

DISCOVERY OF NOVEL GLYCANS AND NUCLEOTIDE SUGAR
BIOSYNTHETIC PATHWAYS IN BACILLUS SPECIES

by

ZI LI

(Under the Direction of Maor Bar-Peled)

ABSTRACT

The spore-forming Gram-positive bacteria of the genus *Bacillus* is present in diverse natural environments. Among *Bacilli*, we show that members of the *Bacillus cereus sensu lato* including *B. cereus*, *B. anthracis*, and *B. thuringiensis* produce a myriad of glycan structures. This thesis presents the discovery and elucidation of a few biosynthetic pathways leading to the formation of these glycans in spores, flagella, and biofilm. In *Bacillus thuringiensis israelensis* ATCC 35646, we found flagellin to be O-glycosylated with pseudaminic acid, and biochemically characterized seven enzymes encoded by the operon PSE, that forms the precursor CMP-pseudaminic acid. Among the seven enzymes, the first two showed unusual enzymatic activities, in the conversion of UDP-D-GlcNAc to UDP-6-deoxy-D-GlcNAc-5,6-ene and then to UDP-4-keto-6-deoxy-L-AltNAc. In other bacterial system studied, these two different activities are catalyzed by a single enzyme.

This work also identified a novel exopolysaccharide (EPS) produced during spore formation in many members of *Bacillus cereus sensu lato* species. This EPS abbreviated pZx, may have roles in spore adherence, dispersal, or anti-aggregation, depending on the pH. The chemical structure of this EPS was characterized by NMR and GC-SM and consists of two

uncommon acetamido sugars (XylNAc and GlcNAcA) that form a backbone of $[-3)\text{XylNAc}4\text{OAc}(\alpha1-3)\text{GlcNAcA}4\text{OAc}(\alpha1-3)\text{XylNAc}(\alpha1-)]_n$. The biosynthetic operon XNAC for the production of UDP-XylNAc, UDP-GlcNAcA and the glycosyltransferases were shown to be involved in the formation of this EPS.

Lastly, this work found that spores of *B. cereus* ATCC 14579 and ATCC 10876 are decorated with two rare sugar epimers of C3-methyl-6-deoxy-hexoses. We also identified C3CM operon that is involved in the formation of such decoration and elucidated the metabolic pathway that forms CDP-glucose and transforms it in three enzymatic steps to CDP-C3-methyl-6-deoxy-gulose and CDP-C3-methyl-6-deoxy-allose. Subsequently, we used genetic approaches to find that this operon is required for the addition of these sugars to spore glycoproteins. Mutants impaired with genes in the C3CM operon gave spores that germinate faster. Knowledge of the glycans found on spores, flagella and the extracellular matrix provides a foundation for investigating the roles of glycosylation in motility, survival and pathogenicity, and may explain some physiological differences observed between members of the *Bacilli* species.

INDEX WORDS: *Bacillus*, glycobiology, enzymology, CMP-pseudaminic acid, sporulation, polysaccharides, biofilm, biosurfactant, adherence, exopolysaccharides, spore

DISCOVERY OF NOVEL GLYCANS AND NUCLEOTIDE SUGAR
BIOSYNTHETIC PATHWAYS IN BACILLUS SPECIES

by

ZI LI

BS, Cornell University, 2011

BS, China Agricultural University, CHINA, 2011

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2017

© 2017

Zi Li

All Rights Reserved

DISCOVERY OF NOVEL GLYCANS AND NUCLEOTIDE SUGAR
BIOSYNTHETIC PATHWAYS IN BACILLUS SPECIES

by

ZI LI

Major Professor:	Maor Bar-Peled
Committee:	Zachery Lewis
	Wolfgang Lukowitz
	Robert Maier
	Michael Tiemeyer

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2017

TABLE OF CONTENTS

CHAPTER	Page
1 INTRODUCTION AND LITERATURE REVIEW	1
1.1 <i>Bacillus</i> Species Overview	1
1.2 <i>Bacillus</i> Cell Wall Structure in Vegetative State	6
1.3 Sporulation and Spore Structure	10
1.4 Exopolysaccharide and Role in Biofilm Formation.....	16
1.5 References.....	21
2 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	32
2.1 Abstract.....	33
2.2 Introduction.....	34
2.3 Experimental Procedure.....	37
2.4 Results.....	44
2.5 Discussion.....	64
2.6 References.....	68

3	DISCOVERY OF A UNIQUE EXTRACELLULAR POLYSACCHARIDE IN MEMBERS OF THE PATHOGENIC BACILLUS THAT CAN CO-FORM WITH SPORES	71
3.1	Abstract	72
3.2	Introduction	72
3.3	Results	77
3.4	Discussion	100
3.5	Experimental Procedure	104
3.6	References	112
4	A FOUR-GENE OPERON IN BACILLUS CEREUS PRODUCES TWO RARE SPORE-DECORATING SUGARS.	116
4.1	Abstract	117
4.2	Introduction	118
4.3	Results	120
4.4	Discussion	145
4.5	Experimental Procedure	148
4.6	References	159
5	CONCLUSION AND FUTURE DIRECTION	162
5.1	Introduction and Significance	162
5.2	Summary of Results	163
5.3	Future Direction	167
5.4	References	169

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1 *Bacillus* Species Overview

The *Bacillus* species are rod-shaped facultative anaerobes Gram-positive spore-forming bacteria. They belong to the family *Bacillaceae*. *Bacillus* is ubiquitous in nature, while commonly found in soil, some species inhabit in raw milk, water and ponds, caterpillars, and animal feces (1), and based on genomic sequences of different environmental niches some *Bacillus* were found in soil below lake or associated with roots and leaf surface (2).

Some *Bacilli* are associated with human. For example, *Bacillus subtilis* is found in the human gastrointestinal tract; and some *Bacillus*, particularly the members of the *B. cereus sensu lato* like *B. anthracis* and *B. cytotoxic* are harmful to human and other species. The *Bacillus cereus* group, also called *Bacillus cereus sensu lato* is a large bacterial species group that consists seven closely related bacilli of ecological, medical, and industrial importance. This group includes the insect pathogen *Bacillus thuringiensis*, which has been successfully applied as bio-pesticide (3); *Bacillus cereus sensu stricto*, which is associated with foodborne illness (4); *Bacillus mycoides* (5) and *Bacillus pseudomycoides* known to form characteristic rhizoidal colonies on agar (6); *Bacillus anthracis* which is the causing agent of anthrax (7); *Bacillus weihenstephanensis*, which is the psychrotolerant (capable of growing at temperatures close to freezing) (8); and *Bacillus cytotoxicus*, a thermotolerant organism (9).

To date, the genome of over 300 members of this group was sequenced, and their sequences have high levels of similarity (10). Despite their genetic similarity, some strains can

be distinguished based on their phenotypes. For instance, *B. mycoides* and *B. pseudomycoides* show distinct rhizoid growth whereas the rest of the species have round to irregular colonies (11); *B. weihenstephanensis* and *B. cytotoxicus* strains have different growth temperature range (9).

Due to the medical and economic impact, these *B. cereus sensu lato* species were attractive to study (12). For example, the crystalline protein toxin produced by *B. thuringiensis* with insecticidal properties has been used as environmental-friendly insecticides (13,14); *B. anthracis* is able to form anthrax toxin along with lethal toxin to cause tissue necrosis of human and livestock (15); The endospore formed by *B. anthracis* is highly resistant to heat, drying, and many disinfectants, as the result, many countries have been studying this organism as biological weapon (16); *B. cereus* is associated mainly with food poisoning, however, there are increasing reports showing that it can cause fatal non-gastrointestinal-tract infections like Endophthalmitis (an inflammation of the internal coats of the eye) (17).

Here I will briefly review three main species in the *Bacillus cereus* group that are more related to my research: *B. anthracis*, *B. cereus* and *B. thuringiensis*.

1.1.1 *Bacillus anthracis*

Bacillus anthracis is the causative agent of anthrax, a common disease of livestock and, occasionally, of humans. It is the only obligate pathogen within the genus *Bacillus* (18). Anthrax is a zoonotic disease, which is mainly associated with animals but can be transmitted from animals to people (19). The common form of the anthrax disease is cutaneous anthrax, which is presented by forming large, black skin lesions, whereas pulmonary anthrax and gastrointestinal anthrax were also observed at a low rate. In 1877, Robert Koch was the first to grow *B. anthracis* in pure culture and demonstrated its ability to form endospores that cause anthrax disease when

injected into animals. The first vaccine against *B. anthracis* with live organisms was developed and experimented on cows by Louis Pasteur (20). A number of anthrax vaccines have been developed ever since, however they are only for preventive use in livestock and humans (20). Antibiotic treatment like penicillin could be effective post-infection if the disease is diagnosed soon enough (21).

B. anthracis is an important organism to public health. In many countries like India where agriculture is the main source of income, anthrax disease occurred in domestic animals can cause significant economic loss especially for the areas with low vaccination rates (19). Even though anthrax disease can be cured by timely antibiotic treatment, it could be fatal in several cases because of lack of proper diagnosis (19). Because of the toxicity of the toxin produced by *Bacillus anthracis*, the spores are also used in biological warfare and bioterrorism (5). The Centers for Disease Control and Prevention (CDC) classified *B. anthracis* in high priority: Category A in the bioterrorism agent categories mainly because *B. anthracis* spores can persist in the environment for many years and it is very difficult to kill them (22). Despite the fact that anthrax is serious infectious disease, it is argued that anthrax is not an ideal biological weapon because a large number of spores in a fine powder form are needed to infect human. Nevertheless, more emphasis on *B. anthracis* is given for the potential emerging problem for public health.

B. anthracis utilizes two large plasmids (pXO1 and pXO2) to encode genes involved in virulence toxin and capsule. Strains lacking either of these plasmids have greatly reduced virulence. *Bacillus anthracis* secretes two virulence factors: a toxin and a poly- γ -d-glutamic acid capsule (γ -DPGA). The toxin has three components: a protective antigen, a lethal factor, and an edema factor (23). The channel-forming protective antigen oligomerizes to create a pre-channel

that forms toxic complexes upon binding the two other components, lethal factor and edema factor (24). These factors are translocated into the host cytosol to destroy white blood cells and increase cyclic AMP levels (25). The poly- γ -d-glutamic acid capsule, on the other hand, protects *B. anthracis* from phagocytosis of host cells by macrophages. Thus even though it is not toxic on its own, the capsule of *B. anthracis* is also a virulence factor for infection. In conjunction with anthrax toxin, the capsule allows *B. anthracis* to grow virtually unimpeded in the infected host (12,26).

1.1.2 *Bacillus cereus*

B. cereus may be the most common aerobic spore existing in many types of soil and in sediments, dust, and plants (17,27). In 1887, it was first discovered growing on agar plate left in a cow shed (28). More recently, *B. cereus* has been identified as the most common contaminant in pharmaceutical manufacturing (29).

B. cereus is naturally found in decaying organic matter, fresh and marine waters, and vegetables (30). Due to the adhesive nature of its spores, *B. cereus* is also frequently present in food production environments (31), which enables the bacterium to spread to improperly cooked food. Contaminated food transmits *B. cereus* spores into the intestinal tract of the human, or to livestock consumed by the human, causing foodborne illness (32).

Upon ingestion of *B. cereus* spores, toxins including hemolysin BL (Hbl), nonhemolytic enterotoxin (Nhe) and cytotoxin K (CytK) are produced. These toxins can cause illnesses such as severe nausea, vomiting, and diarrhea, depending on the specific strain (33,34). These toxins are pore-forming toxins that create pores on the cellular membrane of the small intestine, leading to the loss of cellular membrane potential and diarrheic symptoms (34).

Because they are so common in human surroundings, some *Bacillus cereus* strains are opportunistic human pathogens (35). Some *B. cereus* isolates are occasionally associated with infections, causing periodontal diseases and other more serious infections including endophthalmitis (eye), bacteremia (blood), endocarditis (heart's inner lining), cutaneous infections (skin, nails, and hair), pneumonia (lungs), and meningitis (brain and spinal cord membranes) (36). In the past 15 years, several *B. cereus* strains were reported to cause severe anthrax-like disease in humans (37,38) and apes (39). Some of these virulent strains while retaining *B. cereus* diagnostic phenotypes harbor plasmids similar to the toxin and capsule virulence plasmids pXO1 and pXO2 present in *B. anthracis* (36,40,41). Thus, increasing attention has been given to these pathogenic strains.

1.1.3 *Bacillus thuringiensis*

B. thuringiensis was first discovered in 1901 by Japanese scientist Ishiwata Shigetane and then isolated from flour moth caterpillars in 1911 (42). Currently, *B. thuringiensis* is best known for the use as an insecticide and to exist in the gut microflora of insects (43,44).

Like other species in the *B. cereus* group, *B. thuringiensis* can form endospores. Upon sporulation, crystals of insecticidal δ -endotoxins are produced in the mother cell compartment. These crystals are predominantly comprised of one or more proteins (Cry and Cyt toxins) (45). These toxins have specific activities against insect species of Lepidoptera (moths and butterflies), Diptera (flies and mosquitoes), Coleoptera (beetles), Hymenoptera (wasps, bees, ants and sawflies) and nematodes (46). When insects ingest these crystals, the alkaline nature of digestive tracts solubilizes the crystals to make them accessible to proteases (47). The toxins are cleaved by the proteases from the crystals and then inserted into the insect gut cell membrane,

paralyzing the digestive tract by forming a pore (13). This prevents the insect from eating, causing the host to death.

Because of the insecticidal nature, *B. thuringiensis* is widely used to control agricultural pests since the 1920s by simply spraying spores on the crops (48). Moreover, the *cry* genes, which encode Cry protein from *B. thuringiensis*, have been genetically engineered into various crops for them to obtain insect tolerance, such as tobacco, potato, maize, corn, and cotton (49). Interestingly, the Cry proteins have variable specificity to a variety of mosquitoes including *Anopheles gambiae*, the principal vector of malaria (50). As the result, *B. thuringiensis* is a potential biological agent to control mosquito-transmitted diseases.

1.2 *Bacillus* Cell Wall Structure in Vegetative State

1.2.1 Wall structure and related glycan

Bacillus cells proliferate by symmetric cell divisions to produce identical vegetative cells. Vegetative cell structure consists an inner membrane and a thick peptidoglycan layer that serve to maintain the cell shape (51). A single cell of *Bacillus anthracis* consists of cytoplasm surrounded by the plasma membrane and a cell wall structure (See Fig. 1.1) that includes peptidoglycan, S-layer, and capsules.

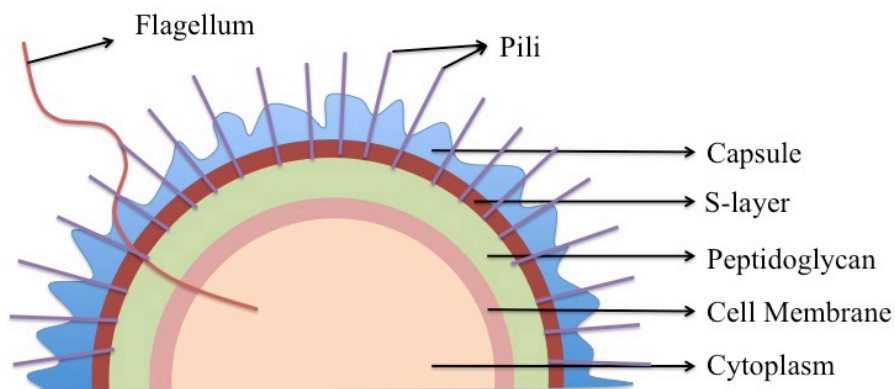


FIGURE 1.1. Illustration of a bacterial cell structure. (Structure based on Biology 10e Textbook (chapter 4, Pg: 63))

Peptidoglycan is essential for the maintenance of cellular viability and shape (52). Both S-layer and capsules, as the outermost surfaces of cells, can act as virulence factors to interact with hosts during infection (53). *Bacillus* strains have a very thick layer of peptidoglycan, which is composed of glycan chains of alternating *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) that are frequently cross-linked to each other by short peptides (54). This rigid polymer surrounds the bacterial cell and therefore determines the shape of the bacterial cells. It also protects the cells against the turgor pressure exerted by the cytoplasm (52,55).

The peptidoglycan layer in Gram-positive bacteria is also often modified with other cell-surface structures called secondary cell wall polymers (SCWPs). SCWPs include anionic polymers teichoic acid (TA) and neutral polysaccharide (SCWP polysaccharide). Both teichoic acids and the SCWP polysaccharide were also reported to be involved in biofilm formation (56,57). In *B. subtilis*, TA is a poly-phosphate polymer and plays an important role in membrane integrity (58,59). The carbohydrate composition varies in different species in *Bacillus subtilis*: strain 168, for example, has a poly-glycerol-phosphate backbone whereas a ribitol phosphate backbone is common in strain W23. TA is commonly decorated with glucose attached to the glycerol phosphate polymers (59). TA is not found in *B. cereus*, *B. anthracis* or *B. thuringiensis*, however, one early study reported the existence of teichoic acid in one strain of *B. cereus*, although the identity of that specific strain remains questionable (60).

SCWP polysaccharides are covalently bound to the MurNAc moieties of the peptidoglycan. In some Gram-positive bacteria like *Staphylococcus aureus*, SCWPs are essential and shown to be involved in the bacterial virulence (61). In many other Gram-positive bacteria, SCWPs are found to bind SLH (S-layer homology) domains to anchor those surface proteins

non-covalently to the cell wall (62). The structures of SCWPs found in *B. anthracis* (Repeat of 4)GlcNAc(β 1–6)GlcNAc(α 1–4)ManNAc(β 1) and *B. cereus* (repeat of 4)GlcNAc(β 1–6)GalNAc(α 1–4)ManNAc(β 1) have been fully characterized (63).

S-layers are present on the surfaces of many bacteria. They are self-assembled spontaneously as a bidimensional crystalline layer covering the entire cell surface (64). In *B. anthracis*, S-layer protein is made of Sap and EA1, which are encoded by an operon of two genes, *sap* and *eag* (65). S-layer proteins in many species have important roles in growth and survival, maintenance of cell integrity, and the interaction with the host and its immune system (66). For example in *B. anthracis*, S-layer protein has adhesin activity, to associate with BslA protein and binds to HeLa cells (67). In *Bacillus coagulans*, the crystalline layer formed by S-layer protein is a permeability barrier (68). Many glycosylated S-layer proteins have been discovered in numerous bacterial, S-layer in *Bacillus* also has glycoprotein component such as rhamnan and diaminouronic acid (69).

1.2.2 Surface structure

In the vegetative state, *Bacillus* species has two surface apparatus to facilitate movement: flagella and pili (See Fig. 1.1). Pili are thin protein filaments on the cell surface. In *B. cereus* or *B. anthracis* vegetative cell, two precursor proteins, BcpA and BcpB, assemble pili, which is assisted by a pilus-associated sortase enzyme (70,71). In some Gram-negative bacteria such as *Neisseria meningitides*, the pilus is a key virulence factor as being the major adhesion, contributing to specificity for the human host recognition (72). The structure of pili in *Neisseria* was well characterized, revealing that these structures are glycosylated post-translationally with an O-linked trisaccharide, Gal(β 1–4)Gal(α 1–3)2,4-diacetimid-2,4,6-trideoxyhexose (73).

However, whether pili are glycosylated or involved in virulence in *Bacillus* strains remains unknown.

Flagella, another important surface structure, have more complex structure. They facilitate two forms of active movement of bacteria, swimming and swarming motility (74,75). Swimming is exhibited by single cells with short flagellated rods, whereas swarming is a collective movement of swarm cells with flagellum that is three to four times longer, and also forty times more flagellated than single swimming cells (76). The structure of flagella includes three architectural domains: the basal body, which is embedded in the cell envelope; the hook, which is a flexible joint connected to the basal body; and the filament, which acts as a helical propeller. The filament consists of a repeating protein monomer called flagellin. Flagellin is a globular protein monomer that stacks helically in the form of a hollow cylinder to build the filament of the flagellum (34,77,78). The N- and C-terminal sequences of flagellin are somehow highly conserved across bacterial species including Gram-negative *Helicobacter* and Gram-positive *Bacillus*; however the central regions of those species have a highly variable length as well as amino acid composition (79). Several studies have shown that flagella are essential for motility, adhesion, host interactions, and secretion during many of the life stages of Gram-negative bacteria (80,81).

In some Gram-negative bacteria, flagellin is glycosylated (82,83). In *Campylobacter* spp. and *Helicobacter pylori*, the central domain of flagellin is post-translationally modified by the addition of pseudaminic acid (5,7-bis(acetylamino)-3,5,7,9-tetradecoxy-L-glycero- α -L-manno-2-nonulopyranosonic acid, Pse) (84,85) that are *O*-linked to serine or threonine residues (86). Our understanding of flagellin glycosylation in *bacillus* is minimal. Studies of Gram-positive *Bacillus* sp. PS3 flagellin have shown that it is *O*-glycosylated and no *N*-linked glycan was detected from

PNGase F digest (87,88), although the sugar involved in this modification was not identified. Recently we have found that in *B. thuringiensis israelensis* the flagellin is glycosylated with Pse as well, however, the role of Pse in *Bacillus* is unclear. In Gram-negative bacteria, mutants lacking Pse glycosylation of flagellin are affected in both their motility and their ability to infect their host (89,90). However, it is not clear why the loss of Pse gives rise to immobility and reduces infectivity, specifically when flagellin of a large family Enterobacteriaceae is not glycosylated (82,83), including mobile and infectious strains *Salmonella* and *Escherichia coli*. It is also unclear if glycosylation of flagellin is required for flagella assembly, for quality control, for protection against host proteolytic enzymes, to avoid host immune surveillance, or to be recognized by host receptors.

1.3 Sporulation and Spore Structure

1.3.1 Sporulation

Common to members of the *B. cereus* group is the formation of endospores. In an adaptive response to nutritional stress, vegetative cells divide asymmetrically resulting in two genetically identical cells with different cell fates (91-93). The mother cell lyses upon the completion of sporulation to release mature spores, which are highly resilient to heat and chemicals and have the ability to survive for long duration (94). Oxygen is also needed for sporulation in many species including *B. anthracis* (95), *B. thuringiensis* (96) and more recently *B. cereus* (97). This could be used as a constraint to control sporulation, which has important application for epidemiology in pathogenic strains. Although many studies have been performed on spores and sporulation of *B. cereus*, most sporulation studies were focusing on *Bacillus subtilis* due to complete genome accessibility and easy handling (98). Here I will provide a brief

overview of different stages of sporulation and the molecular mechanisms behind the initiation of sporulation processes. This knowledge was obtained mainly from studies of *B. subtilis*.

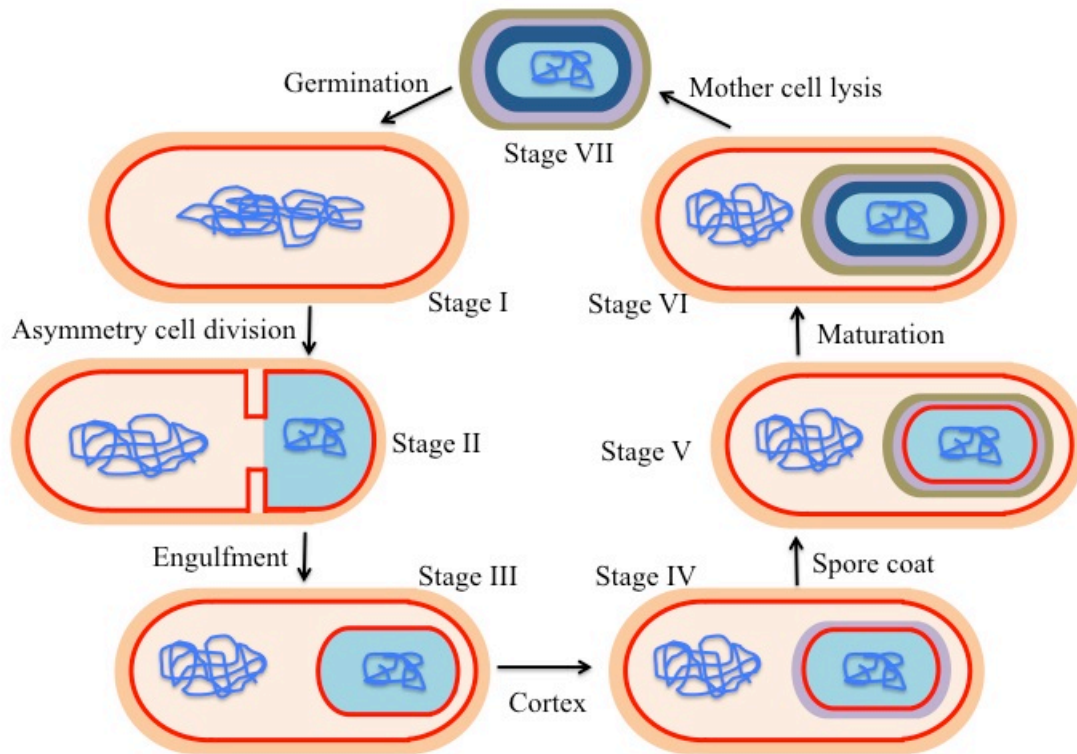


FIGURE 1.2. Illustration of morphological landmarks of sporulation in *Bacillus subtilis*. (Figure modified based on reference (91)). Stage I, chromosome condensation; stage II, asymmetry cell division into mother cell and forespore compartments; stage III, engulfment of forespore by mother cell; stage IV, formation of cortex; stage V, formation of spore coat; stage VI, maturation of forespore; stage VII, mother cell lysis and mature spore release. Mature spore is able to germinate to vegetative cell, and undergo sporulation starting back from stage I.

A simplified diagram in Fig. 1.2 illustrates the morphological landmarks of sporulation represented in *Bacillus subtilis* (91). Before stage I, chromosomes are replicated, but no obvious morphology is present. During stage I, chromosome starts to condense, followed by stage II where the cell divides asymmetrically with the appearance of a polar septum. In stage III, the mother cell engulfs the forespore by curving the polar septum to swallow the forespore. After engulfment, the forespore is bound by the double membrane to form an oval organelle inside the mother cell cytosol. Next in stage IV and V, cortex and spore coat are assembled on the outside

of the forespore, respectively. Then during stage VI, the forespore matures, which becomes more homogeneous, and electron-dense. Meanwhile, cortical peptidoglycan also continues to be synthesized or modified. In the final stage VII, the mother cell lyses releasing the mature, largely dormant spore into the environment (91,99-101).

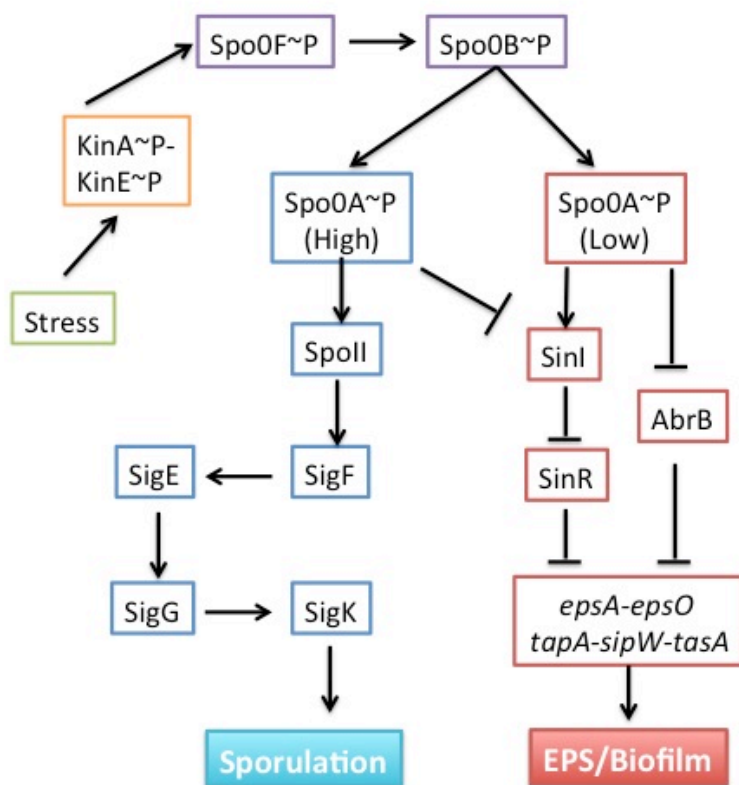


FIGURE 1.3. Network of the different genetics pathways related to sporulation and biofilm in *Bacillus subtilis*. (Figure modified based on reference (102))

The transition of *Bacillus* from vegetative growth to sporulation is largely controlled by the master transcriptional regulator Spo0A, which also regulates biofilm formation (Fig. 1.3) (103). When cells are under stresses, five autophosphorylating histidine kinases (KinA-KinE) respond to the environmental changes, leading to the activation of downstream phosphotransferases Spo0F and Spo0B to phosphorylate Spo0A to an active form (referred to as ‘Spo0A~P’) (99,104,105). The specific environmental cue and the molecular ligand that activate

those histidine kinases remain elusive. Other histidine kinases and phosphatases including RapA, B, E and H, and Spo0E together counter-balance the production of Spo0A~P, to determine the timing of sporulation initiation (106). The ratio of Spo0A~P versus non-phosphorylated Spo0A is also tightly regulated to shift between different life forms: Higher levels of Spo0A~P promote sporulation whereas lower levels of Spo0A~P result in biofilm formation through a promotion of matrix production (103,107).

The sporulation is driven by a cascade of compartment-specific sigma factors. After asymmetric division, several SpoII proteins are activated leading to the activation of the first sporulation-specific sigma factor, SigF (99,108). SigF is also a forespore-specific sigma factor, although the mechanism behind the forespore-specific activation of SigF is not well known. Similar to SigF, The next sporulation-specific sigma factor in the cascade is SigE. SigE is produced as a pro-SigE precursor specifically in the mother cell compartment prior to asymmetric division under the control of Spo0A~P (109). It is matured by the process of SpoIIGA protease and the mature form of SigE then allows the transcription of the SigE regulon, which includes genes necessary for engulfment (110). At the end of engulfment where the forespore is a free floating cell in the mother cell cytosol, a channel between the mother cell and forespore formed by SpoIIAA-SpoIIAH and SpoIIQ proteins under the control of SigE and SigF respectively (111). The mother cell is able to nurture the forespore through the transfer of what are likely small molecules that enable the forespore to continue expressing genes necessary for sporulation (112). At this time, the forespore-specific SigG is activated under the transcriptional control of SigF and is dependent on the arrival of metabolites delivered from the mother cell (99). The activation of SigG leads to the activation of the SigK, which is specific in the mother cell (113). Similar to SigE, SigK is produced as an inactive pro-SigK protein then matured by

the cleavage of SpoIVFB (114,115). Followed by the completion of engulfment, forespore will be matured through the cortex and coat assembly and released from the mother cell (99,116).

1.3.2 Spores structure and known glycan

Bacterial spores are one of the most resilient cell types in nature and are able to survive under both controlled laboratory conditions and natural environment for several decades (117-119). Spores in *Bacillus* have a very different structure than vegetative cells. In *Bacillus anthracis* for example, starting from the inside and proceeding outward the spore layers include the central core, inner membrane, germ cell wall, cortex, out membrane, coats, and exosporium (120) (See Fig. 1.4). The germ cell wall will become the peptidoglycan layer of the outgrowing cell after germination, and the cortex will participate in the maintenance of the dehydrated state of the spore (121). Lastly, the spore coat, which consists of at least 70 proteins, is synthesized in the mother cell and assembled around the forespore to protect the bacterial genome from stresses (121,122). The exosporium component protein BclA, for example, is produced in the mother cell and subsequently assembled on the developing spore (123,124).

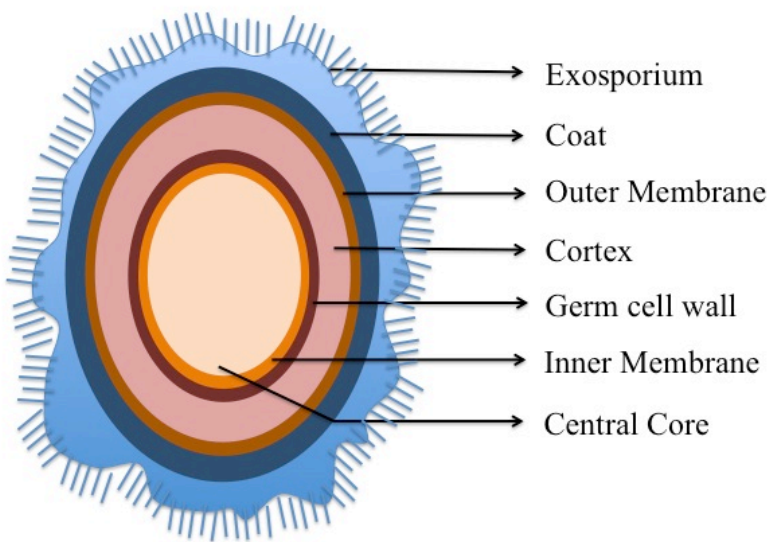


FIGURE 1.4. Illustration of spore structure in *Bacillus anthracis* (Figure modified based on reference (116)).

Several of these structures of spores are composed of glycans, including cortex, germ cell wall, and exosporium. Both cortex and germ cell wall lying under the cortex consist of peptidoglycan. Although the exact detailed peptidoglycan structure of spores are still unknown, the germ cell wall was suggested to have a structure similar to the peptidoglycan of vegetative cells whereas cortex peptidoglycan has several spore-specific modifications, including the presence of the modified sugar muramic-delta-lactam and a low level of peptide cross-links between the glycan strands (125). Muramic-delta-lactam serves as a specificity factor for spore germination lytic enzymes and the cross-linking of spore peptidoglycan has an influence on the rate of spore germination and outgrowth (125). The peptidoglycan layer of the cortex is essential for the formation of a dormant spore and the degradation of this peptidoglycan during spore germination is also important for spore core expansion and subsequent outgrowth (126,127).

Unlike *B. subtilis*, in some species of the *Bacillus cereus* group, an additional spore's outermost structure exists. This structure, called exosporium, surrounds the spore coat layer and is a loose-fitting, balloon-like structure. Exosporium structure exists in strains belonging to both *Bacillus anthracis* and *Bacillus cereus* but not observed in the *B. subtilis* spores, except for one strain isolated from the human gastrointestinal tract (128,129). The exosporium is made of proteins, some of which are modified by glycan structures, i.e. glycoproteins (130-135). Among the exosporium glycoproteins, ExsH and the collagen-like protein, BclA were characterized (136,137). These two glycoproteins in *Bacillus* species have different sugar composition and glycosylation pattern and it was suggested that this might facilitate specific recognition of different cell receptors (138). ExsH contains O-linked glycan including GalNH₂, GlcNH₂, Rha, and 3-Me-Rha sugar residues, but lacks sugar modifications common to BclA including 2-Me-Rha and 2,4-Me-Rha residues (137).

In addition, the nature of the glycan(s) attached to the different domain of BclA structural features also varies. In *B. anthracis*, BclA is decorated at the collagen-like regions with two short O-glycans: 3-*O*-Me-Rha(α 1–2)Rha(α 1–3)GalNAc-ol and Ant(β 1–3)Rha(α 1–3)Rha(α 1–2)Rha(α 1–3)GalNAc-ol with an unusual *N*-acylated monosaccharide in the terminal non-reducing position, named anthrose (136). The C-terminal domain of BclA from *B. cereus* is glycosylated with 2-*O*-methyl-rhamnose and 2,4-*O*-methyl-rhamnose sugar residues (139).

In *Bacillus anthracis*, it was shown that the BclA protein plays a central role in pathogenesis because it promotes interaction of spores with the host phagocytic cell, and BclA was also proposed to facilitate transport of the spores to sites of spore germination and bacterial growth (140,141). Although the contribution of the glycan moiety on these biological roles of BclA is not fully understood, it is noteworthy that the Rha derivatives in *B. anthracis* are involved in specific recognition of the macrophage receptor CD14 (141,142). We cannot rule out the possibility that the Rha along with two unique terminal sugar residues on BclA protein are also involved in similar host recognition mechanism.

1.4 Exopolysaccharide and Role in Biofilm Formation

Bacillus and other bacteria in nature are rarely single planktonic cells (143). It has become more evident that bacteria function and grow within a population called biofilm (10,102,144). Biofilm is defined in *B. subtilis* as communities of tightly associated bacteria encased in an extracellular matrix (102) and is often attached to a surface. Recent deep transcriptomic analyses of *B. cereus* proposed that a single *Bacillus* strain in biofilm consists of at least 6 or more cell types, including vegetative, non-flagellated and flagellated-motile cells, ‘cannibal cell’, sporulating cells and mature spores (145-147). Biofilm formation in *B. subtilis* begins with the differentiation from short, motile cells with flagella to long chains of non-motile

cells. These non-motile cells adhere to each other and the surface by secreting an extracellular matrix (148-150), which is essential to the integrity of the biofilm as it holds the community together (151-153). Upon the maturing of biofilm, the community contains not only the matrix producers, but also motile cells, sporulating cells, and spores (102). In laboratory conditions, biofilms have a limited life span and they eventually disassemble to release spores from the matrix giving them the potential to dispersal and germination upon favorable conditions (102,154).

Bacterial biofilms, in general, require special attention for the food industry, as they can be a source of persistent contamination leading to food spoilage and to the transmission of diseases (155). Indeed, *Bacillus cereus* cells embedded in the biofilm are more protected against sanitizers (156,157). In bacterial species that were studied, biofilm provides means for the bacterial cell to attach and adhere to a variety of surfaces (both natural and man-made) (158,159). In *Staphylococcus* and *Pseudomonas*, it was suggested that cells embedded in biofilm promote the survival of bacteria by forming a niche where bacteria can evade recognition by the host immune system (160,161); however, less is known regarding the myriad roles of biofilm in *Bacillus cereus* group. Biofilms form an extracellular polymeric matrix that is comprised of exopolysaccharides (EPS), proteins, lipids, and nucleic acids (102,144). A rigid biofilm structure from the motile model bacterium, *B. subtilis*, has extremely liquid and gas repellent properties; and mutation in EPS result poor-developed biofilm suggesting the involvement of EPS on biofilm formation (162). Below I will provide a brief overview of the molecular mechanisms behind the production of the extracellular polymeric matrix in *Bacillus subtilis* biofilm formation.

The major EPS required for biofilm are synthesized by the products of the *epsA-epsO* operon (102,149,163), which consists of sixteen genes (149). Another unlinked gene *yhbB* is also characterized as essential for biofilm formation (164); as the result, based on the current knowledge in *B. subtilis*, both *yhbB* and *epsA-O* operon are essential for the production of EPS and biofilms (164). The protein components required for the biofilm are encoded by *tapA-sipW-tasA* operon (165). Both *epsA-epsO* and *tapA-sipW-tasA* operons are repressed by a master regulator SinR (Fig. 1.3). At lower levels of Spo0A~P, SinI, which is a SinR antirepressor, becomes activated and forms SinI-SinR protein-protein complex causing SinR-mediated repression (166), thus both matrix operons are activated to produce the extracellular polymeric matrix for biofilm formation. SinR is produced in all cells but is only inactivated by SinI in a fraction of cells, thus only a subpopulation of cells secretes extracellular polymeric matrix (167,168). On the other hand, at higher levels of Spo0A~P, SinI is repressed to suppress the biofilm formation cascade and promotes the sporulation process (99). Another matrix gene repressor AbrB is also repressed by Spo0A~P (169). Like SinR, AbrB represses both the *epsA-epsO* and *tapA-sipW-tasA* operons (163,170-173). The presence of two very similar Spo0A-regulated repressors, SinR and AbrB could suggest tight regulation of biofilm formation by coordinating expression of all of the matrix genes.

As EPS is essential in biofilm formation, it is also involved in many biological roles including the formation of pellicles and surface structures in *B. subtilis* (164); adherence to variety of surfaces in many species (158,159), physical scaffolding for bacterial cells by maintaining the integrity of biofilm three-dimensional structure, and roles on virulence for pathogenic strains (174).

Different types of EPS have been categorized depending on the location. EPS captured inside the biofilm has a role in constructive, adherence one cell to another, or absorptive aspect. But EPS secreted outside the biofilm may have a role in attachment or detachment (surfactant) of cells to the substratum (175-178). In some *Gram*-negative bacteria such as *Pseudomonas aeruginosa*, EPS plays important roles in structure maintenance and antibiotic resistance of biofilm, for instance, Psl, Pel and even virulence factor such as alginate (175-178). It is also worth noticing that some EPS has a function of inducing macrophage activation probably due to the binding of macrophage mannose receptors (179). The structure of some of these important EPS is well characterized. Psl consists of a repeating pentasaccharide containing D-mannose, D-glucose, and L-rhamnose (180); Alginate is a random linear polymer of variably acetylated 1,4-linked β -D-mannuronic acid and its C5 epimer α -L-guluronic acid (181); levan is composed of β -D-fructans with extensive and irregular branching (182). The structure of Pel still remains unknown (183), although the locus contains genes involved in the biosynthesis and transport for these polysaccharides are well characterized (180,184). The EPS in Gram-positive *Bacillus* species remains largely uncharacterized and the current knowledge regarding such EPS is mainly based on the study in *B. subtilis*. Although *B. subtilis* EPS and the genes encoding its formation were extensively studied, the structure still remains elusive (152). A report of EPS in *B. cereus* AR156 strain showed that the EPS could be involved in induced systemic response (ISR) to the colonization on *Arabidopsis*, suggesting more biological roles of EPS are yet to discovered (185). My work in this dissertation shows the existence of novel EPS produced by *Bacillus cereus* group. It contains rare sugar composition XylNAc, and the chemical structure is fully characterized, as well as the operon involved in the biosynthesis of the nucleotide sugar

precursors. The biological role of the EPS is not fully addressed however evidence suggests that this EPS may be involved in spore movement, adherence, and detachment.

Before I started my Ph.D. work, the lab has discovered an operon called XNAC operon, which produces two uncommon nucleotide sugars UDP-XylNAc and UDP-GlcNAc. And the operon is conserved only in *Bacillus cereus sensu lato* species. However, where these sugars are located in *Bacillus* was not identified, neither the full structure of the glycan containing these two sugars. We hypothesized that the molecule containing these sugars has profound roles in *Bacillus* survival in the host environment, yet no evidence was built to support it. My research started by locating the molecule containing XylNAc and GlcNAcA in the extracellular matrix of *Bacillus thuringiensis* and *Bacillus cereus* species. The structure of this molecular, which is called pzX, was fully characterized by NMR and GC-MS and the biological roles of pzX were also tested to be surfactant and anti-aggregant during sporulation. Studies of mutations in genes of XNAC operon also proved that the formation of pzX is the result of this operon. While this work was progressing, I also found that a unique sugar, pseudaminic acid is decorating the flagellin protein, located on the flagella apparatus in *Bacillus thuringiensis*. A seven-gene operon called PSE operon involved in the formation of the nucleotide precursor of pseudaminic acid was subsequently characterized. Lastly, while pursuing the studies of *Bacillus* spores, I identified two additional uncommon sugars, 3-C-methyl-6-deoxy hexoses, decorating the spores of *Bacillus cereus*. The structures of these two sugars and the operon producing these sugars, C3CM operon, were fully characterized. Mutation study in C3CM operon was carried out to prove that the 3-C-methyl-6-deoxy hexoses glycosylation of spores requires this operon and the mutant in C3CM operon had faster germination rate compared to the wild type.

In the next three chapters, my research which was focused on the discovery of novel glycans in various *Bacillus* species among the *B. cereus* group and their biosynthetic pathway of nucleotide sugar precursors will be presented. In chapter 2, I will show the pseudaminic acid glycosylation of flagellin in *B. thuringiensis* and biosynthetic pathway of CMP-pseudaminic acid. In chapter 3, I will show the discovery of a novel exopolysaccharide with rare sugar XylNAc and its role in spore movement and attachment in *B. cereus* group species. And in chapter 3, I will show the biosynthetic pathway of two CDP-3-C-methyl-6-deoxy sugars involved in the formation of surface glycans on *B. cereus* spores.

1.5 References

1. Salmond, G. P. C., and Whittenbury, R. (1985) Biology of Microorganisms, 4th Edition - Brock, Td, Smith, Dw, Madigan, Mt. *Nature* **314**, 49-49
2. Bodenhausen, N., Horton, M. W., and Bergelson, J. (2013) Bacterial communities associated with the leaves and the roots of *Arabidopsis thaliana*. *PloS one* **8**, e56329
3. de Maagd, R. A., Bravo, A., and Crickmore, N. (2001) How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. *Trends Genet* **17**, 193-199
4. Logan, N. A. (2012) *Bacillus* and relatives in foodborne illness. *Journal of applied microbiology* **112**, 417-429
5. Drewnowska, J. M., and Swiecicka, I. (2013) Eco-Genetic Structure of *Bacillus cereus* sensu lato Populations from Different Environments in Northeastern Poland. *PloS one* **8**
6. Nakamura, L. K. (1998) *Bacillus pseudomyoides* sp. nov. *International journal of systematic bacteriology* **48**, 1031-1035
7. Mock, M., and Fouet, A. (2001) Anthrax. *Annu Rev Microbiol* **55**, 647-671
8. Lechner, S., Mayr, R., Francis, K. P., Pruss, B. M., Kaplan, T., Wiessner-Gunkel, E., Stewart, G. S. A. B., and Scherer, S. (1998) *Bacillus weihenstephanensis* sp. nov. is a new psychrotolerant species of the *Bacillus cereus* group. *International journal of systematic bacteriology* **48**, 1373-1382
9. Guinebretiere, M. H., Auger, S., Galleron, N., Contzen, M., De Sarrau, B., De Buyser, M. L., Lamberet, G., Fagerlund, A., Granum, P. E., Lereclus, D., De Vos, P., Nguyen-The, C., and Sorokin, A. (2013) *Bacillus cytotoxicus* sp nov is a novel thermotolerant species of the *Bacillus cereus* Group occasionally associated with food poisoning. *International journal of systematic and evolutionary microbiology* **63**, 31-40
10. Li, Z., Hwang, S., and Bar-Peled, M. (2016) Discovery of a Unique Extracellular Polysaccharide in Members of the Pathogenic *Bacillus* That Can Co-form with Spores. *The Journal of biological chemistry* **291**, 19051-19067
11. Nakamura, L. K. (1998) *Bacillus pseudomyoides* sp. nov. *International journal of systematic bacteriology* **48 Pt 3**, 1031-1035
12. Turnbull, P. C. B. (1996) *Medical Microbiology. 4th edition.*, Galveston, TX

13. Schnepf, E., Crickmore, N., Van Rie, J., Lereclus, D., Baum, J., Feitelson, J., Zeigler, D. R., and Dean, D. H. (1998) *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiol Mol Biol R* **62**, 775-+
14. Soberon, M., Gill, S. S., and Bravo, A. (2009) Signaling versus punching hole: How do *Bacillus thuringiensis* toxins kill insect midgut cells? *Cell Mol Life Sci* **66**, 1337-1349
15. Voth, D. E., Hamm, E. E., Nguyen, L. G., Tucker, A. E., Salles, I. I., Ortiz-Leduc, W., and Ballard, J. D. (2005) *Bacillus anthracis* oedema toxin as a cause of tissue necrosis and cell type-specific cytotoxicity. *Cell Microbiol* **7**, 1139-1149
16. Gizewski, P. (1999) Biological weapons - Limiting the threat. *Int J* **54**, 715-717
17. Bottone, E. J. (2010) *Bacillus cereus*, a Volatile Human Pathogen. *Clinical microbiology reviews* **23**, 382-+
18. Spencer, R. C. (2003) *Bacillus anthracis*. *Journal of clinical pathology* **56**, 182-187
19. Goel, A. K. (2015) Anthrax: A disease of biowarfare and public health importance. *World journal of clinical cases* **3**, 20-33
20. Sternbach, G. (2003) The history of anthrax. *The Journal of emergency medicine* **24**, 463-467
21. Goldenberg, A., Shmueli, G., Caruana, R. A., and Fienberg, S. E. (2002) Early statistical detection of anthrax outbreaks by tracking over-the-counter medication sales. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 5237-5240
22. (2002). in *An Assessment of the CDC Anthrax Vaccine Safety and Efficacy Research Program*, Washington (DC). pp
23. Feld, G. K., Brown, M. J., and Krantz, B. A. (2012) Ratcheting up protein translocation with anthrax toxin. *Protein science : a publication of the Protein Society* **21**, 606-624
24. Ezzell, J. W., Abshire, T. G., Panchal, R., Chabot, D., Bavari, S., Leffel, E. K., Purcell, B., Friedlander, A. M., and Ribot, W. J. (2009) Association of *Bacillus anthracis* capsule with lethal toxin during experimental infection. *Infection and immunity* **77**, 749-755
25. Thoren, K. L., and Krantz, B. A. (2011) The unfolding story of anthrax toxin translocation. *Molecular microbiology* **80**, 588-595
26. Ezzell, J. W., and Welkos, S. L. (1999) The capsule of *Bacillus anthracis*, a review. *Journal of applied microbiology* **87**, 250-250
27. Ribeiro, N. F. F., Heath, C. H., Kierath, J., Rea, S., Duncan-Smith, M., and Wood, F. M. (2010) Burn wounds infected by contaminated water: Case reports, review of the literature and recommendations for treatment. *Burns : journal of the International Society for Burn Injuries* **36**, 9-22
28. Frankland, G. C. F., P. F. (1887) Studies on Some New Micro-Organisms Obtained from Air. *hilosophical Transactions of the Royal Society B: Biological Sciences*. **178**, 257-287
29. Sandle, T. (2014) The Risk of *Bacillus cereus* to Pharmaceutical Manufacturing. *American Pharmaceutical Review* **17**, 56
30. Jensen, G. B., Hansen, B. M., Eilenberg, J., and Mahillon, J. (2003) The hidden lifestyles of *Bacillus cereus* and relatives. *Environmental microbiology* **5**, 631-640
31. Arnesen, L. P. S., Fagerlund, A., and Granum, P. E. (2008) From soil to gut: *Bacillus cereus* and its food poisoning toxins. *FEMS microbiology reviews* **32**, 579-606
32. Ghosh, A. C. (1978) Prevalence of *Bacillus cereus* in the faeces of healthy adults. *The Journal of hygiene* **80**, 233-236
33. Guinebretiere, M. H., Broussolle, V., and Nguyen-The, C. (2002) Enterotoxigenic profiles of food-poisoning and food-borne *Bacillus cereus* strains. *Journal of clinical microbiology* **40**, 3053-3056
34. Kotiranta, A., Lounatmaa, K., and Haapasalo, M. (2000) Epidemiology and pathogenesis of *Bacillus cereus* infections. *Microbes and infection / Institut Pasteur* **2**, 189-198
35. de Boer, A. S., and Diderichsen, B. (1991) On the safety of *Bacillus subtilis* and *B. amyloliquefaciens*: a review. *Appl Microbiol Biotechnol* **36**, 1-4

36. Hoffmaster, A. R., Hill, K. K., Gee, J. E., Marston, C. K., De, B. K., Popovic, T., Sue, D., Wilkins, P. P., Avashia, S. B., Drumgoole, R., Helma, C. H., Ticknor, L. O., Okinaka, R. T., and Jackson, P. J. (2006) Characterization of *Bacillus cereus* isolates associated with fatal pneumonias: strains are closely related to *Bacillus anthracis* and harbor *B. anthracis* virulence genes. *Journal of clinical microbiology* **44**, 3352-3360
37. Miller, J. M., Hair, J. G., Hebert, M., Hebert, L., Roberts, F. J., and Weyant, R. S. (1997) Fulminating bacteremia and pneumonia due to *Bacillus cereus* (vol 35, pg 504, 1997). *Journal of clinical microbiology* **35**, 1294-1294
38. Avashia, S. B., Riggins, W. S., Lindley, C., Hoffmaster, A., Drumgoole, R., Nekomoto, T., Jackson, P. J., Hill, K. K., Williams, K., Lehman, L., Libal, M. C., Wilkins, P. P., Alexander, J., Tvaryanas, A., and Betz, T. (2007) Fatal pneumonia among metalworkers due to inhalation exposure to *Bacillus cereus* containing *Bacillus anthracis* toxin genes. *Clinical Infectious Diseases* **44**, 414-416
39. Klee, S. R., Ozel, M., Appel, B., Boesch, C., Ellerbrok, H., Jacob, D., Holland, G., Leendertz, F. H., Pauli, G., Grunow, R., and Nattermann, H. (2006) Characterization of *Bacillus anthracis*-like bacteria isolated from wild great apes from Cote d'Ivoire and Cameroon. *Journal of bacteriology* **188**, 5333-5344
40. Bottone, E. J. (2010) *Bacillus cereus*, a volatile human pathogen. *Clinical microbiology reviews* **23**, 382-398
41. Hoffmaster, A. R., Ravel, J., Rasko, D. A., Chapman, G. D., Chute, M. D., Marston, C. K., De, B. K., Sacchi, C. T., Fitzgerald, C., Mayer, L. W., Maiden, M. C., Priest, F. G., Barker, M., Jiang, L., Cer, R. Z., Rilstone, J., Peterson, S. N., Weyant, R. S., Galloway, D. R., Read, T. D., Popovic, T., and Fraser, C. M. (2004) Identification of anthrax toxin genes in a *Bacillus cereus* associated with an illness resembling inhalation anthrax. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 8449-8454
42. Roh, J. Y., Choi, J. Y., Li, M. S., Jin, B. R., and Je, Y. H. (2007) *Bacillus thuringiensis* as a specific, safe, and effective tool for insect pest control. *J Microbiol Biotechnol* **17**, 547-559
43. Broderick, N. A., Raffa, K. F., and Handelsman, J. (2006) Midgut bacteria required for *Bacillus thuringiensis* insecticidal activity. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 15196-15199
44. Swiecicka, I., and Mahillon, J. (2006) Diversity of commensal *Bacillus cereus* sensu lato isolated from the common sow bug (*Porcellio scaber*, Isopoda). *Fems Microbiol Ecol* **56**, 132-140
45. Bravo, A., Gill, S. S., and Soberon, M. (2007) Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control. *Toxicon* **49**, 423-435
46. Schunemann, R., Knaak, N., and Fiuza, L. M. (2014) Mode of Action and Specificity of *Bacillus thuringiensis* Toxins in the Control of Caterpillars and Stink Bugs in Soybean Culture. *ISRN microbiology* **2014**, 135675
47. Dean, D. H. (1984) Biochemical Genetics of the Bacterial Insect-Control Agent *Bacillus-Thuringiensis* - Basic Principles and Prospects for Genetic-Engineering. *Biotechnol Genet Eng* **2**, 341-363
48. Lemaux, P. G. (2008) Genetically engineered plants and foods: A scientist's analysis of the issues (Part I). *Annual review of plant biology* **59**, 771-812
49. Key, S., Ma, J. K. C., and Drake, P. M. W. (2008) Genetically modified plants and human health. *Journal of the Royal Society of Medicine* **101**, 290-298
50. Ibrahim, M. A., Griko, N. B., and Bulla, L. A., Jr. (2013) The Cry4B toxin of *Bacillus thuringiensis* subsp. *israelensis* kills Permethrin-resistant *Anopheles gambiae*, the principal vector of malaria. *Experimental biology and medicine* **238**, 350-359
51. Ticknor, L. O., Kolsto, A. B., Hill, K. K., Keim, P., Laker, M. T., Tonks, M., and Jackson, P. J. (2001) Fluorescent amplified fragment length polymorphism analysis of Norwegian *Bacillus cereus* and *Bacillus thuringiensis* soil isolates. *Applied and environmental microbiology* **67**, 4863-4873

