Hot Start II; New England Biolabs). The PCR thermocycle conditions were 1X 98°C denaturation cycle for 30 sec followed by 25X cycles (each of 8 sec denaturation at 98°C; 25 sec annealing at 50°C; 30 sec elongation at 72°C), and finally 4°C. A similar PCR reaction was used to amplify the expression plasmid (pET28b-Tev) using a specific inverse-PCR primer set (SY118: GGCATGGCGCCCTGAAAATACAGGTTTTTC and SY119: GCGGCCGCACTCGAGCACCACCACCACC) located near the NcoI and HindIII sites, respectively) with 25 sec annealing cycle at 58°C and 3 min elongation at 72°C. After PCR, a portion (4 μ l each) of the amplified plasmid and insert were mixed, digested with 10U DpnI (15 min, 37°C), and then transformed into DH10B competent cells. Clones were selected on LB agar containing kanamycin (50 μ g/ml) and positive clones were verified by PCR and by DNA sequencing using primers (T7 promoter and T7 terminator) flanking the gene insert. DNA sequences of the cloned genes were deposited in Genbank (with respective accession numbers KR012645 and KR012646). The recombinant DNA coding sequence was cloned to yield a recombinant protein fused at the N-terminal to a short peptide linker of 6 histidines (His₆) followed by a TEV recognition amino acid sequence.

Expression and Purification of Recombinant Pdeg and Preq

E. coli strains containing either pET28b-His₆-TEV-Bc3750 (Pdeg), pET28b-His₆-TEV-Bc3749 (Preq), or a non-related gene in pET plasmid that served as control, were cultured in separate flasks of 20 ml LB medium containing kanamycin (50 μ g/ml) and chloramphenicol (35 μ g/ml). The cells were grown for 20 hours while shaking at 250 rpm at 37°C. A 5 ml of each culture was transferred to a separate flask containing 250 ml LB medium supplemented with the same concentrations of antibiotics, and grown at 37°C while shaking at 250 rpm until the cell densities reached $A_{600} = 0.6$. The cells were then induced by the addition of isopropyl β -D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM and grown for 20 hours at 18°C while shaking at 225 rpm. Each culture was then centrifuged (6,000 \times g for 10 min at 4°C), and cell pellet was washed with 20 ml cold double-distilled water and recentrifuged. The cell pellets derived from the IPTG-induced culture were resuspended in 10 ml cold lysis buffer [50 mM Tris-HCl pH 8.3, 30% glycerol, 50 mM KCl, 20 mM β -mercaptoethanol (BME)]. The resuspended cells were kept on ice and sonicated (10 sec pulse and 20 sec off) using a Misonics S4000 Sonicator for 12 cycles at 30% amplitude. Following sonication, the lysed cell solution was centrifuged (6,000 \times g for 15 min at 4°C), and the supernatant (labeled S6) was collected, supplemented 2 mM BME, and centrifuged (20,000 \times g for 30 min at 4°C). The resulting supernatant (labeled S20) was collected, divided to smaller aliquots, flash frozen in liquid nitrogen, and then stored at -80°C until Ni-purification.

The His-tagged recombinant proteins (Pdeg, Preq, and the control unrelated 6His-protein) were purified over a Fast-Flow Ni²⁺-Sephacrose column [2 ml of resin packed in a 1 cm inner diameter x 10 cm long column). Before purification, each column was equilibrated with Buffer A (50 mM Tris-HCl pH 8, 10% glycerol, 100 mM NaCl, 5 mM BME). The frozen S20 fraction of recombinant Pdeg (Bc3750), Preq (Bc3749), or control was rapidly thawed, and 7.5 ml was loaded onto the Ni-column. Unbound protein was washed from the column with 7.5 ml of Buffer A followed by 15 ml Buffer E1 (Buffer A containing 10 mM imidazole-HCl pH 8). Subsequently, the column was washed with 20 ml Buffer E4 (Buffer A containing 60 mM imidazole-HCl pH 8), and 7.5 ml Buffer E7 (Buffer A containing 250 mM imidazole-HCl pH 8, and 5 mM BME). The activity of recombinant Pdeg and Preq typically eluted with E7 buffer. After elution, these fractions were supplemented with an additional 0.75 ml 80% glycerol, and Pdeg was further supplemented with an additional 133 μ M NADP⁺. Following supplement addition, the protein fractions were flash frozen in liquid nitrogen and stored at -80°C. The concentration of purified protein was determined at 280 nm using a theoretical extinction coefficient (<http://web.expasy.org/protparam/>) of $\epsilon = 23,630$, and $30,285 \text{ m}^{-1} \text{ cm}^{-1}$, based on the amino acid sequence of Pdeg (Bc3750) and Preq (Bc3749), respectively. The molecular mass of purified protein was estimated after separation on SDS-PAGE and Coomassie blue staining. The potential size of the purified enzyme in native condition was estimated by chromatography on a gel filtration column (Superdex 75, 1x30 cm; GE Healthcare) and UV detection at 215 nm. The Superdex column was equilibrated with cold buffer [10 mM Tris-HCl (pH 7.0), 150 mM NaCl] at a flow rate of 0.5 ml/min. Size of protein was determined using known molecular weight of marker proteins as standards.

Enzymatic Assays and Product Analysis of Pdeg and Preq

For Pdeg, the activity was examined by time-resolved ¹H-NMR spectroscopy and by HILIC-HPLC with UV or electro spray ionization mass spectrometry (ESI-MS). For HPLC-based assays, the reaction assay, in a total volume of 50 μ l, contained 50 mM Tris-HCl pH 8.0, 1 mM UDP-GlcNAc, 1 mM NADP⁺, and 400 pM of recombinant Pdeg. The reaction was terminated by heating for 2 min at 95°C. Chloroform was then added as described [18] and an aliquot (30 μ l) of the upper layer phase was removed and mixed with 57 μ l acetonitrile and 3 μ l 0.5 M ammonium-acetate pH 4.35. A portion (10 μ l) of this mixture was analyzed using HILIC chromatography coupled to an ESI-MS/MS. HPLC-ESI-MS/MS was carried out on the Shimadzu ESI-IT-TOF using a Nexera LC-30AD pump, autosampler (Sil30), and column heater (at 37°C). The column used for separation was an Accucore 150-amide HILIC column (150x4.6 mm, 2.6 μ m particle size, ThermoScientific). The chromatography gradient conditions for the HPLC-ESI-MS were composed of two solvents: Solvent A, 40 mM ammonium-acetate, pH 4.35 (A) and solvent B, acetonitrile (ACN). The column was equilibrated at 0.7 ml/min with 25% A and 75% B. After injection, chromatography conditions were: from 0–2 min, flow rate 0.7 ml/min at 25% A/75% B, followed by a 23 min gradient to 65% A, then a 10 min gradient to 50% A. Subsequently, column was washed for 10 min with 25% A/75% B prior to the next injection. To detect analytes, the MS detector was operated in the negative mode with detector voltage 1.75 kV. In addition to ESI-MS analysis, another portion of each enzyme assay (10 μ l, prepared as above) was subjected to analysis by an HPLC-UV system. This chromatographic system included an Agilent 1260 pump, autosampler, column heater (at 40°C), and a diode array detector (UV). For chromatographic separation, a narrow-bore Accucore 150-amide HILIC column (250x2.1 mm, 2.6 μ m particle size) was used with a gradient elution method consisting of 40 mM ammonium acetate buffered to pH 4.35 (solvent A) and acetonitrile (solvent B). After equilibration at 0.25 ml/min with 25% A (75% B), the gradient elution

method was programmed to run at 25% A and 0.25 ml/min from 0 to 10 min, increase (linearly) to 39% A from 10 to 30 min, increase to 50% A (and decrease flow rate to 0.2 ml/min) from 30 to 35 min, run at 50% A from 35 to 40 min, decrease to 25% A (and increase flow rate to 0.25 ml/min) from 40 to 45 min, and run at 25% A from 45 to 60 min. Peaks of enzymatic products detected by A_{261} (max for UDP-sugars) were collected for further analyses.

For the NMR-based assays of recombinant Pdeg, a reaction in a volume of 180 μ l consisted of 120 μ l of D_2O and 60 μ l of water-based reagents [50 mM Tris-HCl, pH 8.0, 1 mM UDP-GlcNAc, 400 pM of purified Pdeg, and 1 mM DSS (2,2-dimethyl-2-silapentane-5-sulfonate) that served as internal NMR standard]. The reaction mixture was transferred to a 3-mm NMR tube, and the products were analyzed by 1H -NMR spectroscopy for up to 3 hours at 25°C using a Varian DirectDrive 600-MHz spectrometer equipped with a cryogenic probe. For the NMR-based Preq assay, 105 pM of purified recombinant Preq was added along with 9 μ l of 10 mM NADPH to a 180 μ l Pdeg reaction. The products were analyzed for up to 1 hour at 25°C in the same 600-MHz spectrometer. Data were acquired before the addition of the enzyme as time zero (t_0). After adding the enzyme, acquisition was started after ~3 min to allow the spectrometers operating parameters to be optimized. Sequential one-dimensional proton spectra with presaturation of the water resonance were acquired over the course of the reaction. All NMR spectra were referenced to the resonance of DSS set at 0.00 ppm. Processing of the data was performed with MestreNova (Mestrelab Research).

NMR Analysis of Enzymatic Products

Individual enzyme reaction products were collected based on their UV absorbance during chromatography with HILIC HPLC-UV. After lyophilization, the UDP-sugar products were dissolved in 100% D_2O and characterized by NMR. 2D-NMR spectra were obtained at room temperature on Varian INOVA 600 and 900 MHz spectrometers. Each purified UDP-sugar was identified using COSY [19], TOCSY [20] (80 ms mixing time), HSQC [21], and HSQC-TOCSY [22] with 80 ms mixing time.

Characterization and Kinetic Analyses of the Recombinant Enzymes Pdeg and Preq

Various assays were carried out to determine a suitable temperature, pH, and buffer for optimal activity of both recombinant enzymes. To establish an ideal temperature for optimal activity of Pdeg, a standard reaction assay was performed at various temperatures between 4°C and 42°C for 2 hours. The standard reaction assay contained 400 pM of purified recombinant Pdeg, 1 mM UDP-GlcNAc, 0.5 mM $NADP^+$, and 50 mM Tris-HCl pH 8. The amount of product formed over time was determined by HPLC-UV. The optimal temperature for recombinant Preq was determined by performing a standard reaction assay at various temperatures. To obtain UDP-4-keto-6-deoxy-D-GlcNAc substrate for Preq kinetic reactions, a larger scale Pdeg reaction was first carried out. This reaction consisted of 50 mM Tris-HCl pH 8, 4 mM UDP-GlcNAc, 1 mM $NADP^+$, and 1.6 nM purified recombinant Pdeg. The reaction was incubated overnight for maximum conversion of the product to substrate and the concentration of the UDP-4-keto-6-deoxy-D-GlcNAc product was determined after elution from a HILIC column with UV detection.

The activities of Pdeg and Preq in different buffers (Tris-HCl and sodium phosphate) at pH values between 6 and 10 were also investigated. Assays were performed at 22°C for 15 min (Pdeg) or 75 sec (Preq). The assays specific to Pdeg contained 400 pM of recombinant enzyme, 0.5 mM UDP-GlcNAc, 0.2 mM $NADP^+$, and 50 mM of a given pH-adjusted buffer. Assays specific to Preq contained 105 pM of recombinant enzyme, 565 μ M UDP-4-keto-6-deoxy-GlcNAc,

1 mM NADPH, and 50 mM of a given buffer. The amount of product formed (as detected via HPLC-UV) was used to determine optimal assay conditions.

Inhibition assays were performed by first mixing a potential inhibitor compound, recombinant enzyme, and buffer. After incubating on ice for 5 min, reaction substrates and additives were added to the reaction vial. The inhibition assays for Pdeg consisted of 50 mM Tris-HCl pH 8, 0.2 mM inhibitor, 0.5 mM UDP-GlcNAc, and 400 pM recombinant enzyme. The inhibition assays for Preq consisted of 50 mM Tris-HCl pH 8, 0.2 mM inhibitor, 565 μ M UDP-4-keto-6-deoxy-GlcNAc, 1 mM NADPH, and 105 pM recombinant enzyme. Assays were performed at 22°C for 15 min (Pdeg) or 75 sec (Preq), and product formation was determined with HPLC.

Initial kinetic assays were performed in triplicates with different substrate concentration and the final kinetic assays for Pdeg were performed at 22°C for 15 min, since, between 10 and 20 min, product formation is linear across time. The final Preq kinetic studies were performed at 22°C, and the reaction was terminated after 105 sec. Between 60 and 120 sec, product formation of Preq is linear across time. The final kinetic studies to determine the k_{cat} and K_m of Pdeg, reaction assays consisted of various concentrations of UDP-GlcNAc (10, 20, 40, 80, 100, 160, 200, 300, 400, 500, 600, 700, 800, 900, and 1000 μ M) with a fixed amount of co-factor (200 μ M NADP⁺ and NADPH) and 400 pM of recombinant Pdeg. For Preq assays, the reaction included various concentrations of UDP-4-keto-6-deoxy-GlcNAc (23, 68, 113, 203, 248, 293, 339, 384, 429, 474, 519, 564, 609, 655, 700, 745, 790, 835, 880, and 925 μ M) with a fixed amount of co-factor (2 mM NADPH) and 100 pM of recombinant Preq. The products formed were analyzed by HPLC and used to determine initial velocities. The kinetic parameters were derived by fitting enzyme kinetic curves with GraphPad Prism 5.

For Pdeg substrate specificity studies, assays in a final volume of 50 μ l consisted of 50 mM Tris-HCl pH 8, 0.5 mM NADP⁺, 0.5 mM NADPH, 2 mM of various UDP-sugars (UDP-galactose, UDP-GalNAc, UDP-GlcNAc, and UDP-glucose), and 400 pM recombinant Pdeg and were performed for 4 hours at 22°C. For each reaction, amounts of product and substrate remaining were determined after chromatography via HILIC-HPLC detection.

Results

Identification of QuiNAc in *Bacillus*

In our systematic survey of prokaryotic glycans isolated from various *Bacillus* species, we unexpectedly identified an alditol-acetate 6-deoxy-2-N-acetylhexosamine sugar residue-derivative eluted from a GC-column at 29.7 min (S1 Fig). The electron ionization mass spectrometry (EI-MS) of this peak showed prominent fragment ions at m/z 302, 260, 201, 145, 129, 103, and 85 (see insert in S1 Fig) identical with those found for alditol acetates derivatives of a QuiNAc std. To the best of our knowledge, little was known about QuiNAc formation in gram-positive bacteria, and it inspired us to further investigate the metabolic pathway involved in biosynthesis of QuiNAc in *Bacillus*.

Identification of Pdeg and Preq from *Bacillus cereus* ATCC 14579

To identify potential genes encoding enzymes involved in the formation of activated-QuiNAc, we first performed a BLAST search using amino acid sequences of known bacterial UDP-GlcNAc 4,6-dehydratases. This led us to identify *B. cereus* ATCC 14579 Bc3750 (herein referred to as Pdeg) as a candidate. Interestingly, the *Bacillus* Pdeg protein shares a high amino-acid sequence homology with functional proteins determined to have not only a C4,6-dehydratase but also a C5-epimerase activity. As shown in Table 1A, Pdeg protein has a high amino acid sequence identity (46%) with functional UDP-GlcNAc-5-inverting-C4,6-dehydratase, FlaA1 (PseB, in *Campylobacter jejuni*), from *Helicobacter pylori* [23], a 41% amino acid sequence

Table 1. Comparative analysis of *Bacillus cereus* ATCC 14579 proteins Pdeg (Bc3750) and Preq (Bc3749) with known genes encoding UDP-sugar 4,6-dehydratase (A) and 4-reductase (B) from bacterial sp.

A.	Accession (name)	E-Value	AA Identity	Function in Species
Pdeg	YP_203523	3e ⁻¹⁸	27%	UDP-N-acetyl-glucosamine 4,6-dehydratase (<i>Vibrio Fischeri</i>)
Pdeg	CJ1293 (PseB)	4e ⁻⁷⁷	47%	UDP-GlcNAc-specific C4,6 dehydratase/C5 epimerase (<i>Campylobacter jejuni</i>)
Pdeg	HP0840	5e ⁻⁸³	46%	UDP-GlcNAc-inverting 4,6-dehydratase (<i>Helicobacter pylori</i>)
Pdeg	3VVB_A (CapE)	8e ⁻⁷⁵	41%	UDP-GlcNAc 5-inverting 4,6-dehydratase (<i>Staphylococcus aureus</i>)
Pdeg	4TQG (Mg434)	6e ⁻⁵⁵	35%	UDP-GlcNAc 5-inverting 4,6-dehydratase (<i>Megavirus chilensis</i>)
Pdeg	YP_471773.1	8e ⁻⁴⁸	34%	UDP-N-acetyl-glucosamine 4,6-dehydratase (<i>Rhizobium etli</i>)
Pdeg	CAL35237.1	5e ⁻⁴⁵	34%	PglF (<i>Campylobacter jejuni</i>)
Pdeg	YP_207256.1	8e ⁻⁴⁸	37%	PglD (<i>Neisseria gonorrhoeae</i>)
Pdeg	AHB32380.1	1e ⁻⁴⁸	34%	WeeK (<i>Acinetobacter baumannii</i>)
B.	Accession	E-Value	AA Identity	Function in Species
Preq	Q6T1X6	5e ⁻⁴⁴	33%	GDP-4-keto-6-deoxy-D-mannose 4-reductase (<i>Aneurinibacillus thermoaerophilus</i>)
Preq	1BXK	2e ⁻¹⁴	24%	dTDP-glucose 4,6-dehydratase (<i>Salmonella Enterica</i> serovar Typhimurium)
Preq	YP_470339.1	2.4	29%	UDP-2-acetamido-2,6-dideoxy-D-xylo-4-hexulose 4-reductase (<i>Rhizobium etli</i>)
Preq	AAF23991	2.9	28%	4-reductase (<i>P. aeruginosa</i> O6)

doi:10.1371/journal.pone.0133790.t001

identity with CapE of *Staphylococcus aureus* [24] and 35% similarity to Mg434 from the giant virus *Megavirus chilensis* [25]. It is worth noting that Pdeg has much lower amino acid sequence identities, 27%, 34%, 37%, and 34% with several functional UDP-N-acetyl-glucosamine-C4,6-dehydratase (lacking 5-epimerase) from *Vibrio Fischeri* [26], with PglF, involved in UDP-diNAcBac biosynthesis pathways from *Campylobacter jejuni* [27], with PglD from *Neisseria gonorrhoeae* [28], with WEEK from *Acinetobacter baumannii* [29], respectively, as shown in Table 1A. Here, we provide biochemical evidence showing that despite higher sequence similarity to a dual 5-epimerase, C4,6-dehydratase enzymes, *Bacillus* protein Pdeg is a 4,6-dehydratase with no detectable 5-epimerase activity. This suggests that protein sequence homology is not sufficient to determine activity of an unknown gene. Flanking Pdeg (Bc3750) in *Bacillus cereus* ATCC 14579 is gene Bc3749 (herein referred to as Preq, Fig 1B). Preq and its adjacent gene unit Pdeg are overlapped by 4 nucleotides. Such gene overlap and reading frame offset appeared conserved in other *Bacilli* including, for example, *Bacillus weihenstephanensis* FSL R5-860, *Bacillus cereus* ATCC 10876, *Bacillus thuringiensis* serovar israelensis ATCC 35646, and *Bacillus thuringiensis* serovar kurstaki str. HD-1. It is therefore possible that the expression of the Pdeg and Preq gene unit is under the same regulation mechanism. The Preq protein homology to other known enzymes is not clear. As shown in Table 1B, Preq shares, for example, 33% amino acid sequence identity with functional GDP-4-keto-6-deoxy-D-mannose 4-reductase from *Aneurinibacillus thermoaerophilus* [30] and lower sequence homology (24%) with dTDP-glucose 4,6-dehydratase (Rmlb) from *Salmonella Enterica* serovar Typhimurium [31]. Also, it shares lower sequence homology (29%) with functional UDP-2-acetamido-2,6-dideoxy-D-xylo-4-hexulose-4-reductase from *Rhizobium etli* [32] even though its function, as we will described, is the same with *Bacillus* protein Preq. Below we provide biochemical evidences of Pdeg (Bc3750) and Preq (Bc3749) proteins for their sequential ability to convert UDP-GlcNAc to UDP-4-keto-6-deoxy-GlcNAc and to UDP-QuiNAc.