52. Atrih, A., Bacher, G., Allmaier, G., Williamson, M. P., and Foster, S. J. (1999) Analysis of peptidoglycan structure from vegetative cells of *Bacillus subtilis* 168 and role of PBP 5 in peptidoglycan maturation. *Journal of bacteriology* **181**, 3956-3966
53. Mignot, T., Denis, B., Couture-Tosi, E., Kolsto, A. B., Mock, M., and Fouet, A. (2001) Distribution of S-layers on the surface of *Bacillus cereus* strains: phylogenetic origin and ecological pressure. *Environmental microbiology* **3**, 493-501
54. Vollmer, W., Blanot, D., and de Pedro, M. A. (2008) Peptidoglycan structure and architecture. *FEMS microbiology reviews* **32**, 149-167
55. Bott, K. F. (1994) *Bacillus-Subtilis and Other Gram-Positive Bacteria - Biochemistry, Physiology, and Molecular-Genetics* - Sonenshein, A.I., Hoch, J.A., Losick, R. *Science* **263**, 546-548
56. Moscoso, M., Garcia, E., and Lopez, R. (2006) Biofilm formation by *Streptococcus pneumoniae*: Role of choline, extracellular DNA, and capsular polysaccharide in microbial accretion. *Journal of bacteriology* **188**, 7785-7795
57. Lebeer, S., Verhoeven, T. L. A., Velez, M. P., Vanderleyden, J., and De Keersmaecker, S. C. J. (2007) Impact of environmental and genetic factors on Biofilm formation by the Probiotic strain *Lactobacillus rhamnosus* GG. *Applied and environmental microbiology* **73**, 6768-6775
58. Neuhaus, F. C., and Baddiley, J. (2003) A continuum of anionic charge: Structures and functions of D-Alanyl-Teichoic acids in gram-positive bacteria. *Microbiol Mol Biol R* **67**, 686-+
59. Swoboda, J. G., Campbell, J., Meredith, T. C., and Walker, S. (2010) Wall Teichoic Acid Function, Biosynthesis, and Inhibition. *Chembiochem : a European journal of chemical biology* **11**, 35-45
60. Murazumi, N., Sasaki, Y., Okada, J., Araki, Y., and Ito, E. (1981) Biosynthesis of Glycerol Teichoic-Acid in *Bacillus-Cereus* - Formation of Linkage Unit Disaccharide on a Lipid Intermediate. *Biochemical and biophysical research communications* **99**, 504-510
61. Weidenmaier, C., Peschel, A., Xiong, Y. Q., Kristian, S. A., Dietz, K., Yeaman, M. R., and Bayer, A. S. (2005) Lack of wall teichoic acids in *Staphylococcus aureus* leads to reduced interactions with endothelial cells and to attenuated virulence in a rabbit model of endocarditis. *Journal of Infectious Diseases* **191**, 1771-1777
62. Mesnage, S., Fontaine, T., Mignot, T., Delepierre, M., Mock, M., and Fouet, A. (2000) Bacterial SLH domain proteins are non-covalently anchored to the cell surface via a conserved mechanism involving wall polysaccharide pyruvylation. *Embo J* **19**, 4473-4484
63. Candela, T., Maes, E., Garenaux, E., Rombouts, Y., Krzewinski, F., Gohar, M., and Guerardel, Y. (2011) Environmental and Biofilm-dependent Changes in a *Bacillus cereus* Secondary Cell Wall Polysaccharide. *Journal of Biological Chemistry* **286**, 31250-31262
64. Mesnage, S., Tosi-Couture, E., Gounon, P., Mock, M., and Fouet, A. (1998) The capsule and S-layer: Two independent and yet compatible macromolecular structures in *Bacillus anthracis*. *Journal of bacteriology* **180**, 52-58
65. Kern, J., Wilton, R., Zhang, R. G., Binkowski, T. A., Joachimiak, A., and Schneewind, O. (2011) Structure of Surface Layer Homology (SLH) Domains from *Bacillus anthracis* Surface Array Protein. *Journal of Biological Chemistry* **286**, 26042-26049
66. Fagan, R. P., and Fairweather, N. F. (2014) Biogenesis and functions of bacterial S-layers. *Nature Reviews Microbiology* **12**, 211-222
67. Kern, J., and Schneewind, O. (2010) BslA, the S-layer adhesin of *B. anthracis*, is a virulence factor for anthrax pathogenesis. *Molecular microbiology* **75**, 324-332
68. Sara, M., and Sleytr, U. B. (1987) Molecular-Sieving through S-Layers of *Bacillus-Stearothermophilus* Strains. *Journal of bacteriology* **169**, 4092-4098
69. Sidhu, M. S., and Olsen, I. (1997) S-layers of *Bacillus* species. *Microbiol-Sgm* **143**, 1039-1052
70. Budzik, J. M., Marraffini, L. A., and Schneewind, O. (2007) Assembly of pili on the surface of *Bacillus cereus* vegetative cells. *Molecular microbiology* **66**, 495-510
71. Budzik, J. M., Oh, S. Y., and Schneewind, O. (2008) Cell Wall Anchor Structure of BcpA Pili in *Bacillus anthracis*. *Journal of Biological Chemistry* **283**, 36676-36686

72. Power, P. M., Roddam, L. F., Dieckelmann, M., Srikhanta, Y. N., Tan, Y. C., Berrington, A. W., and Jennings, M. P. (2000) Genetic characterization of pilin glycosylation in *Neisseria meningitidis*. *Microbiol-Uk* **146**, 967-979
73. Stimson, E., Virji, M., Makepeace, K., Dell, A., Morris, H. R., Payne, G., Saunders, J. R., Jennings, M. P., Barker, S., Panico, M., Blench, I., and Moxon, E. R. (1995) Meningococcal Pilin - a Glycoprotein Substituted with Digalactosyl 2,4-Diacetamido-2,4,6-Trideoxyhexose. *Molecular microbiology* **17**, 1201-1214
74. Kearns, D. B., and Losick, R. (2003) Swarming motility in undomesticated *Bacillus subtilis*. *Molecular microbiology* **49**, 581-590
75. Mukherjee, S., and Kearns, D. B. (2014) The Structure and Regulation of Flagella in *Bacillus subtilis*. *Annual Review of Genetics, Vol 48* **48**, 319-340
76. Senesi, S., Celandroni, F., Salvetti, S., Beecher, D. J., Wong, A. C. L., and Ghelardi, E. (2002) Swarming motility in *Bacillus cereus* and characterization of a *fliY* mutant impaired in swarm cell differentiation. *Microbiol-Sgm* **148**, 1785-1794
77. Maki-Yonekura, S., Yonekura, K., and Namba, K. (2003) Domain movements of HAP2 in the cap-filament complex formation and growth process of the bacterial flagellum. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 15528-15533
78. Namba, K., Yamashita, I., and Vonderviszt, F. (1989) Structure of the Core and Central Channel of Bacterial Flagella. *Nature* **342**, 648-654
79. Beatson, S. A., Minamino, T., and Pallen, M. J. (2006) Variation in bacterial flagellins: from sequence to structure. *Trends in microbiology* **14**, 151-155
80. Young, G. M., Schmiel, D. H., and Miller, V. L. (1999) A new pathway for the secretion of virulence factors by bacteria: The flagellar export apparatus functions as a protein-secretion system. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 6456-6461
81. Ottemann, K. M., and Miller, J. F. (1997) Roles for motility in bacterial-host interactions. *Molecular microbiology* **24**, 1109-1117
82. Logan, S. M. (2006) Flagellar glycosylation - a new component of the motility repertoire? *Microbiology* **152**, 1249-1262
83. Hayakawa, J., and Ishizuka, M. (2012) Flagellar Glycosylation: Current Advances. in *Glycosylation* (Petrescu, D. S. ed., CC
84. Schirm, M., Schoenhofen, I. C., Logan, S. M., Waldron, K. C., and Thibault, P. (2005) Identification of unusual bacterial glycosylation by tandem mass spectrometry analyses of intact proteins. *Anal Chem* **77**, 7774-7782
85. Thibault, P., Logan, S. M., Kelly, J. F., Brisson, J. R., Ewing, C. P., Trust, T. J., and Guerry, P. (2001) Identification of the carbohydrate moieties and glycosylation motifs in *Campylobacter jejuni* flagellin. *Journal of Biological Chemistry* **276**, 34862-34870
86. Schirm, M., Arora, S. K., Verma, A., Vinogradov, E., Thibault, P., Ramphal, R., and Logan, S. M. (2004) Structural and genetic characterization of glycosylation of type a flagellin in *Pseudomonas aeruginosa*. *Journal of bacteriology* **186**, 2523-2531
87. Hayakawa, J., Kondoh, Y., and Ishizuka, M. (2009) Cloning and characterization of flagellin genes and identification of flagellin glycosylation from thermophilic *Bacillus* species. *Bioscience, biotechnology, and biochemistry* **73**, 1450-1452
88. Hayakawa, J., Kambe, T., and Ishizuka, M. (2009) Amino acid substitutions and intragenic duplications of *Bacillus* sp. PS3 flagellin cause complementation of the *Bacillus subtilis* flagellin deletion mutant. *Bioscience, biotechnology, and biochemistry* **73**, 2348-2351
89. Parker, J. L., Day-Williams, M. J., Tomas, J. M., Stafford, G. P., and Shaw, J. G. (2012) Identification of a putative glycosyltransferase responsible for the transfer of pseudaminic acid onto the polar flagellin of *Aeromonas caviae* Sch3N. *MicrobiologyOpen* **1**, 149-160

90. Guerry, P., Ewing, C. P., Schirm, M., Lorenzo, M., Kelly, J., Pattarini, D., Majam, G., Thibault, P., and Logan, S. (2006) Changes in flagellin glycosylation affect *Campylobacter* autoagglutination and virulence. *Molecular microbiology* **60**, 299-311
91. Errington, J. (2003) Regulation of endospore formation in *Bacillus subtilis*. *Nature Reviews Microbiology* **1**, 117-126
92. Stragier, P., and Losick, R. (1996) Molecular genetics of sporulation in *Bacillus subtilis*. *Annu Rev Genet* **30**, 297-341
93. Piggot, P. J., and Hilbert, D. W. (2004) Sporulation of *Bacillus subtilis*. *Current opinion in microbiology* **7**, 579-586
94. Inglesby, T. V., Henderson, D. A., Bartlett, J. G., Ascher, M. S., Eitzen, E., Friedlander, A. M., Hauer, J., McDade, J., Osterholm, M. T., O'Toole, T., Parker, G., Perl, T. M., Russell, P. K., Tonat, K., and Biodefense, W. G. C. (1999) Anthrax as a biological weapon - Medical and public health management. *Jama-J Am Med Assoc* **281**, 1735-1745
95. Roth, N. G., Lively, David H., and Hodge, Howard M. . (1955) INFLUENCE OF OXYGEN UPTAKE AND AGE OF CULTURE ON SPORULATION OF *BACILLUS ANTHRACIS* AND *BACILLUS GLOBIGII*. *Journal of bacteriology* **69**, 455-459
96. Sarrafzadeh, M. H., and Navarro, J. M. (2006) The effect of oxygen on the sporulation, delta-endotoxin synthesis and toxicity of *Bacillus thuringiensis* H14. *World J Microb Biot* **22**, 305-310
97. Abbas, A. A., Planchon, S., Jobin, M., and Schmitt, P. (2014) Absence of oxygen affects the capacity to sporulate and the spore properties of *Bacillus cereus*. *Food microbiology* **42**, 122-131
98. Driks, A. (2002) Maximum shields: the assembly and function of the bacterial spore coat. *Trends in microbiology* **10**, 251-254
99. Tan, I. S., and Ramamurthi, K. S. (2014) Spore formation in *Bacillus subtilis*. *Environmental microbiology reports* **6**, 212-225
100. Errington, J. (1993) *Bacillus subtilis* sporulation: regulation of gene expression and control of morphogenesis. *Microbiological reviews* **57**, 1-33
101. Piggot, P. J., and Coote, J. G. (1976) Genetic aspects of bacterial endospore formation. *Bacteriological reviews* **40**, 908-962
102. Vlamakis, H., Chai, Y., Beauregard, P., Losick, R., and Kolter, R. (2013) Sticking together: building a biofilm the *Bacillus subtilis* way. *Nature reviews. Microbiology* **11**, 157-168
103. Hamon, M. A., and Lazazzera, B. A. (2001) The sporulation transcription factor Spo0A is required for biofilm development in *Bacillus subtilis*. *Molecular microbiology* **42**, 1199-1209
104. Burbulys, D., Trach, K. A., and Hoch, J. A. (1991) Initiation of Sporulation in *Bacillus-Subtilis* Is Controlled by a Multicomponent Phosphorelay. *Cell* **64**, 545-552
105. Burbulys, D., Trach, K. A., and Hoch, J. A. (1991) Initiation of sporulation in *B. subtilis* is controlled by a multicomponent phosphorelay. *Cell* **64**, 545-552
106. Perego, M., Hanstein, C., Welsh, K. M., Djavakhishvili, T., Glaser, P., and Hoch, J. A. (1994) Multiple Protein Aspartate Phosphatases Provide a Mechanism for the Integration of Diverse Signals in the Control of Development in *Bacillus-Subtilis*. *Cell* **79**, 1047-1055
107. Fujita, M., and Losick, R. (2003) The master regulator for entry into sporulation in *Bacillus subtilis* becomes a cell-specific transcription factor after asymmetric division. *Genes & development* **17**, 1166-1174
108. Duncan, L., Alper, S., Arigoni, F., Losick, R., and Stragier, P. (1995) Activation of Cell-Specific Transcription by a Serine Phosphatase at the Site of Asymmetric Division. *Science* **270**, 641-644
109. Fujita, M., and Losick, R. (2002) An investigation into the compartmentalization of the sporulation transcription factor sigma(E) in *Bacillus subtilis*. *Molecular microbiology* **43**, 27-38
110. Imamura, D., Zhou, R. B., Feig, M., and Kroos, L. (2008) Evidence that the *Bacillus subtilis* SpoIIIGA protein is a novel type of signal-transducing aspartic protease. *Journal of Biological Chemistry* **283**, 15287-15299

111. Meisner, J., Wang, X., Serrano, M., Henriques, A. O., and Moran, C. P. (2008) A channel connecting the mother cell and forespore during bacterial endospore formation. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 15100-15105
112. Camp, A. H., and Losick, R. (2009) A feeding tube model for activation of a cell-specific transcription factor during sporulation in *Bacillus subtilis*. *Genes & development* **23**, 1014-1024
113. Lopez, D., Vlamakis, H., and Kolter, R. (2009) Generation of multiple cell types in *Bacillus subtilis*. *FEMS microbiology reviews* **33**, 152-163
114. Lu, S., Cutting, S., and Kroos, L. (1995) Sporulation Protein Spoivfb from *Bacillus-Subtilis* Enhances Processing of the Sigma-Factor Precursor Pro-Sigma(K) in the Absence of Other Sporulation Gene-Products. *Journal of bacteriology* **177**, 1082-1085
115. Resnekov, O., and Losick, R. (1998) Negative regulation of the proteolytic activation of a developmental transcription factor in *Bacillus subtilis*. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 3162-3167
116. Setlow, P. (2006) Spores of *Bacillus subtilis*: their resistance to and killing by radiation, heat and chemicals. *Journal of applied microbiology* **101**, 514-525
117. Cano, R. J., and Borucki, M. K. (1995) Revival and Identification of Bacterial-Spores in 25-Million-Year-Old to 40-Million-Year-Old Dominican Amber (Vol 268, Pg 1060, 1995). *Science* **268**, 1265-1265
118. Vreeland, R. H., Rosenzweig, W. D., and Powers, D. W. (2000) Isolation of a 250 million-year-old halotolerant bacterium from a primary salt crystal. *Nature* **407**, 897-900
119. Setlow, P. (2007) I will survive: DNA protection in bacterial spores. *Trends in microbiology* **15**, 172-180
120. Setlow, P. (2006) Spores of *Bacillus subtilis*: their resistance to and killing by radiation, heat and chemicals. *Journal of applied microbiology* **101**, 514-525
121. Plomp, M., Carroll, A. M., Setlow, P., and Malkin, A. J. (2014) Architecture and Assembly of the *Bacillus subtilis* Spore Coat. *PloS one* **9**
122. Vasudevan, P., Weaver, A., Reichert, E. D., Linnstaedt, S. D., and Popham, D. L. (2007) Spore cortex formation in *Bacillus subtilis* is regulated by accumulation of peptidoglycan precursors under the control of sigma K. *Molecular microbiology* **65**, 1582-1594
123. Thompson, B. M., and Stewart, G. C. (2008) Targeting of the BclA and BclB proteins to the *Bacillus anthracis* spore surface. *Molecular microbiology* **70**, 421-434
124. Thompson, B. M., Hsieh, H. Y., Spreng, K. A., and Stewart, G. C. (2011) The co-dependence of BxpB/ExsFA and BclA for proper incorporation into the exosporium of *Bacillus anthracis*. *Molecular microbiology* **79**, 799-813
125. Popham, D. L. (2002) Specialized peptidoglycan of the bacterial endospore: the inner wall of the lockbox. *Cell Mol Life Sci* **59**, 426-433
126. Setlow, P. (2003) Spore germination. *Current opinion in microbiology* **6**, 550-556
127. Setlow, P. (2014) Germination of spores of *Bacillus* species: what we know and do not know. *Journal of bacteriology* **196**, 1297-1305
128. Imamura, D., Kuwana, R., Takamatsu, H., and Watabe, K. (2011) Proteins Involved in Formation of the Outermost Layer of *Bacillus subtilis* Spores. *Journal of bacteriology* **193**, 4075-4080
129. Hong, H. A., Khaneja, R., Tam, N. M. K., Cazzato, A., Tan, S., Urdaci, M., Brisson, A., Gasbarrini, A., Barnes, I., and Cutting, S. M. (2009) *Bacillus subtilis* isolated from the human gastrointestinal tract. *Research in microbiology* **160**, 134-143
130. Redmond, C., Baillie, L. W., Hibbs, S., Moir, A. J., and Moir, A. (2004) Identification of proteins in the exosporium of *Bacillus anthracis*. *Microbiology* **150**, 355-363
131. Daubenspeck, J. M., Zeng, H., Chen, P., Dong, S., Steichen, C. T., Krishna, N. R., Pritchard, D. G., and Turnbough, C. L., Jr. (2004) Novel oligosaccharide side chains of the collagen-like region of BclA, the major glycoprotein of the *Bacillus anthracis* exosporium. *The Journal of biological chemistry* **279**, 30945-30953

132. Todd, S. J., Moir, A. J., Johnson, M. J., and Moir, A. (2003) Genes of *Bacillus cereus* and *Bacillus anthracis* encoding proteins of the exosporium. *Journal of bacteriology* **185**, 3373-3378
133. Lai, E. M., Phadke, N. D., Kachman, M. T., Giorno, R., Vazquez, S., Vazquez, J. A., Maddock, J. R., and Driks, A. (2003) Proteomic analysis of the spore coats of *Bacillus subtilis* and *Bacillus anthracis*. *Journal of bacteriology* **185**, 1443-1454
134. Sylvestre, P., Couture-Tosi, E., and Mock, M. (2002) A collagen-like surface glycoprotein is a structural component of the *Bacillus anthracis* exosporium. *Molecular microbiology* **45**, 169-178
135. Steichen, C., Chen, P., Kearney, J. F., and Turnbough, C. L. (2003) Identification of the immunodominant protein and other proteins of the *Bacillus anthracis* exosporium. *Journal of bacteriology* **185**, 1903-1910
136. Daubenspeck, J. M., Zeng, H. D., Chen, P., Dong, S. L., Steichen, C. T., Krishna, N. R., Pritchard, D. G., and Turnbough, C. L. (2004) Novel oligosaccharide side chains of the collagen-like region of BclA, the major glycoprotein of the *Bacillus anthracis* exosporium. *Journal of Biological Chemistry* **279**, 30945-30953
137. Lequette, Y., Garenaux, E., Tauveron, G., Dumez, S., Perchat, S., Slomianny, C., Lereclus, D., Guerardel, Y., and Faille, C. (2011) Role Played by Exosporium Glycoproteins in the Surface Properties of *Bacillus cereus* Spores and in Their Adhesion to Stainless Steel. *Applied and environmental microbiology* **77**, 4905-4911
138. Lequette, Y., Garenaux, E., Combrouse, T., Dias Tdel, L., Ronse, A., Slomianny, C., Trivelli, X., Guerardel, Y., and Faille, C. (2011) Domains of BclA, the major surface glycoprotein of the *B. cereus* exosporium: glycosylation patterns and role in spore surface properties. *Biofouling* **27**, 751-761
139. Maes, E., Krzewinski, F., Garenaux, E., Lequette, Y., Coddeville, B., Trivelli, X., Ronse, A., Faille, C., and Guerardel, Y. (2016) Glycosylation of BclA Glycoprotein from *Bacillus cereus* and *Bacillus anthracis* Exosporium Is Domain-specific. *The Journal of biological chemistry* **291**, 9666-9677
140. Oliva, C. R., Swiecki, M. K., Griguer, C. E., Lisanby, M. W., Bullard, D. C., Turnbough, C. L., and Kearney, J. F. (2008) The integrin Mac-1 (CR3) mediates internalization and directs *Bacillus anthracis* spores into professional phagocytes. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 1261-1266
141. Oliva, C., Turnbough, C. L., and Kearney, J. F. (2009) CD14-Mac-1 interactions in *Bacillus anthracis* spore internalization by macrophages. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 13957-13962
142. Bozue, J. A., Parthasarathy, N., Phillips, L. R., Cote, C. K., Fellows, P. F., Mendelson, I., Shafferman, A., and Friedlander, A. M. (2005) Construction of a rhamnose mutation in *Bacillus anthracis* affects adherence to macrophages but not virulence in guinea pigs. *Microb Pathogenesis* **38**, 1-12
143. Vu, B., Chen, M., Crawford, R. J., and Ivanova, E. P. (2009) Bacterial extracellular polysaccharides involved in biofilm formation. *Molecules* **14**, 2535-2554
144. Sutherland, I. W. (2001) The biofilm matrix - an immobilized but dynamic microbial environment. *Trends in microbiology* **9**, 222-227
145. Guttenplan, S. B., and Kearns, D. B. (2013) Regulation of flagellar motility during biofilm formation. *FEMS microbiology reviews* **37**, 849-871
146. Faille, C., Benezech, T., Midelet-Bourdin, G., Lequette, Y., Clarisse, M., Ronse, G., Ronse, A., and Slomianny, C. (2014) Sporulation of *Bacillus* spp. within biofilms: a potential source of contamination in food processing environments. *Food microbiology* **40**, 64-74
147. Perchat S, T. A., Poncet S, Lazar N, Li de la Sierra-Gallay I, Gohar M, Lereclus D, Nessler S. (2016) How Quorum Sensing Connects Sporulation to Necrotrophism in *Bacillus thuringiensis*. *PLoS pathogens* **12**
148. Branda, S. S., Chu, F., Kearns, D. B., Losick, R., and Kolter, R. (2006) A major protein component of the *Bacillus subtilis* biofilm matrix. *Molecular microbiology* **59**, 1229-1238

149. Branda, S. S., Gonzalez-Pastor, J. E., Ben-Yehuda, S., Losick, R., and Kolter, R. (2001) Fruiting body formation by *Bacillus subtilis*. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 11621-11626
150. Kobayashi, K. (2007) *Bacillus subtilis* pellicle formation proceeds through genetically defined morphological changes. *Journal of bacteriology* **189**, 4920-4931
151. Branda, S. S., Vik, A., Friedman, L., and Kolter, R. (2005) Biofilms: the matrix revisited. *Trends in microbiology* **13**, 20-26
152. Marvasi, M., Visscher, P. T., and Martinez, L. C. (2010) Exopolymeric substances (EPS) from *Bacillus subtilis*: polymers and genes encoding their synthesis. *FEMS microbiology letters* **313**, 1-9
153. Flemming, H. C., and Wingender, J. (2010) The biofilm matrix. *Nature Reviews Microbiology* **8**, 623-633
154. Kolodkin-Gal, I., Romero, D., Cao, S., Clardy, J., Kolter, R., and Losick, R. (2010) D-amino acids trigger biofilm disassembly. *Science* **328**, 627-629
155. Van Houdt, R., and Michiels, C. (2010) Biofilm formation and the food industry, a focus on the bacterial outer surface. *Journal of applied microbiology* **109**, 1117-1131
156. Nam, H., Seo, H. S., Bang, J., Kim, H., Beuchat, L. R., and Ryu, J. H. (2014) Efficacy of gaseous chlorine dioxide in inactivating *Bacillus cereus* spores attached to and in a biofilm on stainless steel. *International journal of food microbiology* **188**, 122-127
157. Ryu, J. H., and Beuchat, L. R. (2005) Biofilm formation and sporulation by *Bacillus cereus* on a stainless steel surface and subsequent resistance of vegetative cells and spores to chlorine, chlorine dioxide, and a peroxyacetic acid-based sanitizer. *Journal of food protection* **68**, 2614-2622
158. Gotz, F. (2002) Staphylococcus and biofilms. *Molecular microbiology* **43**, 1367-1378
159. Lemon, K. P., Earl, A. M., Vlamakis, H. C., Aguilar, C., and Kolter, R. (2008) Biofilm development with an emphasis on *Bacillus subtilis*. *Current topics in microbiology and immunology* **322**, 1-16
160. Singh, P. K., Schaefer, A. L., Parsek, M. R., Moninger, T. O., Welsh, M. J., and Greenberg, E. P. (2000) Quorum-sensing signals indicate that cystic fibrosis lungs are infected with bacterial biofilms. *Nature* **407**, 762-764
161. Fedtke, I., Gotz, F., and Peschel, A. (2004) Bacterial evasion of innate host defenses - the *Staphylococcus aureus* lesson. *Int J Med Microbiol* **294**, 189-194
162. Epstein, A. K., Pokroy, B., Seminara, A., and Aizenberg, J. (2011) Bacterial biofilm shows persistent resistance to liquid wetting and gas penetration. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 995-1000
163. Kearns, D. B., Chu, F., Branda, S. S., Kolter, R., and Losick, R. (2005) A master regulator for biofilm formation by *Bacillus subtilis*. *Molecular microbiology* **55**, 739-749
164. Branda, S. S., Gonzalez-Pastor, J. E., Dervyn, E., Ehrlich, S. D., Losick, R., and Kolter, R. (2004) Genes involved in formation of structured multicellular communities by *Bacillus subtilis*. *Journal of bacteriology* **186**, 3970-3979
165. Terra, R., Stanley-Wall, N. R., Cao, G. Q., and Lazazzera, B. A. (2012) Identification of *Bacillus subtilis* SipW as a Bifunctional Signal Peptidase That Controls Surface-Adhered Biofilm Formation. *Journal of bacteriology* **194**, 2781-2790
166. Lewis, R. J., Brannigan, J. A., Smith, I., and Wilkinson, A. J. (1996) Crystallisation of the *Bacillus subtilis* sporulation inhibitor SinR, complexed with its antagonist, SinI. *FEBS letters* **378**, 98-100
167. Chai, Y. R., Chu, F., Kolter, R., and Losick, R. (2008) Bistability and biofilm formation in *Bacillus subtilis*. *Molecular microbiology* **67**, 254-263
168. Vlamakis, H., Aguilar, C., Losick, R., and Kolter, R. (2008) Control of cell fate by the formation of an architecturally complex bacterial community. *Genes & development* **22**, 945-953

169. Strauch, M., Webb, V., Spiegelman, G., and Hoch, J. A. (1990) The Spooa Protein of *Bacillus Subtilis* Is a Repressor of the *AbrB* Gene. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 1801-1805
170. Hamon, M. A., Stanley, N. R., Britton, R. A., Grossman, A. D., and Lazazzera, B. A. (2004) Identification of *AbrB*-regulated genes involved in biofilm formation by *Bacillus subtilis*. *Molecular microbiology* **52**, 847-860
171. Diethmaier, C., Pietack, N., Gunka, K., Wrede, C., Lehnik-Habrink, M., Herzberg, C., Hubner, S., and Stulke, J. (2011) A Novel Factor Controlling Bistability in *Bacillus subtilis*: the *YmdB* Protein Affects Flagellin Expression and Biofilm Formation. *Journal of bacteriology* **193**, 5997-6007
172. Chu, F., Kearns, D. B., Branda, S. S., Kolter, R., and Losick, R. (2006) Targets of the master regulator of biofilm formation in *Bacillus subtilis*. *Molecular microbiology* **59**, 1216-1228
173. Strauch, M. A., Bobay, B. G., Cavanagh, J., Yao, F., Wilson, A., and Le Breton, Y. (2007) *Abh* and *AbrB* control of *Bacillus subtilis* antimicrobial gene expression. *Journal of bacteriology* **189**, 7720-7732
174. Ritenberg, M., Nandi, S., Kolusheva, S., Dandela, R., Meijler, M. M., and Jelinek, R. (2016) Imaging *Pseudomonas aeruginosa* Biofilm Extracellular Polymer Scaffolds with Amphiphilic Carbon Dots. *Acs Chem Biol* **11**, 1265-1270
175. Flemming, H. C., Neu, T. R., and Wozniak, D. J. (2007) The EPS matrix: The "House of Biofilm cells". *Journal of bacteriology* **189**, 7945-7947
176. Ma, L., Jackson, K. D., Landry, R. M., Parsek, M. R., and Wozniak, D. J. (2006) Analysis of *Pseudomonas aeruginosa* conditional *Psl* variants reveals roles for the *Psl* polysaccharide in adhesion and maintaining biofilm structure postattachment. *Journal of bacteriology* **188**, 8213-8221
177. Laue, H., Schenk, A., Li, H., Lambertsen, L., Neu, T. R., Molin, S., and Ullrich, M. S. (2006) Contribution of alginate and levan production to biofilm formation by *Pseudomonas syringae*. *Microbiol-Sgm* **152**, 2909-2918
178. Franklin, M. J., Nivens, D. E., Weadge, J. T., and Howell, P. L. (2011) Biosynthesis of the *Pseudomonas aeruginosa* Extracellular Polysaccharides, Alginate, Pel, and *Psl*. *Frontiers in microbiology* **2**, 167
179. Wang, G. D., Zhu, L., Yu, B., Chen, K., Liu, B., Liu, J., Qin, G. Z., Liu, C. Y., Liu, H. X., and Chen, K. S. (2016) Exopolysaccharide from *Trichoderma pseudokoningii* induces macrophage activation. *Carbohydr Polym* **149**, 112-120
180. Byrd, M. S., Sadovskaya, I., Vinogradov, E., Lu, H. P., Sprinkle, A. B., Richardson, S. H., Ma, L. Y., Ralston, B., Parsek, M. R., Anderson, E. M., Lam, J. S., and Wozniak, D. J. (2009) Genetic and biochemical analyses of the *Pseudomonas aeruginosa* *Psl* exopolysaccharide reveal overlapping roles for polysaccharide synthesis enzymes in *Psl* and LPS production. *Molecular microbiology* **73**, 622-638
181. Franklin, M. J., and Ohman, D. E. (1993) Identification of *AlgF* in the Alginate Biosynthetic Gene-Cluster of *Pseudomonas-Aeruginosa* Which Is Required for Alginate Acetylation. *Journal of bacteriology* **175**, 5057-5065
182. Velazquez-Hernandez, M. L., Baizabal-Aguirre, V. M., Bravo-Patino, A., Cajero-Juarez, M., Chavez-Moctezuma, M. P., and Valdez-Alarcon, J. J. (2009) Microbial fructosyltransferases and the role of fructans. *Journal of applied microbiology* **106**, 1763-1778
183. Limoli, D. H., Jones, C. J., and Wozniak, D. J. (2015) Bacterial Extracellular Polysaccharides in Biofilm Formation and Function. *Microbiol Spectr* **3**
184. Franklin, M. J., Nivens, D. E., Weadge, J. T., and Howell, P. L. (2011) Biosynthesis of the *Pseudomonas aeruginosa* extracellular polysaccharides, alginate, Pel, and *Psl*. *Frontiers in microbiology* **2**

185. Jiang, C. H., Fan, Z. H., Xie, P., and Guo, J. H. (2016) *Bacillus cereus* AR156 Extracellular Polysaccharides Served as a Novel Micro-associated Molecular Pattern to Induced Systemic Immunity to Pst DC3000 in Arabidopsis. *Frontiers in microbiology* 7

CHAPTER 2

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 THURINGIENSIS¹

¹ Li, Z., Hwang, S., Ericson, J., Bowler, K., and Bar-Peled, M. 2015. *The Journal of biological chemistry*. 290, 691-704.

Reprinted here with permission of the publisher.

2.1 Abstract

CMP-pseudaminic acid is a precursor required for the O-glycosylation of flagellin in some pathogenic Gram-negative bacteria, a process known to be critical in bacterial motility and infection. However, little is known about flagellin glycosylation in Gram-positive bacteria. Here, we identified and functionally characterized an operon, named *Bti_pse*, in *Bacillus thuringiensis israelensis* ATCC 35646, which encodes seven different enzymes that together convert UDP-GlcNAc to CMP-pseudaminic acid. In contrast, Gram-negative bacteria complete this reaction with six enzymes. The first enzyme, which we named Pen, converts UDP-D-GlcNAc to an uncommon UDP-sugar, UDP-6-deoxy-D-GlcNAc-5,6-ene. Pen contains strongly bound NADP⁺ and has distinct UDP-GlcNAc 4-oxidase, 5,6-dehydratase, and 4-reductase activities. The second enzyme, which we named Pal, converts UDP-6-deoxy-D-GlcNAc-5,6-ene to UDP-4-keto-6-deoxy-L-AltNAc. Pal is NAD⁺-dependent and has distinct UDP-6-deoxy-D-GlcNAc-5,6-ene 4-oxidase, 5,6-reductase, and 5-epimerase activities. We also show here using NMR spectroscopy and mass spectrometry that in *B. thuringiensis*, the enzymatic product of Pen and Pal, UDP-4-keto-6-deoxy-L-AltNAc, is converted to CMP-pseudaminic acid by the sequential activities of a C4''-transaminase (Pam), a 4-N-acetyltransferase (Pdi), a UDP-hydrolase (Phy), an enzyme (Ppa) that adds phosphoenolpyruvate to form pseudaminic acid, and finally a cytidyltransferase that condenses CTP to generate CMP-pseudaminic acid. Knowledge of the distinct dehydratase-like enzymes Pen and Pal and their role in CMP-pseudaminic acid biosynthesis in Gram-positive bacteria provides a foundation to investigate the role of pseudaminic acid and flagellin glycosylation in *Bacillus* and their involvement in bacterial motility and pathogenicity.

2.2 Introduction

Bacillus thuringiensis is a gram-positive bacterium which has been isolated from diverse habitats including soil, water, dusts, plants, cadavers and the intestines of insects (1-5). *B. thuringiensis* belongs to the *Bacillus cereus* group, which includes the two most notably pathogens: *B. cereus* and *B. anthracis*, the causal agents of food poisoning and anthrax, respectively (6). Members of this group have similar genetic backgrounds, but are distinguished by their host specificities and their pathogenicity (7). For example, *B. thuringiensis* is characterized by its ability to form crystalline proteins (Bt-toxin) that are lethal to many insects (8). Indeed, *B. thuringiensis* is used widely to control agricultural pests including *Lepidoptera*, *Diptera*, and *Coleoptera* sp. (9) and may also have a role in controlling *Anopheles gambiae*, the principal vector of malaria (10)

B. thuringiensis as well as other *Bacillus* species are known to exist in the gut microflora of numerous insects (11,12). In this environment, some of these bacteria lose their flagella and become attached to the insects intestinal epithelium (13). Recent cell biology studies have identified *B. thuringiensis* israelensis flagellum structures (7); and genomic analyses revealed that the bacterium has all of the protein components required to make flagellum and promote motility (14-16). Nevertheless, the role of flagellum structures in *Bacillus* and its involvement in delivering the Bt toxin to the insect gut, or in pathogenicity, remains to be determined.

Flagellin, one of the proteins of the flagellum apparatus, is a globular protein monomer that stacks helically in the form of a hollow cylinder to build the filament of the flagellum (17-19). Studies of *Bacillus* sp. PS3 flagellin have shown that it is *O*-glycosylated, although the sugar involved in this modification was not identified (20). The results of preliminary studies in our

laboratory⁵ have indicated that in *B. thuringiensis israelensis* the flagellin is glycosylated with a sugar residue with a mass consistent with a diacetylamino-tetradecoxy-nonulosonic acid.

Several studies have shown that flagella are essential for motility, adhesion, host interactions, and secretion during many of the life stages of Gram-negative bacteria (21,22). In *Campylobacter* spp. and *Helicobacter pylori*, the central domain of flagellin is post-translationally modified by the addition of pseudaminic acid (5,7-diacetamido-3,5,7,9-tetradecoxy-L-glycero- α -L-manno-2-nonulopyranosonic acid, Pse⁴) (23,24) that are *O*-linked to serine or threonine residues (25). Mutants lacking glycosylated flagellin are affected in both their motility and their ability to infect their host (26).

Interest in the function of Pse in Gram-negative bacteria led to the identification of a biosynthetic pathway, named *pse*, where CMP-pseudaminic acid (CMP-Pse) is formed from UDP-GlcNAc (27-29). This pathway involves six enzymes (Fig. 2.1A): A single bifunctional 4,6-dehydratase/5-epimerase (PseB) converts UDP-D-GlcNAc to UDP-4-keto-6-deoxy-L-AltNAc, which is then aminated at C4'' and N-acetylated to form UDP-2,4,6-trideoxy-2,4-NAc-L-altrose. UDP is cleaved from the sugar and pyruvate is added to form Pse. Pse is then activated by the addition of CMP to form CMP-Pse. By contrast, little is known about CMP-Pse formation in Gram-positive bacteria.

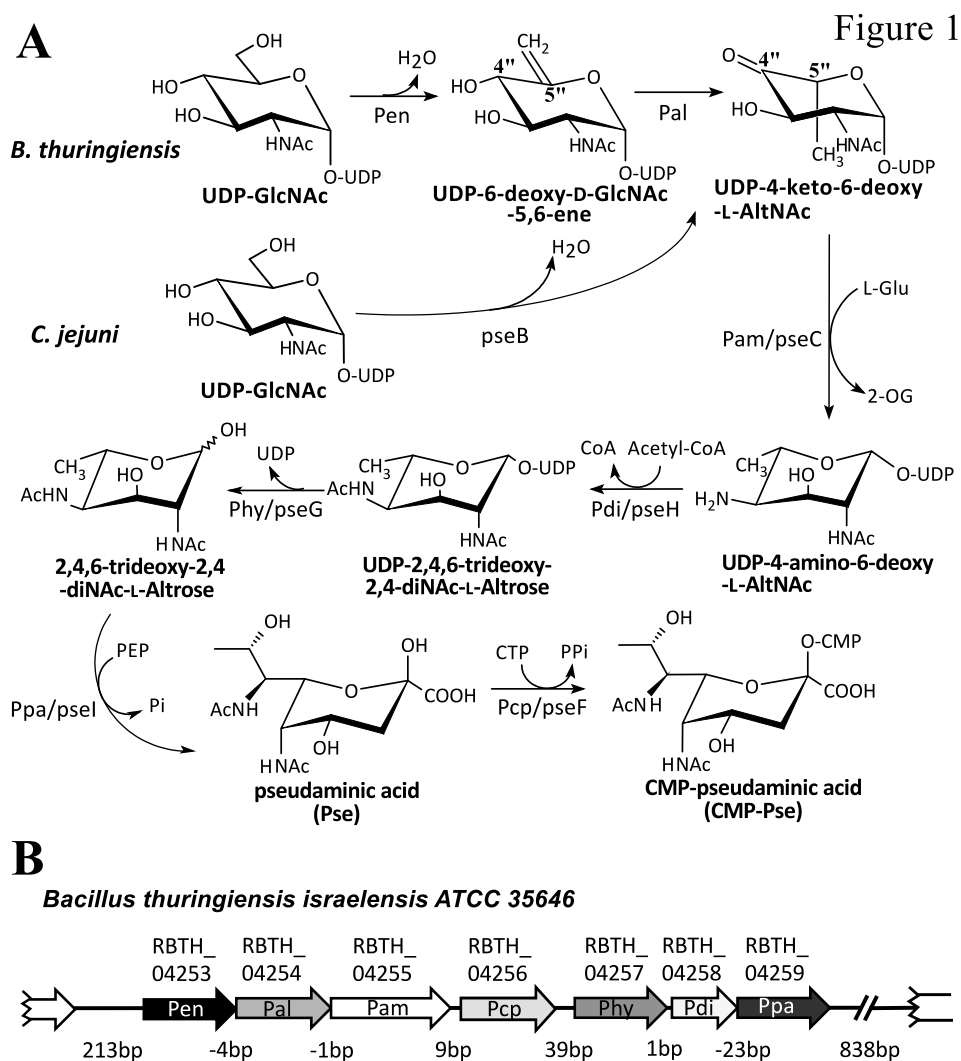


FIGURE 2.1. The proposed biosynthetic pathway of CMP-Pse in the Gram-positive strain *Bacillus thuringiensis israelensis* ATCC 35646 (Bti) requires seven proteins, while six are needed in Gram-negative, *C. jejuni*. Panel **A**. In the *C. jejuni* pathway, a single enzyme PseB converts UDP-2-deoxy-2-acetoamido-D-glucose, UDP-D-GlcNAc, to UDP-4-keto-6-deoxy-L-AltNAc. In the *B. thuringiensis* pathway, two enzymes are needed. First, Pen converts UDP-GlcNAc with enzyme-bound NAD⁺ to UDP-6-deoxy-D-GlcNAc-5,6-ene, then Pal converts it to UDP-4-keto-6-deoxy-L-AltNAc. From this point, both pathways carry out similar enzymatic reactions (although the amino-acid sequence for each specific enzyme is not conserved between species, Table 2.1) leading to the final product of CMP-Pse. Panel **B**. Organization of the seven-genes *pse* operon and flanking regions in *B. thuringiensis israelensis* ATCC 35646. The locus number for each enzyme-encoding gene in the operon is shown.

Here, we report the identification of the *B. thuringiensis pse* operon and the functional characterization of the enzyme cluster that converts UDP-GlcNAc to CMP-Pse. Seven enzymes are required for the conversion of UDP-GlcNAc to CMP-Pse in *B. thuringiensis* (Fig. 2.1B), unlike Gram-negative bacteria that complete this reaction with six enzymes. Two dehydratase-

like activities, which we named Pen and Pal, initiate the Pse pathway in *B. thuringiensis*. We used a combination of NMR spectroscopy and mass spectrometry to show that Pen converts UDP-D-GlcNAc to UDP-2-acetamido-6-deoxy- α -D-xylo-hexopyranose-5,6-ene (herein abbreviated UDP-6-deoxy-D-GlcNAc-5,6-ene) and that Pal converts UDP-6-deoxy-D-GlcNAc-5,6-ene to UDP-2-acetamido-2,6-dideoxy- β -L-arabino-hex-4-ulose (herein abbreviated UDP-4-keto-6-deoxy-L-AltNAc). Only one enzyme, PseB, is required to convert UDP-D-GlcNAc to UDP-4-keto-6-deoxy-L-AltNAc in Gram-negative bacteria. The identification of Pen and Pal and the CMP-Pse biosynthetic pathway in *B. thuringiensis* provides a basis for determining the biological role of Pse in flagellin glycosylation in gram-positive bacterium.

2.3 Experimental Procedure

2.3.1 Strains and culture conditions:

The *Bacillus thuringiensis* israelensis ATCC 35646 used in this study was stored in 16% glycerol at -80 °C, streaked onto agar medium, and grown for 18 h at 30 °C. The media (agar or liquid) used were Luria Bertani (LB) (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl) and LB-Lennox (10 g/L peptone, 10 g/L NaCl, 5 g/L yeast extract). *Escherichia coli* DH10B cells were used for cloning and Rosetta2(DE3)pLysS (Novagen) cells were used to express protein.

2.3.2 Cloning genes of the pse operon:

A single colony of *B. thuringiensis* growing on LB-Lennox agar was suspended in 50 μ l sterile water, heated for 5 min at 95 °C and then centrifuged (10,000 x g, 30 sec). A portion of the supernatant (5 μ l) was used as the source of DNA to amplify each specific gene by PCR. Each primer was designed to include a 15-nucleotide extension with exact sequence homology to the cloning site of the target plasmid at the 5' end. This facilitated cloning between the NcoI and

primer sets ZL271 (5'-cacagccaggatccgttttttaa gataaaaatgttttaattattgg-3'), ZL272 (5'-tcgacttaagca ttatattttctcccctcctgttaatagc-3') and ZL273 (5'-caca gccaggatccgttagatattgaggattttcaattgg-3'), and set ZL274 (5'-tcgacttaagcattactcattttttcaccttcaagttc-3'), respectively. After PCR, portions (2 µl each) of the amplified vector and insert were mixed, and treated for 15 min at 37 °C with 0.5 U FastDigest DpnI (ThermoScientific), and then transformed into DH10B (LifeTechnologies) competent cells. Positive clones were selected on LB agar containing either kanamycin (50 µg/ml, for pET vector) or spectinomycin (50 µg/ml, for pCDF vector). Clones were verified by PCR and DNA sequencing. Plasmids harboring specific genes were sequenced and the DNA sequences were deposited in Genbank (with respective accession numbers KM433664 and KM433665). The resulting recombinant plasmids are: pCDF:His₆*rbth*_04253 yielding N-terminal His₆-tagged Pen; pET28a:*rbth*_04254His₆ yielding C-terminal Pal; pET28b:His₆-Tev-*rbth*_04255 yielding N-terminal His₆-tag followed by Tev-protease and Pam; pET28a:*rbth*_04256His₆ yielding C-terminal His₆-tagged Pcp; pET28a:*rbth*_04257His₆ yielding C-terminal His₆-tagged Phy; pCDF:His₆*rbth*_04258 yielding N-terminal His₆-tagged Pdi; and pET28a:*rbth*_04259His₆ yielding C-terminal His₆-tagged Ppa. Plasmids were extracted by PureLink Quick Plasmid Miniprep Kit (Invitrogen) and transformed into Rosetta2(DE3)pLysS competent cells for recombinant protein expression.

2.3.3 His₆-tagged protein expression and purification:

Rosetta2(DE3)pLysS strains harboring expression plasmids were grown at 37 °C and 250 rpm in 250 ml LB supplemented with chloramphenicol (35 µg/ml) and kanamycin (50 µg/ml) for the pET28a and pET28b vectors, or with chloramphenicol (35 µg/ml) and spectinomycin (50 µg/ml) for the pCDF vector. Gene expression was induced when cell OD₆₀₀ reached 0.6, by

adding 0.5 mM isopropyl-1-thio- β -D-galactopyranoside (IPTG). After induction, cells were grown for 18 h at 18 °C and 250 rpm and then harvested by centrifugation (6,000 x g) for 10 min at 4 °C. The cell pellets were washed with water and then suspended in 10 ml lysis buffer (50 mM Tris-HCl, pH 7.4, 10% (v/v) glycerol, 1mM EDTA, 2 mM DTT, and 0.5 mM PMSF). Cells were lysed by sonication (31) and after centrifugation (6,000 x g for 15 min at 4 °C), the supernatant was supplemented with 1 mM DTT and 0.5 mM PMSF and re-centrifuged at 20,000 g for 30 minutes at 4 °C. An aliquot (5 ml) of the supernatant was applied to a Ni-Sepharose fast-flow column (GE Healthcare Life Sciences, Piscataway, NJ, USA; 2 mL of resin packed in a polypropylene column; inner diameter 1 cm $\times$ 15 cm). Each column was pre-equilibrated with buffer A (50 mM Tris-HCl, pH 8, 10% (v/v) glycerol, 100 mM NaCl). The column was washed with 30 ml of buffer A containing 20 mM imidazole and then with 10 ml of buffer A containing 40 mM imidazole. His-tagged proteins were eluted with 5 ml of buffer A containing 250 mM imidazole. The eluates containing these proteins were divided into small aliquots, flash frozen in liquid nitrogen, and kept at -80 °C. The concentration of each protein was determined (32) from the $A_{280\text{nm}}$ (with epsilon (ϵ) = 14,200 cm⁻¹ M⁻¹ for His₆*rbth_04253*, ϵ = 40,120 cm⁻¹ M⁻¹ for *rbth_04254*His₆). Proteins were separated by SDS-12.5% PAGE and visualized by staining with Coomassie blue.

2.3.4 Enzyme reactions:

The activity of recombinant His₆*rbth_04253* (refer herein as Pen) was examined by HILIC-HPLC with UV or electrospray ionization-mass spectrometry (ESI-MS) detection, and by time-resolved ¹H NMR spectroscopy. For HPLC-based assays, the total reaction volume was 50 μ l and included 50 mM Tris-HCl pH 7.6, 1 mM UDP-GlcNAc, and about 1.5 μ g of purified Pen. Reactions were proceed for 1 h at 30 °C, followed by enzyme inactivation (95 °C for 2 min) and

extraction with 50 μ l chloroform (30). An aliquot (20 μ l) of the upper aqueous layer was removed and mixed with acetonitrile (40 μ l) and 0.5M ammonium-acetate pH 5.3 (2 μ l). A portion (20- μ l) was analyzed using HILIC chromatography coupled to a ESI-MS/MS mass spectrometer or a UV diode array detector. ESI-MS/MS was performed using a Shimadzu LC-MS/MS IT-TOF MS system operating in the negative ion mode with a Nexera UFPLC LC-30AD pump, autosampler (Sil30), and column heater (set to 37 °C). HPLC with UV detection was carried out with an Agilent 1260 pump system equipped with an autosampler, column heater (set at 37 °C), and diode array detector ($A_{261\text{nm}}$). Enzyme reactions were separated on an Accucore 150-amid HILIC column (150 $\times$ 4.6 mm, 2.6 μ m particle size, ThermoScientific) using 40 mM ammonium-acetate pH 4.3 (solvent A) and acetonitrile (solvent B). The column was equilibrated at 0.4 ml/min with 25% A and 75% B prior to sample injection (20 μ l). Following injection, the HPLC conditions were: 0-1 min, 0.4 ml/min with 25% A/75%B then a gradient to 50% A and 50% B over 24 min. The flow rate was then increased to 0.6 ml/min with a gradient to 25% A and 75% B over 5 min. The column was then washed for 5 min with 25%A/75%B prior to the next injection. HPLC peaks of enzymatic reaction products detected by their $A_{261\text{nm}}$ (max for UDP-sugars) were collected, lyophilized, and suspended in D₂O (99.9%) for NMR analysis or in H₂O for MS/MS analysis.

The activity of recombinant *rbth_04254*His₆ (herein refer as Pal) was examined by HILIC-HPLC-UV, ESI-MS and time-resolved proton NMR. The 50 μ l HPLC-based assays included 50 mM Tris-HCl pH 7.6, about 1 mM of the product (purified by HILIC) formed by Pen and 8.5 μ g of purified Pal. The reaction was proceeded for 1 h at 30 °C, followed by inactivation (95 °C for 2 min) and chloroform extraction.

NMR-based Pen assays in a total volume of 180 μ l were performed in a mixture of 90 μ l D₂O and 90 μ l H₂O containing 50 mM Tris-HCl pH 7.6, 1 mM UDP-GlcNAc, 0.1 mM DSS, (2,2-dimethyl-2-silapentane-5-sulfonate that served as NMR reference) and $\sim$ 3 μ g of purified Pen. The reaction mixture was transferred to a 3-mm tube and time-resolved ¹H NMR spectra then obtained for up to 2 h at 30 °C using a Varian DirectDrive 600-MHz spectrometer equipped with a cryogenic probe. Data was acquired before the addition of the enzyme at time zero (*t*₀). After adding the enzyme, data acquisition was started after $\sim$ 5 min in to provide sufficient time to optimize magnet shimming. Sequential one-dimensional proton spectra (32 scans each) with presaturation of the water resonance were acquired over the course of the enzymatic reaction. All NMR spectra were referenced to the resonance of DSS set at 0.00 ppm.

NMR-based Pal assays in a total volume of 180 μ l were performed in 90 μ l D₂O and 90 μ l H₂O containing 50 mM Tris-HCl pH 7.6, $\sim$ 1 mM of the product formed by Pen, 0.1 mM DSS, and about 17 μ g of purified Pal. Time-resolved ¹H NMR acquisition was as described above.

2.3.5 Kinetic studies:

The linear dependence of enzyme concentration with respect to initial velocity was established by changing the protein concentration and maintaining the substrates at constant concentration. *K_m* values were determined using regression analysis (non-linear) with Prism software from plots of initial velocities versus varied substrate concentrations (10, 20, 40, 60, 80, 100, 150, 300 and 500 μ M UDP-GlcNAc).

2.3.6 Gel filtration assays:

Purified Pen or Pal protein (in volume of 300 μ l) was mixed with 50 μ M UDP-GlcNAc and chromatographed on a Superdex 75 column (1 cm i.d. x 30 cm, GE) at a flow rate of 0.5 ml/min, using 50 mM Tris-HCl pH 8, 150 mM NaCl buffer. Peaks observed at 215 nm were

collected and assayed for enzyme activity. The column was calibrated using a molecular mass standard (BioRad).

2.3.7 Determination of Pen and Pal enzyme-bound co-factors:

Purified Pen (600 μ l of 0.63 mg/ml) and purified Pal (1.2 ml of 1.62 mg/ml) were heated for 5 min at 95 °C and then centrifuged at 14,000 x g for 1 min. The supernatant was collected and concentrated to 100 μ l using a Speed-Vac (Savant SC110). A 30- μ l aliquot was analyzed by HILIC-ESI-MS/MS for the detection of NAD⁺ or NADP⁺. Both positive mode and negative mode were used with 25% CID energy for MS/MS fragmentation.

2.3.8 NMR spectroscopy used to characterize product structures:

The HPLC peak corresponding to the enzymatic product of Pen was collected, lyophilized, dissolved in D₂O (99.9%), and analyzed by 2D NMR. Proton chemical shifts were assigned by a COrrelation SpectroscopY (COSY) experiment and verified by a TOtal Correlation SpectroscopY (TOCSY) experiment. Protonated carbon chemical shifts as well as multiplicities were determined by a multiplicity-edited Heteronuclear Single Quantum Coherence (HSQC) experiment. The chemical shift of non-protonated carbon C5'' was assigned by a Heteronuclear Multiple Bond Correlation (HMBC) experiment.

To characterize the structure of the products formed by Pal, the reaction assay of Pal was extracted with chloroform and the aqueous phase chromatographed on a Q15 anion-exchange column as described (33). The peak corresponding to the product was collected, lyophilized, suspended in 99.9% D₂O, and analyzed by COSY, TOCSY, HSQC, and HMBC experiments.

2.3.9 LC-MS/MS analysis:

Reaction products were chromatographed on a HILIC column using a Shimadzu CBM-

20A HPLC system equipped with an autosampler, coupled to a Shimadzu LCMS-IT-TOF ESI, mass spectrometer operated in the negative mode. The HILIC chromatography conditions were the same as the HPLC conditions used for the enzymatic reaction section described above. Enzyme products were identified based on their retention time, the mass of their parent ion and their mass spectral fragmentation pattern.

2.4 Results

2.4.1 CMP-pseudaminic acid operon in *B. thuringiensis* consists of seven genes.

We first performed a BLAST search using amino acid sequences of known bacterial enzymes involved in CMP-Pse formation in order to identify *B. thuringiensis* genes that may have a role in CMP-Pse formation. The BLAST approach however, identified several misleading gene targets. For example, the first enzyme in the pseudaminic acid pathway in *Helicobacter pylori* is a bi-functional UDP-GlcNAc 4,6-dehydratase/5-epimerase (PseB) that forms, UDP-4-keto-6-deoxy-L-AltNAc. PseB has 46% amino acid sequence similarity (e-value $3e^{-84}$) to *rbth_05809* of *B. thuringiensis* israelensis ATCC 35646 and to *bc3750* of *Bacillus cereus* ATCC 14579. However, *bc3750* was found not to encode a 5-inverting 4,6-dehydratase activity⁶. *rbth_04253*, (i.e. Pen) has a 40% amino acid sequence similarity to PseB (e-value $7e^{-72}$, see Table 2.1). Adjacent to the Pen is *rbth_04254* (Pal), annotated as a dehydratase (Fig. 2.1B) and has a low sequence identity to PseB (24% see, Table 2.1). *rbth_04255* (Pam), which has low (35%; e-value of $8e^{-71}$) amino acid identity with the functional aminotransferase PseC from *H. pylori*, follows the two annotated dehydratases. *rbth_04258* (Pdi), has low amino acid identity (30%; e-value of $4e^{-11}$) to the functional N-acetyltransferase PseH; whereas *rbth_04257* (Phy) has only 26% amino acid identity (e-value of $3e^{-9}$) to *H. pylori* PseG; *rbth_04259* (Ppa) shares sequence identity (41%; e-value of $1e^{-79}$) to *H. pylori* pseudaminic acid synthase PseI, and

rbth_04256 (Pcp) protein that shares very low sequence identity (29%; e-value 0.001) to functional pseudaminic acid cytidylyltransferase PseF from *Campylobacter jejuni*, while no similarity was detected compared to *H. pylori* PseF (see Table 2.1).

TABLE 2.1. Amino acid sequence similarity between various proteins involved in the synthesis of CMP-pseudaminic acid. Comparison of *pse* metabolic proteins between *B. thuringiensis* (Bti), *C. jejuni* and *H.pylori*

Bti pse	C.jejuni 81-176	E-value	Identity	Function in Bti
<i>rbth_04253</i> Pen	<i>cj1293</i> PseB	2e ⁻⁶⁹	37%	UDP-GlcNAc 4-oxidase/5,6-dehydratase/4-reductase
<i>rbth_04254</i> Pal	<i>cj1293</i> PseB	4e ⁻⁰⁴	24%	UDP-6-deoxy-D-GlcNAc-5,6-ene 4-oxidase/5,6-reductase/5-epimerase
<i>rbth_04255</i> Pam	<i>cj1294</i> PseC	4e ⁻⁷⁷	39%	C4'' aminotransferase
<i>rbth_04258</i> Pdi	<i>cj1313</i> PseH	1e ⁻⁰⁵	22%	C4'' N-acetyltransferase
<i>rbth_04257</i> Phy	<i>cj1312</i> PseG	6e ⁻²⁰	31%	UDP-sugar hydrolase
<i>rbth_04259</i> Ppa	<i>cj1317</i> PseI	4e ⁻⁸⁷	45%	Pse synthase
<i>rbth_04256</i> Pcp	<i>cj1311</i> PseF	0.001	29%	CMP-Pse synthase
Bti pse	H.Pylori 26695	E-value	Identity	Function in Bti
<i>rbth_04253</i> Pen	<i>hp0840</i> PseB	7e ⁻⁷²	40%	UDP-GlcNAc 4-oxidase/5,6-dehydratase/4-reductase
<i>rbth_04254</i> Pal	<i>hp0840</i> PseB	5e ⁻⁰⁸	22%	UDP-6-deoxy-D-GlcNAc-5,6-ene 4-oxidase/5,6-reductase/5-epimerase
<i>rbth_04255</i> Pam	<i>hp0366</i> PseC	8e ⁻⁷¹	35%	C4'' aminotransferase
<i>rbth_04258</i> Pdi	<i>hp0327</i> PseH	4e ⁻¹¹	30%	C4'' acetyltransferase
<i>rbth_04257</i> Phy	<i>hp0326b</i> PseG	3e ⁻⁹	26%	UDP hydrolase
<i>rbth_04259</i> Ppa	<i>hp0178</i> PseI	1e ⁻⁷⁹	41%	Pse synthase
<i>rbth_04256</i> Pcp	<i>hp0326a</i> PseF	No similarity		CMP-Pse synthase

Although the encoded proteins in this *B. thuringiensis* operon share an overall low amino acid sequence identity to known functional genes involved in CMP-Pse synthesis, we decided to clone and characterize these proteins since sugar (Pse) modification of flagella proteins have not been reported previously in *Bacillus*. Below, we describe that seven enzymes are required to form CMP-Pse from UDP-GlcNAc in *B. thuringiensis* (Fig. 2.1B), unlike Gram-negative bacteria that complete this reaction with six enzymes (27).

2.4.2 Biochemical characterization of CMP-pseudaminic acid operon in *B. thuringiensis*.

Each of the seven predicted Pse pathway genes were cloned from *B. thuringiensis* and expressed in *E. coli* as recombinant His₆-tagged proteins. Each protein was purified over a nickel-affinity column. Individual protein bands for recombinant Pen (40 kDa); Pal (37.1 kDa); Pam (47.6 kDa); Pcp (39.5 kDa); Phy (42.8 kDa); Pdi (24 kDa); Ppa (40.7 kDa) were detected by SDS-PAGE (Fig. 2.2A). The ability of the recombinant enzymes to catalyze the various steps in the formation of CMP-Pse was then tested. As UDP-GlcNAc is the only commercially available substrate, we used the product formed by one enzyme as the substrate for the next enzyme. For example, recombinant Pen and Pal proteins were reacted with UDP-GlcNAc and the product was then used as the substrate for Pam, the putative L-Glu:C4''-transaminase. The UDP-4-amino-sugar formed together with Acetyl-CoA was the substrate for Pdi, an Acetyl-CoA:C4''-N-acetyltransferase. The resulting product (UDP-2,4,6-trideoxy-2,4-diNAc-L-altrose) then served as the substrate for the UDP-sugar hydrolase Phy whose reaction product, along with phosphoenolpyruvate (PEP) and Ppa, led to Pse. In the final reaction, Pse was reacted with CTP and recombinant Pcp to generate CMP-Pse (structure shown in Fig. 2.2B). The progression of each reaction was monitored by LC-MS based assays.

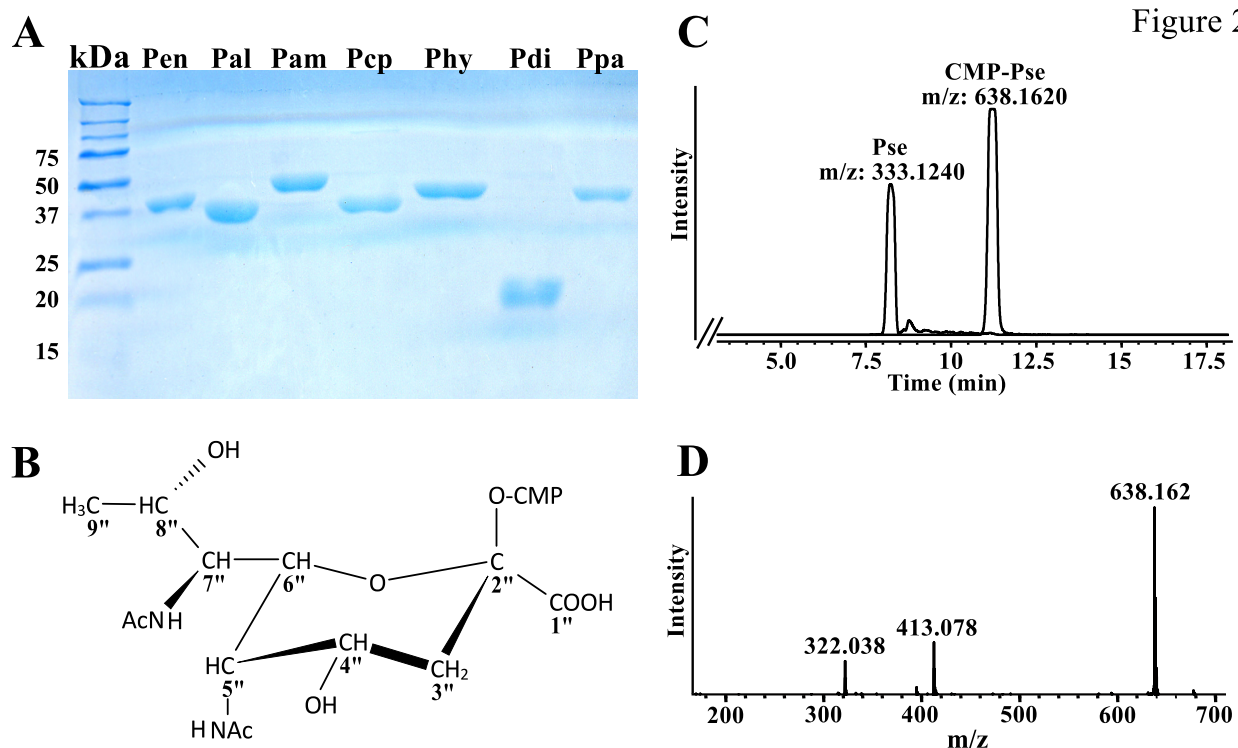


FIGURE 2.2. Purification of the seven *B. thuringiensis israelensis* ATCC 35646 (Bti) recombinant enzymes and their involvement in the biosynthesis of CMP-pseudaminic acid (CMP-Pse). Panel **A**. SDS-PAGE analysis of each of the Bti CMP-Pse biosynthetic enzymes after nickel column purification. Protein standards are shown on the left in kDa. Lane 1, Pen (40 kDa); lane2, Pal (37.1 kDa); lane3, Pam (47.6 kDa); lane4, Pcp (39.5 kDa); lane5, Phy (42.8 kDa); lane 6, Pdi (24 kDa); lane7, Ppa (40.7 kDa). Panel **B**. ¹H and ¹³C NMR spectroscopic analyses of the peak eluted from HILIC column at 11 min gives a product with chemical shift corresponding to CMP-Pse. The structure of CMP-Pse is shown (also see Table 2.2). Panel **C**. LC-MS analysis of the final biochemical step converting CTP + Pse to CMP-Pse, as determined after separation of reaction product by HILIC column. The peaks with [M-H]⁻ ions at *m/z* 638.1 and 333.1, respectively corresponding to the predicted mass of CMP-Pse and Pse. Panel **D**. MS/MS analysis at collision energy 25% of parent ion 638 (CMP-Pse) gives a predicted ion fragments of Pse-phosphate (*m/z* 413) and CMP (322). The structure of CMP-Pse is shown.

HILIC-ESI-MS analysis of the products of the final enzymatic step (Fig. 2.2C) gave two peaks with retention times of 8.2 min and 11.2 min and [M-H]⁻ at *m/z* 333.12 and 638.16, respectively. These two values suggest a mass for 9-carbon sugar and its nucleotide-derivative CMP-diacetyl-amino-tetradeoxy-nonulosonic acid, respectively. NMR analysis (Table 2.2) established that the peak eluting at 11.2 min is indeed CMP-5,7-diacetamido-3,5,7,9-tetradeoxy-L-glycero- α -L-manno-2-nonulopyranosonic acid (CMP-Pse). The enzymatic product is not CMP-legionaminic acid (CMP-Leg), or other CMP-9-carbon sugars such as CMP-Leg 4- and 8-epimers, as the coupling constant between H5'' and H6'' (1.5 Hz) is small, (expected for CMP-

Pse) and not large (10.3 Hz) (expected for CMP-Leg). Fragmentation of the ion at m/z 638.16 gave two ions with m/z 322.04 and 413.08, (Fig. 2.2D) corresponding to CMP and pseudaminic acid-1-phosphate, respectively and is consistent with the presence of CMP-Pse. Like-wise the peak eluting at 8.2 min (Fig. 2.2C) with $[M-H]^-$ at m/z 333.12 is Pse.

TABLE 2.2 NMR data for the sugar moiety of CMP-Pse, product C, H, F
CMP-pseudaminic acid

^1H	Chemical shift (ppm)	^{13}C	Chemical shift (ppm)	$J_{\text{H,H}}$ (Hz)
H3'' _{ax}	1.59	C3''	36.7	$J_{3\text{ax},3\text{eq}}$ (13.5) $J_{3\text{ax},4}$ (12.3)
H3'' _{eq}	2.20	C3''	36.7	$J_{3\text{eq},4}$ (4.6)
H4''	4.22	C4''	64.5	$J_{4,5}$ (4.2)
H5''	4.26	C5''	49.5	$J_{5,6}$ (1.7)
H6''	4.28	C6''	73.4	$J_{6,7}$ (10)
H7''	4.01	C7''	54.5	$J_{7,8}$ (4.8)
H8''	4.10	C8''	69.4	$J_{8,9}$ (6.4)
H9''	1.18	C9''	17.9	
5''NAc-CH ₃	1.98	5''NAc-CH ₃	22.6	
5''NAc-NH	8.49			
7''NAc-CH ₃	1.95	7''NAc-CH ₃	22.7	
7''NAc-NH	8.04			

C: UDP-6-deoxy-D-GlcNAc-5,6-ene

^1H	Chemical shift (ppm)	^{13}C	Chemical shift (ppm)	$J_{\text{H,H}}$ (Hz)
H1''	5.56	C1''	94.23	$J_{1,\text{p}}$ (7.6)
H2''	4.18	C2''	53.25	$J_{1,2}$ (3.1)
H3''	3.71	C3''	70.87	$J_{2,3}$ (9.9)
H4''	4.09	C4''	71.15	$J_{3,4}$ (9.5)
		C5''	154.93	$J_{4,6\text{a}}$ (2.1)
H6 _a ''	4.89	C6''	97.55	$J_{4,6\text{b}}$ (2.1)
H6 _b ''	4.87			$J_{6\text{a},6\text{b}}$ (1.8)
2''NAc-CH ₃	2.07	2''NAc-CH ₃	22.56	
2''NAc-NH	8.36	2''NAc-C=O	175.03	

H: UDP-4-keto-6-deoxy-L-AltNAc-hydrated

¹ H	Chemical shift (ppm)	¹³ C	Chemical shift (ppm)	J _{H,H} (Hz)
H1''	5.56	C1''	97.00	J _{1,P} (8.5)
H2''	4.25	C2''	55.36	J _{1,2} (3.0)
H3''	3.91	C3''	71.96	J _{2,3} (8.0)
		C4''	95.93	
H5''	4.06	C5''	78.44	
H6''	1.37	C6''	18.43	J _{5,6} (6.9)

F: UDP-4-keto-6-deoxy-D-GlcNAc-hydrated

¹ H	Chemical shift (ppm)	¹³ C	Chemical shift (ppm)	J _{H,H} (Hz)
H1''	5.45	C1''	94.20	J _{1,P} (6.9)
H2''	4.10	C2''	52.32	J _{1,2} (3.5)
H3''	3.82	C3''	71.37	J _{2,3} (11)
H5''	4.12	C5''	69.76	
H6''	1.24	C6''	11.31	J _{5,6} (6.2)

We conclude that the seven-gene cluster between *rbth_04253-04259* in *B. thuringiensis israelensis* ATCC 35646 encodes all the enzymes required to convert UDP-GlcNAc to CMP-Pse (see Fig. 2.1A). Thus we called this operon the *Bti_Pse* operon. Since the initial steps of the reaction involve two uncharacterized dehydrate-like enzyme activities, we fully characterized the activities of Pen and Pal.

2.4.3 Characterization of a unique UDP-sugar: UDP-6-deoxy-D-GlcNAc-5,6-ene formed by Pen, a UDP-GlcNAc 4-oxidase/5,6-dehydratase/4-reductase.

Our early analysis indicated that in *B. thuringiensis*, two dehydratase-like proteins are required to initiate conversion of UDP-GlcNAc to a product that can be used for CMP-Pse formation. We therefore studied these two enzymes in more detail. Purified recombinant Pen (His₆ *rbth_04253*) was reacted with UDP-GlcNAc and the products formed were analyzed by HILIC HPLC with UV detection. A new peak with a retention time of 14.5 min (Fig. 2.3A) was detected; a negative control reaction gave no new product. The peak was collected and examined

by ESI-MS and MS/MS. The negative ion mode mass spectrum contained an ion $[M-H]^-$ at m/z 588.06 (Fig. 2.3B). MS/MS analysis of this ion gave two fragments at m/z 323 and 403 consistent with UMP and UDP, respectively, suggesting that the product was a UDP-sugar. The neutral loss of 18 amu suggested the UDP-GlcNAc product lacks the mass of water, implying that Pen is a dehydratase. However, further analysis of the enzyme product demonstrated that the enzyme is not a 'regular' 4,6-dehydratase since no UDP-4-keto-sugar product was detected. Time-resolved 1H -NMR of the reaction showed the conversion of UDP-GlcNAc to a new product (labeled C) with a quartet of signals at 5.56 ppm (Fig. 2.3C and see zoom below). We also observed that this product lacks signals for a C6-methyl and a C4-keto sugar. Most known 4,6-dehydratases form a 4-keto sugar with C6-methyl group (i.e UDP-4-keto-6-deoxy-sugar). Moreover, there was a new signal with a chemical shift of 4.09 ppm, which likely arose from the proton linked to C4'' of the UDP-sugar. This suggested that Pen is a 4-oxidase/5,6-dehydratase/4-reductase and forms a product that has a double bond between C5'' and C6'' (i.e. 5,6-ene). A detailed explanation of the elucidation of the structure of this product is provided below.

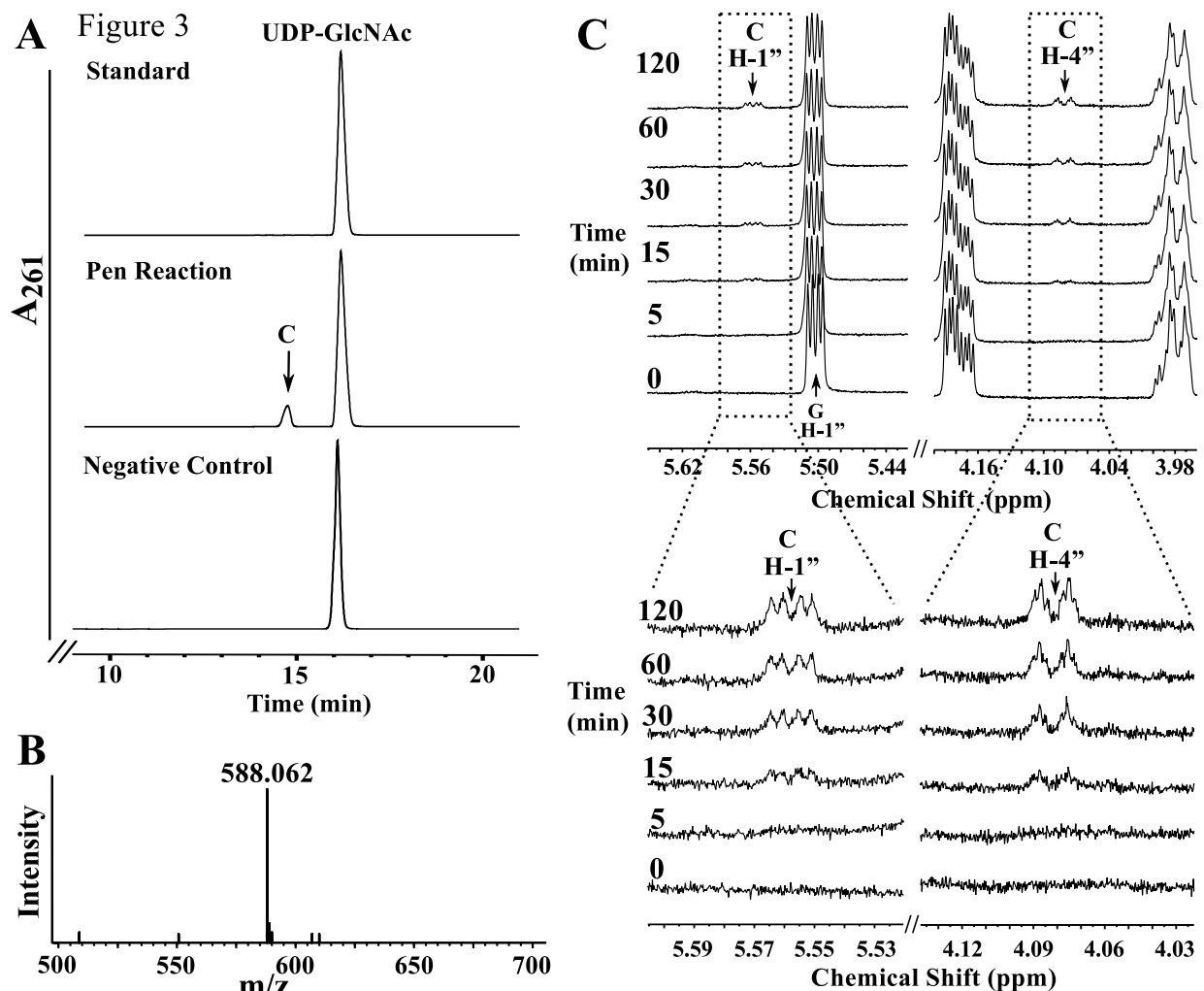


FIGURE 2.3. Analyses of UDP-6-deoxy-D-GlcNAc-5,6-ene, the enzymatic product of Pen by UV-HPLC, Mass spectrometry, and time-resolved 1H NMR. **Panel A.** Purified Pen was reacted with UDP-GlcNAc, (mid-panel) to yield a new UV-peak "C", marked by an arrow. The top UV-chromatogram trace in panel A corresponds to the standard of UDP-GlcNAc, and the lower is negative control showing incubation of irrelevant purified protein with UDP-GlcNAc. The peak C (eluting from HILIC-column at 14.6 min) was collected and characterized by NMR (see Fig. 2.5 and Table 2.2). **Panel B** shows that the enzymatic product C gives a $[M-H]^-$ ion at m/z 588.1 corresponding to UDP-6-deoxy-D-GlcNAc-5,6-ene. **Panel C.** Time-resolved 1H NMR showing Pen-enzymatic conversion of UDP-GlcNAc (G) to UDP-6-deoxy-D-GlcNAc-5,6-ene (C). The selected chemical shift regions of the proton NMR spectrum that are diagnostic for the $H-1''$ anomeric protons of enzymatic reactant and product is shown between 5.44 and 5.64 ppm, and the diagnostic $H-4''$ peak of Pen-product C is shown between 4.18 and 3.98 ppm. A 3 X zoom of the boxed regions is shown below in panel C.

2.4.4 NMR characterization of UDP-6-deoxy-D-GlcNAc-5,6-ene.

Since the structure of the product from the Pen reaction could not be determined from the MS or time-resolved NMR experiments, the peak that eluted from the HILIC column at 14.5 min

was collected and fully characterized by 1D and 2D-NMR spectroscopy. The one-dimensional proton NMR spectrum (Fig. 2.4) indicated that the product was UDP-2-acetamido-6-deoxy- α -D-xylo-hexopyranose-5,6-ene (abbr. UDP-6-deoxy-D-GlcNAc-5,6-ene). Proton and carbon chemical shifts and proton coupling constants of the sugar moiety are listed in Table 2.2. The coupling constant between H1'' and H2'' was small (3.1 Hz), indicating an α -linkage to the phosphate of UDP. The coupling constant between H1'' and the phosphorous of the diphosphate was 7.6 Hz, which is consistent with a D sugar. The coupling constants between H2'' and H3'' (9.9 Hz), along with H3'' and H4'' (9.5 Hz) confirmed that the sugar had the *gluco* configuration. The H2'' chemical shift at 4.18 ppm and the methyl proton resonance of the N-acetyl group (2''Nac) at 1.96 ppm are consistent with an acetamido moiety at the C2''. Two very diagnostic signal peaks for the C6''-methylene protons of product UDP-6-deoxy-D-GlcNAc-5,6-ene, were identified at 4.87 and 4.89 ppm. These peaks were assigned to H6a'' and H6b'' of the 5,6-ene moiety of the product. A COSY experiment showed the correlation of protons that were two and three bonds apart (Fig. 2.5A). Surprisingly connectivity between H4'' and H6'' was also observed in this experiment, even though a correlation between protons that are four bonds apart is typically not observed (34). Proton chemical shifts were assigned and verified by a TOCSY experiment (Fig. 2.5B). Protonated carbon chemical shifts, as well as multiplicities, were established by an HSQC experiment (Fig. 2.5C). The signal of C6'' gave reverse phase from C1'', C2'', C3'', and C4''. This suggested that protons at the 6'' position are methylene protons (-CH₂-) while other protons in the sugar moiety are methine protons (-CH-). The chemical shift of non-protonated carbon C5'' was assigned by a HMBC experiment (Fig. 2.5D), and the connectivity of C5'' to H1'', H4'' and H6'' were observed in the spectrum. These results, when taken together,

provide evidence that Pen is a UDP-GlcNAc 4-oxidase/5,6-dehydratase/4-reductase that converts UDP-GlcNAc to UDP-6-deoxy-D-GlcNAc-5,6-ene.

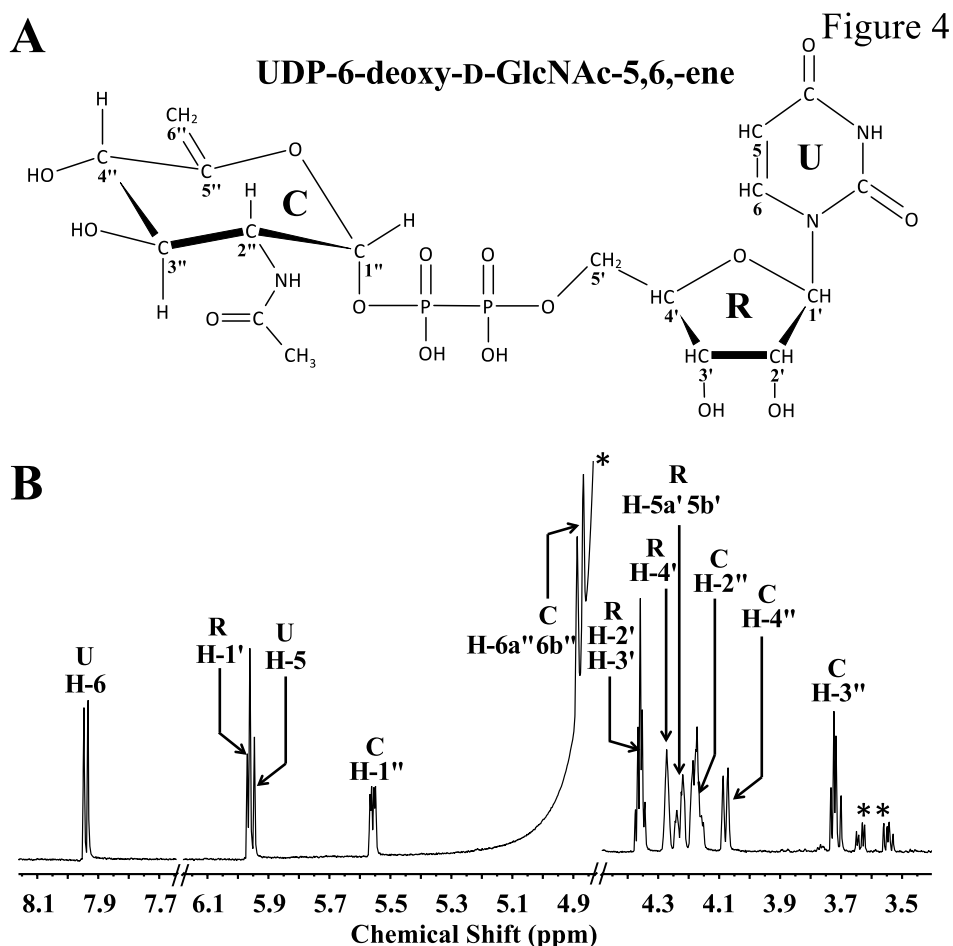


FIGURE 2.4. One-dimensional proton NMR spectrum of Pen-enzymatic product: UDP-6-deoxy-D-GlcNAc-5,6-ene. This provides evidence that the dehydratase is a UDP-GlcNAc 4-oxidase/5,6-dehydratase/4-reductase. Panel **A**. Illustration of UDP-6-deoxy-D-GlcNAc-5,6-ene. The protons (H) and carbons (C) are labeled with a number where the sugar pyranose ring is labeled as “C”; ribose is labeled as “R”; and uracil is labeled as “U”. Panel **B**. ^1H NMR spectrum of UDP-6-deoxy-D-GlcNAc-5,6-ene (selective signals are labeled). Spectrum was cropped to fit the window. Contaminant signals are labeled for H_2O signal (*) and glycerol (**).

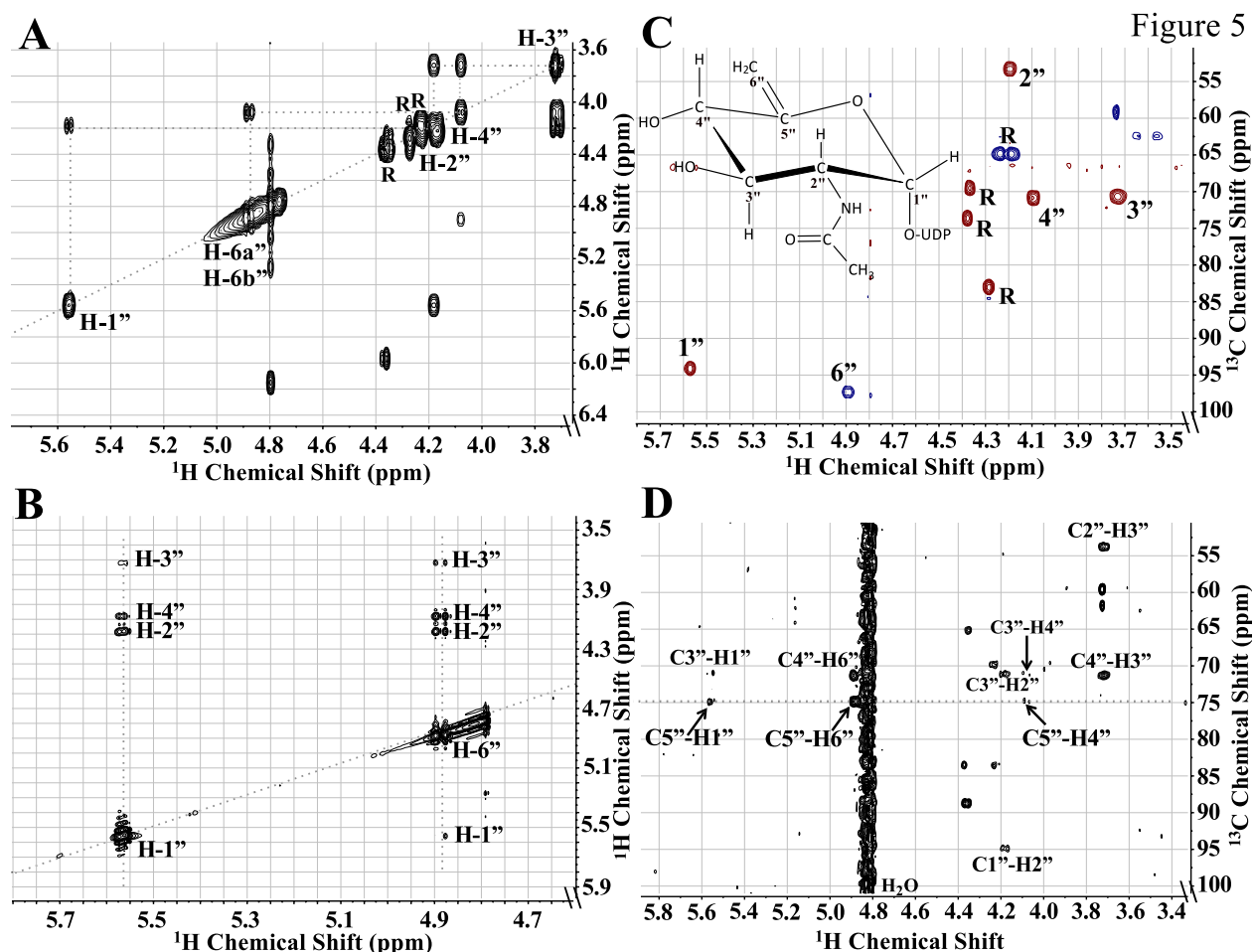


FIGURE 2.5. Two-dimensional NMR characterization of UDP-6-deoxy-D-GlcNAc-5,6-ene derived by enzymatic reaction of Pen with UDP-GlcNAc. **Panel A:** COSY experiment showing the connectivity between protons in the sugar ring moiety that are two to three bonds apart. Specifically note the connection between protons H4'' and H6''. **Panel B:** TOCSY experiment showing the connectivity of H1'' and H6'' to all the other protons in the sugar ring. Note the connection of H6'' to all the other 4 protons: H1'', H2'', H3'' and H4''. **Panel C:** HSQC experiment showing the ^{13}C and ^1H chemical shift of all protonated carbons in the sugar moiety. Note the phase of each CH group indicating 6'' belongs to CH_2 group while others belong to CH group. **Panel D:** HMBC experiment showing the connectivity of carbon to protons that are two to three bonds away, specifically the connection of C5'' to H1'', H4'' and H6''. Note the signal of C5'' was aliased so the actual chemical shift is 154.93 ppm.

2.4.5 Pal, encodes UDP-6-deoxy-D-GlcNAc-5,6-ene 4-oxidase/5,6-reductase/5-epimerase, that converts UDP-6-deoxy-D-GlcNAc-5,6-ene to UDP-4-keto-6-deoxy-L-AltNAc.

We have provided evidence that Pen is a UDP-GlcNAc 4-oxidase/5,6-dehydratase/4-reductase. We now describe the function of Pal, the second enzyme encoded by the *pse* operon of *B. thuringiensis*. The UDP-6-deoxy-D-GlcNAc-5,6-ene formed by recombinant Pen was purified

by HPLC and used as a substrate for recombinant Pal. Two products (labeled K and H in Fig. 2.6A) eluted from HILIC column as a broad UV₂₆₁ peak between 15-17 min and a retention time of ~16.5 min. A negative enzymatic reaction containing unrelated His₆-tagged purified enzyme yielded no product. For initial characterization of the products (K, H), the broad peak was collected and analyzed by ESI-MS and MS/MS. Two ions at m/z 588.01 and 606.03 (Fig. 2.6B) were detected. Interestingly, the new UDP-sugar product (K) has a mass identical to UDP-6-deoxy-D-GlcNAc-5,6-ene. Their different elution times (16.5 versus 15.5 min, respectively,) suggest that the product and substrate are chemically different. MS/MS analysis of each parent ion gave two ion fragments having m/z 323 and 403 values that are consistent with UMP and UDP, respectively. This indicated that the products were UDP-sugars.

To gain further insight into Pal-enzyme activity we monitored the reaction by time-resolved ¹H-NMR (Fig. 2.6C). Two products with a quartet of signals at 5.74 ppm and 5.555 ppm in the anomeric region of the NMR spectrum appeared over the reaction time. The anomeric signal at 5.555 ppm (see 60 min time point) was clear but overlapped at early reaction time points (5-30 min) with the anomeric signal of the substrate, UDP-6-deoxy-D-GlcNAc-5,6-ene (labeled C, H-1''), at 5.56 ppm. The H-4'' proton signal of the substrate (labeled C, H-4'') decreased over time while the amounts of product depicted by other proton signals (labeled H and K) increased. Two signals appeared at the 6-deoxy region around 1.55 ppm and 1.37 ppm. These signals correspond to C6-methyl protons (H-6'') of product K [UDP-2-acetamido-2,6-dideoxy-β-L-arabino-hex-4-ulose (abbr. UDP-4-keto-6-deoxy-L-AltNAc)] and H (hydrated form of K), sugar moiety, respectively. Proton and carbon chemical shifts and proton coupling constants of H are listed in Table 2.2.

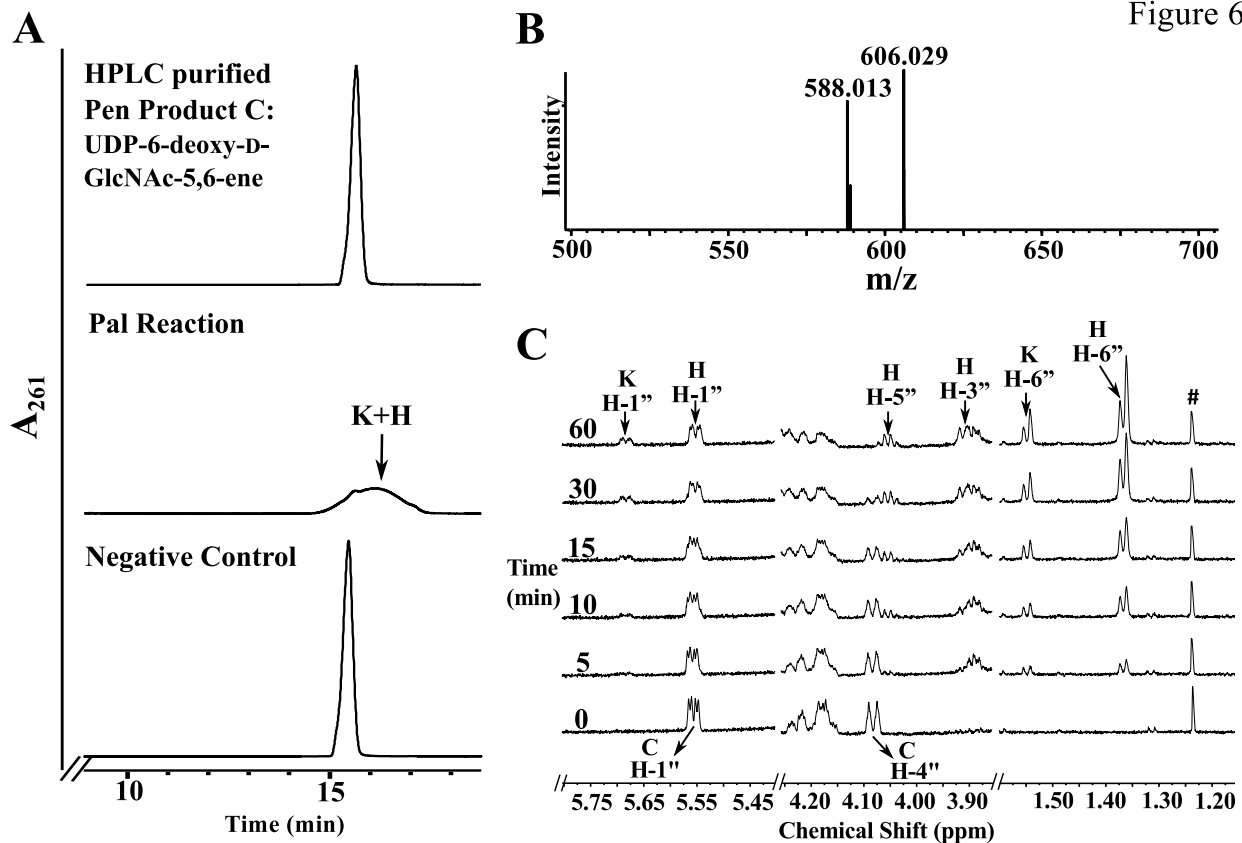


Figure 6

FIGURE 2.6. Analyses of UDP-4-keto-6-deoxy-L-AltNAc, the enzymatic product of Pal by UV-HPLC, *mass spectrometry*, and *time-resolved* ^1H NMR. Panel **A**. Purified Pal was reacted with UDP-6-deoxy-D-GlcNAc-5,6-ene (“C”, the purified enzymatic product of Pen) to yield at least two new UV-broad HPLC peaks labeled “K”, “H” and marked by arrows. The top HPLC-trace in Panel **A** corresponds to purified product “C” standard, and the bottom trace is the negative control: an assay of purified irrelevant protein with “C”. The K and H peaks (eluted from HILIC column at 16 min) were collected and characterized by NMR (see Table 2.2). Panel **B**. MS analysis in negative mode gives two $[\text{M}-\text{H}]^-$ ions at m/z 588.1 and 606.0 corresponding to two forms of UDP-4-keto-6-deoxy-L-AltNAc: 4-keto and its hydrated forms, respectively. Panel **C**. Time-resolved ^1H NMR showing Pal-enzymatic conversion of UDP-6-deoxy-D-GlcNAc-5,6-ene (C) to UDP-4-keto-6-deoxy-L-AltNAc. As the reaction proceeds, two molecular species are produced: a 4-keto form and a 4-keto-hydrated form of UDP-4-keto-6-deoxy-L-AltNAc. The selected chemical shift region of the proton NMR spectrum that corresponds to the H-1'' anomeric proton of enzymatic reactant and product is shown between 5.80 and 5.40 ppm. Note the H-4'' peak of the substrate (“C”) shown at 4.09 ppm is converted to new product K and H (the 4-keto and its hydrated form). The far right Panel (between 1.60 and 1.30 ppm) shows the methyl groups (H-6'') belonging to the hydrated (H) and the keto (K) forms of the product UDP-4-keto-6-deoxy-L-AltNAc.

2.4.6 Pen and Pal together produce UDP-4-keto-6-deoxy-D-GlcNAc and UDP-4-keto-6-deoxy-L-AltNAc.

Another interesting aspect of these two *Bacillus* enzymes was the formation of an additional product when the enzymes were incubated together rather than in a sequential manner

(see illustration in Fig. 2.8). When Pen and Pal were combined and incubated with UDP-GlcNAc an additional product was formed besides K and H. We used ^1H NMR spectroscopy to monitor the reaction of recombinant enzymes with the substrate, UDP-GlcNAc, over 8 h to determine the nature of this new product (Fig. 2.7A, labeled F). Two signals at 5.74 ppm and 5.555 ppm corresponding to the anomeric protons of K and H appeared within 1 h and started to disappear after 3 h. Similarly, two signals at 1.55 ppm and 1.37 ppm corresponding to the 6-deoxy protons of K and H also disappeared after 3 h. A new signal at 5.47 ppm developed gradually over the 8 h period as did a signal at 1.24 ppm. We have assigned these signals to the anomeric proton and H6'' of the sugar moiety of UDP-2-acetamido-2,6-dideoxy- α -D-xylo-hex-4-ulose (abbr. UDP-4-keto-6-deoxy-D-GlcNAc, F), which is the 5-epimer of UDP-4-keto-6-deoxy-L-AltNAc. UDP-4-keto-6-deoxy-D-GlcNAc was formed only when both Pen and Pal were added to the reaction; it was not detected when the two enzymes were added sequentially (Fig. 2.7B). To verify that UDP-4-keto-6-deoxy-D-GlcNAc was produced enzymatically, we reacted UDP-6-deoxy-D-GlcNAc-5,6-ene with Pal for 8 h. No UDP-4-keto-6-deoxy-D-GlcNAc was formed. Thus, UDP-4-keto-6-deoxy-D-GlcNAc is likely formed by an enzymatic reaction. The chemical shift (Table 2.2) suggested that product F is the 4-keto hydrated form of UDP-4-keto-6-deoxy-D-GlcNAc. A biosynthetic pathway for the Pen and Pal reactions is proposed (Fig. 2.8).

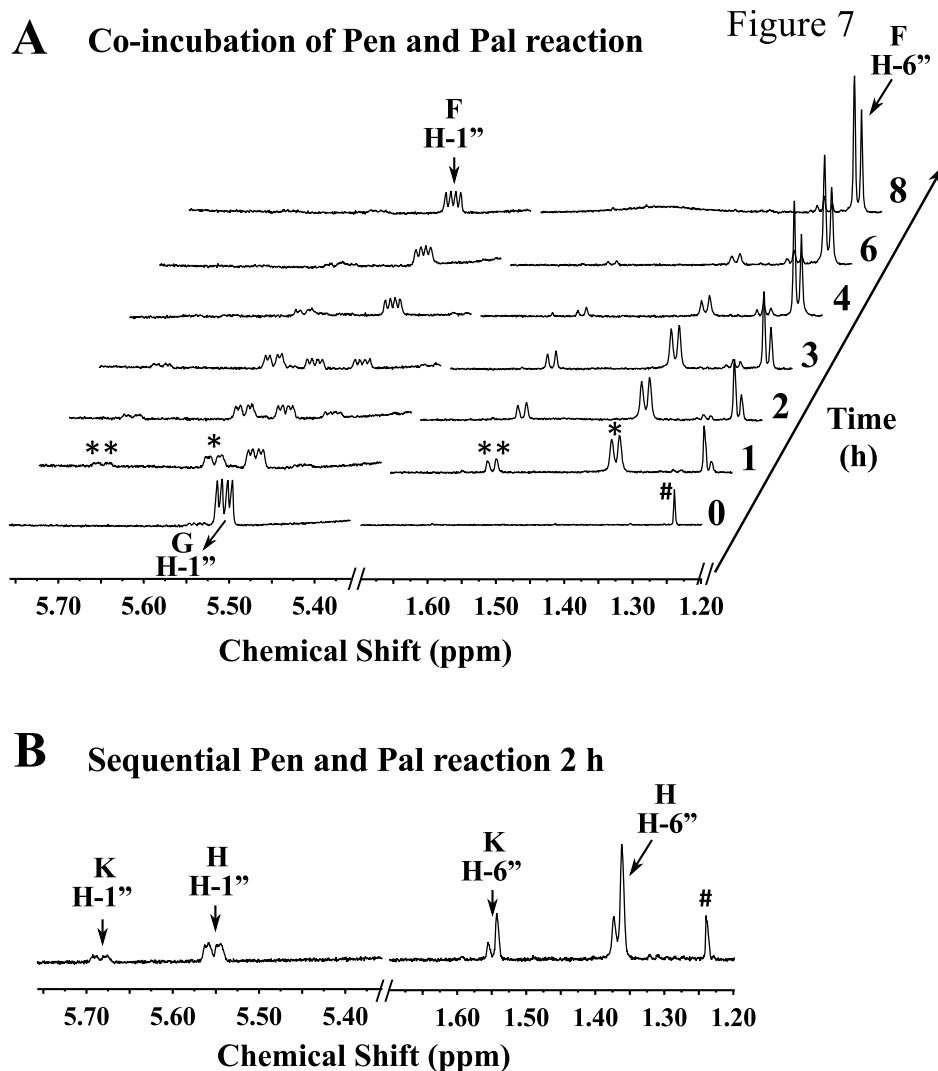


FIGURE 2.7. Co-incubation of Pen and Pal with the UDP-GlcNAc (G) yields four molecular species as determined by time-resolved ^1H NMR. Panel **A**. When co-incubation reaction was carried out up to 8h, in addition to the keto and hydrated forms of UDP-4-keto-6-deoxy-L-AltNAc, another predominant hydrated form of UDP-4-keto-6-deoxy-D-GlcNAc was observed. The selected chemical shift region for the anomeric protons of reactant and product is shown between 5.75 and 5.35 ppm, and the diagnostic H-6'' peak of the dual enzyme final products "F" is shown between 1.70 and 1.20 ppm. The peak annotated * belongs to a signal of Pal's hydrated product ("H") as an intermediate; whereas the peaks annotated ** belongs to signal peaks of Pal's keto product ("K"). Panel **B**. 2 h reaction of the single enzyme Pal incubated with UDP-6-deoxy-D-GlcNAc-5,6-ene, showing only UDP-4-keto-6-deoxy-L-AltNAc was produced but no UDP-4-keto-6-deoxy-D-GlcNAc.

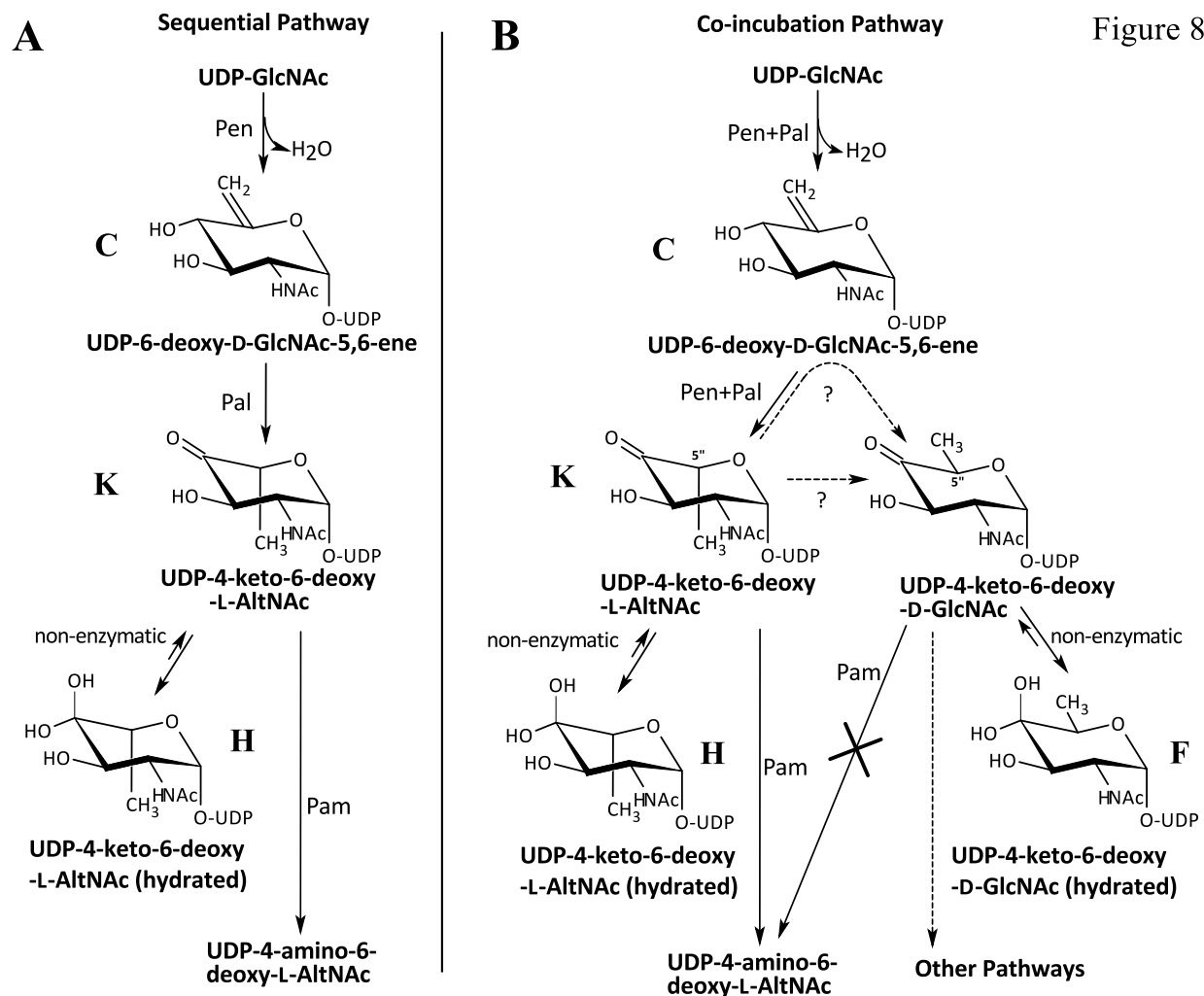


FIGURE 2.8. The two proposed biosynthesis pathways involving Pen and Pal. In a sequential reaction pathway, Pen was incubated with UDP-GlcNAc and Pal was then added. In the co-incubation pathway, both enzymes Pen and Pal were incubated together with UDP-GlcNAc. In the sequential reaction, UDP-GlcNAc (labeled as “G”) was converted to UDP-6-deoxy-D-GlcNAc-5,6-ene (labeled as “C”) by Pen, followed by Pal converting C to keto and hydrated forms of L-configured UDP-4-keto-6-deoxy-L-AltNAc (labeled as “K” and “H”). Note that products K and H were able to undergo C4” transamination reaction by Pam leading to product A: UDP-4-amino-6-deoxy-L-AltNAc. In the co-incubation pathway, UDP-GlcNAc was converted to a mixture of four products, including C, K, H and a novel product F: a D-configured UDP-4-keto-6-deoxy-D-GlcNAc. After a long period of incubation the product would be predominantly F, as the other products including C, K and H would be consumed. Note that product F, UDP-4-keto-6-deoxy-D-GlcNAc, was not a substrate for Pam, hence Pam was unable to C4” transaminate this D-configured UDP-sugar.

2.4.7 UDP-4-keto-6-deoxy-L-AltNAc is a substrate for the Pse metabolic enzyme. UDP-4-keto-6-deoxy-D-GlcNAc is not a substrate.

The sequential activity of Pen and Pal results in the formation of UDP-4-keto-6-deoxy-L-AltNAc, which is a substrate for next enzyme Pam in the Glu:C4”-aminotransferase reaction.

However, the UDP-4-keto-6-deoxy-D-GlcNAc formed when the two enzymes are combined is not a substrate for the transaminase. Thus, the 4-aminotransferase (Pam) is specific for the 5-epimer-4-keto-sugar, UDP-4-keto-6-deoxy-L-AltNAc.

2.4.8 Enzymatic properties of Pen and Pal.

The pH and temperature dependent activity of Pen and Pal and kinetic data for these enzymes is summarized in Table 2.3. Pen has its highest activity in phosphate buffer at pH 7.64; however high activity was also observed in MOPS-NaOH pH 7.6, Tris-HCl pH 7.4, and with the slight acidic condition in MES buffer at pH 6.85. The optimal temperature for Pen activity was at 37 °C; however the activities of Pen ranging from 30 °C to 50 °C were similar and varied by no more than 10%. As for Pal, we were unable to separate the products from the substrates through HILIC column. Consequently, we only measured the optimal pH and temperature for Pal by comparing the peak reduction of the substrates.