Bc3750 (Pdeg) encodes a UDP-GlcNAc-C4,6-Dehydratase

E. coli cells expressing recombinant His6-Bc3750 (Pdeg) were used to isolate and purify the recombinant protein using a Ni-affinity column (Fig 2, lane 2, calculated 38 kDa based on

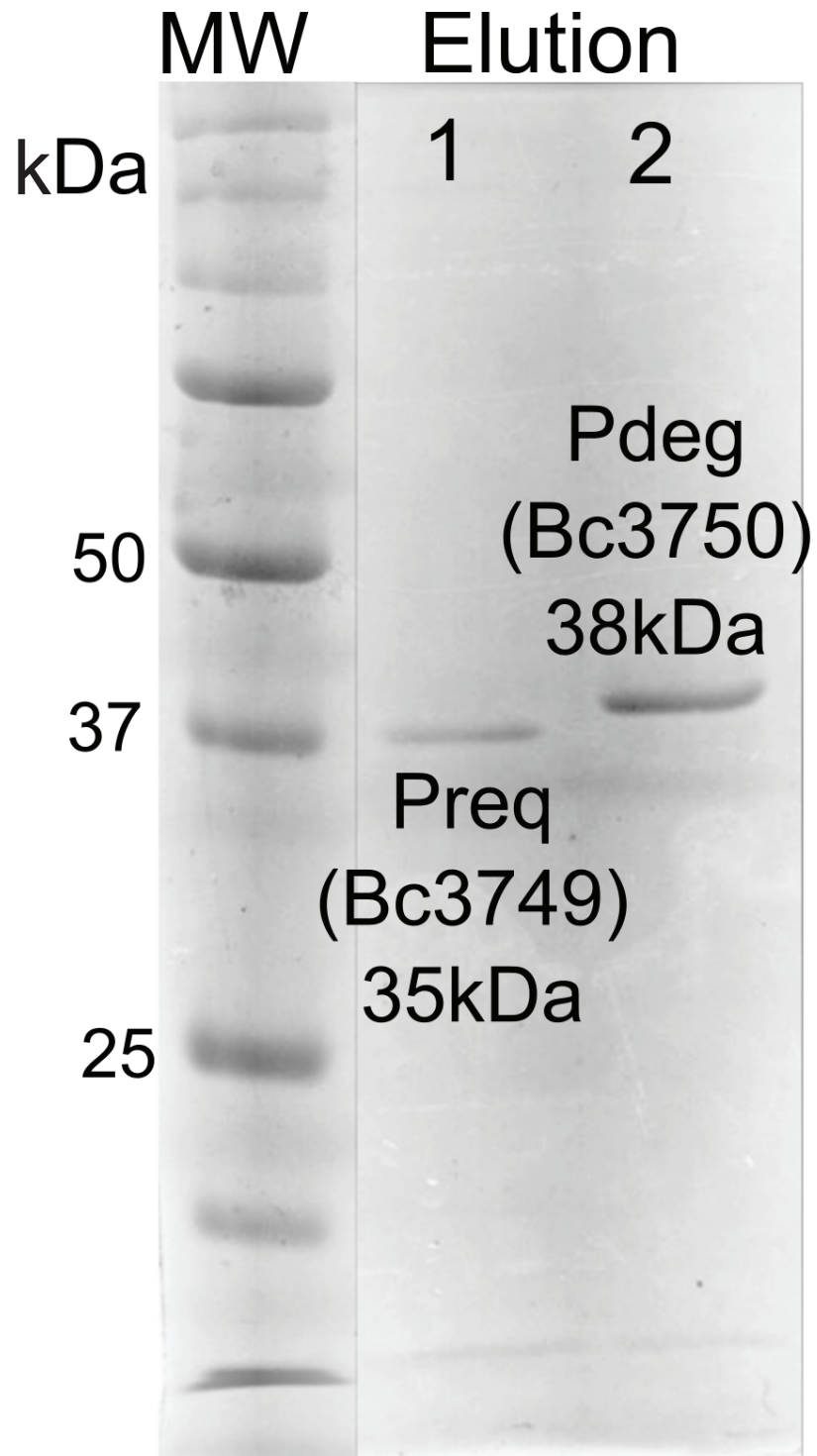


Fig 2. SDS—PAGE analysis of the *B. cereus* ATCC 14579 purified recombinant Bc3750 (Pdeg) and Bc3749 (Preq) proteins involved in the biosynthesis of UDP-QuiNAc. Protein standards are shown on the right in kDa. The final elution fraction (E7) of purified recombinant proteins from affinity column is shown for Bc3749 (lane 1) and Bc3750 (lane 2).

doi:10.1371/journal.pone.0133790.g002

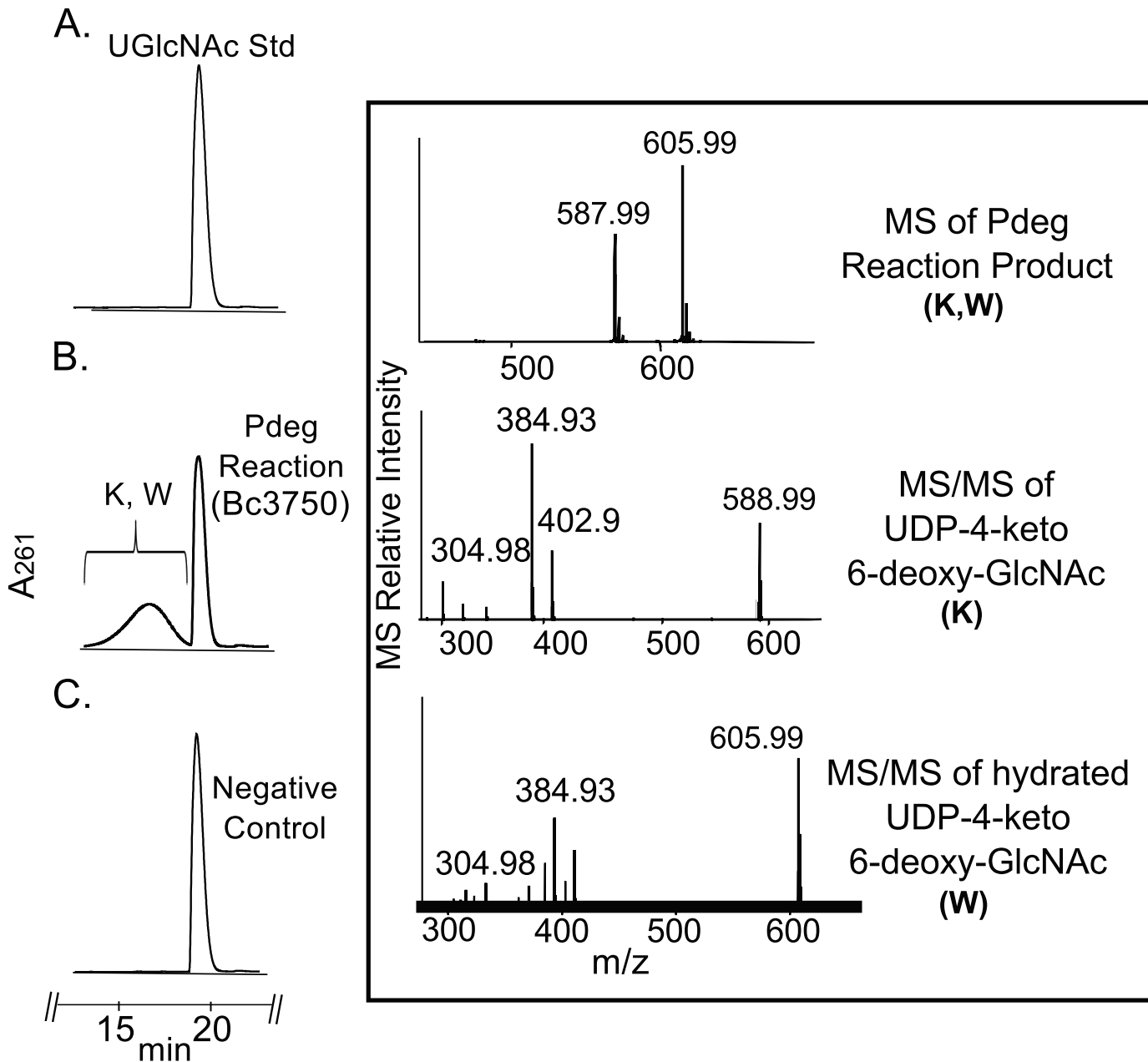


Fig 3. Analysis of recombinant enzyme reaction Pdeg by UV-HPLC and LC-ESI-MS-MS. A. UDP-GlcNAc standard reaction is shown. B. UDP-GlcNAc C4,6-dehydratase reaction, is the conversion of UDP-GlcNAc to UDP-4-keto-6-deoxy-HexNAc. The broad peak (K and W) denotes 4-keto and 4-hydrated-keto form of the UDP-4-keto-6-deoxy-sugar. Boxed top panel shows the product ions, K and W, m/z 587.99 and 605.99, respectively. MS-MS analysis of parent ions K and W gave fragment ions at m/z 402.9, 384.9 and 304.9 that are consistent with $[UDP-H]^-$, $[UDP-H_2O-H]^-$, and $[UMP-H_2O-H]^-$, respectively. (Boxed the second and third panel) C. Pdeg negative control reaction carried out with unrelated protein.

doi:10.1371/journal.pone.0133790.g003

amino-acid sequence). Initial enzyme characterization of purified Pdeg was determined by a UV-HPLC and ESI-mass spectrometry. HILIC analysis of the enzymatic products formed when purified Pdeg was reacted with UDP-GlcNAc showed the appearance of a new broad peak with a retention time of 16 min (Fig 3 panel B, labeled K and W). This peak was not detected in a reaction with unrelated protein (Fig 3 panel C). When the Pdeg enzymatic reactions were chromatographed and analyzed in the negative mode by ESI-MS, the broad peak

Table 2. NMR data for the sugar moiety of UDP-4-keto-6-deoxy-GlcNAc (W) and UDP-QuiNAc (Q).