Comparative analyses of the kinetics between PseB from *C. jejuni* and *H. pylori* (35,36) with Pen from *B. thuringiensis* showed that Pen K_m for UDP-GlcNAc has a similar affinity for the substrate (Table 2.3). However, Pen turnover (k_{cat}/K_m) was 209, which is much higher than PseB, suggesting that the *B. thuringiensis* Pen converts UDP-GlcNAc into a product more efficiently.

To determine if the 4-oxidase activities of Pen and Pal require NAD^+ or $NADP^+$, and to address if Pen enzyme functions as a dimer or forms a functional complex with Pal, we performed a series of experiments (Fig. 2.9) using size-exclusion chromatography on Superdex 75. The proteins were pre-incubated with and without substrate (UDP-GlcNAc). The proteins were either loaded on the column individually or were loaded in a mix together. Protein peaks were then collected and tested for activity and analyzed by SDS-PAGE. Pen eluted from the

column in two peaks (peak a, b, Fig. 2.9A). One peak migrated as a protein of 31.9 kDa (presumably a monomer) while the second peak eluted in the region for a protein with a mass of about 192.8 kDa. The earlier peak (presumably a pentamer, a tetramer or a dimer of dimers) but not the later peak had enzymatic activity when supplied with UDP-GlcNAc (Fig. 2.9B). Conversely, Pal eluted from the column as a single peak (peak d, Fig. 2.9A), migrating 87.6 kDa presumably as a dimer. This peak was only enzymatically active when supplemented with NAD⁺ (Fig. 2.9B). No Pal activity was obtained by adding NADP⁺. Taken together we propose that Pal activity involves the co-factor NAD⁺ for oxidation/reduction cycle of the enzyme product. The nature of the oligomeric state of Pen and Pal however, is speculative and the chromatographic mobility could be influenced by the shape of the protein.

TABLE 2.3. Pen and Pal enzyme properties

	Pen	Pal	PseB (C.j.) ¹	PseB (H.p.) ²
Optimal pH ³	7.6	7.6	7.0	7.0
Optimal temperature (°C) ⁴	37	30	42	37
K_m (mM) ⁵	0.143±0.006	ND	0.050	0.159
V_{max} (nMs ⁻¹)	15.01±0.29	ND	--	--
k_{cat} (min ⁻¹)	30.02±0.58	ND	1.51	5.7
k_{cat}/K_m (mM ⁻¹ min ⁻¹)	209.9±12.9	ND	30.6	35.9
protein monomer (kDa) ⁶	40	37.1	37.4	37.4

¹Kinetic data of PseB in *Campylobacter. jejuni* from (35)

² Kinetic data of PseB in *Helicobacter. pylori* from (36)

³ For Pen, optimal pH assays were determined using phosphate buffer in which Pen yielded highest activity compared to MOPS-NaOH, Tris-HCl, and MES buffer. For Pal, optimal pH assays were determined using Tris-HCl buffer.

⁴ Optimal temperature assays were determined using Tris-HCl for both Pen and Pal.

⁵ The reaction was determined by HPLC-UV after 5 min 30 °C incubation for Pen.

⁶ The active forms of proteins eluted from size exclusion column were for Pen, 192.8 kDa and for Pal, 87.6 kDa, presumably suspected to be a tetramer and a dimer, respectively.

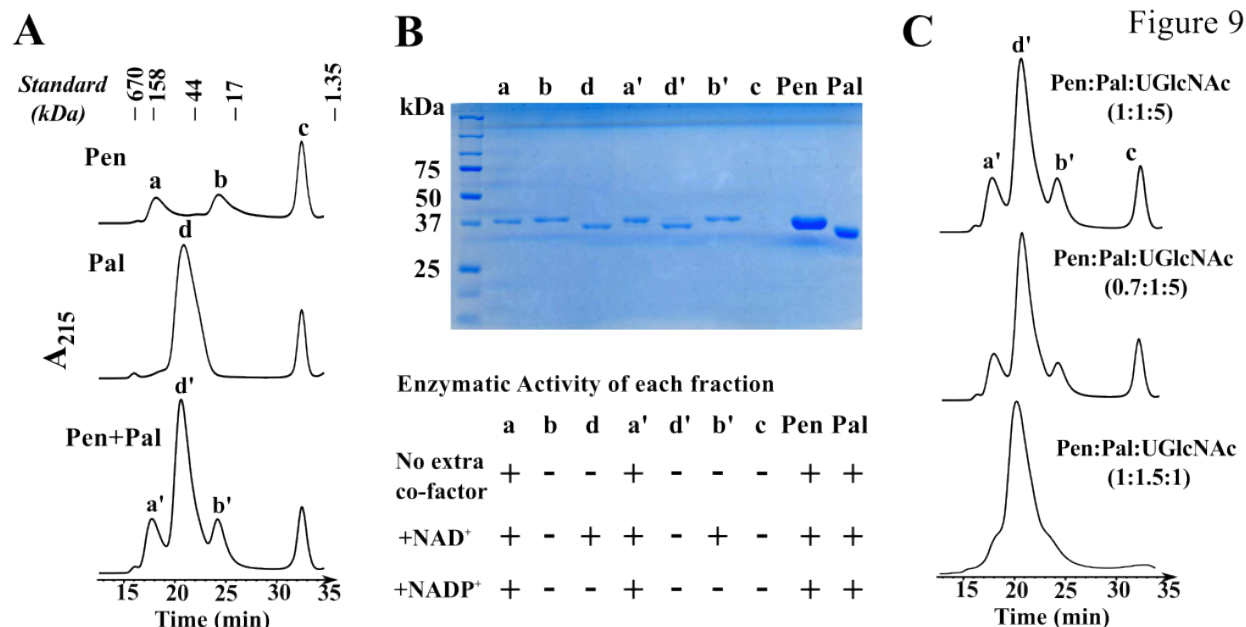


FIGURE 2.9. The nature of the oligomeric state of Pen and Pal as estimated by size exclusion chromatography. Panel A. Individual proteins (Pen or Pal) or combined proteins (Pen plus Pal), were incubated with UDP-GlcNAc on ice prior to chromatography on Superdex-75 column. The UV₂₁₅ chromatogram trace in the top panel is showing that Pen migrates on Superdex75 in two molecular species: peak “a” and peak “b”. Only the peak “a” is enzymatically active. Peak “c” is UDP-GlcNAc. The Mid-panel is showing that Pal migrated as single peak “d”. The bottom-panel shows no apparent interaction between Pen and Pal during co-incubation with UDP-GlcNAc. The migration of standard molecular weight proteins on the Superdex-75 column is indicated on top. Panel B. Each eluted protein fraction (a, b, d, a', b', d') and peak c, was visualized by SDS-PAGE. The last two lanes labeled as “Pen” and “Pal” are controls: purified proteins before size exclusion chromatography. The detailed enzymatic activity of each eluted protein shown below the SDS-PAGE gel was tested in the presence of NAD⁺ and/or NADP⁺. The “+” sign indicates activity. It is not suggesting enhanced activity. Panel C. Different ratios of Pen, Pal enzymes and substrate UDP-GlcNAc were mixed prior to chromatography on Superdex-75 column. No obvious Pen-Pal complex was formed.

We tested if Pen and Pal interacted with each other; however, gel filtration assays showed when both enzymes were mixed together and chromatographed (Fig. 2.9A bottom panel) each enzyme had the same elution pattern as the individual protein. Each peak was collected from the column and resolved on SDS-PAGE (Fig. 2.9B). A single peptide band was observed, (compare lanes a, b, d with a', b', d'). No difference in elution profiles were obtained by adding UDP-GlcNAc or with UDP-GlcNAc + NAD⁺ to the combined enzymes. In addition, when different ratios of Pen, Pal and UDP-GlcNAc were chromatographed (Fig. 2.9C) the peak signal increase

proportionally. Taking together, the results suggest that Pen and Pal do not form a complex under this experimental setup.

No data indicating that Pen requires a cofactor for activity was obtained by size-exclusion chromatography (Fig. 2.9B). Nevertheless, we suspected that this protein binds NAD^+ or NADP^+ so tightly that it would only be released when denatured. Therefore, we heat-inactivated purified Pen as well as Pal and examined the products released by LC-ESI-MS/MS. NADP^+ with $[\text{M}+\text{H}]^+$ ion m/z at 743.9 and diagnostic MS/MS ion fragments (603.9, 621.9, 489.9, 409.9) were detected when Pen was denatured (Fig. 2.10A), whereas NAD^+ with m/z at 664.0 (MS/MS 523.9, 541.9, 232.0, 427.9) was detected when Pal was denatured (Fig. 2.10B). The inability of the size-exclusion column to separate bound- NADP^+ from Pen suggests a strong interaction between the protein and its co-factor. By comparison to Pal, it is likely that within Pen-protein more amino-acid residues are involved in hydrogen-bond connection between the co-factor and the enzyme. Together these experiments provide evidence that NAD^+ is a co-factor required for the oxidase/reductase activity of Pal-enzyme.

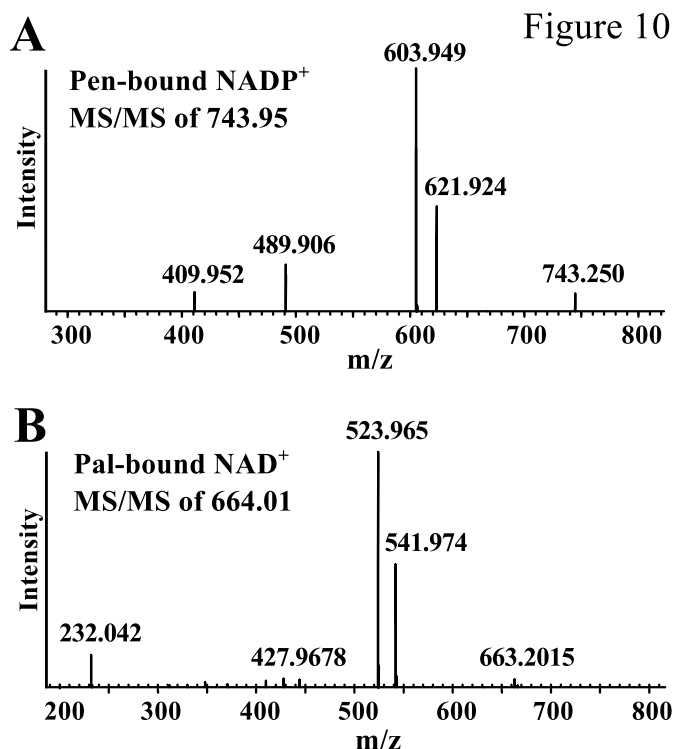


FIGURE 2.10. Pen has an NAD⁺ bound co-factor, and Pal is bound to NADP⁺ as determined after heat inactivation and mass spectrometry analyses. An aliquot of purified Pen or purified Pal was heat-treated, centrifuged, concentrated, and a portion was chromatographed through HILIC column using LC-ESI-MS/MS. A major positive ion at m/z 743.9 was eluted from HILIC column at 10.4 min for the denatured Pen sample. This m/z value corresponding to parent ion NADP⁺ and the diagnostic MS/MS ion fragment's of NADP⁺ is shown in Panel A. A major positive ion at m/z 664.01 was eluted from HILIC column at 8.4 min for the denatured Pal sample. The latter m/z values corresponding to NADP⁺ and the MS/MS fragmentation is shown in Panel B. Std NAD⁺ and NADP⁺ were separately chromatographed and analyzed by LC-ESI-MS/MS.

2.5 Discussion

We have identified an operon in *B. thuringiensis israelensis* ATCC 35646 (Fig. 2.1A, B) which contains seven genes that encode enzymes involved in the conversion of UDP-GlcNAc to CMP-Pse. The product of each enzymatic reaction, UDP-6-deoxy-D-GlcNAc-5,6-ene, UDP-4-keto-6-deoxy-L-AltNAc, UDP-4-amino-4,6-dideoxy-N-acetyl- β -L-altrosamine (UDP-4-amino-6-deoxy-L-AltNAc), UDP-2,4-diacetamido-2,4,6-trideoxy- β -L-altropyranose (UDP-2,4,6-trideoxy-2,4-diNAc-L-altrose), 2,4-diacetamido-2,4,6-trideoxy- β -L-altropyranose (2,4,6-trideoxy-2,4-diNAc-L-altrose), Pse, and CMP-Pse were characterized by LC-MS/MS and NMR spectroscopy. *B. thuringiensis* is unusual as it uses two dehydratase-like enzymes to initiate the conversion of

UDP-GlcNAc to UDP-4-keto-6-deoxy-L-AltNAc. The known functional PseB from the Gram-negative bacteria uses only one enzyme to catalyze this reaction (27,37).

H. pylori and *C. jejuni* both have a single bi-functional 5-epimerase/4,6-dehydratase (PseB) that converts UDP-GlcNAc to UDP-4-keto-6-deoxy-L-AltNAc. *C. jejuni* PseB complexed with NADP⁺ has been proposed to mediate the C4''-oxidation of UDP-GlcNAc, which is followed by series of enzyme-bound intermediates that include 4-keto, dehydration, NADPH-mediated reduction of the ene-group, and a C5''-epimerization to yield UDP-4-keto-6-deoxy-L-AltNAc (37). The two separate *B. thuringiensis* dehydratase-like activities produce the same end product as PseB (Fig. 2.1A), but carry out this reaction differently. We propose that Pen oxidizes C4'' of UDP-GlcNAc via E:NADP⁺ to form the C4''-keto intermediate (Fig. 2.11A). The same enzyme then carries out a 5,6-dehydratase reaction to form the 5,6-ene moiety. E:NADPH then performs a stereospecific C4''-keto reduction to give UDP-6-deoxy-D-GlcNAc-5,6-ene. What elicits the Pen E:NADPH-mediated C4''-reduction rather than the C5''-C6''-reduction as is the case in the 'classical 4,6-dehydratase activity' is not known. Amino-acid sequence comparison shows that, interestingly, Pen has "10 extra" amino acids [aa 242-252] when compared with PseB. We postulate that the extra sequence forms secondary structure that exists to either push away the 5,6-ene moiety or transiently bind the 5,6-ene moiety in order to protect it from the NADPH-reduction step. On the other hand, the extra 10 aa long peptide structure of Pen may exist to facilitate a C4'' specific reduction and maintain the same gluco-sugar ring configuration. Taken together, we propose that Pen is a UDP-GlcNAc 4-oxidase/5,6-dehydratase/4-reductase.

We postulate that Pal catalyzes a NAD⁺-mediated C4''-oxidation to generate a UDP-4-keto-GlcNAc-5,6-ene intermediate (Fig. 2.11B). Following this reaction, the hydride from E:NADH is transferred to the C5''-C6''-ene moiety via the opposite face of the double bond

leading to C5'' reduction and C5'' epimerization to give UDP-4-keto-6-deoxy-L-AltNAc. Hence Pal is a UDP-6-deoxy-D-GlcNAc-5,6-ene 4-oxidase/5,6-reductase/5-epimerase. Interestingly, a long co-incubation of Pen and Pal leads to the production of UDP-4-keto-6-deoxy-D-GlcNAc (Fig. 2.7, 2.8). Thus, it is possible that Pal could mediate 5''-epimerization between UDP-4-keto-6-deoxy-L-AltNAc and UDP-4-keto-6-deoxy-D-GlcNAc.

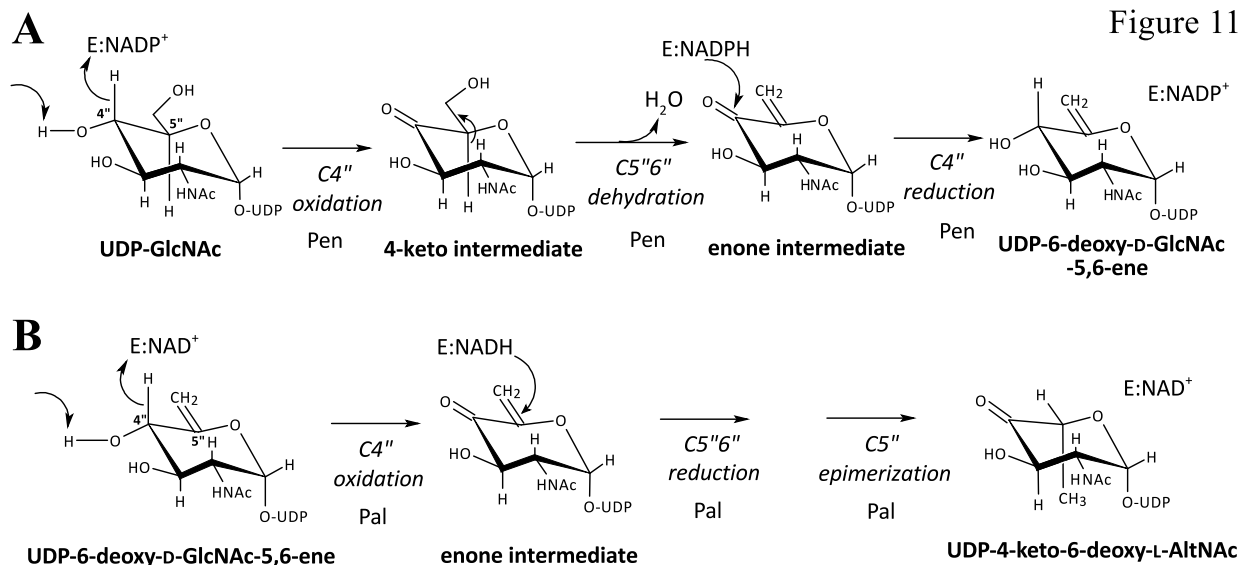


FIGURE 2.11. The proposed enzyme reaction intermediates involved in the conversion of UDP-GlcNAc to UDP-6-deoxy-D-GlcNAc-5,6-ene; and the proposed intermediates in the conversion to UDP-4-keto-6-deoxy-L-AltNAc. Panel **A**. Pen is predicted to first carry out a NADP⁺-dependent C4''-oxidation to form a 4-keto-intermediate. Subsequently, dehydration at C5''-C6'' forms an enone intermediate (5,6-ene moiety). This is followed by NADPH-mediated specific reduction at C4'' to form a product of the same gluco-configuration and the release of a product with a double-bond along C5''=C6'', UDP-6-deoxy-D-GlcNAc-5,6-ene. Hence, we abbreviated the Pen-activities as UDP-GlcNAc 4-oxidase/5,6-dehydratase/4-reductase. Panel **B**. Pal is predicted to first carry out a NAD⁺-dependent C4''-oxidation forming an UDP-4-keto-intermediate. Subsequently, we assume that C5''=C6'' undergoes reduction followed by C5'' epimerization leading to the production of UDP-4-keto-6-deoxy-L-AltNAc. In solution, the product UDP-4-keto-6-deoxy-L-AltNAc is found predominantly in its hydrated form. We abbreviated the Pal-activities as UDP-GlcNAc 4-oxidase/5,6-reductase/5-epimerase.

Pal shares low amino-acid sequence identity to the dTDP-Glc 4,6-dehydratase from *Bacillus Anthracis* Str. Ames. However, Pal is not acting on various NDP-sugars that were tested including UDP-GlcNAc, UDP-GalNAc, and UDP-Glc. Pal shares even lower amino acid sequence identity to PseB and is unable to act on UDP-GlcNAc. Insight into the amino-acid

residues involved in the catalytic activities of Pen and Pal will require the use of site-directed mutagenesis in combination with a description of the enzymes' three-dimensional structures.

In *B. thuringiensis* two-enzymes are required to generate the UDP-4-keto-6-deoxy-L-AltNAc. This reaction also generates UDP-6-deoxy-D-GlcNAc-5,6-ene. It is not known if the UDP-GlcNAc derivative participates in other metabolic pathways. Interestingly, UDP-6-deoxy-D-GlcNAc-5,6-ene is an intermediate formed by TunA during the synthesis of tunicamycin by *Streptomyces* species (38). Whether or not an analogous pathway exists in *B. thuringiensis* remains unknown. Although the TunA enzyme structure was solved (38), very low amino-sequence similarity (24% and e-value 0.004) exists between the Pen and the TunA protein, suggesting these genes are evolutionary not related and perhaps arose from different ancestral genes. With regards to Pal, we are currently unaware of another single activity identical to this enzyme.

Our kinetics data (k_{cat}/K_m) for Pen suggested that it converts UDP-GlcNAc into product more efficiently (Table 2.3) compared with TunA and PseB (27,38). However, the product itself inhibited the Pen forward reaction. This resulted in a maximum ~ 10% conversion of substrate to product. Pal, however, rapidly converts the 5,6-ene to the 4-keto derivative and thereby facilitates almost full conversion of UDP-GlcNAc to UDP-4-keto-6-deoxy-L-AltNAc. The formation of UDP-4-keto-6-deoxy-D-GlcNAc when Pen and Pal were combined together and reacted with UDP-GlcNAc suggests that Pal can be a stand-alone C5''-epimerase, albeit of low efficiency. Incubating *C. jejuni* PseB with UDP-GlcNAc for periods up to 15 h also resulted in the C5'' epimerization UDP-4-keto-6-deoxy-L-AltNAc to UDP-4-keto-6-deoxy-D-GlcNAc (37); the exact mechanism to explain this remains unknown.

Our characterization of the first two enzymes (Pen and Pal) in the Bti-*pse* operon provides insight into the metabolic pathway leading to the formation of Pse in *B. thuringiensis*. However, we cannot exclude the possibility that the products formed by these enzymes also participate in other metabolic pathways. Additional studies are now required to determine if other members of the Bacillaceae also have these enzymes. Understanding the molecular factors that control the flux of UDP-GlcNAc to different metabolic pathways in bacteria is important, as this nucleotide sugar is a precursor used for the formation of wall polysaccharides and peptidoglycan as well as other glycans.

2.6 References

1. Assaeedi, A. S. A., Osman, G. E. H., and Abulreesh, H. H. (2011) The occurrence and insecticidal activity of *Bacillus thuringiensis* in the arid environments. *Aust J Crop Sci* **5**, 1185-1190
2. Sanahuja, G., Banakar, R., Twyman, R. M., Capell, T., and Christou, P. (2011) *Bacillus thuringiensis*: a century of research, development and commercial applications. *Plant Biotechnol J* **9**, 283-300
3. Ichimatsu, T., Mizuki, E., Nishimura, K., Akao, T., Saitoh, H., Higuchi, K., and Ohba, M. (2000) Occurrence of *Bacillus thuringiensis* in fresh waters of Japan. *Curr Microbiol* **40**, 217-220
4. Aly, C., Mulla, M. S., and Federici, B. A. (1985) Sporulation and Toxin Production by *Bacillus-Thuringiensis* Var *Israelensis* in Cadavers of Mosquito Larvae (Diptera, Culicidae). *Journal of invertebrate pathology* **46**, 251-258
5. Mizuki, E., Ichimatsu, T., Hwang, S. H., Park, Y. S., Saitoh, H., Higuchi, K., and Ohba, M. (1999) Ubiquity of *Bacillus thuringiensis* on phylloplanes of arboreous and herbaceous plants in Japan. *Journal of applied microbiology* **86**, 979-984
6. Jensen, G. B., Hansen, B. M., Eilenberg, J., and Mahillon, J. (2003) The hidden lifestyles of *Bacillus cereus* and relatives. *Environmental microbiology* **5**, 631-640
7. Gillis, A., Dupres, V., Delestrait, G., Mahillon, J., and Dufrene, Y. F. (2012) Nanoscale imaging of *Bacillus thuringiensis* flagella using atomic force microscopy. *Nanoscale* **4**, 1585-1591
8. Schnepf, E., Crickmore, N., Van Rie, J., Lereclus, D., Baum, J., Feitelson, J., Zeigler, D. R., and Dean, D. H. (1998) *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiol Mol Biol R* **62**, 775-+
9. Schunemann, R., Knaak, N., and Fiuza, L. M. (2014) Mode of Action and Specificity of *Bacillus thuringiensis* Toxins in the Control of Caterpillars and Stink Bugs in Soybean Culture. *ISRN microbiology* **2014**, 135675
10. Ibrahim, M. A., Griko, N. B., and Bulla, L. A., Jr. (2013) The Cry4B toxin of *Bacillus thuringiensis* subsp. *israelensis* kills Permethrin-resistant *Anopheles gambiae*, the principal vector of malaria. *Experimental biology and medicine* **238**, 350-359
11. Broderick, N. A., Raffa, K. F., and Handelsman, J. (2006) Midgut bacteria required for *Bacillus thuringiensis* insecticidal activity. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 15196-15199

12. Swiecicka, I., and Mahillon, J. (2006) Diversity of commensal *Bacillus cereus* sensu lato isolated from the common sow bug (*Porcellio scaber*, Isopoda). *Fems Microbiol Ecol* **56**, 132-140
13. Margulis, L., Jorgensen, J. Z., Dolan, S., Kolchinsky, R., Rainey, F. A., and Lo, S. C. (1998) The Arthromitus stage of *Bacillus cereus*: intestinal symbionts of animals. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 1236-1241
14. Lovgren, A., Zhang, M. Y., Engstrom, A., and Landen, R. (1993) Identification of two expressed flagellin genes in the insect pathogen *Bacillus thuringiensis* subsp. *alesti*. *Journal of general microbiology* **139**, 21-30
15. Shen, A., and Higgins, D. E. (2006) The MogR transcriptional repressor regulates nonhierarchical expression of flagellar motility genes and virulence in *Listeria monocytogenes*. *PLoS pathogens* **2**, e30
16. Soufiane, B., Xu, D., and Cote, J. C. (2007) Flagellin (fliC) protein sequence diversity among *Bacillus thuringiensis* does not correlate with H serotype diversity. *Anton Leeuw Int J G* **92**, 449-461
17. Maki-Yonekura, S., Yonekura, K., and Namba, K. (2003) Domain movements of HAP2 in the cap-filament complex formation and growth process of the bacterial flagellum. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 15528-15533
18. Namba, K., Yamashita, I., and Vonderviszt, F. (1989) Structure of the Core and Central Channel of Bacterial Flagella. *Nature* **342**, 648-654
19. Kotiranta, A., Lounatmaa, K., and Haapasalo, M. (2000) Epidemiology and pathogenesis of *Bacillus cereus* infections. *Microbes and infection / Institut Pasteur* **2**, 189-198
20. Hayakawa, J., Kondoh, Y., and Ishizuka, M. (2009) Cloning and characterization of flagellin genes and identification of flagellin glycosylation from thermophilic *Bacillus* species. *Bioscience, biotechnology, and biochemistry* **73**, 1450-1452
21. Young, G. M., Schmiel, D. H., and Miller, V. L. (1999) A new pathway for the secretion of virulence factors by bacteria: The flagellar export apparatus functions as a protein-secretion system. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 6456-6461
22. Ottemann, K. M., and Miller, J. F. (1997) Roles for motility in bacterial-host interactions. *Molecular microbiology* **24**, 1109-1117
23. Schirm, M., Schoenhofen, I. C., Logan, S. M., Waldron, K. C., and Thibault, P. (2005) Identification of unusual bacterial glycosylation by tandem mass spectrometry analyses of intact proteins. *Anal Chem* **77**, 7774-7782
24. Thibault, P., Logan, S. M., Kelly, J. F., Brisson, J. R., Ewing, C. P., Trust, T. J., and Guerry, P. (2001) Identification of the carbohydrate moieties and glycosylation motifs in *Campylobacter jejuni* flagellin. *Journal of Biological Chemistry* **276**, 34862-34870
25. Schirm, M., Arora, S. K., Verma, A., Vinogradov, E., Thibault, P., Ramphal, R., and Logan, S. M. (2004) Structural and genetic characterization of glycosylation of type a flagellin in *Pseudomonas aeruginosa*. *Journal of bacteriology* **186**, 2523-2531
26. Parker, J. L., Day-Williams, M. J., Tomas, J. M., Stafford, G. P., and Shaw, J. G. (2012) Identification of a putative glycosyltransferase responsible for the transfer of pseudaminic acid onto the polar flagellin of *Aeromonas caviae* Sch3N. *MicrobiologyOpen* **1**, 149-160
27. Schoenhofen, I. C., McNally, D. J., Brisson, J. R., and Logan, S. M. (2006) Elucidation of the CMP-pseudaminic acid pathway in *Helicobacter pylori*: synthesis from UDP-N-acetylglucosamine by a single enzymatic reaction. *Glycobiology* **16**, 8C-14C
28. Guerry, P., Ewing, C. P., Schirm, M., Lorenzo, M., Kelly, J., Pattarini, D., Majam, G., Thibault, P., and Logan, S. (2006) Changes in flagellin glycosylation affect *Campylobacter* autoagglutination and virulence. *Molecular microbiology* **60**, 299-311
29. Schirm, M., Soo, E. C., Aubry, A. J., Austin, J., Thibault, P., and Logan, S. M. (2003) Structural, genetic and functional characterization of the flagellin glycosylation process in *Helicobacter pylori*. *Molecular microbiology* **48**, 1579-1592

30. Yang, T., Bar-Peled, L., Gebhart, L., Lee, S. G., and Bar-Peled, M. (2009) Identification of galacturonic acid-1-phosphate kinase, a new member of the GHMP kinase superfamily in plants, and comparison with galactose-1-phosphate kinase. *The Journal of biological chemistry* **284**, 21526-21535
31. Gu, X., Glushka, J., Yin, Y., Xu, Y., Denny, T., Smith, J., Jiang, Y., and Bar-Peled, M. (2010) Identification of a bifunctional UDP-4-keto-pentose/UDP-xylose synthase in the plant pathogenic bacterium *Ralstonia solanacearum* strain GMI1000, a distinct member of the 4,6-dehydratase and decarboxylase family. *The Journal of biological chemistry* **285**, 9030-9040
32. Gill, S. C., and Vonhippel, P. H. (1989) Calculation of Protein Extinction Coefficients from Amino-Acid Sequence Data. *Analytical biochemistry* **182**, 319-326
33. Gu, X., Glushka, J., Lee, S. G., and Bar-Peled, M. (2010) Biosynthesis of a new UDP-sugar, UDP-2-acetamido-2-deoxyxylose, in the human pathogen *Bacillus cereus* subspecies cytotoxic NVH 391-98. *The Journal of biological chemistry* **285**, 24825-24833
34. Gunther, H., and Jikeli, G. (1977) H-1 Nuclear Magnetic-Resonance Spectra of Cyclic Monoenes - Hydrocarbons, Ketones, Heterocycles, and Benzo Derivatives. *Chem Rev* **77**, 599-637
35. Creuzenet, C. (2004) Characterization of CJ1293, a new UDP-GlcNAc C6 dehydratase from *Campylobacter jejuni*. *FEBS letters* **559**, 136-140
36. Creuzenet, C., Schur, M. J., Li, J., Wakarchuk, W. W., and Lam, J. S. (2000) FlaA1, a new bifunctional UDP-GlcNAc C6 Dehydratase/ C4 reductase from *Helicobacter pylori*. *The Journal of biological chemistry* **275**, 34873-34880
37. Morrison, J. P., Schoenhofen, I. C., and Tanner, M. E. (2008) Mechanistic studies on PseB of pseudaminic acid biosynthesis: a UDP-N-acetylglucosamine 5-inverting 4,6-dehydratase. *Bioorganic chemistry* **36**, 312-320
38. Wyszynski, F. J., Lee, S. S., Yabe, T., Wang, H., Gomez-Escribano, J. P., Bibb, M. J., Lee, S. J., Davies, G. J., and Davis, B. G. (2012) Biosynthesis of the tunicamycin antibiotics proceeds via unique exo-glycal intermediates. *Nature chemistry* **4**, 539-546

CHAPTER 3

DISCOVERY OF A UNIQUE EXTRACELLULAR POLYSACCHARIDE IN MEMBERS OF THE PATHOGENIC BACILLUS THAT CAN CO-FORM WITH SPORES²

² Li, Z., Hwang, S., and Bar-Peled, M. 2016. *The Journal of biological chemistry*. 291, 19051-19067. Reprinted here with permission of the publisher.

3.1 Abstract

An exopolysaccharide, produced during late stage of stationary growth phase, was discovered and purified from the culture medium of *Bacillus cerues*, *B. anthracis*, and *B. thuringiensis* when strains were grown in a defined nutrient medium that induces biofilm. Two-dimensional NMR structural characterization of the polysaccharide, named pzX, revealed that it is composed of an unusual three amino-sugar sequence repeat of $[-3)\text{XylNAc}4\text{OAc}(\alpha1-3)\text{GlcNAcA}4\text{OAc}(\alpha1-3)\text{XylNAc}(\alpha1-J)_n$. The sugar residue XylNAc had never been previously described in any glycan structure. The XNAC operon that contains the genes for the assembly of pzX is also unique and so far identified only in members of the *Bacillus cereus* sensu lato group. Microscopic and biochemical analyses indicate that pzX co-forms during sporulation, such that upon the release of the spore to the extracellular milieu it becomes surrounded by pzX. The relative amounts of pzX produced can be manipulated by specific nutrient in the medium, but rich medium appears to suppress pzX formation. pzX has unique characteristics: a surfactant property that lowers surface tension, a cell/spore antiaggregant, and an adherence property that increases spores binding to surfaces. pzX in *Bacillus* could represent a trait shared by many spore-producing microorganisms. It suggests pzX is an active player in spore physiology and may provide new insights to the successful survival of the *B. cereus* species in natural environments or in the hosts.

3.2 Introduction

In 1881 Louis Pasteur developed the first vaccine for anthrax, the devastating disease caused by *Bacillus anthracis*, a Gram-positive endospore forming bacterium. Since then, outbreaks of anthrax affecting humans and animals have dramatically decreased (1,2); although infrequent incidents have been reported (3). However, letters containing *B. anthracis* spores sent

by mail to US officials a week after the September 11, 2001 attack, led to public fear of infection and the use of its spores as a bioweapon agent (4). *B. anthracis*, although elicits different disease phenotypes, is closely related in terms of gene content and synteny (5) to other *Bacillus* species collectively named *Bacillus cereus sensu lato* (6) or *Bacillus cereus* group. These highly related *Bacilli* are able to colonize in diverse hosts including insects and mammals, and they are commonly found in soil, water, and depending on the specie in cadavers, vegetation and food. In addition to *B. anthracis*, this group includes *B. cereus*, recognized as a cause of food poisoning toxins, and *B. thuringiensis* that produces insecticidal proteins. In the past 15 years, several *B. cereus* strains were reported to cause severe anthrax-like disease in humans (7,8) and apes (9). Some of these virulent strains while retaining *B. cereus* diagnostic phenotypes harbor plasmids similar to the toxin and capsule virulence plasmids pXO1 and pXO2 present in *B. anthracis* (10-12). Hence, it is no wonder that the need to distinguish these *Bacillus* strains has led to a massive genome sequencing effort around the globe. In addition to 57 *Bacillus* genome sequenced deposited in Genbank until 2009 (13), 94 environmental *Bacillus* genome sequences from the U.S. were added in 2013 (6); followed in 2014 by 122 sequences of French *Bacillus* strains (14) and more recently, in 2015, a Danish genome effort provide sequences of 41 isolates of *Bacillus* (15). These 314 *Bacillus* genome sequences should provide information to discriminate the different sub-groups of *Bacillus*.

Common to members of the *B. cereus* group is the formation of spores that are resilient to heat and chemicals and have the ability to survive for long duration (16). Another trait that raises serious concerns for health organizations is the ability of *Bacillus* to develop and live within a biofilm. Bacterial biofilms, in general, require special attention for the food industry, as they can be a source of persistent contamination leading to food spoilage and to the transmission of

diseases (17). Indeed, *Bacillus cereus* spores and, to a lesser extent, vegetative cells embedded in biofilm are more protected against sanitizers (18,19). In bacterial species that were studied, biofilm provides means for the bacterial cell to attach and adhere to a variety of surfaces (both natural and man-made)(20,21). In *Staphylococcus* and *Pseudomonas*, it was suggested that cells imbedded in biofilm promote the survival of bacteria by forming a niche where bacteria can evade recognition by the host immune system (22,23); however, less is known regarding the myriad roles of biofilm in *Bacillus cereus* group. Biofilms form an extracellular polymeric matrix that is comprised of exopolysaccharides (EPS), proteins, lipids, and nucleic acids (24,25). A rigid biofilm structure from the motile model bacterium, *B. subtilis*, has extremely liquid and gas repellent properties; and mutation in biofilms formation suggested the involvement of EPS (26). While extended genetic research led to identification of numerous biosynthetic and regulatory genes involved in EPS formation in *B. subtilis* biofilm, less is known about biofilm formation in members of the pathogenic *B. cereus* group. Hence, despite differences in their pathogenicity, the exact nature of biofilm produced by this group of *Bacillus* and the repertoire of polysaccharide molecules made remain largely unknown.

In 2010 Gu et al. (27) identified two enzymes (Fig. 3.1A, B) and their corresponding genes in *Bacillus* food pathogen, and showed their involvement in the sequential conversion of UDP-N-acetyl-glucosamine to UDP-N-acetyl-glucosaminuronic acid and then to UDP-N-acetyl-xylosamine (hereafter abbr., UDP-GlcNAcA and UDP-XylNAc, respectively). These enzymes were named UGlcNAcDH for UDP-GlcNAc dehydrogenase and UXylNAcS for UDP-XylNAc synthase. The genes encoding these proteins are located within a conserved six-gene operon (XNAC) that includes, in addition to the nucleotide-sugar biosynthetic enzymes, two genes likely encoding glycosyltransferases and two small proteins of unknown functions. Based on

bioinformatics analyses (Fig. 3.1C), the XNAC operon is unique to bacteria of the *B. cereus* group and thus far appears to be lacking in other bacterial phyla. Despite the functional activities of enzymes in the XNAC operon no *Bacillus* glycan composed of GlcNAcA or XylNAc were reported. Hence, it remained unknown if XNAC-genes are expressed and if such glycan is truly made.

Here we provide the first evidence of a glycan that consists of two uncommon amino sugars, GlcNAcA and XylNAc. This glycan, named pzX, is made in *Bacillus* strains belonging to the *B. cereus* group. The formation of this exopolysaccharide occurs when cells are induced to grow in a medium composition known to trigger biofilm formation in *B. subtilis*, and it is released to the extracellular milieu during the release of mature spores.

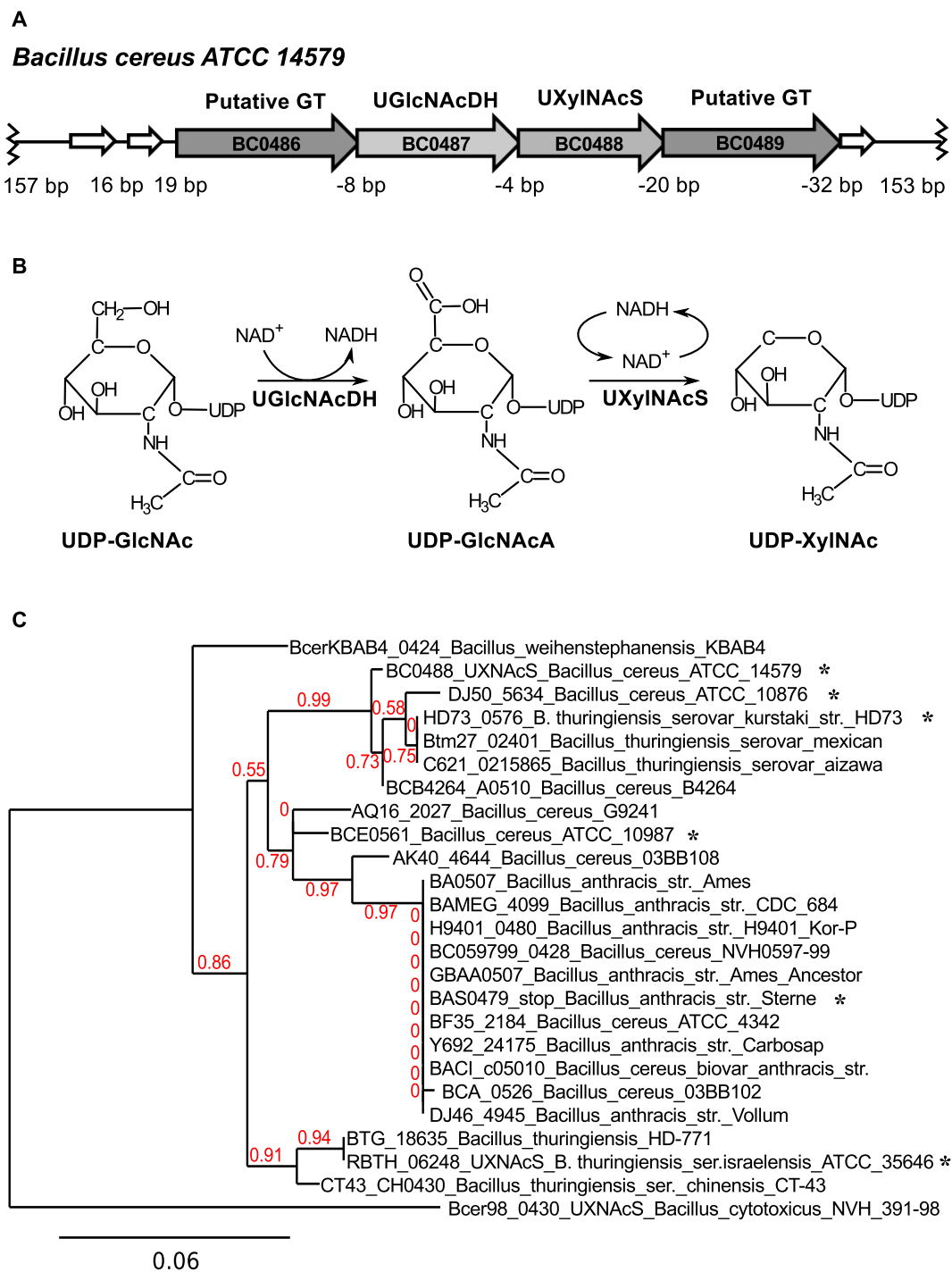


FIGURE 3.1. Comparative analyses of XNAC operon. (A) Organization and conserved synteny of genes within the XNAC operon. (B) The biochemical pathway for the conversion of UDP-GlcNAc to UDP-XyINAc by UGlcNAcDH and UXyINAcS (27) within the XNAC operon are illustrated. A phylogenetic tree of UXyINAcS (C) depicts that XNAC operon is only present in members of the *Bacillus cereus* sensu lato group. Strains marked with * are the lines that were used to analyze pzX in this study.

3.3 Results

3.3.1 Growth condition promoting pzX glycan synthesis

Following the finding of genes and the enzymes involved in UDP-GlcNAcA and UDP-XylNAc synthesis (27), the identification of glycan(s) consisting of these unusual amino sugars was unsuccessful when *Bacillus* cells were growing under typical laboratory conditions. We therefore decided to grow *Bacillus* in different growth media in order to find conditions that stimulate transcription of genes in the XNAC operon and promote the synthesis of XylNAc-containing molecule. RNA isolation followed by RT-PCR analyses concluded that genes likely involved in XylNAc-glycan(s) were not transcribed when *Bacillus* was grown in rich medium like LB or BHI. In contrast, genes of the XNAC operon were highly transcribed (Fig. 3.2A) when cells were grown in Msgg, a medium used to induce biofilm formation in *B. subtilis* (28). Under this medium we were able to identify a polymer that contains XylNAc, herein named pzX (fondly for Zi's XylNAc polymer). pzX was found outside the cells in the culture medium. We found that acidification of the medium led to a selective precipitation of pzX from other components, and such segregation protocol significantly aided in the isolation and further characterization of pzX. Crude pzX obtained from *Bacillus thuringiensis* strain *israelensis* (Bti) was hydrolyzed by TFA to monosaccharides and derivatized to alditol-acetate prior to separation by GC and analyses by EI-MS (Fig. 3.2B). Among the neutral monosaccharides residues observed in crude pzX were arabinose (Ara), glucose (Glc) and galactose (Gal) eluting from GC column at 23, 33.5, 34 min, respectively. A peak annotated X (30.6 min) later identified as XylNAc was also observed only when cells were grown in Msgg medium but not in BHI medium (Fig. 3.2B compare upper and lower panel). A few other amino-sugars residues were observed in the crude pzX including GlcNAc, GalNAc, and ManNAc (not shown). No

detectable XylNAc was observed when the *Bacillus thuringiensis* cells were grown in several other rich media including for example LB.

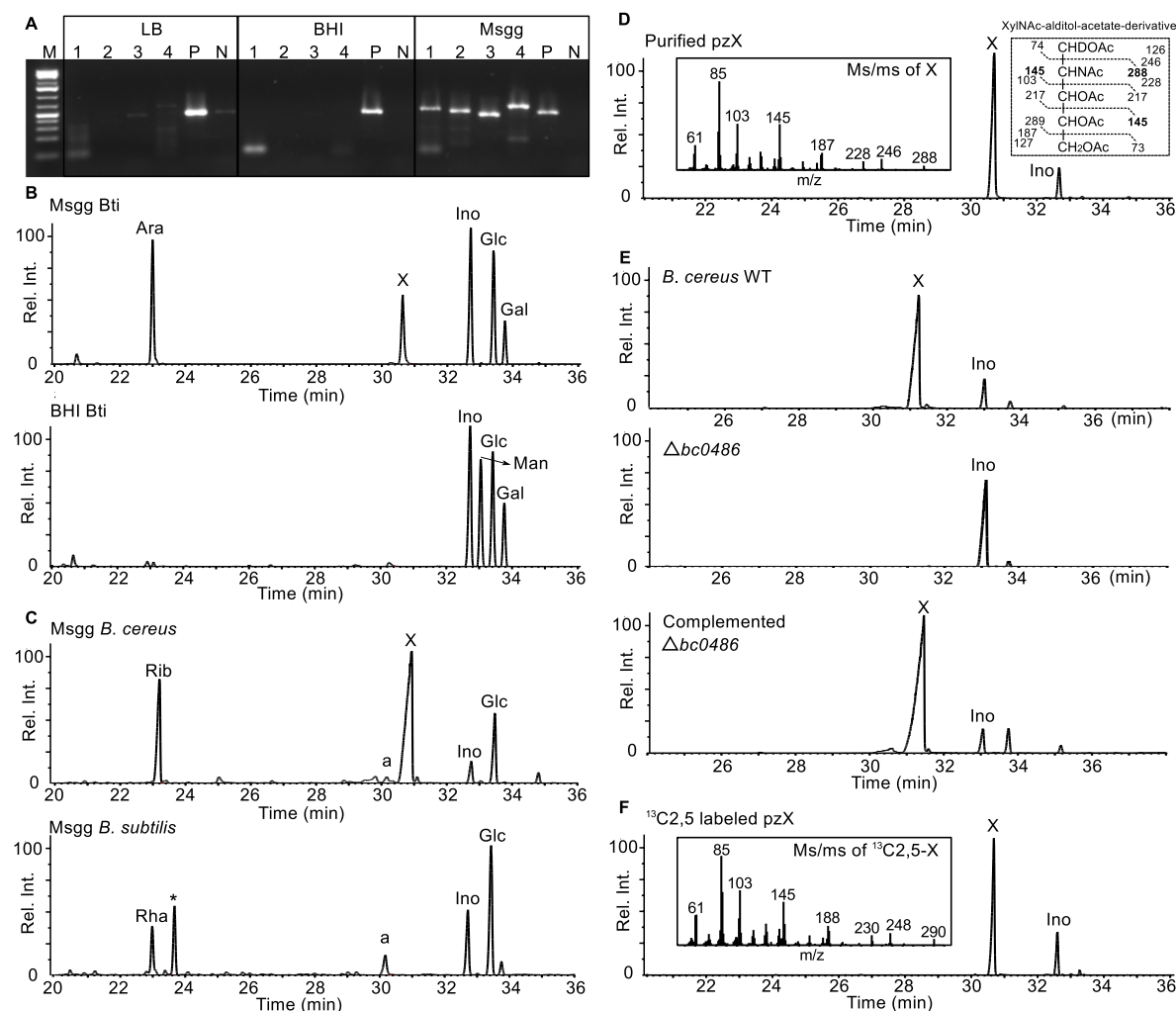


FIGURE 3.2. pZX is a secreted polysaccharide and induced to form when strains belonging to *Bacillus cereus* group are growing in Msgg nutrient medium. **A.** Genes of the XNAC operon are co-transcribed and the expression is induced during growth of *B. thuringiensis* in Msgg liquid medium but not in LB or BHI medium. RT-PCR analysis for the expression of gene encoding: lane 1, putative glycosyltransferase; lane 2, UGlcNAcDH; lane 3, UXylNAcS; lane 4, putative glycosyltransferase; lane P, sigA positive control; lane N, negative control (RT reaction without added primers). GC-MS analyses of alditol-acetate derivatives of **(B)** crude pZX of *B. thuringiensis* israelensis grown in Msgg and BHI medium (lower panel); of **(C)** *B. cereus* and *B. subtilis* (lower panel) grown in Msgg medium; of **(D)** purified pZX^{bc14579}. Note a single peak X (denote XylNAc) of borodeuterated XylNAc-alditol-acetate-derivative. In Panel C, peak * denotes an unknown sugar peak, and peak a denotes an unknown but non-sugar peak. In Panel D, the EI-MS with *m/z* ion fragments of peak XylNAc (X) is shown in the left boxed area; and the right boxed insert shows the predicted primary and secondary MS fragments of C1-deuterated alditol-acetate of derived XylNAc. Panel **(E)** shows the GC-MS spectra of *B. cereus* wild-type (WT), glycosyltransferase deletion (*Δbc0486*) and complemented strains (complemented *Δbc0486*). The GC-MS of **(F)** shows non-deuterated XylNAc-alditol-acetate-derivative obtained from *B. cereus* cells fed with [C₂-¹³C]glycerol and yielded ¹³C-2,5 labeled pZX^{bc14579}. Note the addition of 2 mass units to alditol-acetate derivative due to incorporation of two ¹³C per XylNAc (e.g. 288 → 290).

To determine if pzX synthesis is specific to *Bacillus* sp harboring XNAC operon, we examined different *Bacillus* strains grown in Msgg and BHI media. The data (Table 3.1) show that *Bacillus* strains lacking XNAC operon for instance, *B. megaterium* (Bm) and *B. subtilis* (Bs), did not produce pzX whether grown in Msgg or BHI medium. However, strains harboring XNAC operon like *B. thuringiensis* strain *berliner* ATCC 10792 (Btb), *B. thuringiensis* strain *kurstaki* (Btk), *B. thuringiensis* strain *israelensis* 4Q5 ATCC 35646 (Bti), *B. cereus* ATCC 10876, *B. cereus* ATCC 14579 and *B. anthracis* 34F2, all made pzX glycan but only when cells were grown in Msgg. For reason that remains unclear some strains like Bti, Btb, Btk produce much less pzX-polymer when compared to for example, *B. cereus* ATCC 14579 (Fig. 3.2C; Table 3.1). We thus decided to concentrate our effort on elucidating the structure of the pzX from *B. cereus* ATCC 14579 (herein abbreviated pzX^{bc14579}). Crude pzX^{bc14579} preparation was bound to anion-exchange column and was eluted with ammonium formate. Alditol acetate analyses of this fraction (Fig. 3.2D) show predominantly a single neutral monosaccharide, XylNAc. The electron ionization mass spectrometry (EI-MS) of this peak has primary ion fragments (see left insert in Fig. 3.2D) at *m/z* 288 and 145 along with secondary ion fragments at 246, 228, 126, 187, 127 and 103 (right insert in Fig. 3.2D). The retention time and *m/z* values are identical with those found for alditol acetates derivatives of a XylNAc standard derived from UDP-XylNAc (27). Hence the major neutral sugar in pzX is XylNAc. Since previous works showed that UDP-XylNAc is made by the enzymatic actions of Bc0487 and Bc0488 (27), we addressed if the putative glycosyltransferase (Bc0486) that co-resides in the operon facilitates pzX formation. To this end, we generated Bc0486 mutant (Δ bcb0486) and complemented strains. GC-MS analysis of glycosyltransferase deletion strain (Δ bcb0489) strain reveals no production of

pzX in *B. cereus* (Fig. 3.2E middle panel) while the complementation strain does (Fig. 3.2E lower panel).

Table 3.1. Comparison of pzX production among *Bacillus* strains

<i>Bacillus</i> strains	XNAC operon in genome	^a XylNAc production in Msgg medium	^b XylNAc production in BHI medium
<i>B. thuringiensis israelensis</i> ATCC 35646	Yes	6.8	ND
<i>B. thuringiensis berliner</i> ATCC 10792	Yes	5.1	ND
<i>B. thuringiensis kurstaki</i> HD73	Yes	7.5	ND
<i>B. cereus</i> ATCC 14579	Yes	228	ND
<i>B. cereus</i> ATCC 10876	Yes	493	ND
<i>B. anthracis</i> 34F2	Yes	154	ND
<i>B. subtilis</i> PY79	No	ND	ND
<i>B. megaterium</i> QMB1551	No	ND	ND

a. The relative amount of crude pzX in the medium was estimated after GC-MS analyses by peak integration and ratio of XylNAc and the internal std. inositol.

b. ND not detected under our standard procedure.

3.3.2 NMR Analysis of pzX^{bc14579} reveals a trisaccharide repeating units of [XylpNAc(4OAc) α -1,3 GlcpNAcA(4OAc) α -1,3 XylpNAc]_n

To gain initial insight into carbon flux leading to pzX and follow the metabolism of glycerol to pzX we fed culture with either a ¹³C-labeled glycerol at carbon-2 [C2-¹³C] or at all three carbons [¹³C₃]. Feeding *B. cereus* ATCC 14579 grown in Msgg/[C2-¹³C]glycerol, where the carbon source glycerol was replaced with glycerol labeled at C2 with ¹³C, was shown to incorporate the heavy carbon into C2 and C5 position of the XylNAc residue (see left insert in Fig. 3.2F of non-deuterated alditol acetate XylNAc derivative). The ¹³C labeling was also incorporated to monosaccharaide residues of other glycans at carbon 2 (C2) and carbon 5 (C5) like hexoses (Glc, Man, Gal), HexNAc's (GlcNAc, ManNAc, GalNAc), pentose (ribose) and 6-deoxy-hexose (Rha). Feeding the culture with all labeled glycerol (i.e [¹³C₃]glycerol) yielded cells with fully labeled sugars in the six carbon of hexose and 6-deoxy-hexose and the five

carbon pentose. The ^{13}C -labeled pzX was purified and used subsequently to determine its structure by NMR.

Our studies shown below will provide evidence that the proposed structure of pzX (Fig. 3.3A) is a glycan composed of a repeating trisaccharide unit made of three sugar residues that are linked α -1,3 one to another. For the purposes of NMR assignment and description of the pzX glycan we labeled the repeating three sugar residues as A, B and C: **A** refers to XylNAc-4-O-acetate (XylNAc-4OAc), **B** refers to XylNAc residues, and **C** to GlcNAcA-4-O-acetate (GlcNAcA-4OAc). The structure of pzX fits with our current knowledge and genes within the XNAC operon, where the UDP-GlcNAcA is the precursor for UDP-XylNAc (27) and the two glycosyltransferases likely utilize these nucleotide-sugars for the assembly of GlcNAcA and XylNAc into pzX.

3.3.3 pzX^{bc14579} consists of three sugar residues

Initial analyses of purified pzX^{bc14579} by NMR gave broad peak widths likely due to the high molecular weight and viscosity of this molecule. To reduce the size and obtain sharper spectra, pzX^{bc14579} was sonicated and NMR analysis was carried out at 65 °C. One-dimensional ^1H -NMR revealed three clear regions (Fig. 3.3B) each consists of overlapping proton peaks: Protons (5.05-5.26 ppm) belonging to the sugar anomeric region (H-1) together with downfield shifted protons (H-4s) that derived from O-acetylation (see discussion below); the resonances of sugar ring protons between 3.2 and 4.2 ppm; and proton signals between 1.9 and 2.2 ppm of terminal methyl group protons (typically found within acetate moiety). One-dimensional ^{13}C NMR of ^{13}C fully labeled pzX^{bc14579} gave four distinct spectral regions that are consistent with the structure as follows: the anomeric carbon region (~100 ppm) shows three major glycosyl residues (Fig. 3.3C, see boxed window), each showing expected doublet peak as a result of

anomeric peak split by ^{13}C -labeled carbon-2 of $\text{pzX}^{\text{bc14579}}$; and the three other regions of the ^{13}C spectrum show several resonances near ~ 175 ppm, near ~ 23 ppm and several signals between 80 and 50 ppm, and each fits to the chemical shifts of the carbonyl ($\text{C}=\text{O}$), the methyl (CH_3) groups of acetate, and the carbons from the sugar ring, respectively. To complete the assignment of sugar residues and confirm the structure of $\text{pzX}^{\text{bc14579}}$, two-dimensional NMR including COSY, TOCSY, NOESY, HMQC, HSQC, HMBC, HMQC-TOCSY and HMQC-NOESY were acquired. The complete chemical shifts assignments of $\text{pzX}^{\text{bc14579}}$ and of de-O-acetylated form of $\text{pzX}^{\text{bc14579}}$ are provided in Table 3.2 and Table 3.3, respectively.

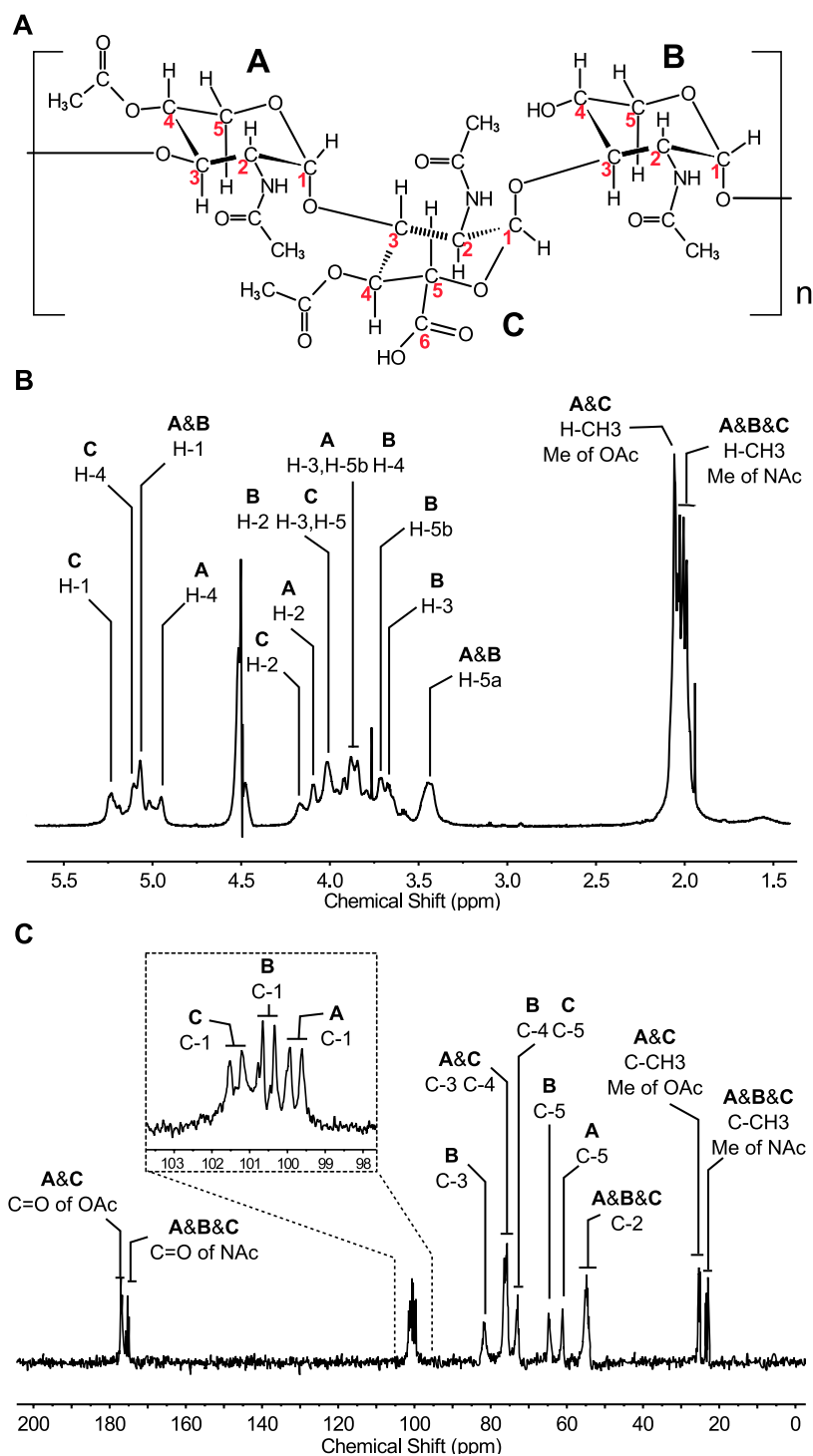


FIGURE 3.3. A proposed chemical structure of pzX. **A.** pzX consists of XylpNAc(4OAc) (denote A) α -1,3 linked to GlcpNAcA(4OAc) (denote C) that is α -1,3 linked to XylpNAc (denote B). One-dimensional proton NMR spectrum (**B**) of purified pzX^{bc14579}; and one-dimensional carbon-13 NMR spectrum (**C**), is shown. Protons belonging to each anomeric region are labeled “H-1” and the signals and chemical shifts for the three sugar residues carbon anomeric region (C-1) are shown in the boxed insert.

Table 3.2. Chemical shifts of pzX polymer components. Data was acquired in NMR 600Hz at 65 °C in D2O. DSS was used as reference.

Chemical shifts in ppm	A 4-OAc-XylNAc		B XylNAc		C 4-OAc-GlcNAcA	
Ring Position	¹ H	¹³ C	¹ H	¹³ C	¹ H	¹³ C
1	5.05	99.41	5.06	100.14	5.24	100.88
2	4.07	53.92	3.99	54.43	4.15	54.79
3	3.85	76.10	3.62	81.45	4.00	75.57
4	4.95	75.42	3.84	72.74	5.10	76.22
5a	3.81	60.70	3.68	64.33	3.94	73.09
5b	3.43		3.40		--	--
6	--	--	--	--	--	176.41
Substitution						
C=O O-Ac	--	175.22	--	--	--	174.85
CH3 O-Ac	2.04	23.00	--	--	2.03	23.20
C=O N-Ac	--	176.61	--	176.61		176.61
CH3 N-Ac	1.97	25.00	1.98	25.14	1.99	25.23

Table 3.3. Chemical shifts of de-O-Ac pzX. Data was acquired in NMR 600Hz at 65 °C in D2O. DSS was used as reference.

Chemical shifts in ppm	A XylNAc		B XylNAc		C GlcNAcA	
Ring Position	¹ H	¹³ C	¹ H	¹³ C	¹ H	¹³ C
1	5.12	101.47	5.06	101.72	5.21	100.99
2	3.97	55.06	3.95	54.98	4.01	54.76
3	3.61	82.12	3.72	81.01	3.71	81.38
4	3.81	72.92	3.81	73.06	3.87	74.22
5a	3.65	64.50	3.64	64.50	3.73	75.33
5b	3.45		3.45			
6	--		--		--	178.73
Substitution						
C=O N-Ac	--	176.66	--	176.66		176.75
CH3 N-Ac	2.01	24.857	2.01	24.857	1.97	25.23

3.3.4 pzX^{bc14579} sugar residues A, B are identified as 2-deoxy-2-N-acetyl-xylose

The initial carbon-proton correlated two-dimensional NMR signals were not prominent and gave poor signal intensity. This problem was overcome after generating in vivo pzX^{bc14579} samples that were ¹³C-labeled. pzX^{bc14579} was either selectively ¹³C-labeled in carbons 2 and 5 or

fully ^{13}C -labeled in all carbons. This labeling was obtained by purifying pzX^{bc14579} from *B. cereus* cell fed with [C2- ^{13}C]glycerol or fully labeled [^{13}C]glycerol, respectively.

Two-dimensional HMQC analyses show the protons directly bonded to their carbons. Analyses of the [2, 5 ^{13}C]pzX^{bc14579} sample show three C-5/H-5 signals (Fig. 3.4A) as well as a number of C-2/H-2 signals (Fig. 3.4B). The three C-5/H-5 signals have carbon chemical shifts between 60 and 75 ppm. The C-2/H-2 signals (A, B, C) have noticeably upfield carbon chemical shift (~54 ppm) and downfield proton (H-2) chemical shifts around 4 ppm. These three upfield carbon chemical shifts are consistent with a nitrogen attached to carbon 2. The downfield proton chemical shifts suggest that the C-2 nitrogen is acetylated; i.e. C-2 is attached to an N-acetyl group. Taken together, each of the three A, B, C residues contains a C-2 that has an N-acetyl substituent; i.e. all three residues are C-2 N-acetylglycoses. Fig. 3.4 also shows that the C-5 of residues A and B each has two attached protons indicating that both residue A and B are pentoses. Analyses of de-O-acetylated pzX^{bc14579} sample (see below) strongly support that the C5-pentose ring for both residue A and B is in a xylo-configuration based the large (~10-11) J coupling constant between H-4 and H-5b (Table 3.3). Hence, the NMR data indicate that A and B are XylNAc (2-deoxy-2-N-acetyl-xylose) which is in agreement with the GC-MS data (Fig. 3.2D).

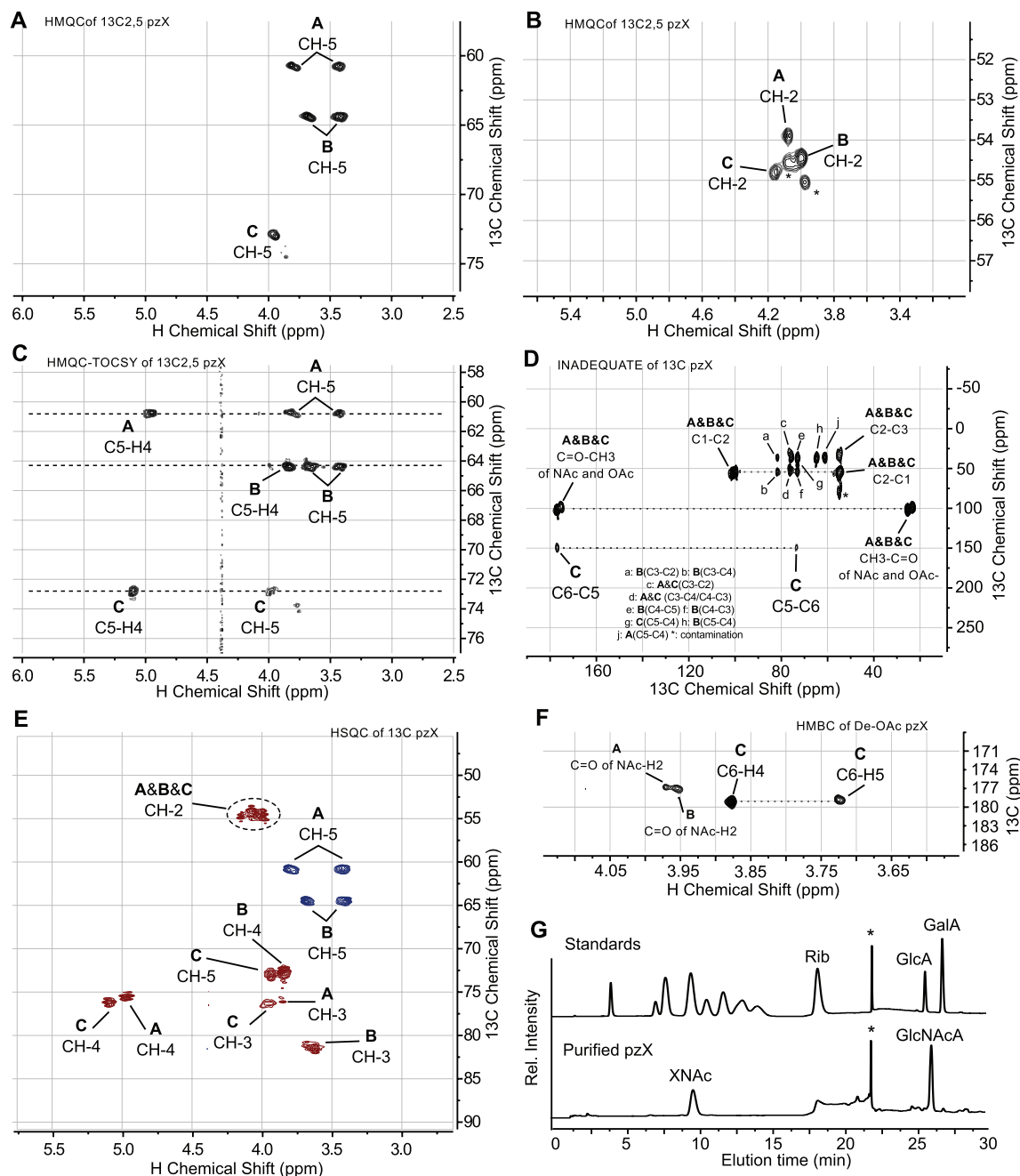


FIGURE 3.4. 2D-NMR revealed pzX consists of three sugars with XylNAc and GlcNAc configurations. The partial HMQC NMR spectrum of $^{13}\text{C}_{2,5}$ labeled pzX showing **A**) protons bond to carbon 5, in residues A, B, C [CH-5]; **B**) protons bond to carbon two in residues A, B, C [CH-2 region]. The partial HMQC-TOCSY spectrum (**C**) of $^{13}\text{C}_{2,5}$ labeled pzX depicts a unique H4 chemical shifts of the proton in residue A and C, implying 4-O substitution. An INADEQUATE ^{13}C - ^{13}C NMR experiment (**D**) of ^{13}C fully labeled pzX shows residue C has a connectivity of C5 to another carbon C6, with carbon chemical shift of a carbonyl. The NMR multiplicity-edited HSQC experiment (**E**) of pzX depicts carbon five (C5) of residue C only have one proton (phase red) whereas C5 of residue A and B have two protons (phase blue) implying residue A and B have *xylo*-configuration with two hydrogens linked to C5 and residue C is a 6-ring sugar with carboxylate at C6. The HMBC experiment of de-O-acetylated pzX (**F**) illustrates connectivity of C6 to H4 and H5 of residue C. Dionex HPLC- PAD spectrum (**G**) of TFA-hydrolyzed pzX (lower panel) gave two peaks, one eluted as neutral sugar (~9 min) and the other (~26 min) eluted as charged sugar. Label * indicates buffer contamination.

3.3.5 pzX^{bc14579} sugar residues C is identified as 2-deoxy-2-N-acetyl-glucuronic acid

The C-5/H-5 signal in HMQC spectra (Fig. 3.4A) for residue C shows its carbon is linked to a single proton implying that C-5 is bonded to another nuclei. Four lines of experimental evidence provide support that residue C is N-acetyl-glucosaminuronic acid (GlcNAcA) residue. The first proof was obtained by a ¹³C carbon-carbon connection INADEQUATE experiment (Fig. 3.4D) with ¹³C fully labeled pzX^{bc14579}, [¹³C]pzX. INADEQUATE experiment shows a pair of signals for C-5 of and a C-6 carbon with chemical shift of 176.4 ppm that fits resonance of CO likely a carboxylate (COO⁻) group (Table 3.2). This data along with C2/H2 HMQC established that C is a N-acetylhexaminouronic residue. The subsequent multiplicity-edited HSQC NMR experiment (Fig. 3.4E) illustrated the C-5/H-5 groups of residue A and B have different phase (blue) indicating that these carbons are linked to two hydrogens, as expected for XylNAc sugar residue. In contrast, the C-5/H-5 group of residue C has opposite phase (Fig. 3.4E, red) supporting the INADEQUATE and HMQC experiments that show that its C-5 carbon is bonded to a single hydrogen as well as to C-4 and C-6. The HMBC spectrum of the de-O-acetylated sample (Fig. 3.4F) also shows that residue C has cross-peaks from C-6 to H-4 and to H-5 consistent with it being a N-acetylhexaminouronic residue. To determine if residue C is in gluco- or galacto- configuration, pzX was TFA-hydrolyzed; and sample was analyzed by TOCSY. As expected, the acetate of OAc and NAc groups were cleaved upon the acid treatment to yield an amino monosaccharide (i.e. 2-deoxy-2-amino-hexouronic acid), of which the large coupling constants of J_{4,5}=8.3, J_{2,3}=9.4, J_{3,4}=8.2 Hz are consistent for residue with gluco-configuration and the small coupling of J_{1,2}=3.6 Hz supports an alpha sugar. Hence, the NMR data along with sugar analyses by HPLC (Fig. 3.4G) indicate that residue C is GlcNAcA (i.e. N-acetyl-

glucosaminuronic acid). Further analysis of HMQC-TOCSY spectra of the partially ^{13}C -labeled pzX^{bc14579} provided additional chemical shifts (see Table 3.2) that are consistent with GlcNAcA.

3.3.6 Residues A and C of pzX^{bc14579} are 4-O-acetylated.

The unusual higher than expected downfield chemical shifts of H-4s (~5 ppm) in residues A and C (Fig. 3.4 C, E) imply a substitution at position 4 of those residues. The carbon and proton C-4/H-4 in both residue A and C have increased C/H chemical shifts of ~ 75/5 ppm (Fig. 3.4C,E) consistent with O-acetylation when compared to unacetylated C-4 with C/H shifts of ~70/3.7 ppm. The correlation between H4 and a carbonyl carbon in the HMBC of the partially ^{13}C -labeled [C2, C5 ^{13}C] pzX^{bc14579} further confirmed that the substitution on the C-4 position of residue A and C is O-acetate (Fig. 3.5 panel A). The carbons from the carbonyl region (~175 ppm) show cross peaks to proton H-4 in both residue A and C (Fig. 3.5A on the left side). Additional cross-peaks from the same carbonyl carbon to the methyl hydrogens (Fig. 3.5A, right side) indicates that an acetate group is attached to carbon C-4 of residue A and C. Hence, residue A is XylNAc-4-O-acetate (XylNAc-4OAc) and residue C is GlcNAcA-4OAc. Further support for 4-OAc bonded to C-4 in residue A and C is shown after chemical hydrolysis of the ester group.

The ester linkage of O-acetate (C-4-OAc) group should be more susceptible to de-acetylation when compared with the amide N-linked acetate (C-2-NAc). Hence, base treatment of pzX^{bc14579} should remove the O-acetate group from C-4 but keep the N-acetate linked to C-2 intact. Indeed, HSQC NMR analysis of the de-O-acetylated pzX^{bc14579} sample shows that the original signals for C-4/H-4 protons of residues A and C (Fig. 3.4E) were shifted to the more conventional position with chemical shifts around 3.8 ppm (marked by arrow in Fig. 3.5 panel B). Comparison of one-dimensional proton NMR between before and after de-O-acetylation (Fig.

3.5C) also indicates a significant reduction in intensity of the methyl group region (dotted boxes) of the O-Ac group. As expected, no obvious shift in the protons linked to C-2 regions of A, B, and C residues (Fig. 3.5B, C-2/H-2 ~55/~4 ppm) is consistent with the expectation that the C-2-N-acetate groups are untouched after KOH treatment.

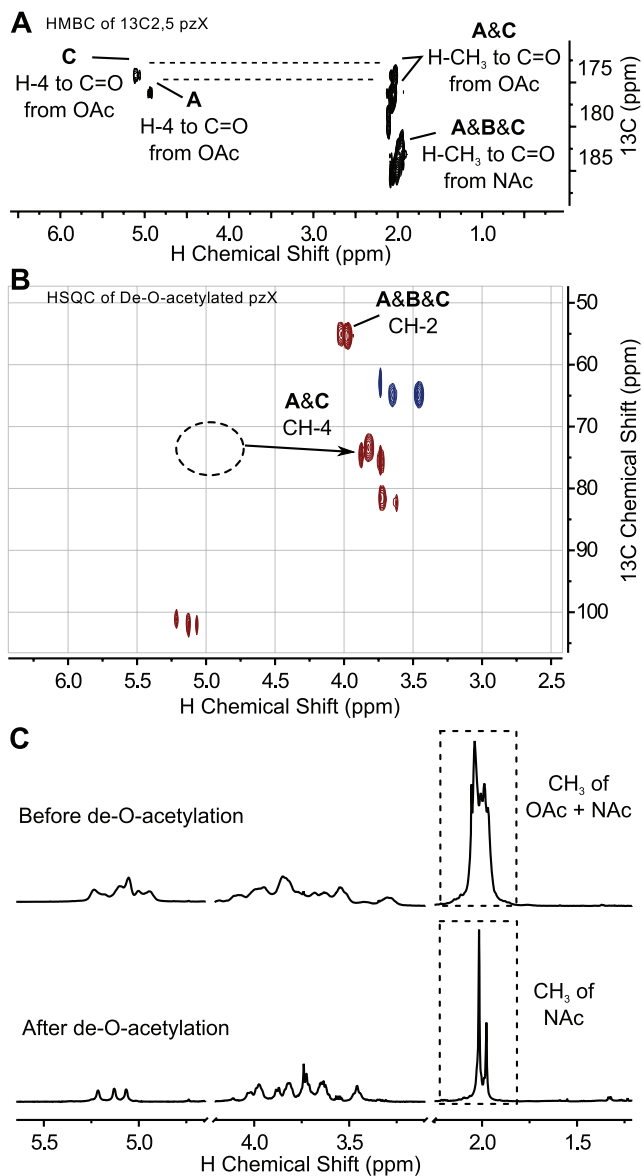


FIGURE 3.5. Residue A and C in pzX are 4-O-acetylated. Partial HMBC NMR spectrum (A) of ^{13}C -2,5 labeled pzX illustrates connectivity of carbon from OAc carbonyl group to protons of OAc methyl group and to the H4 in residues A and C. The partial multiplicity-edited HSQC experiment (B) of de-O-acetylated pzX illustrates the CH-4 proton belonging to A and C residues are shifted from their position (dotted circle denotes the chemical shifts before de-O-acetylation) to a regular non-substituted sugar ring region (pointed with arrow). Comparative analysis of partial spectrum of one-dimensional proton NMR is shown (C) before and after de-O-acetylation. Note the methyl region of acetate moieties (dotted rectangular box) is significantly reduced upon removal of OAc group.

3.3.7 Residue A, B, and C are alpha 1-3 linked to each other.

The position of the glycosidic linkages was determined initially by per-O-methylation analysis using GC-MS (Fig. 3.6 A, B). The free hydroxyl groups of intact pzX^{bc14579} were methylated and the per-O-methylated glycan was hydrolyzed. The resulting monosaccharides were reduced, acetylated, and the partially methylated alditol acetates (PMAA) were separated by GC-MS. A single peak eluting at 27.3 min gave primary and secondary ion fragments (Fig. 3.6B) with the following m/z values: 274, 231, 229, 171, 159 and 117. The m/z 159 and 117 ion fragments are diagnostic for NAc group attached to carbon 2 of the sugar residue, and the ion fragments 171 and 231 along with the other m/z values indicates this peak is a 3-O-acetyl-4-O-methyl-2-methyl-2-acetoamido-2-deoxy-xylitol. These PMAA analyses support that XylNAc residues on pzX^{bc14579} are 1-3 linked, however, linkage information related to the acidic residue C, GlcNAcA was not provided. Thus, further NMR experiments were conducted to complete the linkage analyses.

The linkage positions between the sugar residues were confirmed using HMBC and HSQC-NOESY spectra (see Fig. 3.6C, D, and E). In NOESY, if sugar residues were connected in α -configuration one would expect a strong NOE from H1 to H2 (for both XylNAc and GlcNAc configuration) because both proton are facing up and close in proximity. In the TOCSY spectrum however, one would expect for α -configuration to have a weak cross peak between H1 to H2, but a very strong cross peak in β -linked sugar residues. HSQC-NOESY spectrum of de-O-acetylated pzX shows strong NOEs between H1 and H2 within the same sugar residue (Fig. 3.6C), whereas weak cross peaks between H1 and H2 were detected in the HSQC-TOCSY spectrum (data not shown) suggesting both protons are close in proximity but with small coupling, thus each sugar residue are in α -configuration. The HSQC-NOESY experiment shows

additional cross peaks between H-1 of residue A and H-3 of residue C (Fig. 3.6D) indicating residue A and C are 1-3 linked (**aC1-cH3**) (Fig. 3.6D). Similarly, residue B and A, and residue C and B are 1-3 linked due to inter-glycosidic cross peaks between C-1 and H-3 (**cC1-bH3**; **bC1-aH3**). The same combinations of inter-glycosidic cross peaks were detected in HMBC experiment (Fig. 3.6E). Both spectra together with PMAA analyses support the fact that sugar residues A, C and B are linked A(1-3)C(1-3)B. In toto the NMR data provide strong evidence for the sugar sequence and configurations of the glycosidic linkages of pzX^{bc14579} glycan to be → XylNAc4OAc(α 1 → 3)GlcNAcA4OAc(α 1 → 3)XylNAc(α 1 →

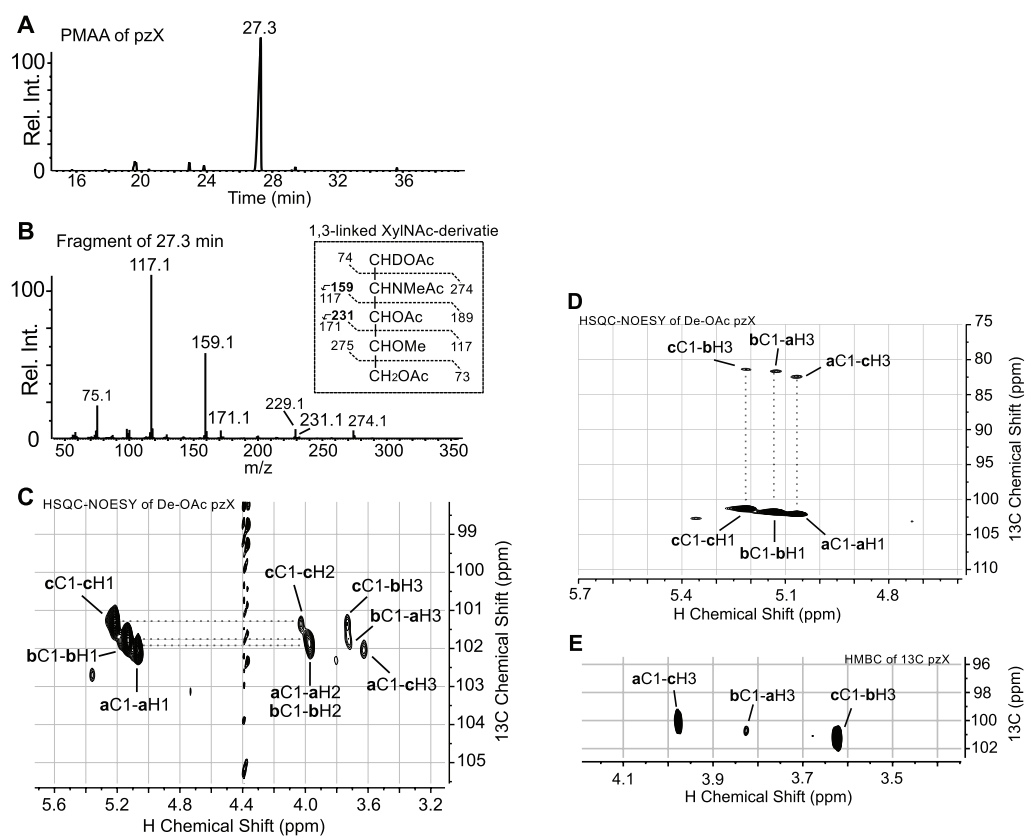


FIGURE 3.6. PMAA and NMR analyses of pzX revealed sugars are linked alpha 1-3 and sugar sequence is A-C-B. GC-MS (**A**) of partially permethylated alditol acetate (PMAA) has a single peak at 27.3 min with *m/z* ion fragments shown in panel (**B**). The boxed insert in B indicates the primary and secondary MS ion fragments depicting that XylNAc is linked at position 3. The partial HSQC-NOESY NMR spectrum (**C**) of de-O-acetylated pzX shows NOE cross peaks of H1- H2 in all three residues, indicative of alpha linked residues; and (**D**) shows 1-3 linkage and connectivity of anomeric carbon (C1) to proton along the glycosidic linkages with cross peaks of C1 of residue A to H3 of residue C, C1 of residue C to H3 of residue B, C1 of residue B to H3 of residue A. A carbon-proton inter-residue cross peaks (**E**) were detected in the HMBC experiments of ¹³C fully labeled pzX further confirming the linkage and sugar order of pzX is -A(α-1,3)B(α-1,3)C(α-1,3)A-.

3.3.8 pzX isolated from *Bacillus anthracis* Sterne 34F2 and *B. cereus* ATCC 10876 is similar to pzX¹⁴⁵⁷⁹

Initial GC-MS analyses of crude pzX isolated from the medium of Mmsgg-grown *B. anthracis* and *B. cereus* 10876 showed predominantly the XylNAc residue (Fig. 3.7 panel A). Further purification and analyses by proton-NMR (Fig. 3.7B) and TOCSY two-dimensional NMR (Fig. 3.7C) provided evidence that the purified EPS from the *B. cereus* ATCC 10876 and *B. anthracis* 34F2 have similar structures to the pzX from *B. cereus* ATCC 14579. The *B. anthracis* 34F2 strain used in this study is the one isolated by Max Sterne in the 1930s, which is the strain that was used to develop vaccine to anthrax in animals (2). The conserved chemical structure among the pzX from these Bacilli characterized in this report thus far suggests that it has a common role that is shared by all members of the *B. cereus* group.

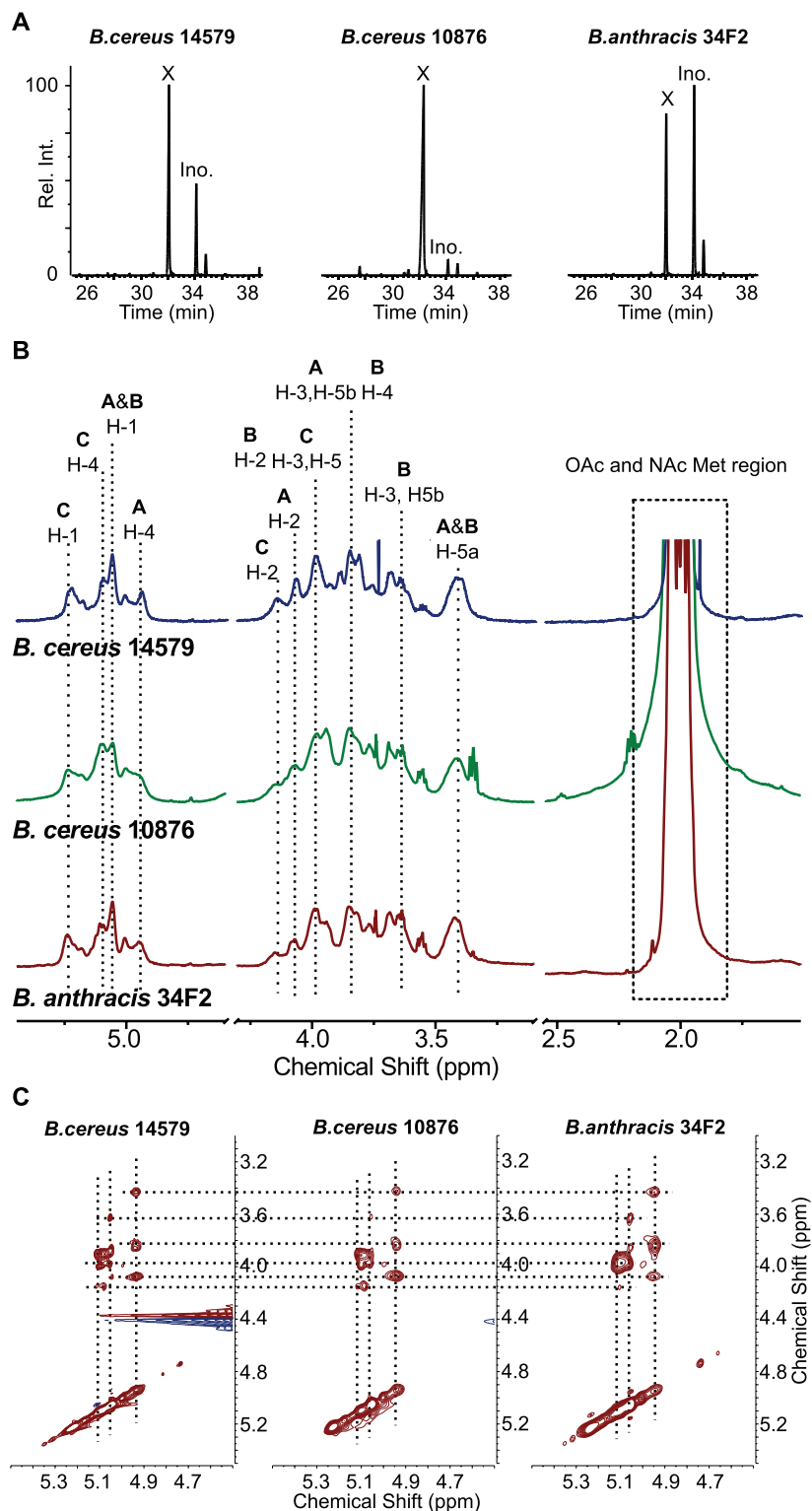


FIGURE 3.7. GC and NMR spectral overlay of pzX from *B. cereus* ATCC 14579, *B. cereus* ATCC 10876 and *B. anthracis* Sterne 34F2. **A.** GC-MS analyses of alditol-acetate derivatives of crude pzX of three *Bacillus* strains grown in Mmsg. **B.** 1D proton NMR and **C.** TOCSY experiment spectral overlay of purified pzX from three *Bacillus* strains.

3.3.9 pzX is synthesized during late stationary growth phase and released into extracellular milieu

In order to address the timing of pzX^{bc14579} synthesis and determine if it is further metabolized, we monitored the amount of pzX daily for 14 days in cells grown in Msgg medium. pzX^{bc14579} starts to accumulate around the second day, with highest amounts being produced during day 4-to 6 to a level of approximately 10 µg/ ml culture. Between days 6 to 14 in the culture, no further accumulation or degradation of pzX was observed, suggesting it is not catabolized. Microscopy analyses (Fig. 3.8A) show that once cells shifted to grow on Msgg, vegetative cells continued to grow and replicate (8 h). Between 12 h to 24 h, at which point cell density of 3 was reached, the cells entered aggregation phase and asynchronous sporulation, which consisted of mixed cell types with less than 10% spores. At 48 h, maximum cell density was reached (A_{600} , 5.5), and the cells entered the beginning of de-aggregation phase and started to release ~20% mature spores. At day 3 as cell aggregation continued to disintegrate, the absorbance was reduced to 4.8 as more mother cell lysed during sporulation, and on days 4-6, almost all cells (>90%) in the culture completed sporulation and released dormant heat-stable spores as confirmed by the ability of cell survival after 80 °C (Fig. 3.8B). The pattern of pzX^{bc14579} accumulation overlays with the release of spores into extracellular milieu when cultures are grown in Msgg (Fig. 3.8B). pzX formation was not detected when cells were grown in rich medium that does not support sporulation like BHI, not even after 14 days in culture.

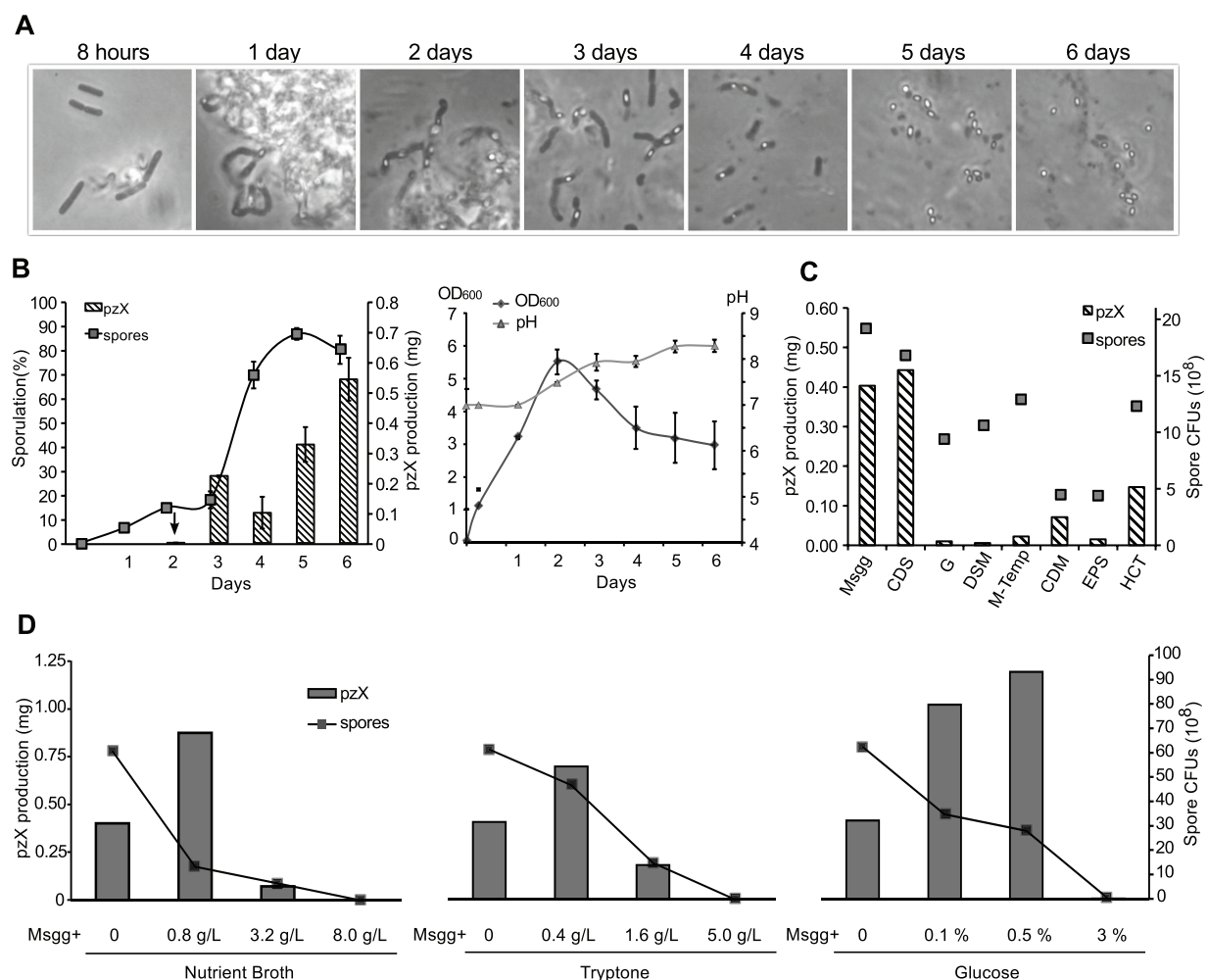


FIGURE 3.8. Analyses of pzX produced during sporulation. Cells of *B. cereus* ATCC 14579 grown in Msgg from 8 hours to 6 days were analyzed by phase-contrast microscopy (**A**) and the relative amount of spores and amount of pzX (**B**, left panel), as well as culture A₆₀₀ and pH (**B**, right panel) are indicated. Various amounts of pzX and spores were produced in other media (**C**). Variable amount of pzX is produced when Msgg medium is supplemented by increasing amounts of rich medium components or glucose (**D**). Note high level of nutrients media did not support mature spore formations nor did pzX formation. Note: CDS is CDS-Glc medium; M-Temp is modified Tempest medium.

The timeline of pzX^{bc14579} formation was further studied using different media known to enhance biofilm or sporulation of *Bacillus* sp. All media at day 1 produced cell aggregates. This was followed by de-aggregation phase and mature spore release after 5 days similar to Msgg medium. The media tested (Fig. 3.8C) included CDS-Glc, G, modified Tempest, CDM, DSM, EPS, and HCT, all produced pzX as determined by GC-MS analyses albeit in different amounts.

Interestingly, rich medium DSM, which is known to support spore formation, produced much smaller amounts of pzX^{bc14579} compared to Msgg. This finding along with the previous observation that rich media like BHI, LB, and TSB prevent pzX^{bc14579} formation prompted us to test if rich medium suppresses pzX synthesis. We examined cultures grown in Msgg (control) and cultures grown in Msgg supplemented with various amounts of rich medium components. GC-MS analyses of crude pzX^{bc14579} (Fig. 3.8D) showed that moderate increased addition of nutrient broth (comprised of beef extract and peptone), tryptone and low glucose to Msgg culture broth do affect the amount of pzX. However the addition of excessive dosage of rich medium (e.g. 8 g/L nutrient broth or 5 g/L tryptone), that is comparable in amounts to DSM or HCT medium, or high level of glucose (3%) prevents the formation of both pzX and spores. Collectively, the data suggest that pzX is released to extracellular milieu predominantly when spores are made. It appears that the production of pzX is regulated by several factors that include the type and amount of nutrients (e.g. glucose), the environment (where more pzX is made in shaking vs. still cultures), and lastly the completion of sporulation.

3.3.10 pzX has surfactant, adherence and antiaggregant properties

Further investigation into the role of pzX in the extracellular medium revealed that pzX has a surfactant property with the ability to increase the growth diameter of a colony when grown in agar by over 30% (Fig. 3.9A). The surfactant property was determined by the drop collapse method, and the data show that when pzX is dissolved in deionized water it decreases the surface tension of water from 84 to 34° (Fig. 3.9B). The dose response curves show that the surface tension decreases continuously with increasing concentration of pzX. The decrease in surface tension is not affected if pzX is deacetylated or autoclaved (Fig. 3.9C) suggesting that the glycan is heat stable and maintains its property even under alkaline or hydrolytic conditions that may

strip off its O-acetate side-chain decoration. We also observed surfactant activity with *B. subtilis* suggesting that this property of the glycan is not unique to *Bacillus cereus sensu lato* group.

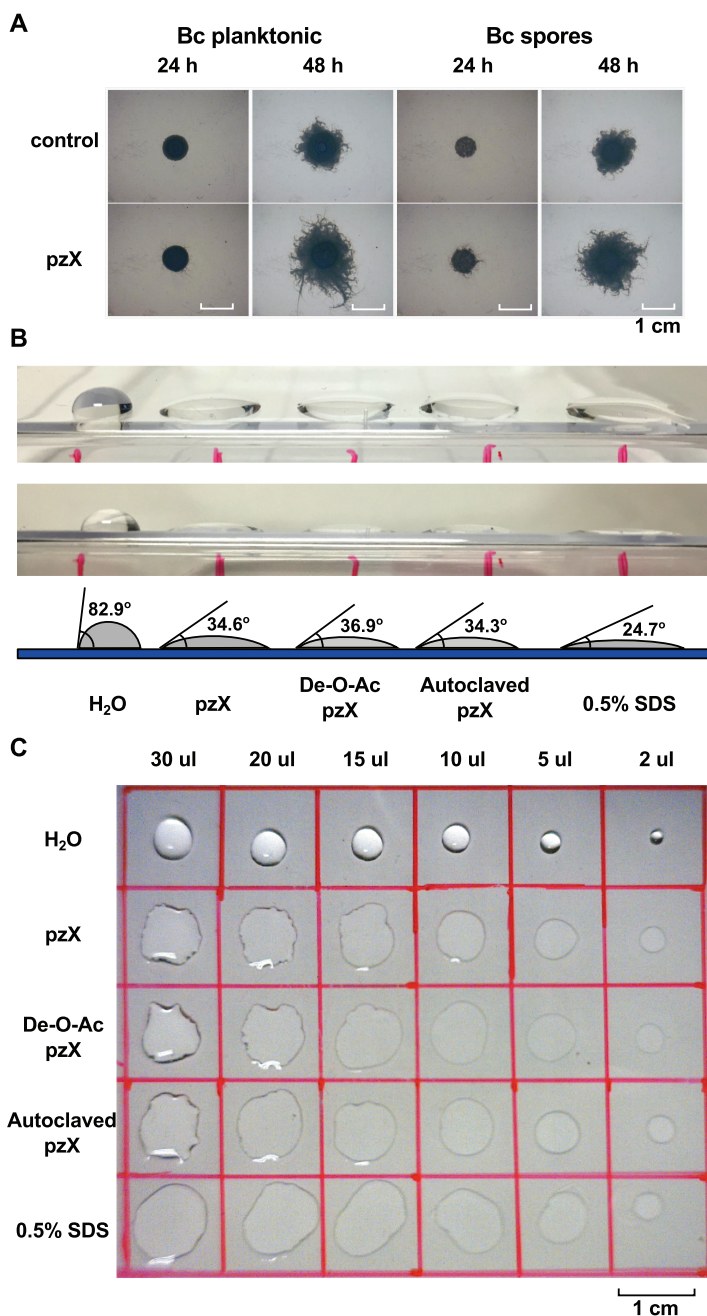


FIGURE 3.9. pzX has surfactant properties. *B. cereus* ATCC 14579 planktonic cells and spores spotted on Mmsg agar plates previously layered with pzX (lower) or water (upper panel, control) developed a wider colony morphology within 48 hours compared to control (A). Drop of water collapsed on a flat polystyrene surface previously spread with native pzX, de-O-acetylated, or autoclaved pzX compared to the surface with water or 0.5% SDS as control. The contact angle of each sample is diagrammed below (B). Various amount of water (30, 20, 15, 10, 5, 2 μ l) collapsed on polystyrene surface spread with native pzX, de-O-acetylated, or autoclaved pzX compared to the surface with water or 0.5% SDS as control (C).

Since no apparent degradation of pzX^{bc14579} is observed after spores are released, we investigated the role of pzX as a molecule that increases/decreases adherence of the spores to surfaces. Figure 3.10 panel A left shows that in the presence of pzX 56% of spores adhered to the surface of a defined soil-like material (vermiculite) when compared with control showing 16% of adherence. The adherence to vermiculite is dose-dependent (Fig. 3.10A right), suggesting that with increasing amount of pzX in the extracellular milieu of the spores, a higher number of cells will adhere to environmental surfaces.

To investigate the role of pzX after spore germination we carried out the following experiments. First we determined that pzX itself is unable to induce spore germination (data not shown). To examine pzX post spore germination we incubated germinated *Bacillus* spores with no nutrient for 2 days in the presence or absence of pzX. The phase-contrast microscopy analyses (Fig. 3.10B) suggest that germinated *B. anthracis* spores tend to aggregate, however, pzX reduces aggregation (antiaggregant). The CFU counts (Fig. 3.10C) show a decrease during the first 4 hours, indicating an initial aggregation immediately post germination. However, post the 4 hours window, the CFU count is increased to its initial level, likely due to the antiaggregant activity of pzX.

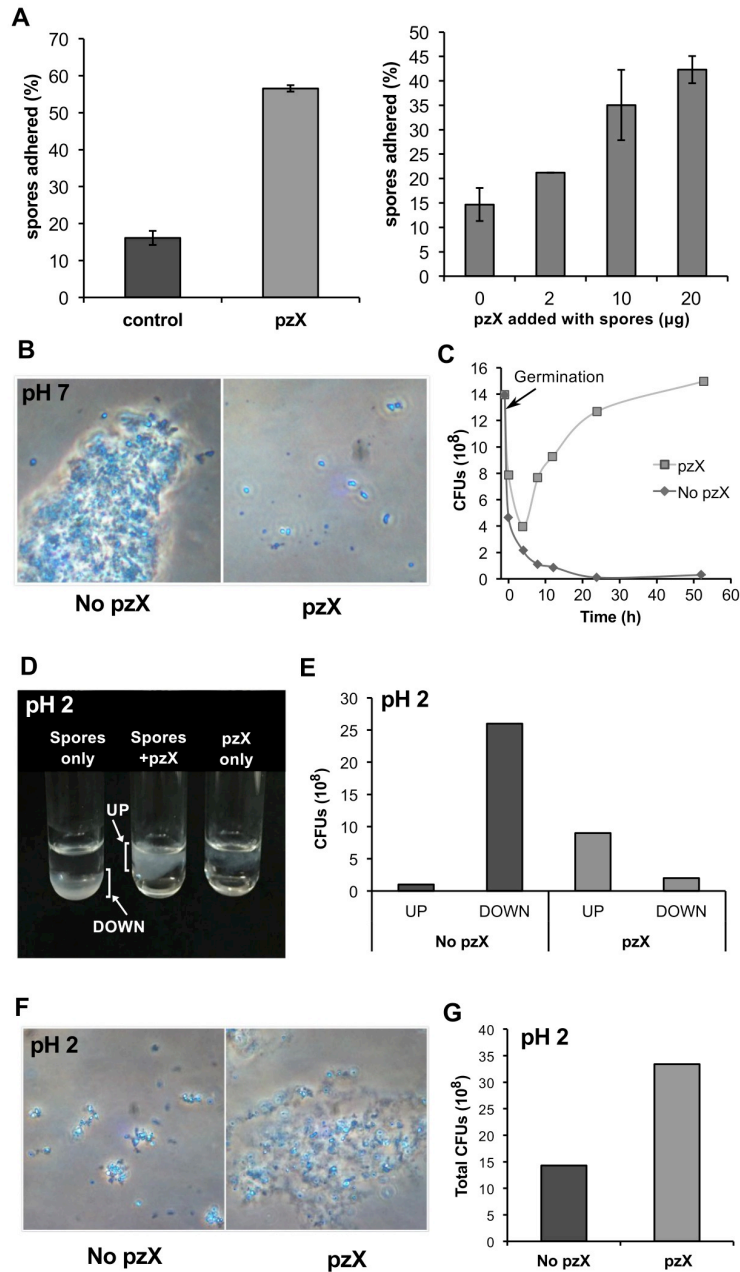


FIGURE 3.10. Other properties of pzX. Panel A shows the role of pzX in spore adherence. Purified spores of *B. cereus* ATCC 14579 wild type were mixed with or without purified pzX; dried on vermiculites; washed and the remnant attached spores were counted by CFUs (**A, left**). Adherence is concentration dependent (**A, middle**). In addition, culture medium of *B. cereus* WT and $\Delta bc0486$ mutant grown for 5 days in Mmsg were directly added to vermiculites without any purification, and the adherence was compared (**A, right**). Panels B, C show pzX as an antiaggregant. Germinated *B. anthracis* Sterne spores were incubated with or without 20 μg pzX for 2 days, and phase-contrast microscopy shows pzX disperses spores (**B**) and in addition enhances number of CFUs (**C**). Panels D, E, F, G provide evidence for pzX role in a film-like matrix formation at environment with low pH. *B. anthracis* spores mixed with or without pzX show the formation of cotton fluff-like floating structure (**D**). Note the UP and DOWN bracket indicates region above or below the ‘fluff’. CFUs from the UP and DOWN regions show that spores are associated with the fluff pzX structure (**E**). Phase-contrast microscopy shows a film-like structure of pzX with spores embedded (**F**). Total CFU of spores after vortexing (panel D) and plating on BHI agar shows that spores within this structure are not aggregated (**G**).

Since the chemical and physical property of pzX is to precipitate below certain pH, and the data so far suggest that pzX co-forms with spores, we wonder how spores behave at such environment. To test this, *Bacillus anthracis* spores were challenged with pH 2 for 24 h in the presence or absence of pzX. Spores without pzX aggregate and settle to the bottom (Fig. 3.10D) as expected. However, we were surprised to notice that the spores with pzX formed a cotton fluff-like structure that floated on the top of the culture (see Fig. 3.10D bracket labeled UP). Analyses of spores above or below the floated fluff show that the cotton structure consists a significant number of spores (Fig. 3.10E). Phase-contrast microscopic analyses (Fig. 3.10F) reveal that the fluff-pzX forms a film-like structure that harbors the spores. Interestingly in such matrix configuration, the spores do not aggregate (see Fig. 3.10G).

3.4 Discussion

Bacterial pathogenic species belonging to *Bacillus cereus* group have an advantage in nature due to their capacity to survive for many years as dormant spores despite harsh changes in the environment. Past work utilized purified spores to examine their chemical and physical properties in adaption and fitness to environment or to examine infection processes, but here we provide new evidence that dormant spores are in fact, co-made within an extracellular matrix containing an unusual polysaccharide, pzX.

The XNAC operon consists of two nucleotide-sugar biosynthetic enzymes that were biochemically shown to form the precursors UDP-GlcNAcA and UDP-XylNAc (27), and two glycosyltransferases that are likely involved in pzX synthesis. Genomic data (Fig. 3.1C) reveal that all strains belonging to this group carry the XNAC operon in their chromosomes and therefore have the capacity to synthesize pzX. We further examined this capacity by analyzing six random *Bacillus* members of this group representing strains from different phylogenetic

clades (see * in Fig. 3.1C), and showed they all produced pzX, while non-group members like *B. megaterium* or *B. subtilis* that are lacking the XNAC operon in their chromosome did not produce pzX (Table 3.1). The collective biochemical analyses including ^{13}C -carbon labeling, sugar methylation, GC-MS and NMR analyses provide strong evidences that pzX backbone consists of a repeating trisaccharide sequence of $[-3)\text{XylNAc}(\alpha 1-3)\text{GlcNAcA}(\alpha 1-3)\text{XylNAc}(\alpha 1-)]_n$ and with a side chain decoration of *O*-acetate group link to C-4 of XylNAc and to C-4 of its acidic amino sugar, GlcNAcA. Hence, the chemical structure of pzX agrees with the two UDP-sugar biosynthetic enzymes encoded by the XNAC operon (Fig. 3.1A). Taken together, the conserved pzX chemical structure among the three *Bacillus* strains (*B. cereus* ATCC 10876, *B. cereus* ATCC 14579, *B. anthracis* Sterne 34F2) as characterized by NMR (Fig. 3.7) and the conserved XNAC operon let us to propose that pzX has a common role that is shared by all members of the *B. cereus* group.