Hydrated-UDP-4-keto-6-deoxy-GlcNAc (W)			UDP-QuiNAc (Q)		
¹ H	Chemical Shift (ppm)	Coupling Constant (Hz)	¹ H	Chemical Shift (ppm)	Coupling Constant (Hz)
H1 ^W	5.44	J _{1,p} (7.1)	H1 ^Q	5.44	J _{1,p} (7.18)
H2 ^W	4.11	J _{1,2} (3.5)	H2 ^Q	3.99	J _{1,2} (3.6)
H3 ^W	3.81	J _{2,3} (10)	H3 ^Q	3.74	J _{2,3} (10)
H4 ^W			H4 ^Q	3.25	J _{3,4} (10)
H5 ^W	4.11		H5 ^Q	3.99	J _{4,5} (9.8)
H6 ^W	1.23	J _{5,6} (6.2)	H6 ^Q	1.29	J _{5,6} (6.0)
NAc-H ^W		2.07	NAc-H ^Q	2.06	

doi:10.1371/journal.pone.0133790.t002

(K, W) gave two major ions with *m/z* 587.99 and 605.99 (Fig 3 Box top panel). These *m/z* values likely correspond to [M-H]⁻ for a UDP-4-keto-6-deoxy-HexNAc and the hydrated form of the UDP-4-keto-6-deoxy-HexNAc. MS-MS analysis of K parent ion gave fragment ions at *m/z* 402.9, 384.93 and 304.98 that are consistent with [UDP-H]⁻, [UDP-H₂O-H]⁻, and [UMP-H₂O-H]⁻, respectively. The neutral loss of 185 mass unit from *m/z* 588 implies a mass for a 4-keto-6-deoxy-HexNAc sugar. These initial analyses led us to suspect that the newly formed Pdeg enzymatic product is a UDP-4-keto-6-deoxy-HexNAc. However, the sugar configuration could not be determined by MS analyses, so the product peak was collected by HPLC and analyzed by NMR spectroscopy. Chemical shift assignments obtained by one-dimensional and two-dimensional NMR experiments and coupling constants (Table 2 and Fig 4) indicated that the enzymatic product is UDP-4-keto-D-GlcNAc. The 4-keto-6-deoxy-GlcNAc H1^W anomeric region of the proton spectrum contains a quadruplet signal with chemical shifts of 5.44 ppm. The distinct chemical shift of the anomeric proton and the coupling constant values of 3 Hz for J_{H1^W, H2^W} and 7.1 Hz for J_{H1^W, p} are consistent with an α-linkage to the phosphate of UDP. The chemical shifts for each H6^W has a value of 1.23 ppm. The methyl protons resonance of the N-acetylated group (C-2^W-NAc) at 2.07 ppm is consistent for a C2 acetamido moiety of the product. The D-configuration of the linked sugar was established based on coupling constants. An L-sugar would have larger coupling between α-phosphate and H1^W proton and also larger coupling between H2^W and H3^W [33]. Further support for the presence of the UDP-4-keto-6-deoxy-D-GlcNAc product was established using two-dimensional NMR experiments. COSY and TOCSY experiments confirmed the assignments for H2^W, H3^W, and H5^W in this 4-keto-6-deoxy sugar. The J_{H2^W, H3^W} coupling constant of 10 Hz and the J_{H3^W, H5^W} coupling constant of 10 Hz is consistent with a gluco-configuration. ¹³C-HSQC and HSQC-TOCSY experiments established carbon assignments and the presence of an N-acetylated carbon.

We next used time-resolved ¹H NMR to monitor the conversion of UDP-GlcNAc to UDP-4-keto-6-deoxy-GlcNAc with purified recombinant Pdeg (Fig 5). The C4,6-dehydratase reaction generates the C4^W-hydrated form of UDP-4-keto-6-deoxy-GlcNAc (W in Fig 5). When the recombinant UDP-GlcNAc 4,6-dehydratase is added, additional signals corresponding to the anomeric proton of C4^W-hydrated form of UDP-4-keto-6-deoxy GlcNAc (WH-1^W) appear and are accompanied by a decrease in the intensity of the signals corresponding to the anomeric protons of GlcNAc (GH-1^W) (Fig 5B). Also, the 6-deoxy proton signal (WH-6^W) starts showing up at 1.23 ppm, while other signals close to the 1.17 ppm do not change (See Fig 5D). Time-resolved NMR also detected a chemical shift change of protons of uracil and the NAc methyl group (Fig 5A and 5C). Together, these data confirm that Bc3750 (Pdeg) encodes a

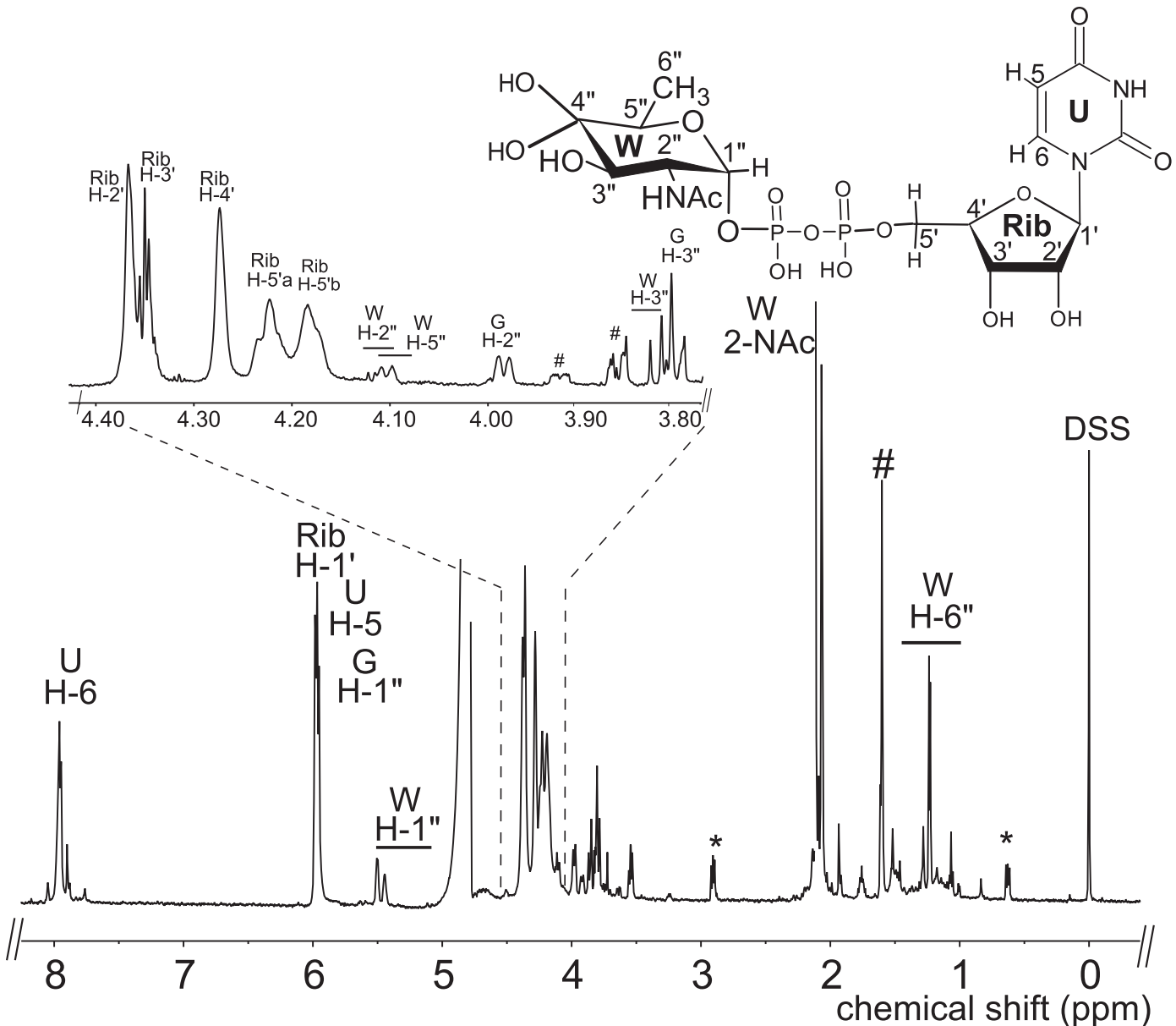


Fig 4. Analysis of Pdeg recombinant enzyme reaction products by $^1\text{H-NMR}$ indicates formation of hydrated-UDP-4-keto-6-deoxy-D-GlcNAc. The product peak of the Pdeg reaction was collected and analyzed at 600 MHz NMR. Full proton spectrum of the Pdeg product hydrated-UDP-4-keto-6-deoxy-D-GlcNAc. Expanded proton spectra between 3.8 and 4.4 ppm that shows the sugar ring. The short line above NMR 'peaks' denotes specific chemical shifts belonging to a UDP-4-keto-6-deoxy-D-GlcNAc. Symbol (#) denotes column contamination and symbol (*) denotes DSS.

doi:10.1371/journal.pone.0133790.g004

UDP-GlcNAc-C4,6-dehydratase that converts UDP-D-GlcNAc to UDP-4-keto-6-deoxy-GlcNAc (Fig 1A).

Bc3749 (Preq) Encodes C4''-reductase and forms UDP-QuiNAc

To determine the function of Preq, *E. coli* cells expressing the gene were used to isolate and purify the recombinant protein (Fig 2, lane 1; calculated 35 kDa). During the initial characterization of the Preq protein, we determined that activity required NADPH and UDP-4-keto-6-deoxy-sugar, as no activity was observed without NADPH or with non-keto-sugars. As

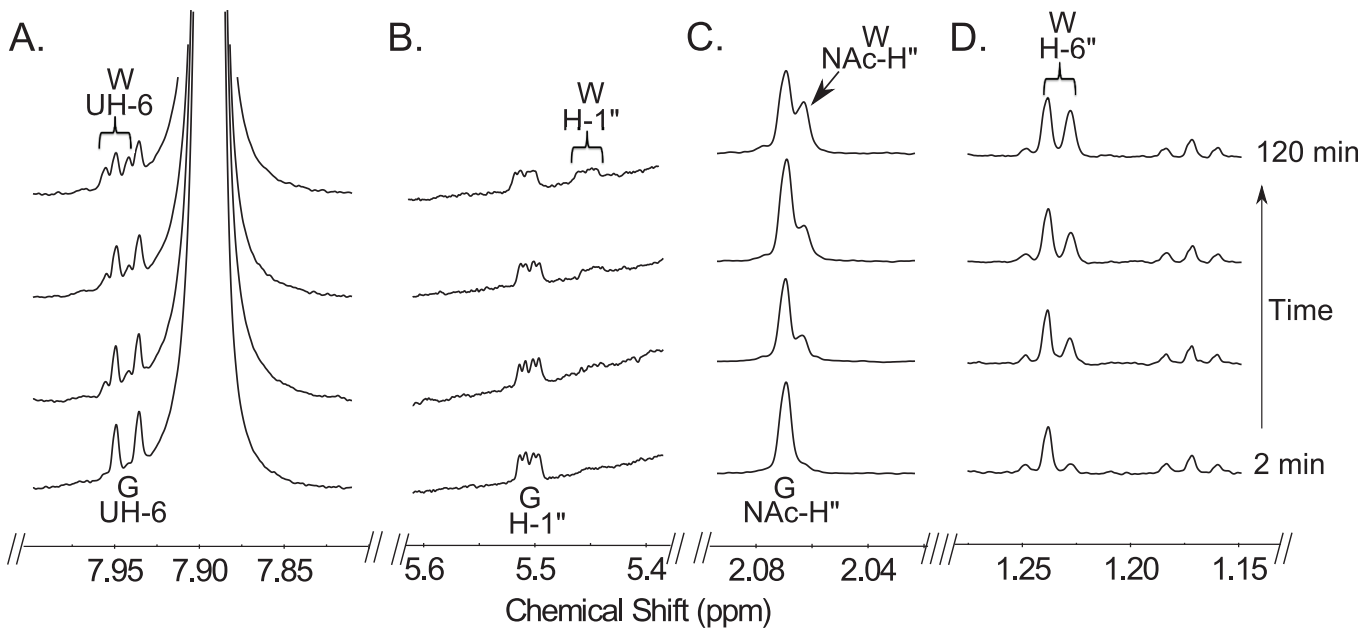


Fig 5. Time-resolved ¹H-NMR analysis of Pdeg reaction showing the conversion of UDP-GlcNAc to hydrated-4-keto-UDP-suagr. Spectra were collected for the first 120 min of the reaction that was conducted at 25°C and included Pdeg and 1 mM of UDP-GlcNAc (time 0). Four selected regions of the UDP-4-keto-6-deoxy-GlcNAc formed over time, can be observed with a diagnostic UH-6 (A), anomeric proton (B), acetyl methyl proton (C), and 6-deoxy proton H6' (D). Proton signals of UDP-GlcNAc and product, UDP-4-keto-6-deoxy-HexNAc, are labeled as GUH-6, GH-1'', GNAc-H'' for the substrate UDP-GlcNAc, WUH-6, WH-1'', WNAc-H'' and WH-6'' for the hydrated form, respectively.

doi:10.1371/journal.pone.0133790.g005

shown in Fig 6 (panel B, labeled Q) purified recombinant Preq in the presence of NADPH readily converted UDP-4-keto-6-deoxy-GlcNAc to a new UV-peak migrating on a HILIC column with a retention time of 14.9 min (Fig 6, panel B). This peak was not formed with unrelated protein (negative control, Fig 6, panel C). Negative ion LC-ESI-MS analyses of the new product gave a major ion at m/z 590 (Fig 6, Box top panel), suggesting the keto moiety is reduced. MS-MS analysis of m/z 591 gave ion fragments at m/z 402.94, 384.93 and 304.98 consistent with $[UDP]^-$, $[UDP-H_2O]^-$, and $[UMP-H_2O]^-$ respectively. The neutral loss of a 187 mass unit implies a mass for a 6-deoxy-HexNAc sugar. One-dimensional (Fig 7) and two-dimensional NMR (Fig 8) spectroscopic analyses of the isolated product provided compelling evidence that recombinant enzyme, Preq, formed UDP-D-QuiNAc. The chemical shift assignments for UDP-QuiNAc, are summarized in Table 2. NMR signals at 56.57 ppm for C_2 of the sugar moiety established that it is substituted by nitrogen atoms. A ¹H signal at 2.06 ppm also indicated that the C_2 position of the sugar was N-acetylated. The $J_{H_2'', H_3''}$ and $J_{H_3'', H_4''}$ coupling constants of 10 Hz and 10 Hz, respectively, are consistent with a gluco sugar. The chemical shift of 5.44 ppm of the anomeric proton of the QuiNAc residue and $J_{H_1'', p}$ and $J_{H_1'', H_2''}$ coupling constants of 7.18 and 3.0 Hz, respectively, are consistent with an α -linkage. Collectively, we concluded that Preq encodes an enzyme that reduces UDP-4-keto-6-deoxy GlcNAc to form UDP-QuiNAc.

Time-resolved ¹H NMR used to monitor the reaction progress of Preq is shown in Fig 9. As time progresses, the peak corresponding to the anomeric proton (WH-1'', see Fig 9A) of the UDP-4-keto-sugar substrate is decreased, whereas the peak of the enzymatic product, UDP-QuiNAc (QH-1''), is increased. It is also important to note that the UDP-4-keto intermediate (the product of recombinant Pdeg) is also disappearing during the formation of UDP-QuiNAc.

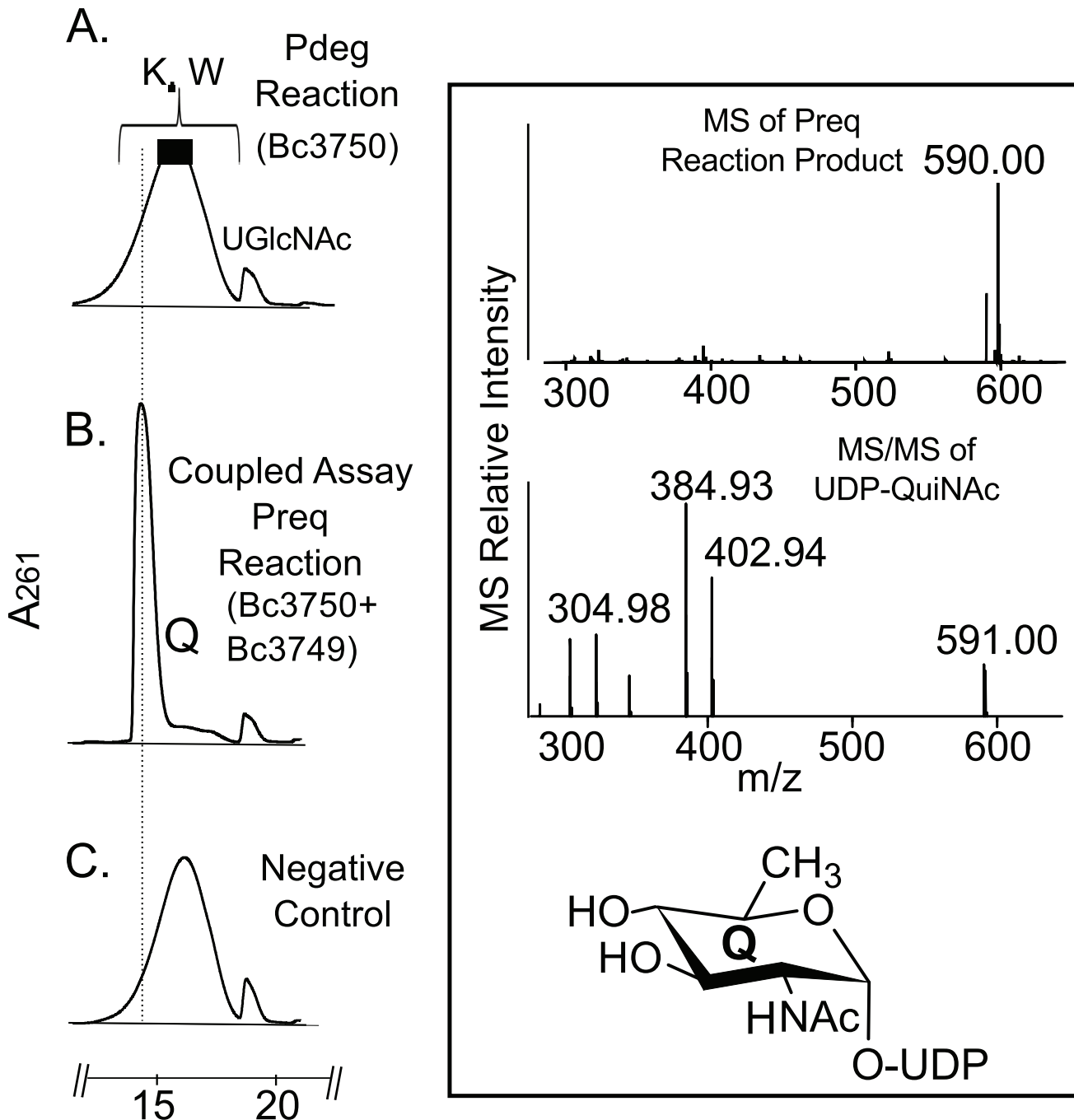


Fig 6. Analysis of Preq recombinant enzyme reaction by UV-HPLC and LC-ESI-MS-MS. A. Pdeg reaction shows the conversion of UDP-GlcNAc to UDP-4-keto-6-deoxy-HexNAc before recombinant Preq addition. B. Following Pdeg reaction, purified Preq was added and reacted with UDP-4-keto-6-deoxy-GlcNAc in the presence of NADPH to give product Q, UDP-QuiNAC, with m/z 590 and ms/ms of 402.94, 384.93, and 304.98 (Boxed top and second panel) C. Preq negative control reaction carried out with unrelated protein.

doi:10.1371/journal.pone.0133790.g006

The distinct chemical shifts for the proton signals belonging to the methyl group of the substrate and product are also an indication of the enzymatic reaction's progress with the 6-deoxy (WH-6") substrate peak decreasing at 1.23 ppm and the product methyl peak (QH-6") of the QuiNAC increasing at 1.29 ppm (Fig 9B). Diagnostic peak QH-4" of UDP-QuiNAC also