In addition to biochemical analyses, we have shown that the XNAC operon is directly involved in pzX formation in members of the *Bacillus cereus* sensu lato. For example, deletion of the glycosyltransferase (*Abc0486*) yielded no pzX (Fig. 3.2E) in *B. cereus* ATCC 14579. Similarly deletion of glycosyltransferases in *B. thuringiensis israelensis* (Bti) strain also showed no production of pzX (data not shown). This genetic data provide strong evidence that the glycosyltransferase encoded by XNAC operon is involved in the synthesis of pzX.

Although *Bacillus* pathogens have been studied extensively for many years, it is not surprising that pzX was not detected since typical laboratory growth media is discarded during spore isolation, and specific and extensive analytical methodology will be required to identify, purify and characterize minute polysaccharides. pzX is formed when *Bacillus* cells are grown in a defined nutrient medium, like Msgg, however, other media also support its formation (see Fig.

3.8C). The data so far suggest that pzX is accumulated during sporulation, however, the relative amount may be regulated by the amount of nutrient available. For example, when the defined medium Msgg was supplemented with 0.5% glucose, higher amount of pzX was produced in comparison to cells grown in Msgg alone. Higher amount of pzX can also be made when Msgg is supplemented with small amount of undefined medium (see Fig. 3.8D). Other defined nutrient media differing in composition that support sporulation also support pzX formation (Fig. 3.8C). On the contrary, media that do not support sporulation produce no pzX. For instance, pzX was not detected when strains were grown in rich medium (like BHI and LB) even after prolonged time (2-6 days of incubation) or when Msgg was supplemented with a rich medium component or with 3% glucose, as those media did not support mature spore formations (Fig. 3.8D). Analyses of the putative promoter DNA region, 5' to Bc0484, revealed sequences that potentially can be recognized by the sporulation specific sigma factors, SigG and SigE. Clearly, future studies will be required to identify the molecular mechanism that controls pzX formation.

Interestingly, little pzX was observed when *B. cereus* ATCC 14579 were grown in Difco Sporulation medium (DSM). This observation raises two questions: do all sporulation media make identical spores or are some spores made in certain medium chemically or physically different from others? Since DSM has excessive nutrients that probably do not mimic natural environments *Bacillus* would encounter, we propose that under natural settings spores are made with pzX. Alternatively, DSM may consist of a factor that suppresses pzX formation. We are currently pursuing a study aimed at identification of nutrient factors controlling XNAC operon gene expression, the accumulation of intermediate metabolites, and the amounts of pzX formed. Further studies are clearly needed to elucidate the molecular mechanism that senses, controls and regulates the level of pzX accumulation. The relationship between nutrients and pzX may

identify new regulatory factors that tune (enhance or suppress) the flux of metabolism to pzX formation, and factors that control EPS common to only species belonging to *B. cereus* sensu lato group.

The biological role of pzX as biosurfactant (see Fig. 3.9) is complex. Some surfactants are toxic, but we have seen no such toxic effects of few laboratory *E. coli* and fungal strains that were co-cultivated on agar with pzX. The surfactant could enhance the spread of germinated spores and this may give an advantage over other bacteria fighting for limited nutrient resources. We have shown that pzX precipitates at low pH, most likely due to protonation of its acidic sugar GlcNAcA. The precipitated pzX is dissolved when pH is higher than 4-5. While several factors are known to protect the spore including for example, spore outer coat layer and the exosporium (29), we cannot exclude the possibility that glycan (pzX) is an additional factor in spore biology. It is therefore probable that under acidic environments (e.g. human stomach, soil) precipitated film pzX like-structure (Fig. 3.10F) gives additional protection to the spore. Contrary to animal, the midgut cavity of some larvae is alkaline, and under such high pH values the pzX is not degraded. Despite the fact that the O-acetate could be cleaved under alkaline condition, our data show that the biosurfactant property is still maintained even when pzX is de-acetylated.

The pzX adherence assays (Fig. 3.10A) show that spores mixed with pzX adhere significantly stronger to an artificial soil surface (vermiculite) in comparison to spores alone. This is difficult to interpret, and our current working model is that pzX enclosed and trapped spores to the soil surface during desiccation. Brief contact with water may not be sufficient to re-hydrate pzX and release spores. This adhesion property may have a more profound affect under acidic soil environment where larger amounts of dormant spores may cling to a soil particle

rather than being washed away. Further research will be required to determine if pzX adherence property plays a role during interaction or persistence with the microbial host.

Following sporulation we found no evidence that pzX undergoes hydrolysis despite the fact that the mother cell releases many hydrolytic enzymes. Furthermore, the chemical nature of pzX and its distinctive α -1,3-linked amino-sugars suggest that it may protect cells from common lytic enzymes like lysozyme, an enzyme that cleaves β -1,4-linked sugars. Taken together, the results of the present study reveal a novel extracellular polysaccharide that is formed during sporulation and likely to play multiple roles in the interaction of the spore with its environment. Many questions still await answers to fully understand the biology of this novel polysaccharide we named pzX.

3.5 Experimental Procedures

3.5.1 Strains and culture conditions

The bacteria used in this study (Table 3.1) were stored in BHI containing 16% glycerol at -80 °C. *Bacillus* strains were routinely grown in BHI medium, LB, or Msgg medium (28). *In addition*, various sporulation and biofilm-inducing media were used: CDS-Glc (30), EPS (31), HCT (32), modified Tempest medium (33) with 5 mM phosphate and 0.1% glucose, modified-G medium (34), the modified Schaeffer (35) also known as Difco sporulation media (DSM), and CDM medium (36). Spores were separated from medium of 5-day culture by centrifugation (8,000 g, 10 min); washed with sterile deionized water (ddw), suspended and pelleted again before storage (up to 1 month) to give a total yield of 10^9 per ml. The number of spores in liquid culture was determined after heat treatment at 80 °C; conditions that activate spores and kill any remnant vegetative cells (37). An aliquot of a culture was heat-treated for 20 min at 80 °C, cooled to 25 °C and vortexed, and serial dilutions were plated on BHI-1.5% agar plates. Total

CFUs and the percentage of spores from vegetative cells were estimated by comparing samples that were heat and non-heat treated.

3.5.2 RNA isolation and RT-PCR

B. thuringiensis cells were grown in different liquid media (e.g. BHI, LB, and Msgg) to an A_{600} of 4-5.5; and cell amounts equivalent to A_{600} of 4 were pelleted at 6,000 g for 1 min at 4 °C, resuspended in 800 µl of lysozyme solution (10 mM Tris-HCl pH 8, 1 mM EDTA, 10 mg/ml lysozyme), and incubated at room temperature for 10 min. Each sample was supplemented with 80 µl 10xEB (0.3 M NaOAc pH 5.2, 50 mM EDTA, 5% sarkosyl and 1.42 M β -mercaptoethanol), and incubated at 65 °C for 2 min, followed by addition of 1 volume of one-phase 70 °C preheated acidic phenol (phenol:AcE buffer (50 mM NaOAc pH 5.1, 5 mM EDTA) 1:1 v/v). After brief vortex and incubation at 65 °C for 7 min, each sample was centrifuged (10,000 g for 5 min at 4 °C); the aqueous layer was collected and mixed with 1 volume of chloroform. Nucleic acid partition to upper phase was collected and precipitated with cold ethanol and 30 mM NaOAc pH 5.2 at -20°C. The concentration of crude RNA was determined from the absorbance at 260 nm. To digest remnant genomic DNA contamination, an aliquot of 5 µg crude RNA was treated with DNase I. The resulting RNA was ethanol-precipitated and about 250 ng of RNA was used for further reverse transcriptase (RT) and PCR reactions. For cDNA synthesis the 20 µl RT reaction was consisted of 250 ng of RNA, 5 µM random hexamers as primer, buffer, 0.2 µM dNTP's and 1 unit of reverse transcriptase (SuperScript III, Invitrogen). Negative control RT reactions were done without added primer. Transcripts of genes in XNAC operon and Sigma A as positive control (sigA, rpoD BC4289) were amplified using a 25 µl PCR reaction that included 2 µl of RT reaction, buffer, dNTPs, 1 unit of Taq DNA polymerase (Fermentas) and 0.4 µM of each gene specific sense and antisense primers (ZL035-5'-

ctagcattcatctgtattcttttctcc; ZL036-5'-gcagatattgtttattcgttgatgggtg; ZL037-5'-
tcatacactcaccttctttataattcc; ZL038-5'-atctacaaaaccgtttggatttactc; ZL039-5'-
ttagccattgttttcttcatgaaacc; ZL040-5'-cgttatTTtaattattatggccaagagc; ZL041-5'-
ctataaaacatctcttttctattgccatc; ZL042-5'-attatttagatctaagagtaccacttcac; ZL043 -5'-
ctattctaagaaatccttaagacgcttac; ZL044-5'-caattcgtattccagttcatatgggtg). Following PCR, a portion (5
μl) of each RT-PCR reaction was loaded on 1% agarose-TAE gel casted with 10 μg/ml ethidium
bromide, separated by gel electrophoresis and UV-imaged using gel imager.

3.5.3 Isolation, purification and analyses of XylNAc-glycan, pzX

A portion of *Bacillus* strain cells grown in BHI (0.5 ml) was used to inoculate 50 ml
Msgg medium and culture was incubated for 4 days at 30 °C, shaking at 200 rpm. The culture
was centrifuged (10,000 g, 10 min, 4 °C) and the medium was filtered prior to lowering the pH
to 2. After centrifugation (8,000 g, 1 h, 4 °C), the precipitated crude pzX polymer was washed;
resuspended with deionized water, and the sugar composition was analyzed by GC-MS. For
purification, the crude pzX solution was chromatographed on Q-sepharose (GE, 5 ml column)
pre-equilibrated with 5 mM ammonium formate, and eluted by stepwise gradient of increasing
concentration of ammonium formate. The purified pzX eluted at 400 mM ammonium formate
was lyophilized, dialyzed against deionized water, and used for further analyses. To obtain ¹³C-
labeled glycans, the glycerol in Msgg medium was substituted with carbon-2 labeled glycerol,
[C2-¹³C]glycerol, or with uniformly labeled glycerol, *[¹³C₃]glycerol* (Cambridge Isotope
laboratories). Following 4 days of incubation, cultures were processed and pzX was purified as
above. The size of pzX was determined by gel filtration on Superdex 75 column (1 cm i.d. x 30
cm, GE). An aliquot of pzX (0.5 ml) was injected via Agilent HPLC system equipped with UV

detector (200 nm) to the column and chromatography was carried out at a flow rate of 0.5 ml/min using 0.5 M NH_4HCO_3 .

To determine neutral and amino-sugar composition of crude pzX, column fractions, or purified pzX glycan, an aliquot was supplemented with 10 μg inositol; hydrolyzed with 2M *trifluoroacetic acid* (TFA) at 120 °C; and the released monosaccharides were reduced to their alditols (38), and acetylated. The resulting alditol acetate derivatives were analyzed by GC/EI-MS system (Agilent 7890a/5975c), equipped with an autosampler injector (Agilent 7693). A 1 μl sample was injected into GC-column (Equity-1 or DB-5, 30 m $\times$ 0.25 mm, 0.25 μm film thickness) using split mode (1:50) with injector inlet setting of 250 °C (helium at 3 ml/min). Helium was also used as column carrier gas (1 ml/min). After injection, the GC column chamber temperature program was held for 2 min at 80 °C; the temperature was increased to 140 °C at a rate of 20 °C/min, followed by an increase to 200 °C at 2 °C/min; an increase to 250 °C at 30 °C/min; finally the temperature was held at 300 °C for 5 min before next sample injection (run time of $\sim$ 50 min). The MS detector was operating under electron impact (EI) ionization at 70 eV and the temperature of the transfer line between the column end to MS was 250 °C. The temperature of the MS source was 230 °C and the quadrupole 150 °C, respectively. MS data were collected after solvent delay of 5 min, in a continuous scanning mode, recording ion abundance in the range of 50-550 m/z. The spectra were analyzed using Software MSD ChemStation D.02.00.275 (Agilent Technologies). To determine the elution time of authentic XylNAc and the EI-mass fragments formed, UDP-XylNAc was produced enzymatically and purified over Q15-column (27); TFA-hydrolyzed, converted to alditol-acetate, separated by GC using DB-5 column and analyzed EI-MS using above GC conditions.

3.5.4 O-Methylation, de-O-acetylation, and HPLC-PAD analyses

The position of glycosidic linkages was determined by methylation analysis. Dry pzX-glycan was dissolved in 200 μ l dimethylsulfoxide and free OH groups were methylated with methyl iodide in a NaOH-DMSO slurry as catalyst (39) for 7 min at room temperature; the resulting per-O-methylated pzX was extracted with dichloromethane and dried. The O-methylated glycan sample was TFA-hydrolyzed, and the released partially methylated monosaccharides were reduced with NaBD₄; acetylated with acetic anhydride in pyridine and the derived partially methylated alditol-acetate derivatives were analyzed by GC-MS as described above. For MALDI-TOF MS analyses, 1 μ l of per-O-methylated-pzX oligosaccharides was mixed with 1 μ l DHBA solvent (10 mg α -dihydroxybenzoic acid in 0.5 ml MeOH:H₂O v/v) and spotted on MALDI plate and mass spectrometry was carried out in positive mode (Bruker microflex LT MALDI-TOF). To obtain pzX glycan lacking O-acetate groups, purified pzX glycan was incubated for 2 h with 1M KOH at room temperature, dialyzed three times against 2 liter of deionized water, lyophilized and resuspended in D₂O for NMR analyses. For HPLC analysis of pzX sugars, purified pzX was hydrolyzed by 2 M TFA at 120 °C for 2 h, and the excessive TFA was evaporated by air flow followed by three times of 1 ml 2-propanol washes. Samples were dissolved in 100 μ l of water, and 10 μ l of each was separated by Dionex-Carbopac PA1 column and detected by PAD.

3.5.5 Characterization of pzX by NMR

For NMR analyses, approximately 1 mg of the Q-column purified pzX fraction was desalted by dialysis (1,000 MWCO), sonicated with a 1/8-inch microtip probe at amplitude of 45 (S-4000 Misonix Inc, Farmingdale, NY) for 80 cycles each of 30 sec pulse, 30 sec rest, dried by Speed-Vac, and dissolved in 200 μ l of D₂O (99.99% deuterium) supplemented with 1 μ l of 10

mM DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid). The sample was centrifuged and transferred to a 3-mm NMR tube and NMR data were recorded at 65 °C on an Agilent DD2 600 MHz NMR spectrometer equipped with a cryogenic 3 mm probe. Standard pulse sequences were used unless otherwise mentioned. Proton and carbon chemical shifts were referenced to an internal DSS peak set at 0.00 ppm for both proton and carbon spectra. The structure of pzX glycan was analyzed using one- and two-dimensional proton and carbon NMR experiments. The one-dimensional (1D) proton NMR spectrum was recorded using the water-presaturated pulse sequence and obtained with a spectral width of 6 kHz, a 90° pulse field angle (7.5 μ s), a 2.7-sec acquisition time, and a 2-sec relaxation delay (RD). The free induction decays (FIDs) were multiplied by an exponential function with a line-broadening factor of 1.5 Hz before Fourier transformation. The one-dimensional ^{13}C spectra were obtained with a spectral width of 37.9 kHz, a 45° pulse field angle (7.5 μ s), a 0.87-sec acquisition time, and a 1-sec relaxation delay (RD). The free induction decays (FIDs) were multiplied by an exponential function with a line-broadening factor of 3 Hz before Fourier transformation. In addition to ^1H and ^{13}C 1D NMR spectra, a series of homo- and heteronuclear two-dimensional NMR data sets were obtained including HMQC, HSQC, HMBC, INADEQUATE, and HMQC-TOCSY, HMQC-NOESY, HSQC-NOESY. Each experiment was modified to accommodate optimal setup to detect complex glycan structures. One-bond coupling constant was set to 170 Hz and the two-dimensional data were processed using Gaussian functions and zero-filled to a final size of 2000 $\times$ 1000. Data processing and plotting were performed using software MestreNova.

3.5.6 XNAC operon mutants

Gene knockout was achieved by double crossover recombination that designed to replace large middle portion of the target gene with chloramphenicol resistance and mKate2 cassette.

Two DNA fragments flanking bc0486 gene from *B. cereus* ATCC 14579 were PCR amplified by primer sets ZL009: 5'- ggatccgatatcgcccgacgcgaggctggatggc -3'; ZL123 5' – gaattcgatccttatctgtgccccagtttgctagg-3' and ZL405 5'- cagataaggatcgaattctgatttcgggaaataaagtaagc -3'; ZL406 5'-ggcgatatcggatccctaatgcgtgggctgcatttc-3' and they were individually cloned into pZL accepting shuttle plasmid, a derivative of pBCB13 (40) that harbors the thermosensitive ori from pMAD (41) and facilitates a cloning strategy for chromosomal integration at non-permissive temperature and subsequently a double crossover recombination event. The resulting pZL-KO-bc0486 plasmid was transformed to *E. coli* strain INV110 (Dam⁻; Dcm⁻, Invitrogen) to isolate unmethylated plasmid followed by electroporation (42) into wild type *B. cereus* ATCC 14579 competent cells. The double recombinant mutant strains were screened for chloramphenicol (5 µg/ml) resistant and erythromycin (5 µg/ml) sensitive clones and the positive Δbc0486 was further confirmed by PCR screen to validate independent double crossover events. For Δbc0486 complementation, a 1.5 kb fragment containing the Bc0486 homolog from *B. thuringiensis israelensis* ATCC 35646 was amplified by PCR and cloned by in-fusion into pDZ vector, a derivative *E.coli/Bacillus* shuttle plasmid of pDG148-stu (43). The resulting construct carrying constitutively expressed promoter Phspank, functional glycosyltransferase gene and kanamycin resistance gene and was then electro-transformed into Δbc0486 *as previously described*. Positive clones were selected on TSA plates supplemented with 5 µg/ml kanamycin.

3.5.7 Surfactant analysis of pzX

To study the effect of pzX on the growth of *Bacillus*, purified native pzX, de-O-acetylated pzX or autoclaved pzX derived from 100 ml Msgg culture of *B. cereus* was dissolved in 1 ml of sterile deionized water (DDW) and 50 µl was plated on half of Msgg (1.5% agar) plate, while 50 µl of sterile DDW was plated on the other half as control. A 2 µl drop of

vegetative cells or spores prepared from 8 h BHI or 4 days BHI culture, respectively, was spotted on both halves of the plate and the plate was incubated inverted at 30 °C for 2 days. An aliquot of 50 µl 0.1% SDS was also layered on half of Mmsg agar plate as positive control. Pictures were taken by both handheld microscope and dissecting microscope. To measure the surface tension of water in the presence or absence of purified native pzX, de-O-acetylated pzX or autoclaved pzX, a square area of 1.2 cm² of flat polystyrene surface was spread with either 10 µl of DDW or each of the pzX samples, and the liquid was allowed to dry. Various amounts of DDW drops (30, 20, 15, 5, 2 µl) were placed on the center each spread surface, and top view pictures were taken to show the diameter of droplets by handheld microscope. A side view pictures of the 30 µl of water droplets were also taken at 45- and 90-degree angle. The contact angle of a droplet was measured.

3.5.8 Adherence analysis of pzX

A replicate set of 4 pre-autoclaved vermiculites was soaked with 50 µl of *B. cereus* spores mixed with 50 µl purified pzX (autoclaved) or control. Vermiculites were allowed to air-dry for 24 h. The vermiculites were washed twice with sterile deionized water (DDW) each included a quick dip of a vermiculite 5 times into 1 ml sterile DDW. The DDW wash was discarded and the vermiculites were placed in 1 ml fresh sterile DDW before vortexed (2 min) to release remnant adhered spores. An aliquot of the water containing spores was used for serial dilutions, plated on BHI agar plates, and number CFUs were counted. For pzX titration adherence assays, different amount of pzX (0, 2, 10, 20 µg) were mixed with spores and co-dried with vermiculites prior to washing and plating. Additionally, 100 µl of 5-day Mmsg culture including spores and medium was directly co-dried with vermiculites, without any purification of spores or pzX. This adherence test was carried out with *B. cereus* WT and Δbc0486.

3.5.9 Analysis of pzX as antiaggregant

B. anthracis Sterne 34F2 and *B. cereus* ATCC 14579 spores isolated from 5 days Msgg medium were heat activated (80 °C; 20 min) in 0.2 ml of sterile DDW water. The activated spores were incubated for 15 min at 37 °C in 0.2 ml buffer (10 mM Tris-HCl pH 8, 10 mM NaCl). Subsequently, the spore suspension was supplemented with 30 µl of germinant solution (50 mM inosine dissolved in 10 mM Tris-HCl pH 8, 10 mM NaCl), and incubated at 37 °C for 20 min. The germinated spores were washed three times with sterile DDW and incubated in 0.5 ml Msgg salts (Msgg medium lacking glycerol and amino acids) that was supplemented with 20 µg pzX at 30 °C for 2 days. One set of aliquots was used for serial dilutions, and CFUs were counted *before and after germination, and 4, 8, 12, 24, 53 h post germination*. The second set of aliquots was used to examine cell morphology under phase-contrast microscope. The relative amount of intact pzX remained in the medium was also tested by alditol-acetate derivatization and GC-MS analysis.

3.5.10 Analysis of pzX at different pH

B. anthracis Sterne 34F spores isolated from 5 days Msgg medium was suspended in 0.5 ml of sterile DDW and mixed with or without 100 µg pzX. For acid condition, pH was lowered to 2 by adding 50 mM HCl to a final concentration 10 mM in 0.2 ml of solution. The samples was incubated at 30 °C for 24 hours. To test if pzX precipitates and bioencapsulates spores, the top spores/pzX suspension portion and the bottom clear portion of the sample were aliquoted and tested separately by CFU counts and phase-contrast microscopy. Subsequently samples were vortexed to mix and were plated on BHI agar plates to count the total spore numbers by CFUs.

3.6 References

1. Cherkasskiy, B. L. (1999) A national register of historic and contemporary anthrax foci. *Journal of applied microbiology* **87**, 192-195

2. Sterne, M. (1939) The use of anthrax vaccines prepared from avirulent (uncapsulated) variants of *Bacillus anthracis*. *Journal of Veterinary Science and Animal Industry* **13**, 307-312
3. Thapa, N. K., Tenzin, Wangdi, K., Dorji, T., Migma, Dorjee, J., Marston, C. K., and Hoffmaster, A. R. (2014) Investigation and Control of Anthrax Outbreak at the Human-Animal Interface, Bhutan, 2010. *Emerg Infect Dis* **20**, 1524-1526
4. Jernigan, D. B., Raghunathan, P. L., Bell, B. P., Brechner, R., Bresnitz, E. A., Butler, J. C., Cetron, M., Cohen, M., Doyle, T., Fischer, M., Greene, C., Griffith, K. S., Guarner, J., Hadler, J. L., Hayslett, J. A., Meyer, R., Petersen, L. R., Phillips, M., Pinner, R., Popovic, T., Quinn, C. P., Reefhuis, J., Reissman, D., Rosenstein, N., Schuchat, A., Shieh, W. J., Siegal, L., Swerdlow, D. L., Tenover, F. C., Traeger, M., Ward, J. W., Weisfuse, I., Wiersma, S., Yeskey, K., Zaki, S., Ashford, D. A., Perkins, B. A., Ostroff, S., Hughes, J., Fleming, D., Koplan, J. P., Gerberding, J. L., and Investi, N. A. E. (2002) Investigation of bioterrorism-related anthrax, United States, 2001: Epidemiologic findings. *Emerg Infect Dis* **8**, 1019-1028
5. Rasko, D. A., Altherr, M. R., Han, C. S., and Ravel, J. (2005) Genomics of the *Bacillus cereus* group of organisms. *FEMS microbiology reviews* **29**, 303-329
6. Van der Auwera, G. A., Feldgarden, M., Kolter, R., and Mahillon, J. (2013) Whole-Genome Sequences of 94 Environmental Isolates of *Bacillus cereus* Sensu Lato. *Genome announcements* **1**
7. Miller, J. M., Hair, J. G., Hebert, M., Hebert, L., Roberts, F. J., and Weyant, R. S. (1997) Fulminating bacteremia and pneumonia due to *Bacillus cereus*. *Journal of clinical microbiology* **35**, 504-507 correction pg. 1294
8. Avashia, S. B., Riggins, W. S., Lindley, C., Hoffmaster, A., Drumgoole, R., Nekomoto, T., Jackson, P. J., Hill, K. K., Williams, K., Lehman, L., Libal, M. C., Wilkins, P. P., Alexander, J., Tvaryanas, A., and Betz, T. (2007) Fatal pneumonia among metalworkers due to inhalation exposure to *Bacillus cereus* containing *Bacillus anthracis* toxin genes. *Clinical Infectious Diseases* **44**, 414-416
9. Klee, S. R., Ozel, M., Appel, B., Boesch, C., Ellerbrok, H., Jacob, D., Holland, G., Leendertz, F. H., Pauli, G., Grunow, R., and Nattermann, H. (2006) Characterization of *Bacillus anthracis*-like bacteria isolated from wild great apes from Cote d'Ivoire and Cameroon. *Journal of bacteriology* **188**, 5333-5344
10. Bottone, E. J. (2010) *Bacillus cereus*, a volatile human pathogen. *Clinical microbiology reviews* **23**, 382-398
11. Hoffmaster, A. R., Ravel, J., Rasko, D. A., Chapman, G. D., Chute, M. D., Marston, C. K., De, B. K., Sacchi, C. T., Fitzgerald, C., Mayer, L. W., Maiden, M. C., Priest, F. G., Barker, M., Jiang, L., Cer, R. Z., Rilstone, J., Peterson, S. N., Weyant, R. S., Galloway, D. R., Read, T. D., Popovic, T., and Fraser, C. M. (2004) Identification of anthrax toxin genes in a *Bacillus cereus* associated with an illness resembling inhalation anthrax. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 8449-8454
12. Hoffmaster, A. R., Hill, K. K., Gee, J. E., Marston, C. K., De, B. K., Popovic, T., Sue, D., Wilkins, P. P., Avashia, S. B., Drumgoole, R., Helma, C. H., Ticknor, L. O., Okinaka, R. T., and Jackson, P. J. (2006) Characterization of *Bacillus cereus* isolates associated with fatal pneumonias: strains are closely related to *Bacillus anthracis* and harbor *B. anthracis* virulence genes. *Journal of clinical microbiology* **44**, 3352-3360
13. Kolsto, A. B., Tourasse, N. J., and Okstad, O. A. (2009) What Sets *Bacillus anthracis* Apart from Other *Bacillus* Species? *Annu Rev Microbiol* **63**, 451-476
14. Girault, G., Blouin, Y., Vergnaud, G., and Derzelle, S. (2014) High-throughput sequencing of *Bacillus anthracis* in France: investigating genome diversity and population structure using whole-genome SNP discovery. *Bmc Genomics* **15**, 288
15. Derzelle, S., Girault, G., Kokotovic, B., and Angen, O. (2015) Whole Genome-Sequencing and Phylogenetic Analysis of a Historical Collection of *Bacillus anthracis* Strains from Danish Cattle. *PloS one* **10**, e0134699

16. Inglesby, T. V., Henderson, D. A., Bartlett, J. G., Ascher, M. S., Eitzen, E., Friedlander, A. M., Hauer, J., McDade, J., Osterholm, M. T., O'Toole, T., Parker, G., Perl, T. M., Russell, P. K., Tonat, K., and Biodefense, W. G. C. (1999) Anthrax as a biological weapon - Medical and public health management. *Jama-J Am Med Assoc* **281**, 1735-1745
17. Van Houdt, R., and Michiels, C. (2010) Biofilm formation and the food industry, a focus on the bacterial outer surface. *Journal of applied microbiology* **109**, 1117-1131
18. Nam, H., Seo, H. S., Bang, J., Kim, H., Beuchat, L. R., and Ryu, J. H. (2014) Efficacy of gaseous chlorine dioxide in inactivating *Bacillus cereus* spores attached to and in a biofilm on stainless steel. *International journal of food microbiology* **188**, 122-127
19. Ryu, J. H., and Beuchat, L. R. (2005) Biofilm formation and sporulation by *Bacillus cereus* on a stainless steel surface and subsequent resistance of vegetative cells and spores to chlorine, chlorine dioxide, and a peroxyacetic acid-based sanitizer. *Journal of food protection* **68**, 2614-2622
20. Gotz, F. (2002) Staphylococcus and biofilms. *Molecular microbiology* **43**, 1367-1378
21. Lemon, K. P., Earl, A. M., Vlamakis, H. C., Aguilar, C., and Kolter, R. (2008) Biofilm development with an emphasis on *Bacillus subtilis*. *Current topics in microbiology and immunology* **322**, 1-16
22. Singh, P. K., Schaefer, A. L., Parsek, M. R., Moninger, T. O., Welsh, M. J., and Greenberg, E. P. (2000) Quorum-sensing signals indicate that cystic fibrosis lungs are infected with bacterial biofilms. *Nature* **407**, 762-764
23. Fedtke, I., Gotz, F., and Peschel, A. (2004) Bacterial evasion of innate host defenses - the *Staphylococcus aureus* lesson. *Int J Med Microbiol* **294**, 189-194
24. Sutherland, I. W. (2001) The biofilm matrix - an immobilized but dynamic microbial environment. *Trends in microbiology* **9**, 222-227
25. Vlamakis, H., Chai, Y., Beauregard, P., Losick, R., and Kolter, R. (2013) Sticking together: building a biofilm the *Bacillus subtilis* way. *Nature reviews. Microbiology* **11**, 157-168
26. Epstein, A. K., Pokroy, B., Seminara, A., and Aizenberg, J. (2011) Bacterial biofilm shows persistent resistance to liquid wetting and gas penetration. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 995-1000
27. Gu, X. G., Glushka, J., Lee, S. G., and Bar-Peled, M. (2010) Biosynthesis of a New UDP-sugar, UDP-2-acetamido-2-deoxyxylose, in the Human Pathogen *Bacillus cereus* Subspecies cytotoxis NVH 391-98. *Journal of Biological Chemistry* **285**, 24825-24833
28. Branda, S. S., Gonzalez-Pastor, J. E., Ben-Yehuda, S., Losick, R., and Kolter, R. (2001) Fruiting body formation by *Bacillus subtilis*. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 11621-11626
29. Severson, K. M., Mallozzi, M., Bozue, J., Welkos, S. L., Cote, C. K., Knight, K. L., and Driks, A. (2009) Roles of the *Bacillus anthracis* spore protein ExsK in exosporium maturation and germination. *Journal of bacteriology* **191**, 7587-7596
30. Nakata, H. M. (1964) Organic Nutrients Required for Growth and Sporulation of *Bacillus Cereus*. *Journal of bacteriology* **88**, 1522-1524
31. Hsueh, Y. H., Somers, E. B., Lereclus, D., Ghelardi, E., and Wong, A. C. (2007) Biosurfactant production and surface translocation are regulated by PlcR in *Bacillus cereus* ATCC 14579 under low-nutrient conditions. *Applied and environmental microbiology* **73**, 7225-7231
32. Fedhila, S., Daou, N., Lereclus, D., and Nielsen-LeRoux, C. (2006) Identification of *Bacillus cereus* internalin and other candidate virulence genes specifically induced during oral infection in insects. *Molecular microbiology* **62**, 339-355
33. Ellwood, D. C., and Tempest, D. W. (1972) Influence of culture pH on the content and composition of teichoic acids in the walls of *Bacillus subtilis*. *Journal of general microbiology* **73**, 395-402

34. Hornstra, L. M., de Vries, Y. P., de Vos, W. M., and Abee, T. (2006) Influence of sporulation medium composition on transcription of ger operons and the germination response of spores of *Bacillus cereus* ATCC 14579. *Applied and environmental microbiology* **72**, 3746-3749
35. Cliff, J. B., Jarman, K. H., Valentine, N. B., Golledge, S. L., Gaspar, D. J., Wunschel, D. S., and Wahl, K. L. (2005) Differentiation of spores of *Bacillus subtilis* grown in different media by elemental characterization using time-of-flight secondary ion mass spectrometry. *Applied and environmental microbiology* **71**, 6524-6530
36. Purohit, M., Sassi-Gaha, S., and Rest, R. F. (2010) Rapid sporulation of *Bacillus anthracis* in a high iron, glucose-free medium. *Journal of microbiological methods* **82**, 282-287
37. Clements, M. O., and Moir, A. (1998) Role of the gerI operon of *Bacillus cereus* 569 in the response of spores to germinants. *Journal of bacteriology* **180**, 6729-6735
38. York, W. S., Darvill, A. G., McNeil, M., Stevenson, T. T., and Albersheim, P. (1986) Isolation and Characterization of Plant-Cell Walls and Cell-Wall Components. *Methods in enzymology* **118**, 3-40
39. Anumula, K. R., and Taylor, P. B. (1992) A Comprehensive Procedure for Preparation of Partially Methylated Alditol Acetates from Glycoprotein Carbohydrates. *Analytical biochemistry* **203**, 101-108
40. Pereira, P. M., Veiga, H., Jorge, A. M., and Pinho, M. G. (2010) Fluorescent reporters for studies of cellular localization of proteins in *Staphylococcus aureus*. *Applied and environmental microbiology* **76**, 4346-4353
41. Arnaud, M., Chastanet, A., and Debarbouille, M. (2004) New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, gram-positive bacteria. *Applied and environmental microbiology* **70**, 6887-6891
42. Turgeon, N., Laflamme, C., Ho, J., and Duchaine, C. (2006) Elaboration of an electroporation protocol for *Bacillus cereus* ATCC 14579. *Journal of microbiological methods* **67**, 543-548
43. Joseph, P., Fantino, J. R., Herbaud, M. L., and Denizot, F. (2001) Rapid orientated cloning in a shuttle vector allowing modulated gene expression in *Bacillus subtilis*. *FEMS microbiology letters* **205**, 91-97

CHAPTER 4
**A FOUR-GENE OPERON IN BACILLUS CEREUS PRODUCES TWO RARE SPORE-
DECORATING SUGARS³**

³ Li, Z., Mukherjee, T., Bowler, K., Namdari, S., Rono, J., Prestridge, S., Carlton, A. and Bar-Peled, M.
Accepted by *The Journal of biological chemistry*.
Reprinted here with permission of publisher.

4.1 Abstract

Bacterial glycan structures on cell surfaces are critical for cell-cell recognition and adhesion and in host-pathogen interactions. Accordingly, unraveling the sugar composition of bacterial cell surfaces can shed light on bacterial growth and pathogenesis. Here, we found that two rare sugars with a 3-C-methyl-6-deoxy-hexose structure were linked to spore glycan in *Bacillus cereus* ATCC 14579 and ATCC 10876. Moreover, we identified a four-gene operon in *B. cereus* ATCC 14579 that encodes proteins with the following sequential enzyme activities as determined by mass spectrometry and one- and two-dimensional NMR methods: CTP:glucose-1-phosphate cytidylyltransferase, a CDP-Glc 4,6-dehydratase, an NADH-dependent SAM:C-methyltransferase, and an NADPH-dependent CDP-3-C-methyl-6-deoxy-hexose 4-reductase. The last enzyme predominantly yielded CDP-3-C-methyl-6-deoxy-gulose (CDP-cereose) and likely generated a 4-epimer, CDP-3-C-methyl-6-deoxy-allose (CDP-cillose). Some members of the *B. cereus sensu lato* group produce CDP-3-C-methyl-6-deoxy sugars for the formation of cereose-containing glycans on spores, while others such as *Bacillus anthracis*, do not. Gene knockouts of the *Bacillus* C-methyltransferase and the 4-reductase confirmed their involvement in the formation of cereose-containing glycan on *B. cereus* spores. We also found that cereose represented 0.2%–1% spore dry weight. Moreover, mutants lacking cereose germinated faster than the wild type, yet the mutants exhibited no changes in sporulation or spore resistance to heat. The findings reported here may provide new insights into the roles of the uncommon 3-C-methyl-6-deoxy-sugars in cell surface recognition and host-pathogen interactions of the genus *Bacillus*.

4.2 Introduction

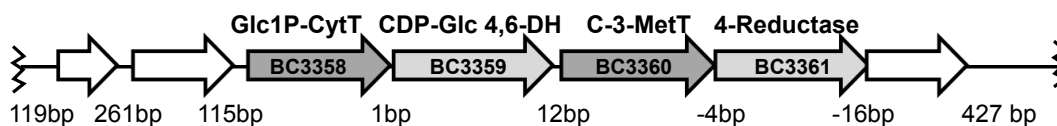
Bacterial cells produce myriad glycan structures that are secreted or attached to cell surfaces. Often, being at the outermost surface of the cells, glycans have significant impact on cell-cell recognition, cell-to-cell adhesion and in host-pathogen interaction (1-4). Glycans also have a profound role in shaping the cell (5). In recent years it was also recognized that spore-producing bacteria like members of *Bacillus cereus sensu lato* group are decorated with glycan structures (6,7). While very few of these glycans were studied, they were found as glycoproteins of the spore exosporium, the outer most surface layer surrounding the spore coat. Among the exosporium glycoproteins that were characterized are ExsH and the collagen-like protein, BclA (8,9). The BclA protein plays a central role in *B. anthracis* pathogenesis by promoting interaction of spores with the host phagocytic cell. This was proposed to facilitate transport of the spores to sites of spore germination and bacterial outgrowth (10,11). These two glycoproteins characterized in the *B. cereus* ATCC 14579 exosporium have different sugar composition and glycosylation patterns (12). The reason for the presence of several surface glycoproteins and each being differently glycosylated on *B. cereus* ATCC 14579 exosporium remains unknown, but various glycosylation patterns could suggest specific recognition to different cell receptors.

Genome sequence analyses have revealed the presence of at least 13 different collagen-like protein-encoding genes and proteomic studies indicate some of them being in the exosporium (13). Whether these are also glycoproteins with glycans that differ in their sugar composition and sequence remains unknown. This feature is critical to study because spore glyco-epitopes could explain host recognition and strain specificity. Thus, in a systematic screen to identify sugar-containing polymers in spores, we have identified by GS-MS analyses new unknown sugar-like residues in spore glycans.

FIGURE 1

A

Bacillus cereus ATCC 14579



B

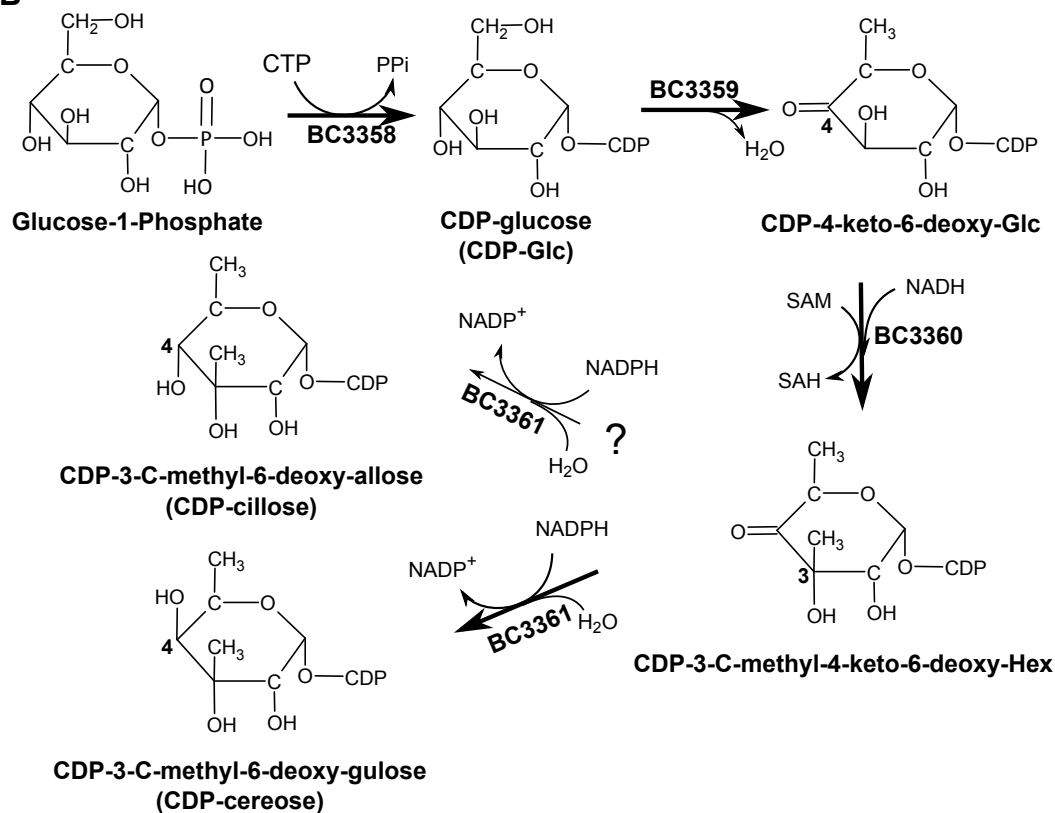


FIGURE 4.1. A proposed biosynthetic step for the formation of CDP-cillose and CDP-cereose in *Bacillus cereus* ATCC 14579. Panel **A**. Organization of the four-genes in the C3CM operon. Panel **B**. Bc3358 converts Glc-1-P and CTP to CDP-glucose. Bc3359, a 4,6-dehydratase, converts CDP-glucose to CDP-4-keto-6-deoxy-glucose. Bc3360, a C3-methyltransferase, converts CDP-4-keto-6-deoxy-glucose to CDP-3-C-methyl-4-keto-6-deoxy-hexose in the presence of NADH and SAM. Bc3361 is a 4-reductase converting predominantly CDP-3-C-methyl-4-keto-6-deoxy-hexose to CDP-3-C-methyl-6-deoxy-gulose (CDP-cereose), and likely at a much reduced amount, to CDP-3-C-methyl-6-deoxy-allose (CDP-cillose).

Here we provide the evidence of two rare sugars, 3-C-methyl-6-deoxy-hexoses, that are displayed on *Bacillus* spores. We showed that the activated forms of the sugars are CDP-3-C-methyl-6-deoxy-hexose isomers as illustrated in Fig. 4.1. The formation of the C-methyl-sugar-

containing glycan on spore surface occurs when cells are induced to grow in a sporulation medium.

4.3 Results

4.3.1 Spores are decorated with two rare sugars, 3-C-methyl-6-deoxy-hexoses

Subtle differences among glycoconjugate structures were proposed to facilitate specific interaction of a pathogen with its host. Hence, one important task is to determine all sugar residues that decorate a glycan. This becomes difficult when different growth environments influence glycan composition. Furthermore, some sugar residues are labile to the common hydrolytic harsh conditions (high temperature and high concentration of acid).

In this study, we have examined sugar composition under different hydrolytic conditions using *Bacillus cereus* ATCC 14579 cells. In addition, *B. cereus* was grown in different media because its glycan do change to altered medium (14). This approach led us to detect two unknown sugar peaks (labeled B and C, see Fig. 4.2A) by GC-MS analyses. The EI/MS spectra of B and C showed identical fragment ions pattern suggesting they are isomers. Calculations of the potential structure of the alditol-acetate derivatives implied that B and C are sugar residues predicted to be 3-C-methyl-6-deoxy-hexoses. The GC chromatogram and MS fragmentation pattern of such structures were not reported in *Bacillus*. The peaks at m/z 245 and 231 (Fig. 4.2A) suggest a cleavage between C3-C4. The peaks at m/z 189, 129, 87 and m/z 203, 143 are likely secondary ion fragments due to loss of m/z 42 (ketene) or 60 (acetate) from m/z 231 and 245, respectively. Interestingly, hydrolyses of samples at very low molarity of TFA, HCl, acetic acid gave different sugar composition profile. However, no peaks for B and C sugar residues were observed when samples were hydrolyzed with 2 or 4 M TFA suggesting they are labile.

FIGURE 2

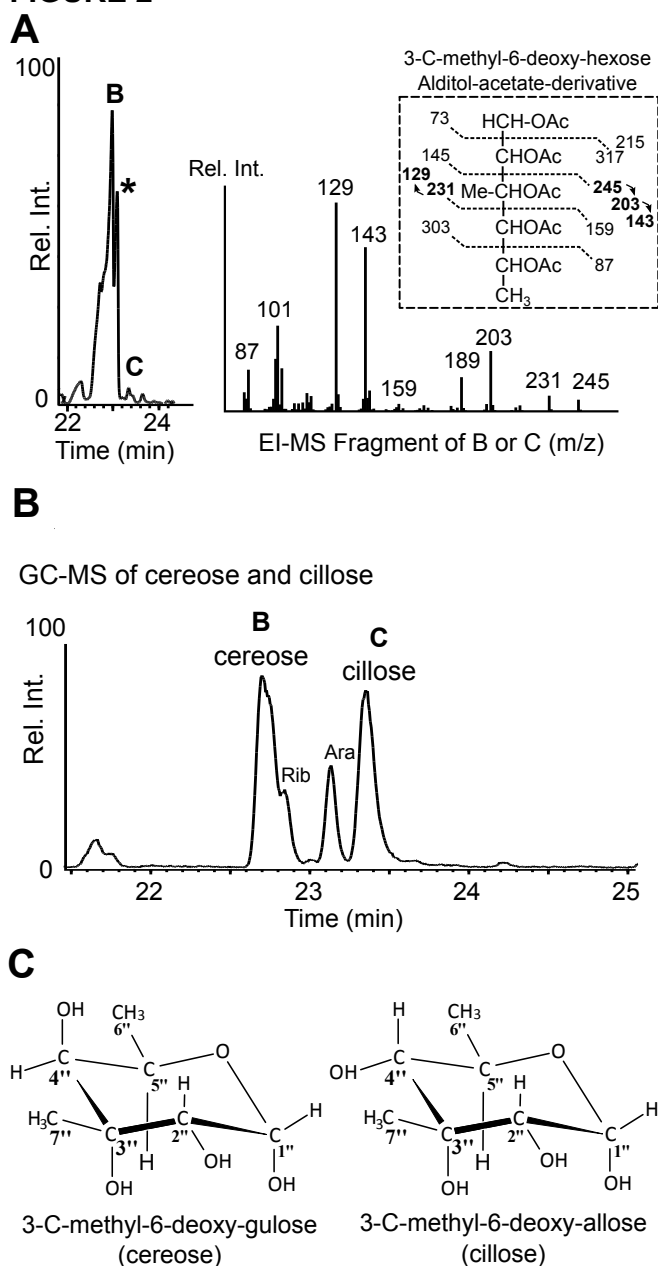


FIGURE 4.2 Two uncommon 3-C-methyl-6-deoxy-hexose sugar residues were detected upon mild hydrolysis of spore glycan from *Bacillus*. Panel **A**. GC-MS of alditol-acetate derivatives of sugars released by mild acid hydrolyses from spore glycan of *B. cereus* ATCC 14579 grown in Msgg medium. The two novel sugar peaks are eluting at 22.7 and at 23.3 min and labeled ‘B’ and ‘C’, respectively. Peak ‘*’ is 6-deoxy hexose. The EI-MS fragmentation pattern of peak ‘B’ and ‘C’ with their major *m/z* ion fragments is shown in the right spectrum; and the boxed insert shows the predicted primary and secondary MS fragments (labeled in bold) of CDP-3-C-methyl-6-deoxy-hexose alditol-acetate derivative. Panel **B**. GC-MS analyses of cereose and cillose standards depict that peak B is cereose; and peak C is cillose. Standards were obtained from mild hydrolysis of corresponding CDP-3-C-methyl-sugars purified from in-microbe based assay (see Fig. 4.4 for CDP-cillose); and after enzymatic assay (see Fig. 4.8 for CDP-cereose). Alditol acetate sugar derivatives peaks of ribose and arabinose are also shown. Panel **C**. Structures of cereose and cillose.

4.3.2 C3CM operon harbors a biosynthetic pathway that includes a methyltransferase to form CDP-3-C-methyl-6-deoxy-hexoses

Because sugar B and C were unknown deoxy-sugars, possibly uncommon, with a potential unique 3'-C modification, we predicted that their formation would require an enzyme having a capability to C-methylate its NDP-sugar precursor. We therefore, used the amino-acid sequence of functional C-3'-methyltransferase from *Micromonospora chalybeata* (15) to BLAST against translated genome of *B. cereus* ATCC 14579 and this led us to identify a potential gene, BC3360, encoding a protein with 31% amino-acid sequence identity. The *M. chalybeata* C-3'-methyltransferase is involved in the formation of dTDP-3-amino-2,3,6-trideoxy-4-keto-3-methyl-D-glucose (dTDP-tetronitrose) (15,16). This unusual C3-methylated amino deoxysugar (tetronitrose), is incorporated from its precursor, dTDP-tetronitrose, to a molecule with anti-bacterial property, kijanimicin (17), and to an antitumor agent, tetrocarcin A (16). While 31% sequence identity is not high (with expect value $1e^{-66}$), we were encouraged to pursue this *Bacillus* homologous protein as a candidate methyltransferase (MetT) because it harbors a few conserved amino acids and domains shared with the *M. chalybeata* C-MetT. Interestingly, flanking the putative *Bacillus* C-methyltransferase are three additional genes predicted to be involved in NDP-sugar synthesis as well (Fig. 4.1A). However, rather than dTDP-biosynthetic genes, as in *M. chalybeata*, the genes flanking the *Bacillus* BC3360 were annotated as CDP-sugar biosynthetic genes.

A proposed biosynthetic pathway of this operon is shown in figure 4.1B. The C3CM operon encodes the following enzymes based on their specific activities that will be described in details in the succeeding sections: CTP:glucose-1-phosphate cytidylyltransferase (Bc3358, Glc1P-CytT), CDP-glucose 4,6-dehydratase (Bc3359, 4,6-dehydratase); SAM:CDP-4-keto-6-

deoxy-glucose C3-methyltransferase (Bc3360, C3-MetT). Subsequently, a CDP-3-C-methyl-4-keto-6-deoxy-hexose 4-reductase forms predominantly the final product CDP-3-C-methyl-6-deoxy-gulose.

4.3.3 In-microbe analyses of genes of the C3CM operon in *E. coli* show synthesis of CDP-3-C-methyl-6-deoxy-allose

Of the two experimental approaches used to determine the function of genes within the C3CM operon, the first one is named “in-microbe”. The in-microbe experimental setup consisted of a combination of up to four recombinant genes of the C3CM operon that were co-transformed into *E. coli*. After IPTG-gene induction, the NDP-sugars were directly extracted from the cell and separated by HILIC-column. Chromatographic peaks were further purified and analyzed by NMR. *E. coli* induced to express the first gene BC3358 gave a distinct peak (Fig. 4.3A i) with an ion $[M-H]^-$ at m/z 564 that gave MS/MS ion fragments at m/z 322 and 241 (Fig. 4.3B i) which are consistent with CDP-hexose, CMP $[M]^-$, and sugar-1-P $[M-H_2O]^-$, respectively. The m/z 564 was not found in control *E. coli* expressing empty plasmid (Fig. 4.3A v). Subsequently, the *E. coli* (Bc3358)-produced NDP-sugars were purified by HPLC and analyzed by NMR spectroscopy. The NMR data confirmed the m/z 564 peak is CDP-Glc (supplemental Fig. 4.S1). The chemical shift assignments for CDP-Glc are summarized in Table 4.1. The large $J_{H2'', H3''}$ and $J_{H3'', H4''}$ coupling constants of 9.9 Hz and 9.8 Hz, respectively, are consistent with a sugar in *gluco*-configuration. The chemical shift of 5.59 ppm of the anomeric proton and $J_{H1'', H2''}$ coupling constants of 3.4 Hz are consistent with a α -linkage. This suggests that enzyme Bc3358 is a CTP:glucose-1-phosphate cytidylyltransferase.

E. coli induced to express both BC3358 and BC3359 gave a broad peak eluting from HILIC column between 16.5 to 18 min. This peak was characterized as CDP-4-keto-6-deoxy-

glucose (CDP-4k-6d-Glc) (Fig. 4.3A ii) with an ion $[M-H]^-$ at m/z 546 that gave by MS/MS an ion fragment of m/z 322 (Fig. 4.3B ii). This suggests that BC3359 encodes a CDP-glucose 4,6-dehydratase. Similarly, *E. coli* induced to express three *Bacillus* genes BC3358, BC3359 and BC3360 produced yet another broad peak with retention time (14-15.5 min) distinct from the previous time shown in Figure 4.3A iii and an ion $[M-H]^-$ at m/z of 560 (Fig. 4.3B iii) that gave MS/MS ion fragments of m/z 322 and 237. The empty vector control yielded no m/z 560 (Fig. 4.3A v). This increase of 14 amu indicates that the combined three *Bacillus* genes yielded a C-methylated CDP-4-keto-6-deoxy-glucose. The MS/MS of the product (m/z 560) gave an ion fragment at m/z 237 (Fig. 4.3B iii), which is consistent with a methyl modification of 4-keto sugar-1-P $[M-H_2O]^-$. Hence, we tentatively suggest that BC3360 encodes a CDP-4-keto-6-deoxy-glucose C3-methyltransferase. Lastly, when all four genes were co-expressed in *E. coli*, a new peak observed (Fig. 4.3A iv) with an ion $[M-H]^-$ at m/z 562 which gave MS/MS ion fragments of m/z 322 and 239 (Fig. 4.3B iv). This increase of 2 amu indicates that the 4-keto moiety of the substrate is reduced. Such peak was not detected in *E. coli* control. Hence, the last gene (BC3361) when expressed in *E. coli* suggested encoding a CDP-3-C-methyl-4-keto-6-deoxy-hexose 4-reductase.