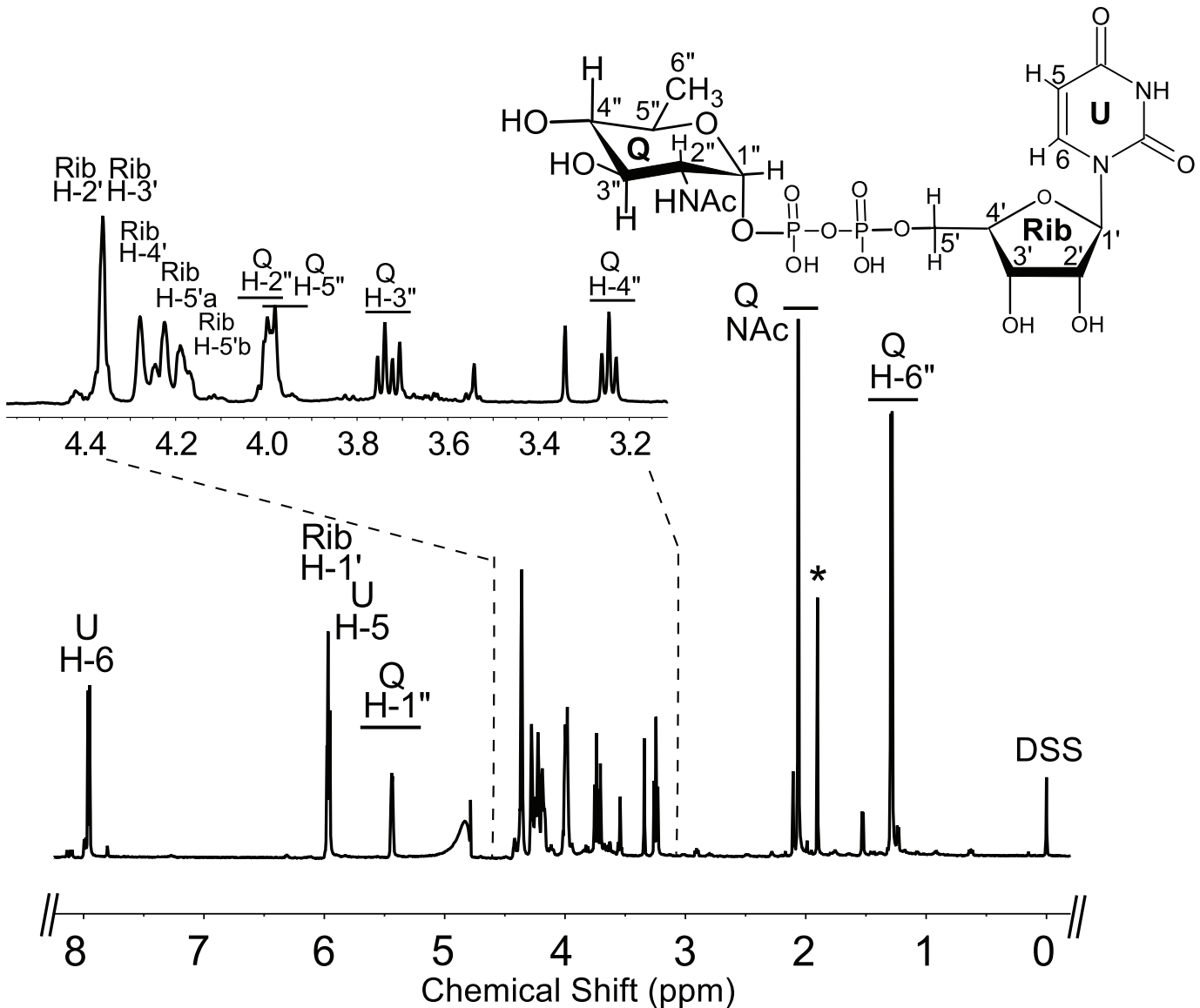


Fig 7. Analysis of Preq reaction product by ¹H-NMR indicates formation of UDP-QuiNac. The product of the Preq reaction (peak Q, in Fig 6 Panel B) was collected and analyzed at 600 MHz NMR. Full proton spectrum of HPLC-collected product UDP-QuiNac. Expanded proton spectra between 3.2 and 4.4 ppm that shows the QuiNac sugar ring. The short line above NMR 'peaks' denotes specific chemical shifts belonging to a UDP-QuiNac.

doi:10.1371/journal.pone.0133790.g007

appears around 3.25 ppm (Fig 9C) while WH-2'' and WH-5'' of 4-keto-6-deoxy GlcNAc disappears (Fig 9D).

Enzymatic Properties of Pdeg and Preq

The recombinant Pdeg 4,6-dehydratase had its highest activity between 22 and 25°C and at pH 8 and 9 irrespective of the buffer used. Similar pH and temperature profiles were observed for Preq. Kinetics parameters for the recombinant Pdeg and Preq activities are summarized in Table 3. Recombinant Pdeg eluted from a Superdex 75 size-exclusion column in the region for a protein with a mass of ~26.7 kDa, suggesting that the enzyme is active as a monomer. Similarly, recombinant Preq eluted from the same column in the region for a protein with mass of

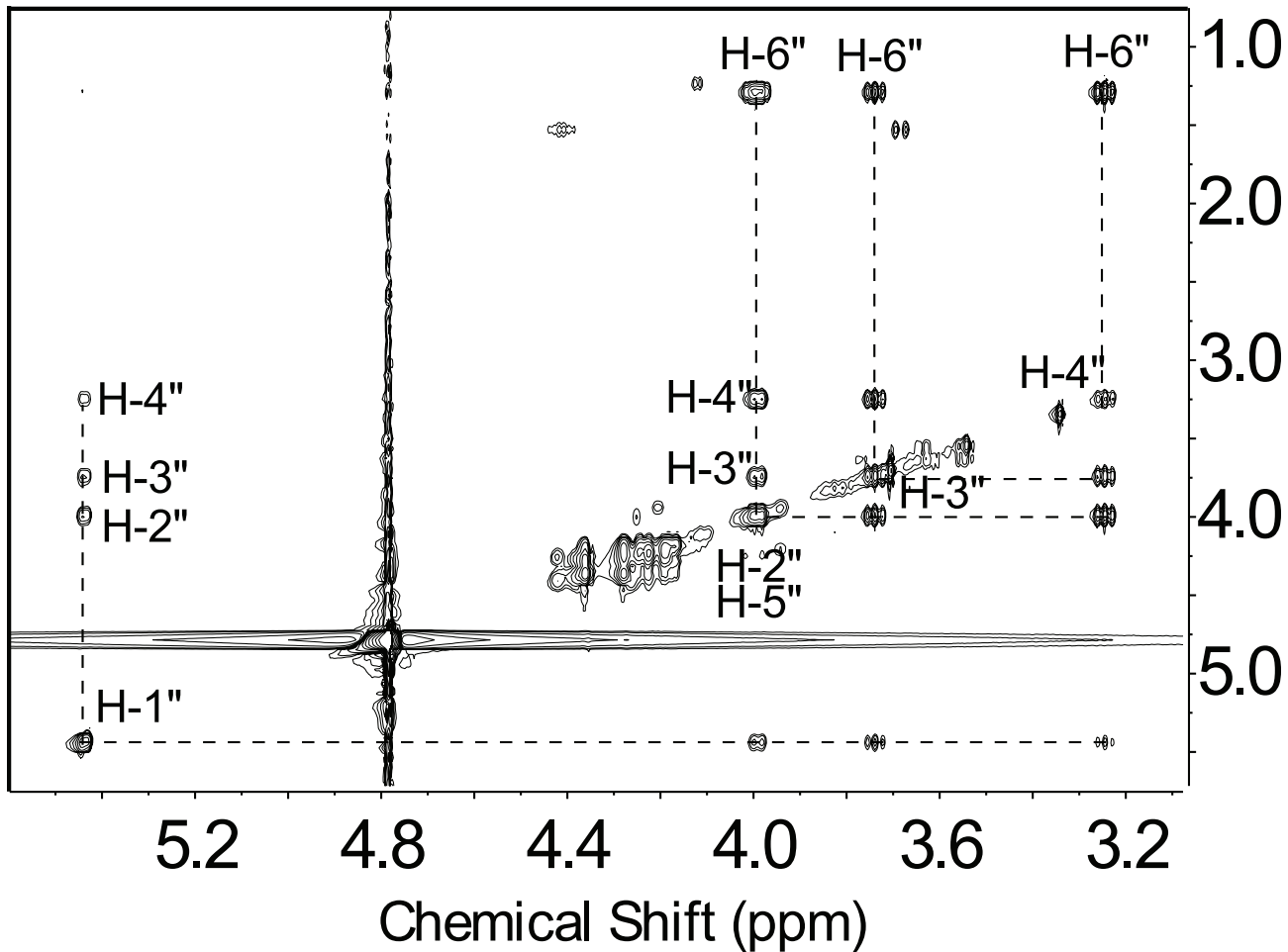


Fig 8. TOCSY spectrum of UDP-QuiNAc shows the coupling among QuiNAc sugar ring protons.

doi:10.1371/journal.pone.0133790.g008

~26.8 kDa, implying this enzyme is active predominantly as a monomer. Further kinetics analyses of the recombinant Pdeg and Preq activities are summarized in [Table 3](#). The apparent K_m values were 486.1 μM and 1,544 μM for UDP-sugar substrates, the V_{max} values were 7.4 $\mu\text{M min}^{-1}$ and 2.3 $\mu\text{M min}^{-1}$, and the k_{cat}/K_m values were 38.3 $\mu\text{M}^{-1}\text{min}^{-1}$ and 15.1 $\mu\text{M}^{-1}\text{min}^{-1}$ with Pdeg and Preq, respectively. None of the UDP-sugars tested including UDP-galactose, UDP-GalNAc, UDP-GlcNAc, and UDP-glucose are substrates for recombinant Pdeg. Inhibition studies showed that UDP-galactose, UDP, UDP-glucose, UDP-GalNAc, and UTP reduced recombinant Pdeg activity by 51%, 77%, 81%, 91%, 98% respectively.

Discussion

In this study, we have identified two genes Preq and Pdeg (Bc3749 and Bc3750) in *B. cereus* ATCC 14579 that encode the enzymes capable of converting UDP-GlcNAc to UDP-QuiNAc ([Fig 1](#)). The first step is initiated by Bc3570 a specific C4,6-dehydratase, Pdeg, that generates the 4-keto sugar-nucleotide intermediate, UDP-4-keto-6-deoxy-GlcNAc. This intermediate exists in two forms: hydrated and keto forms. Existing in two forms appears to be a common feature observed with other 4-keto nucleotide-sugar derivatives including UDP-4-keto-6-deoxy-glucose [[34](#)], UDP-4-keto-6-deoxy-AltNAc [[35](#)], and UDP-4-keto-xylose [[36](#)]. At steady state the ratio

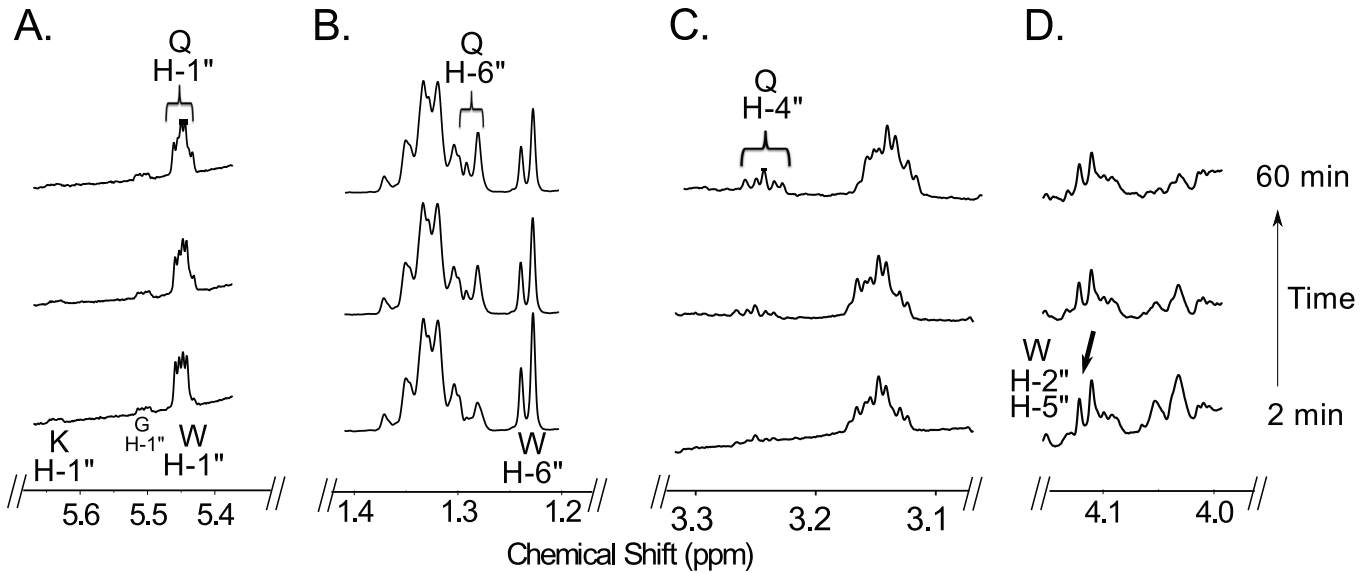


Fig 9. Time-resolved ¹H-NMR analysis of Preq reaction showing the conversion of hydrated-UDP-4-keto-6-deoxy-GlcNAc to UDP-QuiNAC. Spectra were collected for the first 60 min of the reaction that was conducted at 25°C and included Preq and 1 mM NADPH after Pdeg reaction. Four selected regions of the UDP-QuiNAC formed over time, can be observed with a diagnostic anomeric proton QH-1'' (A), 6-deoxy proton QH-6'' (B), appearance of QH-4''. Proton signals of UDP-QuiNAC are labeled as QH-1'', QH-6'', and QH-4'' and disappearance of WH-2'' and WH-5'' (D). Note that the chemical shift of the anomeric proton of UDP-QuiNAC (QH-1'') is very close to the proton of UDP-4-keto-6-deoxy-GlcNAc (slightly shifted). WH-1'' is the hydrated form of UDP-4-keto-6-deoxy-GlcNAc.

doi:10.1371/journal.pone.0133790.g009

Table 3. Enzymatic properties of recombinant Bc3750-His₆ (Pdeg) and Bc3749-His₆ (Preq).

	Pdeg	Preq
Optimal pH ^a	8.0–9.0	8.0–9.0
Optimal temperature (°C) ^a	22–25	22
K _m (μM) ^b	486.1 ^c	1544 ^c
V _{max} (μM·min ⁻¹)	7.4 ^c	2.3 ^c
k _{cat} (min ⁻¹)	1.86x10 ⁴	2.34x10 ⁴
k _{cat} /K _m (μM ⁻¹ min ⁻¹)	38.3 ^c	15.1 ^c
Protein Mass (kDa)	37.2	35.0

^a Optimal pH and temperature assays were determined using Tris-HCl; the activity in pH 9 and 8 vary by less than 10%.

^b The reaction was determined by HPLC-UV after 15 min at 22°C incubation for Pdeg and 105 s at 22°C incubation for Preq. For Pdeg assays, the reaction consisted of various concentration of UDP-GlcNAc (10, 20, 40, 80, 100, 160, 200, 300, 400, 500, 600, 700, 800, 900, and 1000 μM) with fixed amount of co-factor (200 μM NADPH) and 0.75 ng recombinant Pdeg. For Preq assays, the reaction included various concentrations of UDP-4-keto-6-deoxy-GlcNAc (79, 237, 395, 553, 711, 869, 1027, 1185, 1343, 1501, 1659, 1817, 1975, 2133, 2291, 2449, 2607, 2765, 2923, 3081, and 3239 μM) with fixed amount of co-factor (2 mM NADPH) and 4 pg recombinant Preq.

^c The K and V values were derived by fitting enzyme kinetic curves with GraphPad Prism 5. Data were fitted with the best curve and sum of square calculated as 3.31 (for Pdeg) and 0.065 (Preq) and the relative standard deviation (RSDR) of Pdeg reaction and Preq reaction was calculated as 0.40 and 0.05, respectively.

doi:10.1371/journal.pone.0133790.t003

of 4-hydrated to 4-keto is ~ 9:1. Hence, to distinguish the hydrated from the 4-keto-sugar-nucleotide by a typical proton NMR experiment is not obvious since no protons are detected at C-4 (keto or hydrated). Thus, one may report the chemical shifts for the hydrated rather the keto-form. An easy way to detect these two chemical forms is to carry the enzyme reaction in a time-resolved fashion or to analyze the enzyme reaction with a mass spectrometer (see Fig 3). Following the formation of the 4-keto nucleotide sugar, the second step catalyzed by Preq (Bc3749), reduces the C4-keto sugar moiety and converts it in an NADPH-dependent reaction to UDP-QuiNAc (Fig 9). The analog NADH could not be substituted for NADPH.

Despite the high amino acid sequence identity (46%) between the 4,6-dehydratases, Pdeg and PseB (*C. jejuni*), our experimental data shows that Pdeg does not have a 5'-epimerase) activity. While the amino acid sequence of the 4-reductase, Preq, indicated some similarities with proteins annotated as an epimerase, our biochemical analyses of the Preq enzymatic product data showed a 4-reductase activity forming only UDP-QuiNAc. The 4-epimer UDP-D-FucNAc was not produced by Preq under the experimental setup. When considering the data together we propose that the combined activities of Pdeg and Preq (as shown in Fig 1) contributes to synthesis of the UDP-QuiNAc in *B. cereus* and very likely in other *Bacillus* sp like *Bacillus weihenstephanensis* FSL R5-860, *Bacillus cereus* ATCC 10876, *Bacillus thuringiensis* serovar *israelensis* ATCC 35646, and *Bacillus thuringiensis* serovar *kurstaki* str. HD-1 strains.

While this work was in progress, a study in the gram-negative, *Rhizobium etli*, discovered for the first time that this bacterium has two enzyme activities [32] identical to those as described here for the gram-positive bacterium, *Bacillus*. Both the rhizobium and *Bacillus* proteins have the same 4,6-dehydratase and 4-reductase activities to make UDP-QuiNAc. Interestingly, the 4-reductases from *Rhizobium etli* (YP_470339) and *Pseudomonas aeruginosa* O6 (WbpV, AAF23991.1) [37], share no sequence homology with the *Bacillus* enzyme. Furthermore, several UDP-GlcNAc-C4,6-dehydratases lacking 5'-epimerase activity from *C. jejuni*, *N. gonorrhoeae*, and *A. baumannii* share much lower sequence homology with Pdeg enzyme (Table 1A). Current work is underway to determine the crystal structures of these *Bacillus* enzymes.

This illustrates that biochemical analysis is essential to assign function, and protein sequence homology is an insufficient criterion to infer functional specificity. The fact that *Rhizobium* 4-reductase requires NADH for its activity while Preq requires NADPH, could not explain in part the differences in primary sequences. Therefore, it is more likely that the genes encoding these activities in gram-positive and gram-negative bacteria are evolutionarily not related and perhaps arose from different ancestral genes.

Supporting Information

S1 Fig. Glycosyl residue compositions of QuiNAc-containing glycans in *Bacillus* sp. GC spectrum of alditol-acetates derived 6-deoxy-2-N-acetylhexosamine sugar residue eluted from GC-column at 29.7 min. The left box insert shows the electron impact mass spectrum and fragmentation (EI-MS) of this peak including prominent fragment ions at m/z 302, 260, 201, 145, 129, 103, and 85 identical with those found for alditol acetates derivatives of QuiNAc std. The right box insert shows the predicted primary and secondary MS fragments of C1-deuterated alditol-acetate of derived QuiNAc.

(TIFF)

S2 Fig. TOCSY spectrum of UDP-4-keto-6-deoxy-GlcNAc shows the coupling among keto sugar ring protons.

(TIFF)

Author Contributions

Conceived and designed the experiments: SH MB-P. Performed the experiments: SH AA. Analyzed the data: SH AA MB-P. Contributed reagents/materials/analysis tools: MB-P. Wrote the paper: SH AA MB-P.