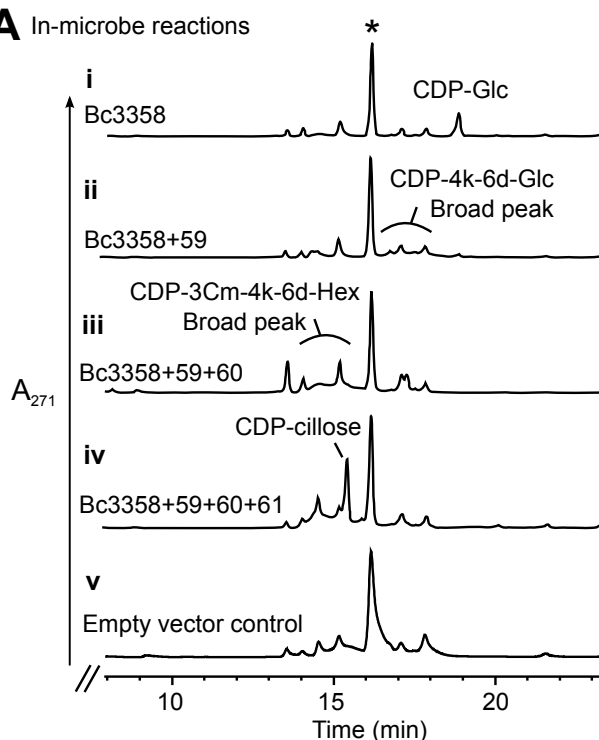
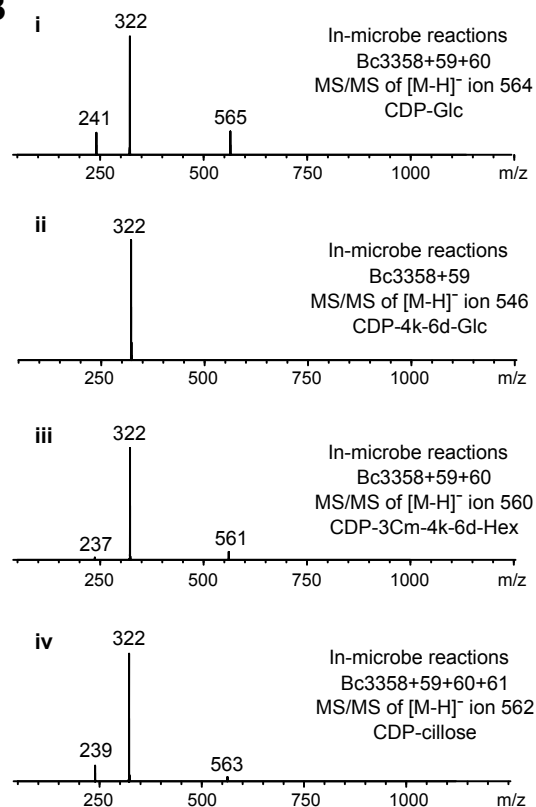
FIGURE 3**A** In-microbe reactions**B**

FIGURE 4.3. In-microbe analyses of the four genes encoded by the C3CM operon. HILIC chromatograms of extracts isolated from *E. coli* cells overexpressing combinations of up to four genes are shown in Panel **A**. *E. coli* expressed a single gene, BC3358 (see **i**) produces CDP-glucose; *E. coli* expressed the combined BC3358+BC3359 genes (**ii**) produces CDP-4-keto-6-deoxy-glucose; *E. coli* expressed the combined BC3358+BC3359+BC3360 genes (**iii**) produces a 3-C-methylated CDP-4-keto-6-deoxy-hexose; *E. coli* expressed the combined four genes (**iv**) produces a new UV-peak, later assigned as CDP-cillose; *E. coli* expressed an empty vector control (**v**) produces no such peaks. A_{271} is the maximum absorption for cytosine-based molecule. The large peak (16 min) labeled * indicates internal reference. Panel **B**. LC-MS/MS analyses of each in-microbe reaction product. (**i**) CDP-glucose with m/z of $[M-H]^-$ 564 and its MS/MS fragment products; the combined BC3358+BC3359 (**ii**) produces CDP-4-keto-6-deoxy-glucose with m/z of $[M-H]^-$ 546 and its MS/MS fragment products; the combined BC3358+BC3359+BC3360 (**iii**) produces a 3-C-methylated CDP-4-keto-6-deoxy-hexose glucose with m/z of $[M-H]^-$ 560 and its MS/MS fragment product; the combined four genes (**iv**) yields CDP-3-C-methyl-6-deoxy-allose with m/z of $[M-H]^-$ 562 and its MS/MS fragment product. The control empty vector had no such mass.

FIGURE S1

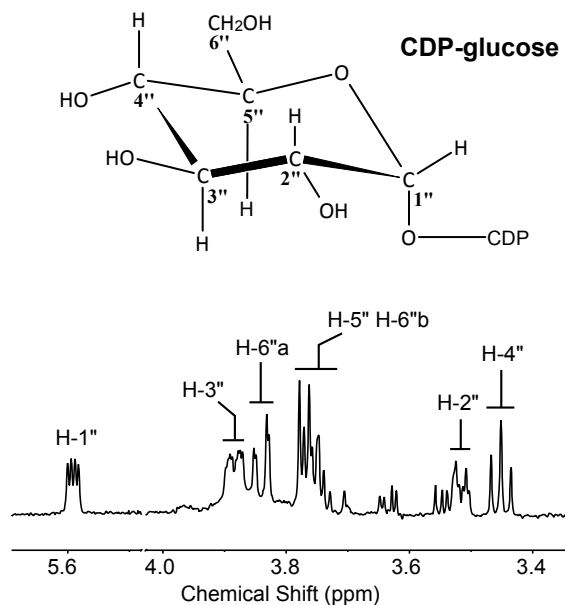


FIGURE 4.S1. Proton NMR analyses of CDP-glucose peak eluting from HILIC-column at 18.5 min in Fig. 4.3 Ai. The chemical shift is shown on Table 4.1.

Table 4.1 chemical shifts of CDP-cillose, CDP-cereose and CDP-glucose

CDP-cillose	H	¹³ C	J coupling
1	5.52	97.70	J _{1,p} =7.18
2	3.56	73.33	J _{1,2} =3.79
3	--	76.28	--
4	3.15	77.69	J _{4,5} =10.10
5	4.08	67.87	J _{5,6} =5.37
6	1.26	19.46	--
7	1.29	23.40	--
CDP-cereose	H	¹³ C	J coupling
1	5.54	98.07	J _{1,p} =7.26
2	3.67	70.67	J _{1,2} =3.68
3	--	75.80	--
4	3.36	77.89	J _{4,5} <1
5	4.56	67.10	J _{5,6} =6.72
6	1.19	18.23	--
7	1.31	25.45	--

CDP-Glucose	H	J coupling
1	5.59	$J_{1,p}=7.1$
2	3.52	$J_{1,2}=3.4$
3	3.88	$J_{2,3}=9.9$
4	3.45	$J_{3,4}=9.8$
5	3.76	$J_{4,5}=9.8$
6a	3.84	$J_{5,6a}=2.1$
		$J_{5,6b}=6.4$
6b	3.75	$J_{6a,b}=12.5$

4.3.4 One- and two-dimensional NMR reveal the chemical structure of CDP-3-C-methyl-6-deoxy-allose

Column purification of the final four-gene product (CDP-3-C-methyl-6-deoxy-hexose) produced by the in-microbe experiment (see peak point by the arrow in Fig. 4.3A iv) was analyzed by NMR spectroscopy (Fig. 4.4A). One-dimensional NMR spectrum (Fig. 4.4B) shows protons H5, H6 belonging to Cyt base ring; anomeric signals belonging to ribose H1' (5.98 ppm) and 3-C-methyl-6-deoxy-allose H1'' (5.52 ppm), as well as ring protons (3 to 4.5 ppm) and diagnostic H5'' proton of 6-deoxysugar. In addition, two distinct signals in the methyl region were observed: the first peak (H7'', 1.29 ppm) is a singlet and was later assigned to the methyl group attached to C3''; the second peak is a doublet and later assigned to the methyl group at C6''. The singlet peak suggested no proton is connected to the C3'' on the sugar ring, while the doublet peak (i.e. split) at H6'' is the result of coupling to H5''. If the C3''-methyl was O-linked to C3'' (like in methanol) we would expect a H3'' signal and a doublet signal of H7'' due to the split of H3''.

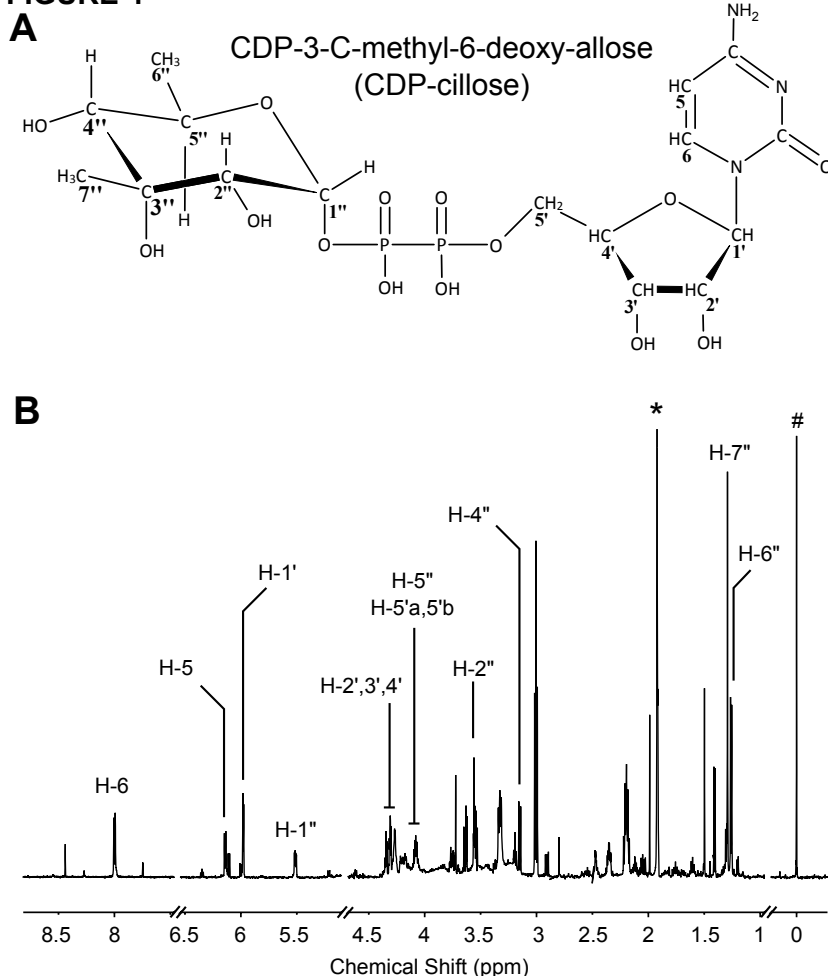
FIGURE 4

FIGURE 4.4. The combined *Bacillus cereus* BC3358+BC3359+BC3360+BC3361 genes expressed in *E. coli* produce CDP-3-C-methyl-6-deoxy-allose: CDP-cillose. Panel **A**. Structure of CDP-cillose. The specific numbered protons (H) and carbons (C) are labeled. Panel **B**. ^1H NMR spectrum of CDP-cillose with selective peak proton signals. Spectrum was cropped to fit the window. Signals labeled with '*' and '#' are ammonium formate and DSS, respectively.

Two-dimensional COSY experiment (Fig. 4.5A) shows the connectivity between H1'' to H2'', H4'' to H5'' and H5'' to H6'' on the cillose ring. The absence of a proton connected to C3'' position and the absence of connectivity of methyl H7'' to any adjacent protons suggests that this methyl group is connected via carbon at 3'' position. In addition, HBMC experiment (Fig. 4.5B) also indicates the connectivity of methyl H6'' to C4'' and C5''; and methyl H7'' to C2'', C3'' and C4'' suggesting a sugar ring with 3-C-methyl-6-deoxy-allose configuration.

FIGURE 5

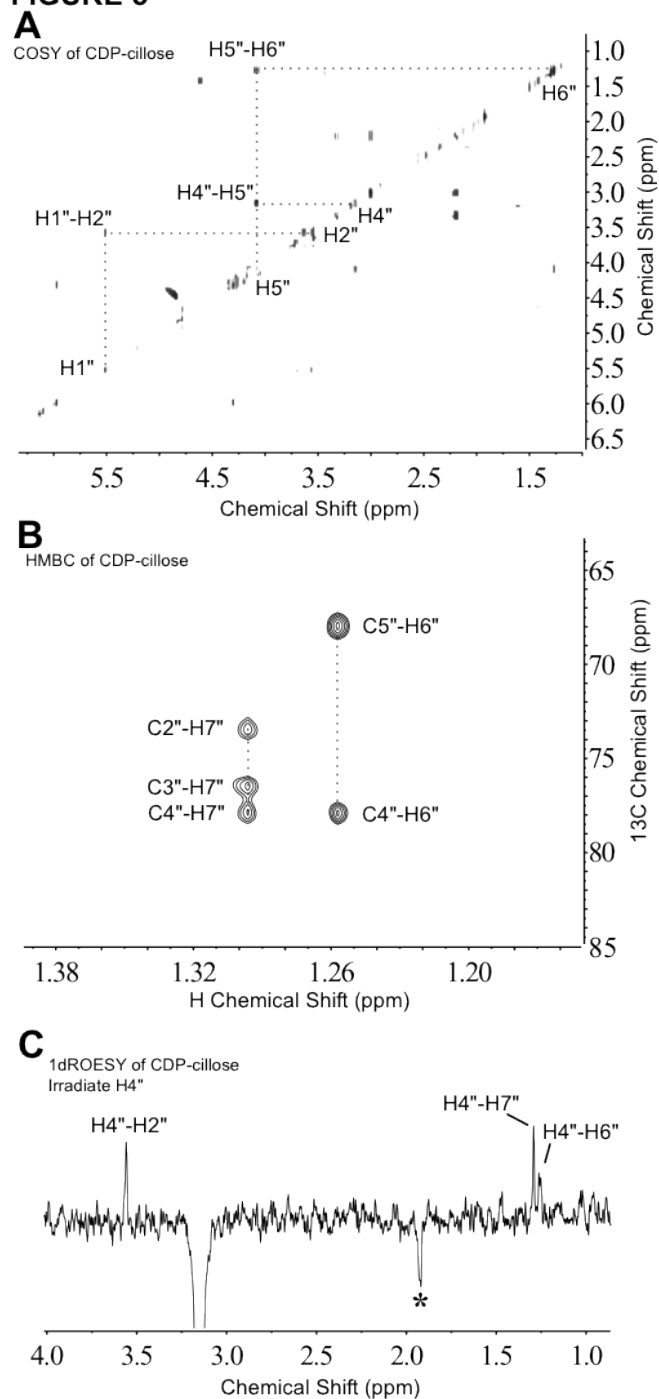


FIGURE 4.5. Detailed NMR characterization of CDP-cillose derived by in-microbe HPLC purified reaction product. Panel **A**. COSY experiment showing the connectivity between protons in the sugar ring moiety that are two to three bonds apart. Note that there is no proton connected through H2'' and H4'' indicating no proton linked to C3''. Panel **B**. selected spectral region of HMBC experiment showing protons on the methyl region connected to adjacent carbons that are two to three bonds away. Note, the methyl group (H7'') is connected to C3'' indicating cillose is C3-methylated. Panel **C**. One-dimensional ROESY experiment with H4'' irradiated. Dipolar couplings (NOE effect) are observed between H4'' to H2'' also indicating allose-configuration.

Complete determination of the stereoisomerism of this nucleotide sugar came by measuring the coupling constants in one-dimensional NMR experiments. The large coupling constant $J_{4,5}=10.1\text{Hz}$ between H4'' and H5'' (Table 4.1) strongly supports allose-configuration and not a gulose-configuration. The distinct chemical shift of the anomeric proton H1'' and the J coupling constant value of 3.79 Hz for $J_{\text{H1'', H2''}}$ are consistent with α -linkage to the phosphate moiety of CDP. Additionally an one-dimensional ROESY experiment was performed (Fig. 4.5C). When H4'' is irradiated, dipolar couplings (NOE effect) between H4'' to H2'', H4'' to H6'' and H4'' to H7'' was observed confirming allose-configuration. Taken together, this nucleotide sugar is characterized as CDP-3-C-methyl-6-deoxy-allose, and we named it CDP-cillose.

In addition to NMR analysis, we also utilized purified CDP-cillose sample as standard for GC-MS analysis. When the CDP-cillose standard was hydrolyzed, derivatized to its alditol-acetate and analyzed by GC-MS, the retention time, EI/MS ion and fragments were identical to the sugar peak "C" in Fig. 4.2B.

4.3.5 *In vitro* assays of C3CM operon reveal that the combined enzymatic product is CDP-3-C-methyl-6-deoxy-gulose

While the first method provided evidence that the four-*Bacillus* genes expressed in *E. coli* produced CDP-cillose, the second approach based on *in vitro* enzymatic assays suggests that the last step gives a different product with altered sugar configuration, as described below. In the second method, each individual recombinant protein was expressed and purified by affinity column (supplemental Fig. 4.S2). Our biochemical data confirm that recombinant His₆-Bc3358 is indeed a CTP: Glc-1-P cytidylyltransferase as determined by LC-MS/MS (Fig. 4.6A) while the control empty-vector gave no product. During the initial characterization of the recombinant enzyme, we determined that the activity requires magnesium, Glc-1-P, and CTP as no activity

was observed without the metal. The enzyme is specific to CTP as a nucleotide because ATP, dTDP, GTP are not substrates under the conditions described under “Experimental procedures”; however a less than 1% conversion was observed with UTP. Gal-1-P is not a substrate. MS analyses of the enzymatic peak product yielded an ion at m/z 564 and MS/MS ion fragment of m/z 322 that is consistent with [CDP-Glc-H] $^-$ and [CMP-H] $^-$, respectively. As CDP-Glc is commercially unavailable, we purified a large amount of it using HILIC column and used it for further characterization of the subsequent enzymes.

FIGURE 6

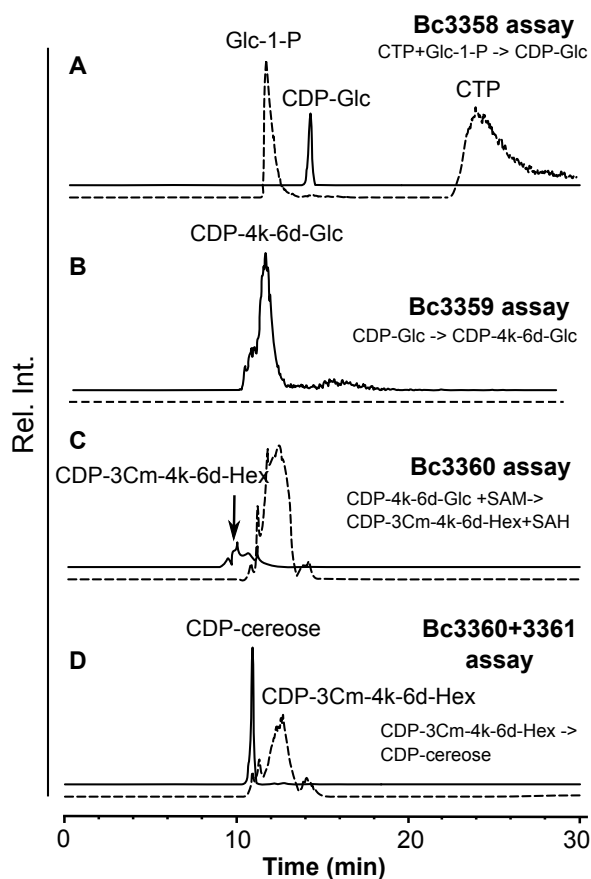


FIGURE S2

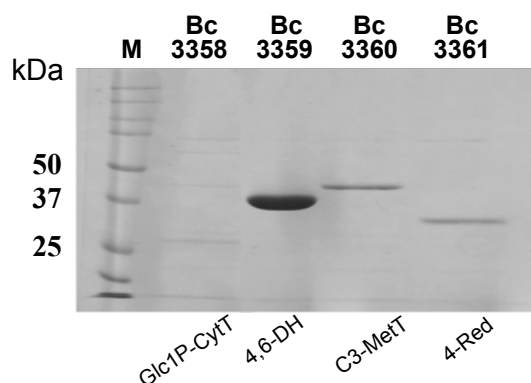


FIGURE 4.S2. SDS-PAGE analysis of each of the C3CM operon biosynthetic enzymes after nickel column purification. Protein standards (M) are shown on the left in kDa. Lane1, Bc3358 (28.8 kDa) CTP:glucose-1-phosphate cytidyltransferase (Glc1P-CytT); lane2, Bc3359 (39.4 kDa) 4,6-dehydratase (4,6-DH); lane3, Bc3360 (46.7 kDa) C-3-methyltransferase (C3-MetT); lane4, Bc3361 (34.7 kDa) 4-reductase (4-Red).

The second enzyme is a specific 4,6-dehydratase. The recombinant enzyme, His₆-Bc3359, readily converted CDP-Glc to a CDP-4-keto-6-deoxy-glucose, while a control reaction yielded no product. Fig. 4.6B shows LC-MS/MS analyses of the activity of Bc3359, where the enzymatic products (labeled CDP-4k-6d-Glc, refer to CDP-4-keto-6-deoxy-Glc) eluted from the HILIC column as a broad peak between 10 and 12 min and with ion $[M-H]^-$ at m/z 546. MS/MS analysis of this peak yielded a fragment of 322 m/z consistent with CMP. To gain further insight into Bc3359 4,6-dehydratase enzyme activity, we monitored the reaction by time-resolved ¹H NMR. The anomeric signal at 5.59 ppm (Fig. 4.7 labeled G, H-1''; see 0-min time point) indicates the chemical shift of the substrate, CDP-Glc. As the enzymatic reaction progressed two new products appeared, each with a quartet of peak signals at 5.56 and 5.74 ppm (Fig. 4.7 labeled K, H-1''' and H, H-1'') in the anomeric region of the NMR spectrum, whereas the H-1'' of the substrate disappeared. Product K is CDP-4-keto-6-deoxy-Glc and H is the hydrated form of K. The H-4'' proton signal of the substrate (labeled C, H-4'') decreased over time reflecting the disappearance of H-4'' when forming 4-keto structure product. Two signals also appeared at the

6-deoxy regions around 1.20 and 1.30 ppm. These signals correspond to C6-methyl protons (H-6'') of products K and H sugar moieties, respectively. The collective MS/MS and NMR data provide evidence that Bc3359 encodes a CDP-Glc 4-6-dehydratase. Further assays were conducted to determine substrate specificity, but dTDP-Glc and UDP-GlcNAc were not substrates, however, less than 1% of UDP-Glc was converted to UDP-4-keto-6-deoxy-Glc as evident by MS/MS.

FIGURE 7

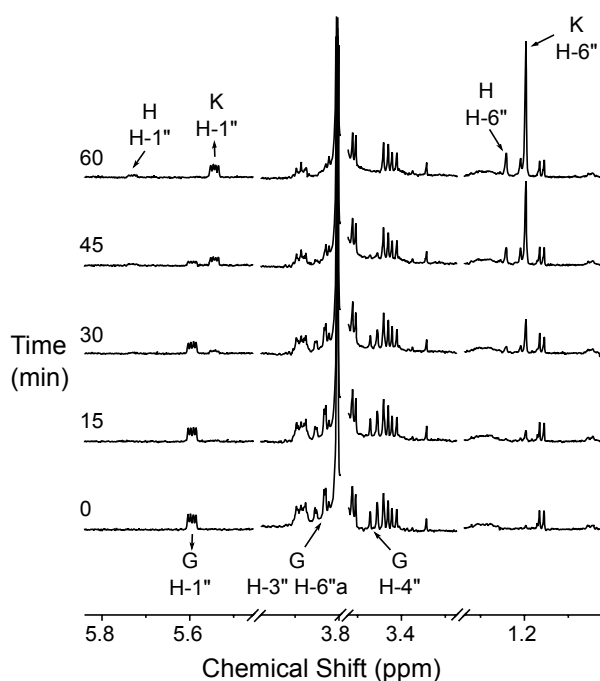


FIGURE 4.7 Time-resolved ^1H NMR showing that recombinant Bc3359 is a CDP-Glc 4,6-dehydratase. The enzymatic conversion of CDP-Glc (G) to CDP-4-keto-6-deoxy-Glc (K) and its hydrated form (H) is shown over time by proton NMR. The chemical shift regions of the proton NMR spectrum that are diagnostic for the H-1'' anomeric protons of enzymatic reactant (G) and products (H and K) is shown between 5.5 and 5.8 ppm. The diagnostic signals for the three methyl protons (H6'') peaks of product K and H are shown between 1.1 and 1.3 ppm.

The next enzyme activity was found as SAM:CDP-4-keto-6-deoxy-hexose C3-methyltransferase (abbr. C3-MetT). Preliminary assays established that the recombinant *Bacillus* Bc3359 C3-MetT requires magnesium and NADH for activity and is SAM-dependent. The purified recombinant C3-MetT (His₆-Bc3360) was reacted with CDP-4-keto-6-deoxy-glucose

(the reaction product of Bc3359). A new broad peak with a retention time between 9.5-10.5 min (Fig. 4.6C) was detected while a control reaction gave no product. The mass spectrum showed an ion $[M-H]^-$ at m/z 560 proposing that the product was also CDP-4-keto-sugar, but the increase of 14 amu when compared with the substrate suggested that the product gained a methyl group. Hence, the product of Bc3360 is CDP-3-C-methyl-4-keto-6-deoxy-hexose. Interestingly, the activity of the C3-MetT was relatively low and only ~10% conversion was obtained. However, adding the last enzyme in the pathway (4-reductase, Bc3361) drove the reaction further to produce a new product. Figure 4.6D shows the enzymatic assay of combined C3-MetT and 4-reductase. Both purified recombinant C3-MetT (His₆-Bc3360) and 4-reductase (His₆-Bc3361) readily reacted with CDP-4-keto-6-deoxy-hexose to yield a final product while negative control reactions gave no product. A peak with a retention time at 11 min (Fig. 4.6D) was detected having an ion $[M-H]^-$ at m/z 562. MS/MS analysis of this ion gave fragments at m/z 322 and 239, consistent with CMP and sugar-1-P derivatives for CDP-3-C-methyl-6-deoxy-hexose.

When substituting Mg^{2+} for Zn^{2+} or Mn^{2+} in Bc3360 assay, no C3-MetT activity detected. When omitting NADH no C3-MetT activity and only low activity observed when NADPH was substituted for NADH or when pyridoxal phosphate (PLP) was used. Lastly, the C3-MetT activity requires SAM as the methyl donor with the highest activity at 22 °C with little activity at 37 °C. Further support for the *in vitro* C3-MetT specific activity came from in-microbe experiment. Engineered *E. coli* cells co-expressing the C3-MetT gene with either a gene encoding an enzyme that forms dTDP-4-keto-6-deoxy-Glc, UDP-4-keto-6-deoxy-Glc or UDP-4-keto-6-deoxy-N-acetylhexosamine were not able to methylate these derivatives. The last enzyme in the pathway is a 4-reductase. It has the highest activity with NADPH rather NADH. In the in-microbe experiments, the 4-reductase activity also appeared specific with no activities towards

dTDP-4-keto-6-deoxy-Glc, UDP-4-keto-6-deoxy-Glc or UDP-4-keto-6-deoxy-N-acetylhexosamine.

4.3.6 One- and two-dimensional NMR reveal the chemical structure of CDP-3-C-methyl-6-deoxy-gulose

Column purified product (product peak in Fig. 4.6D) of the *in vitro* combined enzymatic assays of His₆-Bc3360, His₆-Bc3361 was analyzed by NMR spectroscopy. One-dimensional NMR spectra (Fig. 4.8) showed anomeric signals belonging to the H1'' peak (5.54 ppm), ring protons H2'' and H4'' (3.25 to 3.75 ppm) and diagnostic H5'' proton of 6-deoxysugar (4.56 ppm). Similar to CDP-cillose, two distinct methyl signals were observed: the singlet peak H7'' (1.31 ppm) which represents the methyl group attached to C3''; and the doublet peak H6'' (1.19 ppm) that was assigned to the methyl group at C6''. The chemical shifts acquired are different from CDP-cillose, indicating the enzymatic product from *in vitro* assay of C3CM operon is different from the *E. coli*-based in-microbe reaction. The chemical shift differences between CDP-cillose made in *E. coli* when compared to the *in vitro* enzymatic product led us to acquire more NMR data to determine the structural nature of this material.

FIGURE 8

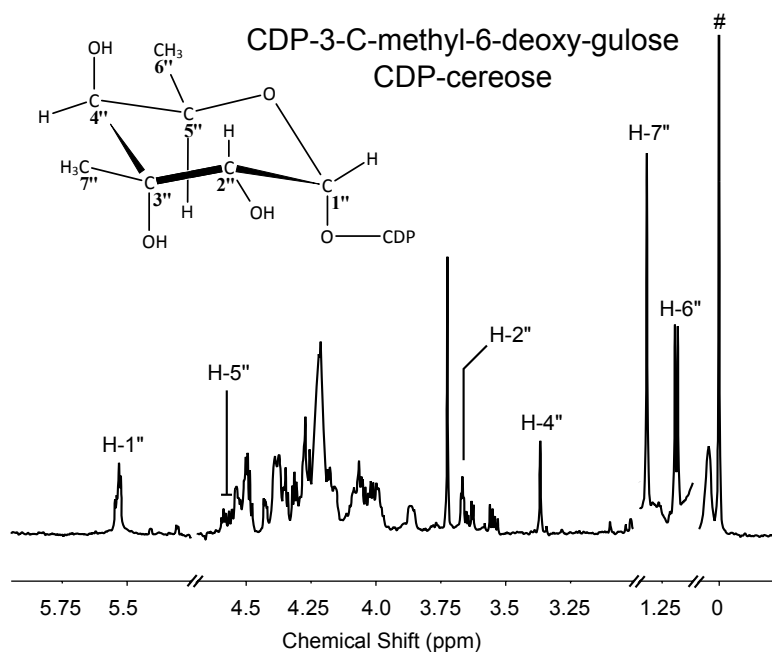


FIGURE 4.8. One-dimensional proton NMR spectrum of the combined four-gene *in vitro* enzymatic product shows that the product is CDP-3-C-methyl-6-deoxy-gulose: CDP-cereose. Selected signals are labeled on the spectrum. Signal labeled with '#' is DSS.

Further two-dimensional COSY and ROESY experiments were performed to determine the stereoisomerism of H4'' (Fig. 4.9). COSY (Fig. 4.9A) shows the connectivity between H1'' to H2'' and H5'' to H6'' on the ring. No proton connected C3'' position as expected; however the coupling between H4'' and H5'' was not observed, indicating the coupling constant between H4'' and H5'' is very small ($J_{4,5} < 1$). Additional ROESY experiment was conducted to confirm that the H4'' is in axial position (Fig. 4.9B). Dipolar couplings were observed between H4'' to H6'' and H4'' to H7'', no NOE effect between H4'' to H2'' (Data not shown), suggesting H4'' is in axial position, consistent with *gulose*-configuration. The NOE observed between H2'' and H7'' (Fig. 4.9B) also indicates that methyl on C3'' is equatorial. Taken together, the product of *in vitro* enzymatic assay from C3CM operon is CDP-3-C-methyl-6-deoxy-gulose. We decided to name it CDP-cereose because while this work was in progress, a study on spore glycoprotein BclA in *Bacillus cereus* ATCC 14579 (18) identified cereose to be attached as a terminal sugar that is β -

1,4-linked to 3-O-methyl-rhamnose on the oligosaccharide of BclA. The complete chemical shifts of CDP-cereose are listed in Table 4.1. We also utilize purified CDP-cereose sample as a standard for the GC-MS analysis and the retention time, EI/MS ion and fragments are identical to the sugar peak “B” in Fig. 4.2B.

To solve the discrepancy between the *E. coli* in-microbe analysis and the *in vitro* enzymatic activities, we decided to examine if the 4-reductase is also a 4-epimerase. However, CDP-cillose was not converted by Bc3361 to CDP-cereose when assays utilized NADPH or NADH as determined by NMR.

FIGURE 9

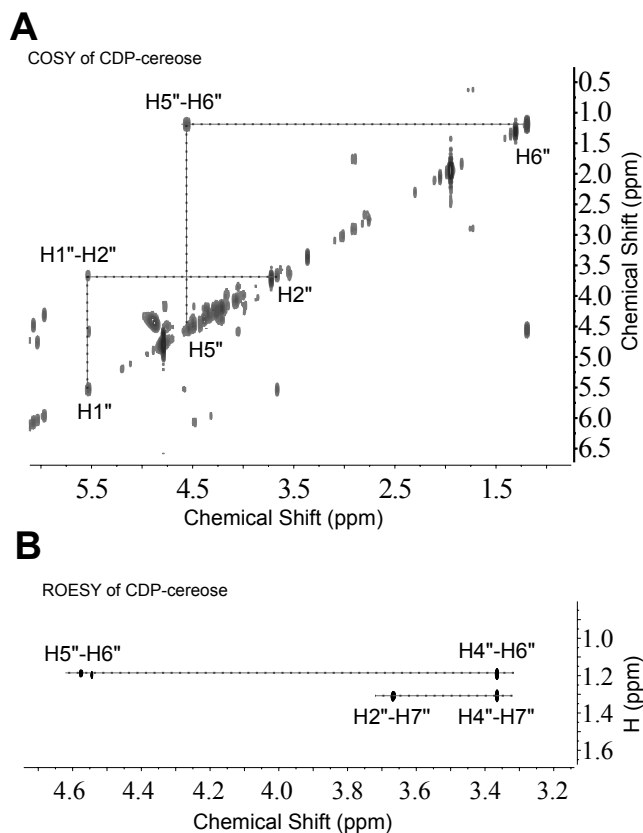


FIGURE 4.9. Detailed NMR characterization of CDP-cereose derived by *in vitro* assays. Panel **A**. COSY experiment showing the connectivity between protons in the sugar ring moiety that are two to three bonds apart. Note that there is no proton connected to H2'' and J coupling between H4'' and H5'' is too small to observe indicating they H4'' is axial position. Panel **B**. selected spectral region of ROESY experiment showing dipolar couplings (NOE effect) between H4'' to H6'', H4'' to H7'', H5'' to H6'' and H2'' to H7'' indicating a gulose-configuration.

4.3.7 *Bacillus* produces predominantly CDP-cereose *in vivo* during sporulation

We next addressed which CDP-3-C-methyl-6-deoxy-hexose sugar(s) *Bacillus* produces *in vivo* and when. For that purpose we analyzed NDP-sugars as well as the expression of genes belonging to the C3CM operon. We first examined if the genes are constitutively expressed or are induced by specific environmental cues. To that end we found that the rich medium BHI, suppressed transcription of the four genes (Fig. 4.10A), while under a defined nutrient medium (Msgg) these genes were transcribed as determined by RT-PCR. It took however, about 48 h for genes of the C3CM operon to be transcribed upon shifting from BHI to Msgg medium.

Figure 10

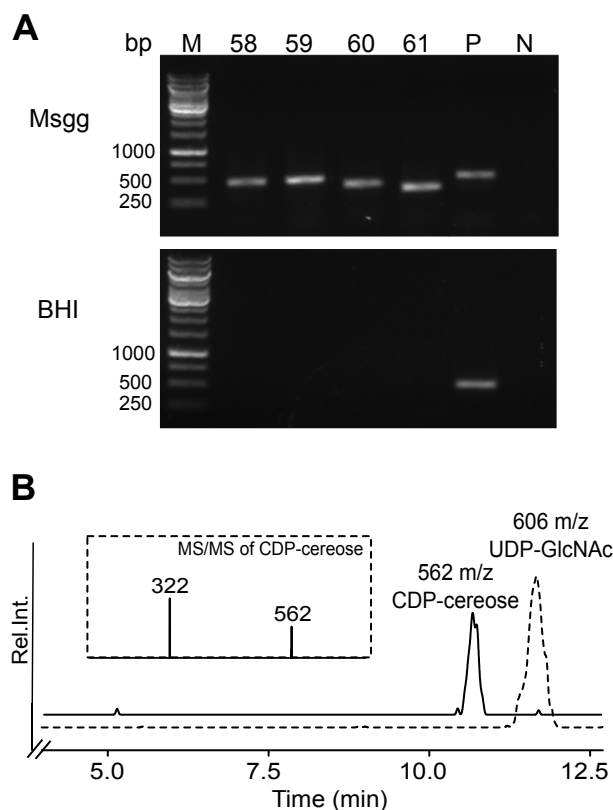


FIGURE 4.10. *Bacillus* produces predominantly CDP-cereose *in vivo* during sporulation. Panel **A**. Genes of the C3CM operon are transcribed during growth in Msgg liquid medium but are not transcribed in rich BHI medium, which maintains vegetative growth. RT-PCR analysis for the expression of gene encoding: lane 58, BC3358; lane 59, BC3359; lane 60, BC3360; lane 61, BC3361; lane P, sigA positive control; lane N, negative control (RT-PCR reaction without RT). Panel **B**. LC-MS/MS analysis of nucleotide sugar extracts from *Bacillus cereus* ATCC 14579 growing in Msgg medium for 2 days. The data was selected to show only [M-H]⁻ ions at *m/z* 562 for CDP-3-C-methyl-6-deoxy-hexose and *m/z* 606 for UDP-GlcNAc as a reference. The MS/MS fragmentation pattern of ion 562 *m/z* is shown in the dotted box.

Because Msgg medium gave rise to both biofilm and spores, we tested if DSM that induces spore formation, also gave rise to accumulation of CDP-3-C-methyl-6-deoxy-hexose. To that end, NDP-sugars were extracted from *Bacillus* growing at either Msgg or DSM. A peak with a retention time at 11 min (Fig. 4.10B) was detected with m/z 562, an ion consistent with CDP-3-C-methyl-6-deoxy-hexose. In both DSM and Msgg the m/z 562 was identified in the first day with highest level at day 2 and reduced amounts in day 4. The HILIC column is unable to separate the CDP-cillose from its 4-epimer, CDP-cereose. To determine the identity of peak 562, the NDP-sugars were isolated, mildly hydrolyzed, derivatized to alditol-acetates and analyzed by GC-MS. The elution time (22.7 min) and MS fragmentation data indicate that the major CDP-sugar is CDP-cereose (Data not shown). Based on the retention time (23.3 min), the amount of cillose derived from CDP-cillose was too low to be designated conclusively. Hence, the combined LC-MS and GC-MS data suggests that *Bacillus* makes predominately CDP-cereose *in vivo* but we cannot exclude the possibility that CDP-cillose is made, albeit in substantially smaller amounts.

4.3.8 C3CM operon is required to form cereose-containing glycan

We first addressed if the C3CM operon is involved in cereose-glycan formation. For this purpose, we deleted the gene encoding Bc3360, the C3-MetT in *B. cereus* ATCC 14579 (Supplementary Fig. 4.S3). The data of spore-derived glycan in Δ BC3360 not surprisingly showed the lack of the cereose sugar residue (Fig. 4.11). This data suggested that genetically C3-MetT of the C3CM operon is involved in cereose-containing spore glycans. We then examined genetically the function of the Bc3361, and whether it is a specific 4-reductase. To that end, we have generated a BC3361 gene deletion strain Δ BC3361. The spores' glycans were analyzed by GC-MS and data showed a significant reduction in the amount of cereose. This strongly suggests

that Bc3361 is indeed involved in the formation of cereose-containing glycan in spores. Further investigation and protein crystallography will be required to fully understand the contribution of Bc3361 to the formation of cillose-containing glycan.

FIGURE S3

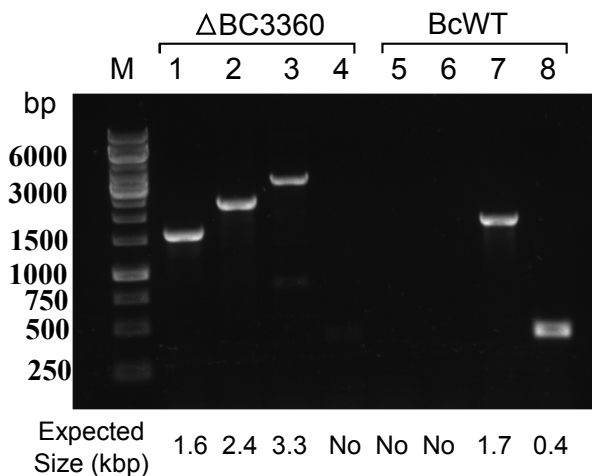


FIGURE 4.S3. PCR verification of mutant Δ BC3360. PCR was conducted on Δ BC3360 and wild type *B. cereus* ATCC 14579 (BcWT) to confirm the deletion of BC3360 and integration of chloramphenicol resistant gene CAT using 4 pairs of primers: Primer set TM003 and ZL053 was designed to encompass a 1.6 kbp fragment (Lane1) that included 3'-portion of BC3359 and the disrupted BC3360 with the CAT gene in the mutant strain Δ BC3360. Similar PCR reaction using BcWT gave no band (Lane5) as expected. Primers TM006 and ZL054 were designed to amplify a 2.4 kbp fragment (Lane2) that contained the disrupted BC3360 with the CAT gene and BC3361 in Δ BC3360, while BcWT gave no band (Lane6). Primer set TM003 and TM006 was designed to amplify DNA region from BC3359 to BC3361 that included the CAT-disrupted BC3360 in the mutant Δ BC3360, and expected to have a band of 3.3 kbp (Lane3,). Similar primer set TM003 and TM006 used with BcWT amplified as expected a DNA of 1.7 kbp (Lane7) that included BC3359, intact BC3360, and BC3361 region; Primer set TM005 and TM006 was designed to amplify DNA region from BC3360 to BC3361: in the mutant Δ BC3360 no PCR DNA band was expected due to partial deletion of BC3360 (Lane 4), while in the BcWT a band of 0.4 kbp was expected (Lane 8,). Primers are listed in supplemental Table 4.S2.

FIGURE 11

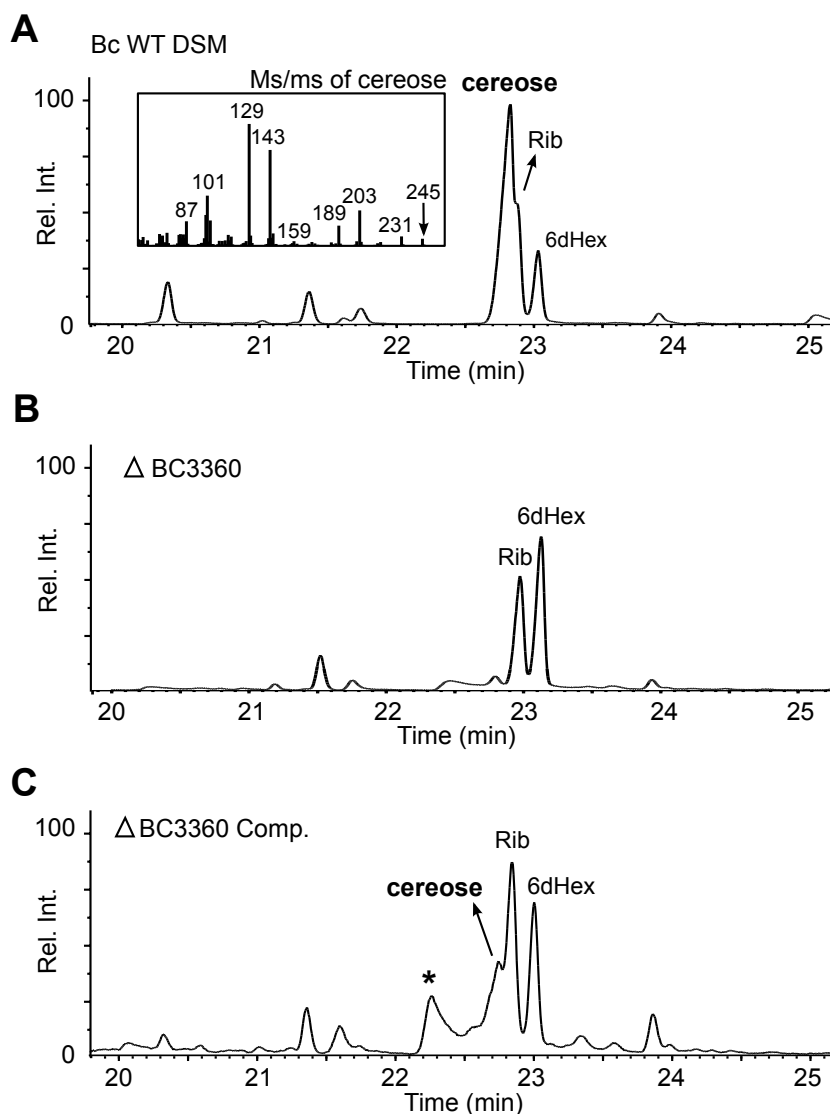


FIGURE 4.11. GC-MS of alditol-acetate derivatives of sugars released by mild acid hydrolyses from spores of: Panel **A.** wild type *B. cereus* ATCC 14579; Panel **B.** Δ BC3360, and Panel **C.** Δ BC3360 complementation strain. All strains were grown on DSM sporulation agar plates. *An unknown contaminant; non-sugar peak.

We next addressed if all *Bacillus* *sp* have the capability to form the cereose-containing glycans. We examined 17 different *Bacillus* strains (Table 4.2) of the *Bacillus cereus sensu lato* group: *B. cereus*, *B. thuringiensis*, and *B. anthracis*. Genomic sequences revealed that four of the seventeen strains do not harbor DNA sequences for the C3CM operon; and these strains showed no cereose residue by GC-MS. But all strains having the C3CM operon produced cereose.

Together, it is evident that the C3CM operon is indeed involved in the formation of cereose containing glycans.

Table 4.2. Different strains in *Bacillus* correlate with C3CM operon and cereose production in spores.

<i>Bacillus</i> strain	C3CM operon in genome	Cereose production in spores
<i>B.thuringiensis israelensis</i> ATCC 35646	No	No
<i>B.thuringiensis berliner</i> ATCC 10792	Yes	Yes
<i>B.thuringiensis kurstaki</i> HD73	Yes	Yes
<i>B.cereus</i> ATCC 14579	Yes	Yes
<i>B.cereus</i> ATCC 10876	Yes	Yes
<i>B.anthraxis</i> 34F2	No	No
<i>B.subtilis</i> PY79	No	No
<i>B.megaterium</i> QMB1551	No	No

4.3.9 C3CM operon may contribute to spore germination

We next address whether mutants of the C3CM operon have an effect on spore properties, sporulation, or spore germination. Upon the initial analysis of Δ BC3360 and Δ BC3361, we did not detect difference in sporulation rate after 8 days of growth on DSM agar plates compared to the wild type. Similarly, the lack of 3C-methyl sugars had no impact on spore stability; and mutant spores pre-incubated at 80°C, 85°C or 90°C had similar heat stability properties as the wild type (Data not shown). However, we observed between 10 to 20 % increase in the germination rate of Δ BC3360 and Δ BC3361 when compared with the wild type (Fig. 4.12). Taken together, cereose glycosylation may contribute to spore germination, but seems not to impact the stability of spores under heat.

FIGURE 12

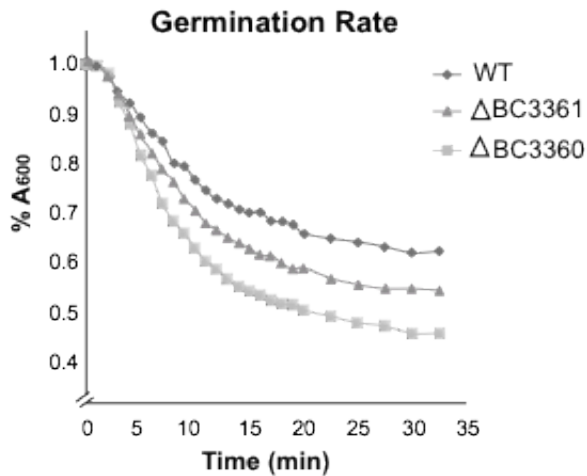


FIGURE 4.12. Mutants Δ BC3360 and Δ BC3361 are germinating faster than parental wild type *B. cereus* ATCC 14579. Germination of spores was determined by the relative decrease in spore optical density at 600 nm over time. The data represent average of two separate experiments.

4.3.10 Timing of cereose-containing glycan expression

In Msgg or DSM, over 90% of *B. cereus* ATCC 14579 cells sporulate at day 3-4. Analyses of glycans isolated from one-day cultures provided no evidence of cereose-containing glycan. Cereose-containing glycan started to appear in day two to three in culture with higher amounts accumulated at day six. To determine if cereose-containing glycans are stable we analyzed spores isolated from culture grown for up to 14 days in Msgg. No degradation of cereose sugar signal was observed between 6 and 14 days, suggesting that cereose containing glycan is not degraded. Similar amount of cereose was observed in spores produced by DSM. In general, the amount of cillose-containing glycans based on GC-MS analyses was not sufficient to be quantified for the timing analysis when compared with cereose-containing glycans.

The amount of cereose in spores was calculated to be 0.2-1% of dry mass. This significant amount may play an important role in spore biology.

4.3.11 Cereose sugar is not exclusive to spore glycoprotein BclA

Following the identification of the CDP-cereose in this report, we further examined whether cereose is found only in BclA and enriched in spores by analyzing the enriched exosporium fraction in the BclA mutant as described earlier (18,19). Cereose indeed appeared in the exosporium preparation, but significant amount of cereose, was also detected in the spores lacking exosporium (Fig. 4.13A), suggesting cereose may not be unique to BclA glycoprotein. In addition, upon examination of the BclA as well as ExsJ mutants from *Bacillus cereus* ATCC 10876 (20) we found cereose is present in spores, indicating that cereose may be attached to other glycans as well (Fig. 4.13B).

Consistent with the above observation, we treated spores with urea and SDS, as the combined reagents are known to release glycoproteins like BclA (21) and BclB (22). Analysis of *B. cereus* ATCC 14579 spores after urea/SDS treatment showed two clear GC-MS signals of cereose and cillose at a ratio of 36:1 respectively (Fig. 13C). This suggests that a large portion of cereose- and cillose-containing glycans still decorates the spores. The explanation to this will require further intensive research, and it could suggest that some of the cereose/cillose glycan structures on the spores are either cross-linked to other macromolecules, urea-resistant glycoproteins or SDS-resistant glycolipids.

FIGURE 13

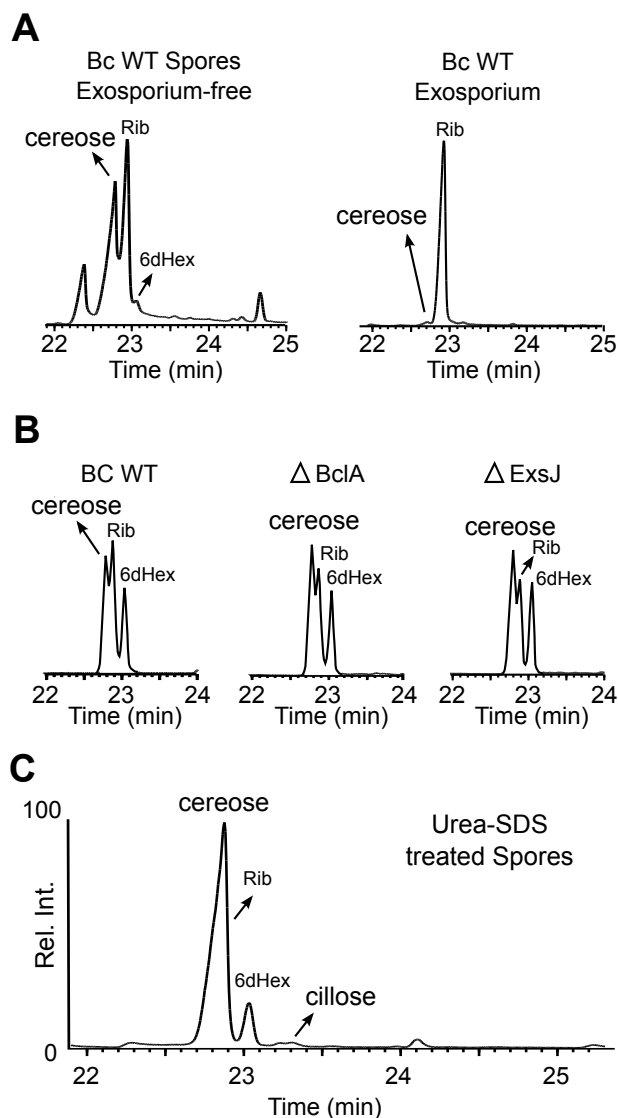


FIGURE 4.13. Cereose-containing glycan is not exclusive to the major spore glycoprotein, BclA. GC-MS of alditol-acetate derivatives of sugars released by mild acid hydrolyses from: Panel **A**. exosporium-free spore of *B. cereus* ATCC 14579 and exosporium fraction; Panel **B**. spores of wild-type *B. cereus* ATCC 10876 and glycoprotein mutants Δ BclA and Δ ExsJ (20); Panel **C**. urea-SDS-treated spores of wild-type *B. cereus* ATCC 14579.

4.4 Discussion

In this study, we have identified and biochemically characterized a new metabolic pathway involved in the formation of uncommon sugar residues in *Bacillus*. The characterized pathway includes four *B. cereus* ATCC 14579 recombinant proteins (Bc3358-Bc3361) and

Figure 4.1 summarizes our proposed biochemical route for the formation and conversion of CDP-glucose to CDP-3-C-methyl-6-deoxy-gulose, named CDP-cereose and potentially its 4-epimer CDP-cillose. The first protein Bc3358 is magnesium- and CTP-dependent enzyme that attaches the cytidyl group from CTP to Glc-1-P to form CDP-glucose. The second enzyme, Bc3359, is a specific 4,6-dehydratase that converts CDP-Glc to a CDP-4-keto-6-deoxy-Glc. 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 (23), UDP-4-keto-6-deoxy-AltNAc (24), UDP-4-keto-6-deoxy-GlcNAc (25), and UDP-4-keto-xylose (26). The third enzyme is a distinct C3-methyltransferase (C3-MetT, Bc3360) that requires magnesium to catalyze the transfer of the methyl group from SAM and link it to C3' of the sugar moiety of the CDP-sugar to form CDP-3-C-methyl-4-keto-6-deoxy-hexose. Like some of other characterized C3-methyltransferases (27,28), Bc3360 is a NADH-dependent. There is no apparent formation of NAD^+ and the enzyme does not appear to function as a reductase in the absence of SAM. It is also able to use NADPH as co-factor with 50% less conversion than NADH as a co-factor. Other co-factor like pyridoxal phosphate (PLP), which maybe suitable for other C3-methyltransferase (29) gave very little conversion (less than 1%). Finally, *in vitro* Bc3361 reduces in an NADPH-dependent manner the 4-keto moiety of the methyl sugar and converts it to CDP-cereose and NADP. The analog NADH could not substitute NADPH in the 4-reductase reaction.

The order of genes within the C3CM operon, the nucleotide space and the overlap amongst the gene is highly similar to other *Bacillus* sp. For example, BC3360 is overlapped by one nucleotide with BC3361 and 1- and 12-nucleotide gap are found between coding region of BC3358 and BC3359; BC3359 and BC3360, respectively. Such gene overlap and reading frame

offset appeared conserved in other *Bacillus* spp including *B. weihenstephanensis* FSL R5-860, *B. cereus* ATCC 10876, *B. thuringiensis* serovar israelensis ATCC 35646, *B. thuringiensis* serovar kurstaki str. HD-1; *B. toyonensis* BCT-7112, and *B. cereus* m1293. It is therefore possible that the expression of the C3CM genes is under the same regulation mechanism. Not all strains of the *B. cereus* group harbor the C3CM operon, for example the *B. anthracis* strains including 34F2, Ames, vollum do not have the C3CM genes in there genome. On the other hand, some *Bacillus cereus* strains which are also human pathogens (supplemental Table 4.S1) like *B. cereus* G4264, G9241 or food and dairy spoilage strains like *Bacillus wiedmannii* do have C3CM operon. This could suggest that cereose decoration on spores may function in assisting spore survival under certain environmental conditions or infecting hosts via specific mechanisms.