References

1. Bretthauer RK. Structure, expression, and regulation of UDP-GlcNAc: dolichol phosphate GlcNAc-1-phosphate transferase (DPAGT1). *Curr Drug Targets*. 2009; 10:477–482. PMID: [19519349](#)
2. Stanley P SH, Taniguchi N. *Essentials of Glycobiology*. 2nd edition., Vol Chapter 8 (Cold Spring Harbor Laboratory Press). 2009.
3. Park JT, Uehara T. How bacteria consume their own exoskeletons (turnover and recycling of cell wall peptidoglycan). *Microbiol Mol Biol Rev*. 2008; 72:211–227, table of contents. doi: [10.1128/MMBR.00027-07](#) PMID: [18535144](#)
4. Raetz CR, Whitfield C. Lipopolysaccharide endotoxins. *Annu. Rev. Biochem*. 2002; 71:635–700. PMID: [12045108](#)
5. Sharon N, Shif I, Zehavi U. The isolation of D-fucosamine (2-amino-2,6-dideoxy-D-galactose) from polysaccharides of *Bacillus*. *Biochem. J*. 1964; 93:210–214. PMID: [4953793](#)
6. Bowser DV, Wheat RW, Foster JW, Leong D. Occurrence of quinovosamine in lipopolysaccharides of *Brucella* species. *Infect Immun*. 1974; 9:772–774. PMID: [4207190](#)
7. Sonesson A, Jantzen E, Bryn K, Larsson L, Eng J. Chemical composition of a lipopolysaccharide from *Legionella pneumophila*. *Arch Microbiol*. 1989; 153:72–78. PMID: [2610584](#)
8. Hrabak EM, Urbano MR, Dazzo FB. Growth-phase-dependent immunodeterminants of *Rhizobium trifolii* lipopolysaccharide which bind trifoliin A, a white clover lectin. *J Bacteriol*. 1981; 148:697–711. PMID: [6170630](#)
9. Forsberg LS, Bhat UR, Carlson RW. Structural characterization of the O-antigenic polysaccharide of the lipopolysaccharide from *Rhizobium etli* strain CE3. A unique O-acetylated glycan of discrete size, containing 3-O-methyl-6-deoxy-L-talose and 2,3,4-tri-O-,methyl-I fucose. *J. Biol. Chem*. 2000; 275:18851–18863. PMID: [10858446](#)
10. Carlson RW, Kalembasa S, Turowski D, Pachori P, Noel KD. Characterization of the lipopolysaccharide from a *Rhizobium phaseoli* mutant that is defective in infection thread development. *J Bacteriol*. 1987; 169:4923–4928. PMID: [3667520](#)
11. Noel KD, Vandenbosch KA, Kulpaca B. Mutations in *Rhizobium-Phaseoli* That Lead to Arrested Development of Infection Threads. *J Bacteriol*. 1986; 168:1392–1401. PMID: [3782040](#)
12. Cava JR, Elias PM, Turowski DA, Noel KD. *Rhizobium-Leguminosarum* Cfn42 Genetic Regions Encoding Lipopolysaccharide Structures Essential for Complete Nodule Development on Bean-Plants. *J Bacteriol*. 1989; 171:8–15. PMID: [2644215](#)
13. Forsberg LS, Noel KD, Box J, Carlson RW. Genetic locus and structural characterization of the biochemical defect in the O-antigenic polysaccharide of the symbiotically deficient *Rhizobium etli* mutant, CE166. Replacement of N-acetylquinovosamine with its hexosyl-4-ulose precursor. *J. Biol. Chem*. 2003; 278:51347–51359. PMID: [14551189](#)
14. Morrison MJ, Imperiali B. The renaissance of bacillosamine and its derivatives: pathway characterization and implications in pathogenicity. *Biochemistry*. 2014; 53:624–638. doi: [10.1021/bi401546r](#) PMID: [24383882](#)
15. Ito M, Aono R, Horikoshi K. Identification of 2-amino-2,6-dideoxy-D-glucose (D-quinovosamine), isolated from the cell walls of the alkaliphilic *Bacillus* sp. Y-25, by 500-MHz 1H NMR spectroscopy. *Carbohydr. Res*. 1993; 242:173–180. PMID: [8495438](#)
16. Arnesen LPS, Fagerlund A, Granum PE. From soil to gut: *Bacillus cereus* and its food poisoning toxins. *Fems Microbiol Rev*. 2008; 32:579–606. doi: [10.1111/j.1574-6976.2008.00112.x](#) PMID: [18422617](#)
17. Bottone EJ. *Bacillus cereus*, a volatile human pathogen. *Clin. Microbiol. Rev*. 2010; 23:382–398. doi: [10.1128/CMR.00073-09](#) PMID: [20375358](#)
18. Gu XG, Glushka J, Lee SG, Bar-Peled M. Biosynthesis of a New UDP-sugar, UDP-2-acetamido-2-deoxyxylose, in the Human Pathogen *Bacillus cereus* Subspecies cytotoxis NVH 391–98. *J. Biol. Chem*. 2010; 285:24825–24833. doi: [10.1074/jbc.M110.125872](#) PMID: [20529859](#)
19. Rance M, Sorensen OW, Bodenhausen G, Wagner G, Ernst RR, Wuthrich K. Improved Spectral Resolution in COSY 1H-NMR Spectra of Proteins Via Double Quantum Filtering. *Biochem. Biophys. Res. Com*. 1983; 117:479–485.

20. Braunschweiler L, Ernst RR. Coherence transfer by isotropic mixing—application to proton correlation spectroscopy. *J Magn Reson.* 1983; 53:521–528.
21. Bodenhausen G, Ruben DJ. Natural Abundance N-15 NMR by Enhanced Heteronuclear Spectroscopy. *Chem. Phys. Lett.* 1980; 69:185–189.
22. Marion D, Driscoll PC, Kay LE, Wingfield PT, Bax A, Gronenborn AM, et al. Overcoming the Overlap Problem in the Assignment of 1H-NMR Spectra of Larger Proteins by Use of 3-Dimensional Heteronuclear ¹H-¹⁵N Hartmann-Hahn Multiple Quantum Coherence and Nuclear Overhauser Multiple Quantum Coherence Spectroscopy—Application to Interleukin-1-Beta. *Biochemistry.* 1989; 28:6150–6156. PMID: [2675964](#)
23. Ishiyama N, Creuzenet C, Miller WL, Demendi M, Anderson EM, Harauz G, et al. Structural studies of FlaA1 from *Helicobacter pylori* reveal the mechanism for inverting 4,6-dehydratase activity. *J. Biol. Chem.* 2006; 281:24489–24495. PMID: [16651261](#)
24. Miyafusa T, Caaveiro JM, Tanaka Y, Tanner ME, Tsumoto K. Crystal structure of the capsular polysaccharide synthesizing protein CapE of *Staphylococcus aureus*. *Biosci Rep.* 2013; 33:
25. De Castro C, Molinaro A, Piacente F, Gurmon JR, Sturiale L, Palmigiano A, et al. Structure of N-linked oligosaccharides attached to chlorovirus PBCV-1 major capsid protein reveals unusual class of complex N-glycans. *Proc. Natl. Acad. Sci. USA.* 2013; 110:13956–13960. doi: [10.1073/pnas.1313005110](#) PMID: [23918378](#)
26. Osipiuk J, Marshall N, Tesar C, Pearson L, Buck K, Joa A. pdb 3PVZ, unpublished.
27. Olivier NB, Chen MM, Behr JR, and Imperiali B. In vitro biosynthesis of UDP-N,N'-diacetylbaucillosamine by enzymes of the *Campylobacter jejuni* general protein glycosylation system. *Biochemistry.* 2006; 45:13659–13669. PMID: [17087520](#)
28. Hartley MD, Morrison MJ, Aas FE, Borud B, Koomey M, Imperiali B. Biochemical characterization of the O-linked glycosylation pathway in *Neisseria gonorrhoeae* responsible for biosynthesis of protein glycans containing N,N'-diacetylbaucillosamine. *Biochemistry.* 2011; 50:4936–4948. doi: [10.1021/bi2003372](#) PMID: [21542610](#)
29. Morrison MJ, Imperiali B. Biosynthesis of UDP-N,N'-diacetylbaucillosamine in *Acinetobacter baumannii*: Biochemical characterization and correlation to existing pathways. *Arch Biochem Biophys.* 2013; 536:72–80. doi: [10.1016/j.abb.2013.05.011](#) PMID: [23747578](#)
30. King JD, Poon KKH, Webb NA, Anderson EM, McNally DJ, Brisson JR, et al. The structural basis for catalytic function of GMD and RMD, two closely related enzymes from the GDP-d-rhamnose biosynthesis pathway. *Febs Journal.* 2009; 276:2686–2700. doi: [10.1111/j.1742-4658.2009.06993.x](#) PMID: [19459932](#)
31. Allard STM, Giraud MF, Whitfield C, Messner P, Naismith JH. The purification, crystallization and structural elucidation of dTDP-D-glucose 4,6-dehydratase (RmIB), the second enzyme of the dTDP-L-rhamnose synthesis pathway from *Salmonella enterica* serovar Typhimurium. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2000; 56:222–225.
32. Li T, Simonds L, Kovrigin EL, Noel KD. In vitro biosynthesis and chemical identification of UDP-N-acetyl-d-quinovosamine (UDP-d-QuiNAc). *J. Biol. Chem.* 2014; 289:18110–18120. doi: [10.1074/jbc.M114.555862](#) PMID: [24817117](#)
33. Bubb WA. NMR spectroscopy in the study of carbohydrates: Characterizing the structural complexity. *Concept Magn Reson A.* 2003; 19A:1–19.
34. Martinez V, Ingwers M, Smith J, Glushka J, Yang T, Bar-Peled M. Biosynthesis of UDP-4-keto-6-deoxy-glucose and UDP-rhamnose in pathogenic fungi *Magnaporthe grisea* and *Botryotinia fuckeliana*. *J. Biol. Chem.* 2012; 287:879–892. doi: [10.1074/jbc.M111.287367](#) PMID: [22102281](#)
35. Li Z, Hwang S, Ericson J, Bowler K, Bar-Peled M. Pen and Pal are nucleotide-sugar dehydratases that convert UDP-GlcNAc to UDP-6-deoxy-D-GlcNAc-5,6-ene and then to UDP-4-keto-6-deoxy-L-AltNAc for CMP-pseudaminic acid synthesis in *Bacillus thuringiensis*. *J. Biol. Chem.* 2015; 290:691–704. doi: [10.1074/jbc.M114.612747](#) PMID: [25414257](#)
36. Gu X, Glushka J, Yin Y, Xu Y, Denny T, Smith J, et al. Identification of a bifunctional UDP-4-keto-pentose/UDP-xylose synthase in the plant pathogenic bacterium *Ralstonia solanacearum* strain GM1000, a distinct member of the 4,6-dehydratase and decarboxylase family. *J. Biol. Chem.* 2010; 285:9030–9040. doi: [10.1074/jbc.M109.066803](#) PMID: [20118241](#)
37. Belanger M, Burrows LL, Lam JS. Functional analysis of genes responsible for the synthesis of the B-band O antigen of *Pseudomonas aeruginosa* serotype O6 lipopolysaccharide. *Microbiology.* 1999; 145 (Pt 12):3505–3521. PMID: [10627048](#)