Enzyme activity of recombinant proteins showed that the last protein a 4-reductase gives a gulose configured sugar residue, cereose. However, in *E. coli* the 4-reductase produces predominantly the 4-epimer with allose configuration, cillose. The discrepancy was not easily discerned, specifically when cillose and cereose are both present in *Bacillus*. One possibility is that the predominant activity of the 4-reductase is to generate CDP-cereose and little activity is expended to form CDP-cillose. The second possibility is that an additional 4-epimerase in *Bacillus* or *E. coli* epimerizes CDP-cereose to CDP-cillose. Or alternatively, the 4-reductase activity could be regulated post-translationally by other factors, resulting in this enzyme having a different activity compared to its activity *in vitro*. The specific mechanism behind this variance is unknown. Further protein crystallography is required to elucidate the enzymology of Bc3361.

Supplemental table 4.S1. Sequence similarity of C3CM operon genes from *B. cereus* ATCC 14579 to other *Bacillus* strains

	<i>Bacillus cereus</i> ATCC 14579	Bc3358 Glc1P CytT	Bc3359 CDP-Glc 4,6- dehydratase	BC3360 C3-MetT	Bc3361 4-reductase
Origin	Sp.	Locus tag, % AA seq ID (e-value)	% AA seq ID (e-value)	% AA seq ID (e-value)	% AA seq ID (e-value)
toxic wild strain to Lepidoptera	<i>B. thuringiensis</i> BMB171	ADH07857 BMB171_C3047 95 (0)	BMB171_C3048 97(0)	BMB171_C3049 100(0)	BMB171_C3050 99(0)
food spoilage	<i>B. cereus</i> AH187 ()	BCAH187_A3399 99(0)	BCAH187_A3400 88(0)	BCAH187_A3401 92(0)	BCAH187_A3402 92(0)
Contaminated flask	<i>B. cereus</i> ATCC 10876	EEK49878 bcere0002_31300 DJ50_149 99(0)	bcere0002_31310 DJ50_148 99(0)	bcere0002_31320 DJ50_147 97(0)	bcere0002_31330 DJ50_145 97(0)
Tissue, animal	<i>B. thuringiensis</i> serovar <i>berliner</i> ATCC 10792	bthur0008_30730 99(0)	bthur0008_30740 92(0)	bthur0008_30750 97(0)	bthur0008_30760 99(partial seq)
pasteurized milk	<i>Bacillus weihenstephanensis</i> FSL R5-860	C175_03708 99(0)	C175_03713 99(0)	C175_03718 100(0)	C175_03723 99(0)
Wound	<i>Bacillus cereus</i> 172560W	bcere0005_29970 98(0) 98(0)	bcere0005_29980 99(0)	bcere0005_29990 97(0)	bcere0005_30000
Blood and pleural fluid of fatal pneumonia male patient	<i>Bacillus cereus</i> B4264	ACK60932 BCB4264_A3373 99% (0)	BCB4264_A3374 99(0)	BCB4264_A3375 98(0)	BCB4264_A3376 97(0)
Animal feed	<i>Bacillus toyonensis</i> BCT-7112	Btoyo_0594 95(0)	Btoyo_0595 85(0)	Btoyo_0596 91(0)	Btoyo_0597 88(0)
Human clinical sample resembling inhalation anthrax	<i>Bacillus cereus</i> G9241	EAL16983 BCE_G9241_3322 96(0)	BCE_G9241_3323 88(0)	BCE_G9241_3324 92(0)	BCE_G9241_3325 92(0)
Food poisoning	<i>Bacillus cereus</i> AH1271	bcere0028_31110 94(0)	bcere0028_31120 86(0)	bcere0028_31130 93(0)	bcere0028_31140 91(6xe-175)
Raw milk	<i>Bacillus wiedmannii</i> FSL W8-0169	AT260_18250	AT260_18255	AT260_18260 93(0)	AT260_18265

4.5 Experimental Procedures

4.5.1 Strains and culture conditions

The strains used in this study were listed in Table 4.2. Each strain was stored in 16% glycerol at -80 °C. The media (agar or liquid) used were Luria Bertani (LB) (10 g/L tryptone, 5

g/L yeast extract, 10 g/L NaCl), BHI, DSM (Difco Sporulation Medium) and Msgg medium (30). *Escherichia coli* DH10B cells were used for cloning and Rosetta2(DE3)pLysS (Novagen) cells were used to express protein.

B. cereus cells were cultured in either liquid or agar plate with Msgg or DSM plates and incubated at 30 °C for 5 days to achieve over 90% sporulation rate, as confirmed by microscopy and CFU counts after 80 °C heat stability treatment. Spores were purified as described (31); and subsequently washed repeatedly with sterile double distilled water (ddw) and finally centrifuged (8,000 x g, 10 min).

4.5.2 Cloning genes of the C3CM operon

A single colony of *B. cereus* growing on BHI agar was suspended in 50 µl sterile ddw, heated for 5 min at 95 °C and then centrifuged (10,000 x g, 30 sec). A portion of the supernatant (5 µl) was used as the source of DNA to amplify each specific gene by PCR. Each primer was designed to include a 15-nucleotide extension with exact sequence homology to the cloning site of the target plasmid at the 5' end. This facilitated cloning between the NcoI and Bpu1101I sites of pET28_Tev1.15 (32) or the SacI site of pCDF-ParaB vector. pCDF-ParaB is a modified pCDF-duet where expression is control by arabinose via the regulator gene *araC*. The pET28_Tev1.15 vector was (32) opened with primer sets ZL169 and KB_T7t, and the individual genes (BC3358, BC3359, BC3360, and BC3361) were PCR-amplified with primer sets KB2 S and KB2 AS; KB5 S and KB5 AS; KB3 S and KB3 AS; and KB4 S and KB4 AS, respectively (See supplemental Table 4.S2), using high fidelity *Pyrococcus* DNA polymerase (0.4 U Phusion Hot Start II; ThermoScientific), including buffer, dNTPs (0.4 µl 10 mM), *B. cereus* DNA template (5 µl), PCR primer sets (1 µl each of 10 µM) in a 20-µl reaction volume. The PCR conditions were a 98 °C denaturation cycle for 30 sec followed by 25 × cycles [8 sec

denaturation at 98 °C, 25 sec annealing at 54 °C, and 20 sec elongation at 72 °C] and 4 °C. After PCR, portions (2 µl each) of the amplified vector and insert were mixed, treated for 15 min at 37 °C with 0.5 U FastDigest DpnI (ThermoScientific), and then transformed into DH10B (LifeTechnologies) competent cells. Positive transformed clones were selected on LB agar containing kanamycin (50 µg/ml for pET vector) or Spectinomycin (100 µg/ml for pCDF vector). Clones were verified by PCR and DNA sequencing. The resulting recombinant plasmids are: pET28b:His₆-Tev-Bc3358#3 yielding N-terminal His₆-Tev-tagged Glc1P-CytT; pET28b:His₆-Tev-Bc3359#10 yielding N-terminal His₆-Tev-tagged CDP-Glc 4,6-DH; pET28b:His₆-Tev-Bc3360#12 yielding N-terminal His₆-Tev-tagged C3-MetT; and pET28b:His₆-Tev-Bc3361#2 yielding N-terminal His₆-Tev-tagged 4-reductase. Plasmids were isolated and purified by PureLink Quick Plasmid Miniprep Kit (Invitrogen) and transformed into Rosetta2(DE3)pLysS competent cells for recombinant protein expression. Once protein activities were confirmed, the DNA sequences of the plasmids harboring specific genes were deposited in Genbank with the following accession numbers: KY445942 for Bc3358, KY445943 for Bc3359, KY432406 for Bc3360, and KY432407 for Bc3361.

Supplemental table 4.S2. Primer list.

ZL169	5'-CATGGCGCCCTGAAAATACAGGTTTTTCGCC-3'
KB T7t	5'-TAGCCGCTGAGCAATAACTAGCATAACCCCTTG-3'
KB2 S	5'-TTTCAGGGCGCCATGaaagctgtattcttgctggtg-3'
KB2 AS	5'-ATTGCTCAGCGGCTAttacaaaccttccacgg-3'
KB5 S	5'-TTTCAGGGCGCCATGgcttcacatcttttgaataag-3'
KB5 AS	5'-ATTGCTCAGCGGCTAttataaattttgaatacatcgattgtgtgc-3'
KB3 S	5'-TTTCAGGGCGCCATGgaacagaaaaaatgccgttttg-3'
KB3 AS	5'-ATTGCTCAGCGGCTAtcatggctcgatcacctc-3'
KB4 S	5'-TTTCAGGGCGCCATGaagaaagttgttgaacaggtg-3'
KB4 AS	5'-ATTGCTCAGCGGCTAGATTACGATttagtgacgtatatcg-3'
ZL190	5'-GCAAAAAACCCCTCAAGACCCGTTTAGAGG-3'
ZL191	5'-TGAAAGGAGGAATAATATCCGGATTGGCG-3'
ZL192	5'-CTTGAGGGGTTTTTTCGggtgatgtcgcgatatagg-3'
ZL193	5'-TATAGTTCCTCCTTTCAgcaaaaaaccttcaagacc-3'
ZL194	5'-CGGGCTCATGAGCGCTTGTTTCGGC-3'
ZL195	5'-AAGTGGCGAGCCCGATCTTCCCCATC-3'

ZL196	5'-AAGCGCTCATGAGCCCGggtgatgtcggcgatatagg-3'
ZL197	5'-GATCGGGCTCGCCACTTgcaaaaaacccctcaagacc-3'
JGI001	5' ccatacccggttttttgggaattcgagctc-3'
JGI002	5'-ggttaattaagctgcgctagtagagagctc-3'
TM001	5'-gtgtagaaccatggaaagtacac-3'
TM002	5'-ctattactttcccatagctcgac-3'
TM003	5'-tcagagctggtaacgtcatc-3'
TM004	5'-gaatacatcgatttgtgtcgtg-3'
TM005	5'-gaagtcgatcatgggtagatac-3'
TM006	5'-tcgatcacctcaacttctgg-3'
TM007	5'-acgaacaaggtaatcggtc-3'
TM008	5'-atcgtatgttggttcccag-3'
ZL347	5'-GCGCGAACAATTCGTATTCC-3'
ZL348	5'-CCTTAAGACGCTTACTACGGC-3'
ZL479	5'-CGATGCATGCCATGGcattcttcgcactaatacccag-3'
ZL480	5'-GAGTTAGGATCGCTAGCcaagagttacaaacaatgtgtgc-3'
ZL481	5'-CAGATAAGGATCGAATTCggaacaccctttgaattattgc-3'
ZL482	5'-GGCGATATCGGATCCctgtttgtgtgcatacgtttctag-3'
ZL489	5'-CGATGCATGCCATGGtgatgtagaagaactttcaacgc-3'
ZL490	5'-GAGTTAGGATCGCTAGCggtgaatgagatggctaggc-3'
ZL491	5'-CAGATAAGGATCGAATTCactagacgcgattcctttctc-3'
ZL492	5'-GGCGATATCGGATCCgatgcatatgcgacagatcatg-3'
ZL007	5'-GCTAGCGATCCTAACTCACATTAATTGCGTTGCGC-3'
ZL008	5'-CCATGGCATGCATCGATAGATCTGTCTAGTTAATGTG-3'
ZL009	5'-GGATCCGATATCGCCCGACGCGAGGCTGGATGGC-3'
ZL123	5'-GAATTCGATCCTTATCTGTGCCCCAGTTTGCTAGG-3'
ZL213	5'-CATAGTAGTTCCTCCTTAAGCTTAATTGTTATCC-3'
ZL214	5'-GGATCCGATCAGACCAGTTTTTAATTTAAGC-3'
ZL483	5'-CTTAAGGAGGAAGTACTATGgaacagaaaaaatgccgtttttgcc-3'
ZL484	5'-ACTGGTCTGATCGGATCCctcatggctcgatcacctcaac-3'
ZL493	5'-CTTAAGGAGGAAGTACTATGaagaaagttgtgtaacaggtgg-3'
ZL494	5'-ACTGGTCTGATCGGATCCttatgtttttgtataaatgatteccacc-3'
ZL053	5'-ATGAACCTTAATAAAAATTGATTTAGACAATTGG-3'
ZL054	5'-TTATAAAAGCCAGTCATTAGGCC-3'

For C3CM operon gene stacking, plasmid pET28_Tev harboring BC3359 was used as a secondary vector for the insertion of Bc3360. Vector pET28b:His₆-Tev-Bc3359 and insert pET28b:His₆-Tev-Bc3360 were amplified with primer sets ZL190 and ZL191; ZL192 and ZL193, respectively (See supplemental Table 4.S2), and cloned as previously stated. The resulting construct pET28b:His₆-Tev-Bc3359+His₆-Tev-Bc3360 was then used as tertiary vector for the insertion of BC3358. Vector pET28b:His₆-Tev-Bc3359+His₆-Tev-Bc3360 and insert

pET28b:His₆-Tev-Bc3358 were amplified with primer sets ZL194 and ZL195; ZL196 and ZL197, respectively, and cloned as previously stated. The resulting construct, pET28b: His₆-Tev-Bc3358+His₆-Tev-Bc3359+His₆-Tev-Bc3360, was verified by PCR and DNA sequencing, co-transformed with pCDF-Parab: Bc3361-His₆ plasmid, and cloned by primer sets JGI001 and JGI002 into Rosetta2(DE3)pLysS-competent cells for C3CM recombinant proteins co-expression.

4.5.3 His₆-tagged protein expression and purification

Rosetta2(DE3)pLysS strains harboring pET expression plasmids were grown at 37 °C and 250 rpm in 250 ml LB supplemented with chloramphenicol (35 µg/ml) and kanamycin (50 µg/ml). Gene expression was induced when cell A₆₀₀ reached 0.6 by adding 0.5 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG). After induction, cells were grown for 18 h at 18 °C and 250 rpm and then harvested by centrifugation (6,000 x g) for 10 min at 4 °C. The cell pellets were washed with ddw and then suspended in 10 ml lysis buffer (50 mM Tris-HCl, pH 7.4, 10% (v/v) glycerol, 1 mM EDTA, 2 mM DTT, and 0.5 mM PMSF). Cells were lysed by sonication (26) and after centrifugation (6,000 x g for 15 min at 4 °C), the supernatant was supplemented with 1 mM DTT and 0.5 mM PMSF and again centrifuged (20,000 g for 30 min at 4 °C). The supernatant was applied to a Ni-Sepharose fast-flow column (GE Healthcare Life Sciences, Piscataway, NJ, USA; 2 ml of resin packed in a polypropylene column; inner diameter 1 cm × 15 cm). Each column was pre-equilibrated with buffer A (50 mM Tris-HCl, pH 8, 4% (v/v) glycerol, 100 mM NaCl). The column was washed with 30 ml of buffer A containing 20 mM imidazole and then with 10 ml of buffer A containing 40 mM imidazole. His-tagged proteins were eluted with 5 ml of buffer A containing 250 mM imidazole. The eluates containing these

proteins were divided into small aliquots, flash frozen in liquid nitrogen, and kept at -80 °C. Proteins were separated by SDS-12.5% PAGE and visualized by staining with Coomassie blue.

4.5.4 Nucleotide-sugar analyses derived from “in-microbe”-based assay

Nucleotide sugars produced by the C3CM operon were extracted as previously described (33). Briefly, Rosetta2(DE3)pLysS strains harboring expression plasmids were grown at 37 °C and 250 rpm in 5 ml LB supplemented with chloramphenicol (35 µg/ml) and kanamycin (50 µg/ml) (pET vector), or chloramphenicol (25 µg/ml) and kanamycin (35 µg/ml) with additional spectinomycin (35 µg/ml) (pET plus pCDF vector co-expression). Gene expression was induced after 2 h by adding 0.5 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG). After induction, cells were grown for 3-4 h at 30 °C and 250 rpm and then harvested by centrifugation (10,000 x g, for 1 min at 2 °C). The cell pellets were washed with 4 volumes of 10 mM sodium phosphate (pH 7.5) and 150 mM NaCl (PBS) and then suspended in 1 volume of ddw. Ten volumes of cold chloroform/methanol (1:1, v/v) was added, and the sample was mixed for 15 min on ice. Samples were centrifuged (12,000 rpm, 3 min, 22 °C), and the upper aqueous phase enriched in nucleotide sugar was collected. Portions of this aqueous phase were chromatographed on HILIC column using HPLC-UV or LC-ESI-MS/MS (liquid chromatography–electrospray ionization–tandem mass spectrometry), using a Shimadzu ESI-MS/MS IT-TOF mass spectrometer system that included a Nexera UFPLC LC-30AD pump, A Sil30AC autosampler, and a column heater at 37 °C). HPLC peaks of nucleotide sugars were detected by UV $A_{271\text{nm}}$ (max for CDP-sugars), collected, and lyophilized prior to analyses by NMR. The purified product of His₆BC3358 (CDP-Glc) was also used as substrate for the downstream enzymatic reactions. The purified product of entire C3CM operon (CDP-cillose or CDP-cereose) was suspended in D₂O (99.9%) for NMR analysis.

For analyses by HPLC-UV or LC-MS/MS, an Accucore 150-amide HILIC column (150 × 4.6 mm, 2.6 µm particle size, ThermoScientific) was used for chromatography with solvent system of 40 mM ammonium-acetate pH 4.3 (solvent A) and acetonitrile (solvent B). The column was equilibrated at 0.4 ml/min with 25% A and 75% B prior to sample injection (20 µl). Following injection, the HPLC conditions were: 0-1 min, 0.4 ml/min with 25% A/75%B then a gradient to 50% A and 50% B over 24 min. The flow rate was then increased to 0.6 ml/min with a gradient to 25% A and 75% B over 5 min. The column was then washed for 5 min with 25%A/75%B prior to the next injection.

4.5.5 *In vitro* Enzyme reactions

The activity of recombinant His₆Bc3358 (refer herein as Bc3358, or Glc1P-CytT) was examined by HPLC-UV and LC-ESI-MS/MS. The total reaction volume of 50 µl was consisted of 50 mM Tris-HCl pH 7.4, 3.3 mM MgCl₂, 1.5 mM CTP, 1 mM Glc-1-P, and up to 15 µl of purified Bc3358. Reactions were proceed for 1 h at 30 °C. An aliquot (20 µl) was mixed with acetonitrile (40 µl) and 0.5M ammonium-acetate pH 5.3 (2 µl) and a portion (30 µl) was chromatographed and detected by the LC-ESI-MS/MS operating in the negative ion mode. Enzyme products were identified based on their retention time, the mass of their parent ion and their mass spectral fragmentation pattern.

The activity of recombinant His₆Bc3359 (refer herein as Bc3359, or 4,6-DH) was examined similarly as above. To a total reaction volume of 50 µl was added 50 mM Tris-HCl pH 7.4, CDP-Glc, and up to 15 µl of purified Bc3359. Reactions were preceded for 2 h at 30 °C, followed by enzyme inactivation and extraction with 50 µl chloroform ([32](#)). The activity of recombinant His₆Bc3360 (refer herein as Bc3360, C3-MetT) was examined using the reaction product of Bc3359 supplemented with 1 mM NADH, 3.3 mM MgCl₂, 1 mM s-

adenosylmethionine (SAM) and up to 15 μ l of purified Bc3360 in total volume of 25 μ l. Reactions were preceded for 2 h at 22 °C. For the co-incubation assays that included both Bc3360 and Bc3361, enzymatic reaction product of Bc3359 was supplemented with 1 mM NADH, 1 mM NADPH, 3.3 mM MgCl₂, 1 mM s-adenosylmethionine (SAM), 4 μ l of purified Bc3360 and 4 μ l of purified Bc3361 in total volume of 25 μ l. Reactions were preceded for 2 h at 30 °C. Aliquots of reactions were prepared as previously described and analyzed by HILIC LC-ESI-MS/MS.

4.5.6 NMR spectroscopy used to characterize the structure of products

Time-resolved NMR for the formation of CDP-4-keto-6-deoxy-Glc by recombinant Bc3359 was carried out in a final volume of 180 μ l, essentially as described (34). The HPLC peaks of CDP-Glc, CDP-cillose and CDP-cereose collected as previously described were lyophilized, dissolved in D₂O (99.9%), supplemented with 1 μ l of 10 mM DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid), and analyzed by NMR spectroscopy (Agilent DD2 600 MHz NMR spectrometer equipped with a cryogenic 3 mm probe). Proton and carbon chemical shifts were referenced to an internal DSS peak set at 0.00 ppm for both proton and carbon spectra. The one-dimensional proton NMR spectrum was recorded using the water-presaturated pulse sequence and obtained with a spectral width of 6 kHz, a 90° pulse field angle (7.5 μ s), a 2.7-sec acquisition time, and a 2-sec relaxation delay (RD). Proton chemical shifts were assigned by a CORrelation Spectroscopy (COSY) experiment. Carbon chemical shifts as well as the connectivity between carbons and protons were determined by a Heteronuclear Multiple Bond Correlation (HMBC) experiment. The configuration of each hydroxyl group on CDP-cillose or CDP-cereose was determined by one-dimensional proton and ROESY (Rotating-frame Overhauser Spectroscopy) experiment. Data processing and plotting were performed using software MestreNova.

4.5.7 RNA isolation and RT-PCR

A 1:100 dilution of over night BHI-culture of *B. cereus* ATCC 14579 were used to inoculate 50 ml liquid medium (BHI or Mmsgg) and cultures were grown for up to 72 h at 30 °C while shaking (200 rpm). Cells were rapidly pelleted by centrifugation (10,000 x g for 1 min at 4 °C), resuspended in 0.8 ml of lysozyme solution ([14](#)) and incubated at room temperature for 10 min. To each sample 80 µl 10xEB (0.3 M NaOAc pH 5.2, 50 mM EDTA, 5% sarkosyl and 1.42 M β-mercaptoethanol) was added. After 3 min at 65 °C one volume of 70 °C pre-heated acidic phenol (phenol:AcE buffer: 50 mM NaOAc pH 5.1, 5 mM EDTA (1:1 v/v)) was added and the mixture was incubated at 65 °C for 7 min. The sample was centrifuged (10,000 g for 5 min at 4 °C); the top aqueous layer was collected and mixed with equal volume of chloroform. Nucleic acid partitioned to the upper phase was collected mixed with 1-volume of 30 mM NaOAc pH 5.2 and 3-volumes of cold ethanol and nucleic acids were precipitated at -20 °C. To digest remnant genomic DNA, an aliquot of 5 µg crude RNA was treated with DNase I. The RNA was ethanol-precipitated, and a portion was used for transcript analyses by reverse transcriptase (RT) and PCR reactions. The RT reaction (20 µl, final) consisted of 250 ng of RNA, 1 µl of 10 µM gene specific reverse primers, buffer, 0.2 µM dNTPs and 1 unit of reverse transcriptase (SuperScript III, Invitrogen). Negative control RT reactions were done without added reverse transcriptase. Transcripts of genes in the C3CM operon, negative controls and positive control (SigA) were amplified each, at final 25 µl, by PCR reaction that included 2 µl of RT reaction, buffer, dNTPs, 1 unit of Taq DNA polymerase (Promega) and 0.4 µM of each gene specific sense and antisense primers (primer set TM001 and TM002 for BC3358; TM003 and TM004 for BC3359; TM005 and TM006 for BC3360; TM007 and TM008 for BC3361; ZL347 and ZL348 for SigA (See supplemental Table 4.S2). Following PCR, a portion (8 µl) of each RT-PCR reaction was loaded

on 1% agarose-TAE gel casted with 10 µg/ml ethidium bromide; separated by gel electrophoresis and UV-imaged using gel imager.

4.5.8 GC-MS analyses for identification of cereose and cillose

As cereose and cillose are labile sugars, mild hydrolysis was used. Purified spores collected either from liquid culture or scraped from agar plate were washed and suspended. Around 1×10^9 of spores were supplemented with 10 µg inositol; hydrolyzed with 0.2 M trifluoroacetic acid (TFA) for 2 h in 70 °C; and the released monosaccharides were reduced to their alditols (35), and acetylated. The resulting alditol acetate derivatives were analyzed by GC/EI-MS system (Agilent 7890a/5975c), equipped with an autosampler injector (Agilent 7693). A 1 µl sample was injected into GC-column (Equity-1 or DB-5, 30 m × 0.25 mm, 0.25 µm film thickness) using split mode (1:50) with injector inlet setting of 250 °C (helium at 3 ml/min); and chromatography was performed as previously described (14). In some cases the GC column used was Rtx-2330 by RESTEK. MS data were collected after a solvent delay of 5 min and ion abundance in the range of 50-550 *m/z* was recorded. The spectra were analyzed using MSD ChemStation. To determine the elution time of authentic cereose and cillose and the EI-mass fragments formed, CDP-cereose and CDP-cillose was produced and purified over HILIC column; TFA-hydrolyzed, converted to alditol-acetate, separated by GC using DB-5 column and analyzed EI-MS using above GC conditions.

4.5.9 Mutant generation in C3CM operon

Gene knockout was achieved by double crossover recombination as previously described (36). Two DNA fragments flanking gene BC3360 and two DNA fragments flanking gene BC3361 were PCR amplified by specific primer sets ZL479 and ZL480; ZL481 and ZL482; ZL489 and ZL490; ZL491 and ZL492, respectively (See supplemental Table 4.S2). The flanking

regions were individually cloned by primer sets ZL007 and ZL008; ZL009 and ZL123, respectively, into pZL-KO shuttle plasmid (36), a derivative of pBCB13 that harbors the thermo-sensitive ORI (origin of replication) from pMAD (37,38). Cloning strategy for chromosomal integration at non-permissive temperature and subsequently a double crossover recombination event was conducted as described before (36) to achieve the resulting *B. cereus* mutant strain Δ BC3360 and Δ BC3361. For complementation of mutants, PCR was used to amplify BC3360 and BC3361 genomic DNA region from wild type *B. cereus* ATCC 14579 with primer sets ZL483 and ZL484; ZL485 and ZL486, respectively. Each DNA fragment was cloned into pDZ vector, a derivative *E. coli* /*Bacillus* shuttle plasmid of pDG148-stu (39), that was generated by primer sets ZL213 and ZL214 as previously described (36).

4.5.10 Germination analysis of mutants

Spores from *B. cereus* wild-type and mutant strains Δ BC3360 and Δ BC3361 were prepared as previously stated. Prior to germination, purified spores were heated at 70 °C for 30 min, and washed again with ddw. Equal amount of spores, (calculated to an optical density A_{600} = 0.8), were resuspended in germination solution (50 mM inosine in 10 mM Tris-HCl pH 8, 10 mM NaCl), and degree of germination was monitored at A_{600} over a time course of 32 min (40) .

4.5.11 Exosporium extraction and Urea-SDS treatment

Purified spores from *B. cereus* wild-type were washed twice in water and the exosporium extraction was conducted as previously described (19). Briefly, spores were subjected to four successive passages through a French press (20,000 psi) and the exosporium-free spores were recovered after centrifugation (3,000 x g, 30 min, 4 °C). Exosporium in the supernatant was pelleted by ultracentrifuge at (120,000 x g, 30 min, 4 °C), treated by mild acid hydrolysis, derivatized to alditol-acetates; and analyzed by GC-MS.

Spores were also suspended in urea-SDS buffer (50 mM Tris-HCl pH 7, 2 % SDS, 8M urea) and boiled for 15 min as described (20,21). The sample was then centrifuged at 10,000 x g for 10 min at 4 °C, and the supernatant was dialyzed against ddw, lyophilized, and followed by mild acid hydrolysis, alditol-acetate derivatization and GC-MS analyses.

4.6 References

1. Jacques, M. (1996) Role of lipo-oligosaccharides and lipopolysaccharides in bacterial adherence. *Trends in microbiology* **4**, 408-410
2. Kohler, H., Rodrigues, S. P., and McCormick, B. A. (2002) Shigella flexneri interactions with the basolateral membrane domain of polarized model intestinal epithelium: Role of lipopolysaccharide in cell invasion and in activation of the mitogen-activated protein kinase ERK. *Infection and immunity* **70**, 1150-1158
3. Tang, H. B., DiMango, E., Bryan, R., Gambello, M., Iglewski, B. H., Goldberg, J. B., and Prince, A. (1996) Contribution of specific Pseudomonas aeruginosa virulence factors to pathogenesis of pneumonia in a neonatal mouse model of infection. *Infection and immunity* **64**, 37-43
4. Day, C. J., Tran, E. N., Semchenko, E. A., Tram, G., Hartley-Tassell, L. E., Ng, P. S. K., King, R. M., Ulanovsky, R., McAtamney, S., Apicella, M. A., Tiralongo, J., Morona, R., Korolik, V., and Jennings, M. P. (2015) Glycan:glycan interactions: High affinity biomolecular interactions that can mediate binding of pathogenic bacteria to host cells. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E7266-E7275
5. Cabeen, M. T., and Jacobs-Wagner, C. (2005) Bacterial cell shape. *Nature Reviews Microbiology* **3**, 601-610
6. Sylvestre, P., Couture-Tosi, E., and Mock, M. (2002) A collagen-like surface glycoprotein is a structural component of the Bacillus anthracis exosporium. *Molecular microbiology* **45**, 169-178
7. Steichen, C., Chen, P., Kearney, J. F., and Turnbough, C. L. (2003) Identification of the immunodominant protein and other proteins of the Bacillus anthracis exosporium. *Journal of bacteriology* **185**, 1903-1910
8. Daubenspeck, J. M., Zeng, H. D., Chen, P., Dong, S. L., Steichen, C. T., Krishna, N. R., Pritchard, D. G., and Turnbough, C. L. (2004) Novel oligosaccharide side chains of the collagen-like region of BclA, the major glycoprotein of the Bacillus anthracis exosporium. *Journal of Biological Chemistry* **279**, 30945-30953
9. Lequette, Y., Garenaux, E., Tauveron, G., Dumez, S., Perchat, S., Slomianny, C., Lereclus, D., Guerardel, Y., and Faille, C. (2011) Role Played by Exosporium Glycoproteins in the Surface Properties of Bacillus cereus Spores and in Their Adhesion to Stainless Steel. *Applied and environmental microbiology* **77**, 4905-4911
10. Oliva, C. R., Swiecki, M. K., Griguer, C. E., Lisanby, M. W., Bullard, D. C., Turnbough, C. L., and Kearney, J. F. (2008) The integrin Mac-1 (CR3) mediates internalization and directs Bacillus anthracis spores into professional phagocytes. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 1261-1266
11. Oliva, C., Turnbough, C. L., and Kearney, J. F. (2009) CD14-Mac-1 interactions in Bacillus anthracis spore internalization by macrophages. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 13957-13962
12. Lequette, Y., Garenaux, E., Combrouse, T., Dias Tdel, L., Ronse, A., Slomianny, C., Trivelli, X., Guerardel, Y., and Faille, C. (2011) Domains of BclA, the major surface glycoprotein of the B. cereus exosporium: glycosylation patterns and role in spore surface properties. *Biofouling* **27**, 751-761

13. Bozue, J. A., Parthasarathy, N., Phillips, L. R., Cote, C. K., Fellows, P. F., Mendelson, I., Shafferman, A., and Friedlander, A. M. (2005) Construction of a rhamnose mutation in *Bacillus anthracis* affects adherence to macrophages but not virulence in guinea pigs. *Microb Pathogenesis* **38**, 1-12
14. Abhyankar, W., Hossain, A. H., Djajasaputra, A., Permpoonpattana, P., Ter Beek, A., Dekker, H. L., Cutting, S. M., Brul, S., de Koning, L. J., and de Koster, C. G. (2013) In Pursuit of Protein Targets: Proteomic Characterization of Bacterial Spore Outer Layers. *J Proteome Res* **12**, 4507-4521
15. Li, Z., Hwang, S., and Bar-Peled, M. (2016) Discovery of a Unique Extracellular Polysaccharide in Members of the Pathogenic *Bacillus* That Can Co-form with Spores. *The Journal of biological chemistry* **291**, 19051-19067
16. Bruender, N. A., Thoden, J. B., Kaur, M., Avey, M. K., and Holden, H. M. (2010) Molecular architecture of a C-3'-methyltransferase involved in the biosynthesis of D-tetronitrose. *Biochemistry* **49**, 5891-5898
17. Martinez, V., Ingwers, M., Smith, J., Glushka, J., Yang, T., and Bar-Peled, M. (2012) Biosynthesis of UDP-4-keto-6-deoxyglucose and UDP-rhamnose in pathogenic fungi *Magnaporthe grisea* and *Botryotinia fuckeliana*. *The Journal of biological chemistry* **287**, 879-892
18. Li, Z., Hwang, S., Ericson, J., Bowler, K., and Bar-Peled, M. (2015) 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*. *The Journal of biological chemistry* **290**, 691-704
19. Hwang, S., Li, Z., Bar-Peled, Y., Aronov, A., Ericson, J., and Bar-Peled, M. (2014) The biosynthesis of UDP-d-FucNAc-4N-(2)-oxoglutarate (UDP-Yelosamine) in *Bacillus cereus* ATCC 14579: Pat and Pyl, an aminotransferase and an ATP-dependent Grasp protein that ligates 2-oxoglutarate to UDP-4-amino-sugars. *The Journal of biological chemistry* **289**, 35620-35632
20. Gu, X., Glushka, J., Yin, Y., Xu, Y., Denny, T., Smith, J., Jiang, Y., and Bar-Peled, M. (2010) Identification of a bifunctional UDP-4-keto-pentose/UDP-xylose synthase in the plant pathogenic bacterium *Ralstonia solanacearum* strain GMI1000, a distinct member of the 4,6-dehydratase and decarboxylase family. *The Journal of biological chemistry* **285**, 9030-9040
21. Du, Y. L., Alkhalaf, L. M., and Ryan, K. S. (2015) In vitro reconstitution of indolmycin biosynthesis reveals the molecular basis of oxazolinone assembly. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 2717-2722
22. Barkovich, R. J., Shtanko, A., Shepherd, J. A., Lee, P. T., Myles, D. C., Tzagoloff, A., and Clarke, C. F. (1997) Characterization of the COQ5 gene from *Saccharomyces cerevisiae*. Evidence for a C-methyltransferase in ubiquinone biosynthesis. *The Journal of biological chemistry* **272**, 9182-9188
23. Thorson, J. S., Lo, S. F., and Liu, H. W. (1993) Molecular-Basis of 3,6-Dideoxyhexose Biosynthesis - Elucidation of Cdp-Ascarylose Biosynthetic Genes and Their Relationship to Other 3,6-Dideoxyhexose Pathways. *Journal of the American Chemical Society* **115**, 5827-5828
24. Maes, E., Krzewinski, F., Garenaux, E., Lequette, Y., Coddeville, B., Trivelli, X., Ronse, A., Faille, C., and Guerardel, Y. (2016) Glycosylation of BclA Glycoprotein from *Bacillus cereus* and *Bacillus anthracis* Exosporium Is Domain-specific. *The Journal of biological chemistry* **291**, 9666-9677
25. Todd, S. J., Moir, A. J., Johnson, M. J., and Moir, A. (2003) Genes of *Bacillus cereus* and *Bacillus anthracis* encoding proteins of the exosporium. *Journal of bacteriology* **185**, 3373-3378
26. Branda, S. S., Gonzalez-Pastor, J. E., Ben-Yehuda, S., Losick, R., and Kolter, R. (2001) Fruiting body formation by *Bacillus subtilis*. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 11621-11626
27. Yang, T., Bar-Peled, L., Gebhart, L., Lee, S. G., and Bar-Peled, M. (2009) Identification of galacturonic acid-1-phosphate kinase, a new member of the GHMP kinase superfamily in plants,

- and comparison with galactose-1-phosphate kinase. *The Journal of biological chemistry* **284**, 21526-21535
28. Yang, T., Bar-Peled, Y., Smith, J. A., Glushka, J., and Bar-Peled, M. (2012) In-microbe formation of nucleotide sugars in engineered *Escherichia coli*. *Analytical biochemistry* **421**, 691-698
 29. York, W. S., Darvill, A. G., Mcneil, M., Stevenson, T. T., and Albersheim, P. (1986) Isolation and Characterization of Plant-Cell Walls and Cell-Wall Components. *Methods in enzymology* **118**, 3-40
 30. Li, Z., Hwang, S., and Bar-Peled, M. (2016) Discovery of a Unique Extracellular Polysaccharide in Members of the Pathogenic *Bacillus* That Can Co-form with Spores. *Journal of Biological Chemistry* **291**, 19051-19067

CHAPTER 5

CONCLUSION AND FUTURE DIRECTION

5.1 Introduction and Significance

Gram-positive bacteria of the genus *Bacillus* are universally distributed in many natural environments, although the primary habitat is soil. *B. cereus*, *B. anthracis* and *B. thuringiensis* are important Bacilli that belong to the *B. cereus* group and have profound impacts on agriculture and public health. Outbreaks of *B. cereus* contamination in ventilation, disinfectant and dialysis equipment in hospitals have been reported; and occasionally food poisoning incidents caused by *B. cereus* contaminated meat, fruits or vegetables have raised concerns about proper microbial control during food processing (1). The serious infectious disease anthrax caused by *B. anthracis* in livestock and potentially human (2) and the environmental friendly pesticide originated *B. thuringiensis* have a significant influence on agriculture worldwide (3).

Bacillus species are extensively studied, although much knowledge of *Bacillus cereus* group species is based on the model system *B. subtilis*, a model system for sporulation and biofilm studies (4,5). Even though *B. subtilis* share some similarity in the genome with *B. cereus* group species, gene organization is not always conserved and many putative genes in *B. cereus* group species were originated from other bacteria and archaea than *B. subtilis* (6). This raises questions about the applicability of existing knowledge to the *B. cereus* group. Many studies have shown that *B. cereus* group bacteria have different structures than *B. subtilis* in the spore exosporium and in secondary cell wall polysaccharides and biofilm etc. (7-10). As the result, there is significant interest in understanding the functions of *B. cereus* genes that are not present

in *B. subtilis*. My work on three operons (Pse, XNAC, C3CM) that only exist in *B. cereus* group bacteria, but not in *B. subtilis* has fulfilled the need of exploring unknown areas of *B. cereus* biology. And the novel glycans identified on spores, flagella and extracellular matrix exclusively for *B. cereus* group may have roles on the phenotypic difference observed between *B. subtilis* and *B. cereus* and such knowledge could be potentially used to the epidemic control and agricultural application of those species.

5.2 Summary of Results

5.2.1 Flagellin glycosylation in *B. thuringiensis* and Pse operon

Flagellin, the protein of the flagellum apparatus has known to be glycosylated with pseudaminic acid in Gram-negative bacteria, such as *Helicobacter pylori* (11). And the glycosylation has a profound role in motility and pathogenicity to hosts (12). Although one report showed that flagellin in *Bacillus* sp. PS3 is O-glycosylated, the sugar composition is still unknown. My preliminary data have indicated, for the first time that in *B. thuringiensis israelensis* the flagellin is also O-glycosylated to serine or threonine residues by a sugar with a mass consistent to a pseudaminic acid (diacetylamino-tetradecoxy-nonulosonic acid) (13).

The pseudaminic acid sugar residue is likely to be transported by a glycosyltransferase with a nucleotide sugar donor CMP-pseudaminic acid (CMP-Pse). This let us to identify an operon, named pse by a BLAST-search against the genes involved in CMP-Pse formation in *H. pylori* (14). The operon in *B. thuringiensis* consists seven genes: rbth_04253-rbth_04259. The biochemical characterization of this operon was achieved by recombinant protein expression in *E. coli* and purification, followed by in vitro enzymatic activity tests. The final product of this operon is indeed CMP-Pse, as confirmed by LC-MS and NMR spectrometry.

Interestingly, compared to the biosynthetic pathway of CMP-Pse in *H. pylori*, there is one additional intermediate product in the early steps of the enzymatic pathway in *B. thuringiensis* that has an uncommon structure of UDP-6-deoxy-D-GlcNAc-5,6-ene. Thus, more work was done focusing on the enzymatic characterization of the first two enzymes, named Pen and Pal. Biochemical analysis via LC-MS and NMR revealed that the first enzyme Pen is UDP-GlcNAc 4-oxidase, 5,6-dehydratase, and 4-reductase, and strongly bound to NADP⁺. It converts UDP-GlcNAc to UDP-6-deoxy-D-GlcNAc-5,6-ene. And the second enzyme Pal is NAD⁺-dependent and has distinct UDP-6-deoxy-D-GlcNAc-5,6-ene 4-oxidase, 5,6-reductase, and 5-epimerase activities. It converts the Pen product to UDP-4-keto-6-deoxy-L-AltNAc. Moreover, when Pen and Pal are co-incubating with substrate UDP-GlcNAc, additional product UDP-4-keto-6-deoxy-D-GlcNAc was produced which was not the substrate for the downstream enzyme of pse operon. The mechanism behind this activity is unknown.

5.2.2 pzX as exopolysaccharide in *B. cereus* group bacteria

An exopolysaccharide, named pzX was found in the species *Bacillus cereus*, *B. anthracis* and *B. thuringiensis* when they were growing in defined sporulation media. pzX was found outside the cell in the medium supernatant, and little was seen on the cell pellet. A combination of biochemical analysis including GC-MS, HPLC and NMR revealed that the structure of pzX contains trisaccharide repeating units of [XylpNAc(4OAc) α -1,3 GlcpNAcA(4OAc) α -1,3 XylpNAc]_n. The amount of pzX produced by the cell is tightly controlled by the nutrient level. Many sporulation media we have tried in the lab showed pzX production in the extracellular polysaccharide extract, however, the amount varies based on different recipes of media. The rich medium like BHI or LB did not yield pzX, rich undefined sporulation medium like DSM produced much less pzX, and medium with poor nutrient level (carbon source) also produced

less pzX. The defined media with a mediate level of nutrient produced the maximum amount of pzX.

Our lab has previously identified an operon, called XNAC operon, which is unique in *B. cereus* group bacteria. This operon produces UDP-XylNAc and UDP-GlcNAcA, which are the nucleotide sugar precursors for pzX formation. Mutants in this operon were generated (See Table 5.1) and the absence of pzX in those mutants was confirmed, suggesting that XNAC operon is indeed involved in the formation of pzX exopolysaccharide. The XNAC operon existed in nearly all species of *B. cereus* group. Although only a few strains have been tested for the appearance of pzX, we assume that all *B. cereus* group species with XNAC operon produce this exopolysaccharide.

Table 5.1. Mutants generated in this dissertation

Mutant name	Originated from	Locus	Operon
<i>Δpal</i>	<i>B. thuringiensis israelensis</i> ATCC 35646	rbth_04254	Pse operon (Chapter 2)
<i>Δphy</i>	<i>B. thuringiensis israelensis</i> ATCC 35646	rbth_04257	Pse operon (Chapter 2)
<i>Δbc0486</i>	<i>B. cereus</i> ATCC 14579	bc0486	XNAC operon (Chapter 3)
<i>Δbc0489</i>	<i>B. cereus</i> ATCC 14579	bc0489	XNAC operon (Chapter 3)
<i>Δbti46</i>	<i>B. thuringiensis israelensis</i> ATCC 35646	rbth_06246	XNAC operon (Chapter 3)
<i>Δbti47</i>	<i>B. thuringiensis israelensis</i> ATCC 35646	rbth_06247	XNAC operon (Chapter 3)
<i>Δbti48</i>	<i>B. thuringiensis israelensis</i> ATCC 35646	rbth_06248	XNAC operon (Chapter 3)
<i>Δbti49</i>	<i>B. thuringiensis israelensis</i> ATCC 35646	rbth_06249	XNAC operon (Chapter 3)
<i>Δbc3360</i>	<i>B. cereus</i> ATCC 14579	bc3360	C3CM operon (Chapter 4)
<i>Δbc3361</i>	<i>B. cereus</i> ATCC 14579	bc3361	C3CM operon (Chapter 4)

Since pzX was produced in sporulation medium, we suspected it had biological roles related to spores. We examined the timing of pzX production and found that it overlaid perfectly with the timing when mature spore released from the mother cell. More biological tests were conducted to show that pzX has surfactant activity which is able to assist spores moving faster; an adherence activity which helps spores to adhere surface (of artificial soil) better; and anti-aggregation property which prevents germinated spores from aggregation upon nutritional stress. However, the biological roles of pzX in the natural setting rather than a laboratory condition remain elusive.

5.2.3 Spore glycan in *B. cereus* and C3CM operon

Upon glycan analysis of spores in *Bacillus cereus* strain ATCC 14579, we discovered on the spores two rare sugar residues with a structure of 3-C-Me-6-deoxy-hexose. The unique carbon-methylation on the sugar ring makes the sugar labile and can only sustain mild hydrolysis during sugar alditol-acetate derivatization compared to other sugars like glucose or GlcNAc. These two sugars are 4-epimers to each other: with 3-C-Me-6-deoxy-gulose and 3-C-Me-6-deoxy-allose configuration. In general, the amount of 3-C-Me-6-deoxy-allose-containing glycans on *B. cereus* spores, based on GC-MS analyses, was very low and barely detected when compared with 3-C-Me-6-deoxy-gulose-containing glycans.

The analyses of these 3-C-Me-6-deoxy-hexoses and their unique C-methyl structures let us perform BLAST-search against known C-methyltransferases in other bacteria. This resulted the discovery of an operon called C3CM operon. C3CM operon contains four genes with a C-methyltransferase and it is able to produce the nucleotide sugar precursors (CDP-3-C-Me-6-deoxy-allose (named later CDP-cillose) and CDP-3-C-Me-6-deoxy-gulose), depending on the type of assays performed. When enzymes encoded by C3CM operon were co-expressed in *E.*

coli with an “in-microbe” analysis, the product was mainly CDP-cillose. On the other hand, when enzymes were individually expressed, purified and tested by *in vitro* assays, the product of C3CM operon was predominately CDP-3-C-Me-6-deoxy-gulose, as evident by LC-MS and NMR analysis. The mechanism behind these variances is unknown. Further gene expression of C3CM operon was conducted to confirm that it was only expressed in sporulation medium like Msgg but not rich non-sporulation medium BHI. *B. cereus* produced endogenously CDP-3-C-Me-6-deoxy-gulose as determined predominantly via LC-MS and GC-MS analysis followed by nucleotide sugar extraction. This might explain the abundant amount of CDP-3-C-Me-6-deoxy-gulose found on the spore glycan compared to cillose. Following the chemical structure of CDP-3-C-Me-6-deoxy-gulose we named it CDP-cereose because while this work was on its way a group (15) identified a sugar residue they named cereose (CDP-3-C-Me-6-deoxy-gulose) decorating the oligosaccharide- attached to BclA.

Mutants of this operon were also generated (Table 5.1) to confirm that no cereose observed in glycan isolated from this strain by the GC-MS analysis, suggesting that C3CM is indeed responsible for the formation of the activated sugar CDP-3-C-Me-6-deoxy-gulose that is incorporated into cereose-containing glycan in *B. cereus* spores. We have also examined several strains of *Bacillus* that harbor the C3CM operon and verified that they all have cereose-containing glycan in the spores. This suggests that the spores of other species in *Bacillus* harboring this operon are also likely decorated with cereose.

5.3 Future Direction

My work has shown that flagellin in *B. thuringiensis* is glycosylated with pseudaminic acid, which is possibly transported by CMP-Pse produced by pse operon. As previously described, the glycosylation of flagellin in Gram-negative bacteria is essential for motility and

infection to hosts. *B. thuringiensis* on the other hand is a pathogen to many insects by producing toxins during sporulation (16). Thus, it is prominent to investigate the role of pseudaminic acid glycosylation in the motility and host (insect) infection of *B. thuringiensis*. Mutants of two genes in this operon were generated for this work (Table 5.1). The absence of pseudaminic acid glycosylation in the mutant lines needs to be confirmed and more biological tests need to be done to further investigate the role of pseudaminic acid glycosylation in *B. thuringiensis*. Moreover, the specific glycosylation site on the flagellin protein and other potential sugar components besides pseudaminic acid also require further research.

As for the exopolysaccharide pzX, while the chemical structure was well characterized, the studies on the biological roles were conducted in the laboratory conditions with less biological relevance, thus it is needed to investigate the roles of pzX in a natural environment for example during infection. In some *Gram*-negative bacteria such as *Pseudomonas aeruginosa*, exopolysaccharide plays important roles in structure maintenance, antibiotic resistance of biofilm, and even virulence on hosts (17-20). As the result, it is prominent to test whether pzX has such roles in *Bacillus* species. In addition, due to the tight control of pzX production in response to surrounding nutrient, this may indicate that this exopolysaccharide is highly regulated by environmental changes, and possibly has a role in the transition between life forms, such as sporulation. In order to study the relationship between pzX to sporulation, More studies on the genetic regulation of XNAC operon will be required to provide developmental detail if its synthesis initiates at the mother cell or endospore compartment of sporulation cell.

Lastly, spores of *B. cereus* have shown to be glycosylated with two rare 3-C-methyl-deoxy sugars, cereose and cillose. Little is known for complete cereose- and cillose-containing glycan structures and the biological roles of these sugar modification. A study on spore

glycoprotein BclA in *Bacillus cereus* ATCC 14579 (15) identified cereose to be attached as a terminal sugar that is β -1,4-linked to 3-O-Me-Rha on the oligosaccharide of BclA. The BclA protein plays a central role in *B. anthracis* pathogenesis as it promotes interaction of spores with the host phagocytic cell as well as spore germination and bacterial outgrowth (21,22). However, the role of glycans on BclA on such functions is unknown. Thus, it is necessary to study the role of cereose glycosylation on BclA and examine whether it contribute to any spore surface recognition during infection or germination. Moreover, upon the analysis of spores of BclA mutants, cereose was still present by the GC-MS analysis, indicating that cereose may not be specific for BclA. As the result, it is prominent to investigate the location and full structure of cereose containing glycan in order to study the biological function of cereose glycosylation.

5.4 References

1. Bottone, E. J. (2010) *Bacillus cereus*, a Volatile Human Pathogen. *Clinical microbiology reviews* **23**, 382-+
2. Goel, A. K. (2015) Anthrax: A disease of biowarfare and public health importance. *World journal of clinical cases* **3**, 20-33
3. Lemaux, P. G. (2008) Genetically engineered plants and foods: A scientist's analysis of the issues (Part I). *Annual review of plant biology* **59**, 771-812
4. Driks, A. (2002) Maximum shields: the assembly and function of the bacterial spore coat. *Trends in microbiology* **10**, 251-254
5. Vlamakis, H., Chai, Y., Beauregard, P., Losick, R., and Kolter, R. (2013) Sticking together: building a biofilm the *Bacillus subtilis* way. *Nature reviews. Microbiology* **11**, 157-168
6. Okstad, O. A., Hegna, I., Lindback, T., Rishovd, A. L., and Kolsto, A. B. (1999) Genome organization is not conserved between *Bacillus cereus* and *Bacillus subtilis*. *Microbiol-Uk* **145**, 621-631
7. Imamura, D., Kuwana, R., Takamatsu, H., and Watabe, K. (2011) Proteins Involved in Formation of the Outermost Layer of *Bacillus subtilis* Spores. *Journal of bacteriology* **193**, 4075-4080
8. Hong, H. A., Khaneja, R., Tam, N. M. K., Cazzato, A., Tan, S., Urdaci, M., Brisson, A., Gasbarrini, A., Barnes, I., and Cutting, S. M. (2009) *Bacillus subtilis* isolated from the human gastrointestinal tract. *Research in microbiology* **160**, 134-143
9. Epstein, A. K., Pokroy, B., Seminara, A., and Aizenberg, J. (2011) Bacterial biofilm shows persistent resistance to liquid wetting and gas penetration. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 995-1000
10. Swoboda, J. G., Campbell, J., Meredith, T. C., and Walker, S. (2010) Wall Teichoic Acid Function, Biosynthesis, and Inhibition. *Chembiochem : a European journal of chemical biology* **11**, 35-45

11. Schirm, M., Arora, S. K., Verma, A., Vinogradov, E., Thibault, P., Ramphal, R., and Logan, S. M. (2004) Structural and genetic characterization of glycosylation of type a flagellin in *Pseudomonas aeruginosa*. *Journal of bacteriology* **186**, 2523-2531
12. Parker, J. L., Day-Williams, M. J., Tomas, J. M., Stafford, G. P., and Shaw, J. G. (2012) Identification of a putative glycosyltransferase responsible for the transfer of pseudaminic acid onto the polar flagellin of *Aeromonas caviae* Sch3N. *MicrobiologyOpen* **1**, 149-160
13. Li, Z., Hwang, S., Ericson, J., Bowler, K., and Bar-Peled, M. (2015) 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*. *Journal of Biological Chemistry* **290**, 691-704
14. Schoenhofen, I. C., McNally, D. J., Brisson, J. R., and Logan, S. M. (2006) Elucidation of the CMP-pseudaminic acid pathway in *Helicobacter pylori*: synthesis from UDP-N-acetylglucosamine by a single enzymatic reaction. *Glycobiology* **16**, 8C-14C
15. Maes, E., Krzewinski, F., Garenaux, E., Lequette, Y., Coddeville, B., Trivelli, X., Ronse, A., Faille, C., and Guerardel, Y. (2016) Glycosylation of BclA Glycoprotein from *Bacillus cereus* and *Bacillus anthracis* Exosporium Is Domain-specific. *The Journal of biological chemistry* **291**, 9666-9677
16. Schnepf, E., Crickmore, N., Van Rie, J., Lereclus, D., Baum, J., Feitelson, J., Zeigler, D. R., and Dean, D. H. (1998) *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiol Mol Biol R* **62**, 775-+
17. Flemming, H. C., Neu, T. R., and Wozniak, D. J. (2007) The EPS matrix: The "House of Biofilm cells". *Journal of bacteriology* **189**, 7945-7947
18. Ma, L., Jackson, K. D., Landry, R. M., Parsek, M. R., and Wozniak, D. J. (2006) Analysis of *Pseudomonas aeruginosa* conditional Psl variants reveals roles for the Psl polysaccharide in adhesion and maintaining biofilm structure postattachment. *Journal of bacteriology* **188**, 8213-8221
19. Laue, H., Schenk, A., Li, H., Lambertsen, L., Neu, T. R., Molin, S., and Ullrich, M. S. (2006) Contribution of alginate and levan production to biofilm formation by *Pseudomonas syringae*. *Microbiol-Sgm* **152**, 2909-2918
20. Franklin, M. J., Nivens, D. E., Weadge, J. T., and Howell, P. L. (2011) Biosynthesis of the *Pseudomonas aeruginosa* Extracellular Polysaccharides, Alginate, Pel, and Psl. *Frontiers in microbiology* **2**, 167
21. Oliva, C. R., Swiecki, M. K., Griguer, C. E., Lisanby, M. W., Bullard, D. C., Turnbough, C. L., and Kearney, J. F. (2008) The integrin Mac-1 (CR3) mediates internalization and directs *Bacillus anthracis* spores into professional phagocytes. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 1261-1266
22. Oliva, C., Turnbough, C. L., and Kearney, J. F. (2009) CD14-Mac-1 interactions in *Bacillus anthracis* spore internalization by macrophages. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 13957-13962