Glycosylation of BclA Glycoprotein from Bacillus cereus and Bacillus anthracis Exosporium Is Domain-specific

Emmanuel Maes, Frédéric Krzewinski, Estelle Garénaux, Yannick Lequette, Coddeville Bernadette, Xavier Trivelli, Annette Ronse, Christine Faille, Yann Guerardel

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 291, NO. 18, pp. 9666–9677, April 29, 2016

Glycosylation of BclA Glycoprotein from *Bacillus cereus* and *Bacillus anthracis* Exosporium Is Domain-specific*[S]

Received for publication, January 27, 2016, and in revised form, February 24, 2016 Published, JBC Papers in Press, February 26, 2016, DOI 10.1074/jbc.M116.718171

Emmanuel Maes[‡], Frederic Krzewinski[‡], Estelle Garenaux[‡], Yannick Lequette[§], Bernadette Coddeville[‡], Xavier Trivelli[‡], Annette Ronse[§], Christine Faille[§], and Yann Guerardel[‡1]

From the [‡]Université de Lille, CNRS, UMR 8576-UGSF-Unité de Glycobiologie Structurale et Fonctionnelle, F-59000 Lille, France and [§]INRA, UMR8207, F-59650 Villeneuve d'Ascq, France

The spores of the *Bacillus cereus* group (*B. cereus*, *Bacillus anthracis*, and *Bacillus thuringiensis*) are surrounded by a paracrystalline flexible yet resistant layer called exosporium that plays a major role in spore adhesion and virulence. The major constituent of its hairlike surface, the trimerized glycoprotein BclA, is attached to the basal layer through an N-terminal domain. It is then followed by a repetitive collagen-like neck bearing a globular head (C-terminal domain) that promotes glycoprotein trimerization. The collagen-like region of *B. anthracis* is known to be densely substituted by unusual *O*-glycans that may be used for developing species-specific diagnostics of *B. anthracis* spores and thus targeted therapeutic interventions. In the present study, we have explored the species and domain specificity of BclA glycosylation within the *B. cereus* group. First, we have established that the collagen-like regions of both *B. anthracis* and *B. cereus* are similarly substituted by short *O*-glycans that bear the species-specific deoxyhexose residues anthrose and the newly observed cereose, respectively. Second we have discovered that the C-terminal globular domains of BclA from both species are substituted by polysaccharide-like *O*-linked glycans whose structures are also species-specific. The presence of large carbohydrate polymers covering the surface of *Bacillus* spores may have a profound impact on the way that spores regulate their interactions with biotic and abiotic surfaces and represents potential new diagnostic targets.

All *Bacillus* spores share a common architecture that consists of a set of concentric layers with a nucleotide-containing inner core surrounded by a peptidoglycan cortex and a spore coat. Species may differ according to the presence or not of an additional loosely fitting envelope enclosing individual spores. Indeed, mature spores from the *Bacillus cereus* group, which includes *B. cereus*, *Bacillus anthracis*, and *Bacillus thuringiensis*, are surrounded by a protein-rich, flexible envelope called the exosporium, whereas other species such as *Bacillus subtilis* and *Bacillus licheniformis* are not (1–3). Exosporia appear as a semipermeable barrier and are thought to exhibit a wide range of functions, including resistance to chemical and enzymatic treatments (4), enhancement of spore adhesion to biotic (5) and abiotic surfaces (6, 7), and germination (8). Some of the properties of the exosporium may also exert an influence on the infection process of pathogenic *B. anthracis* strains by protecting spores from macrophage-induced degradation (4, 9) and by targeting them toward phagocytic cells (5). As revealed by electron microscopy, the exosporium is made up of an external hairlike nap that sits on top of a paracrystalline basal layer (10, 11). The supporting basal layer of *B. anthracis* was shown to contain about 20 different proteins (12, 13) that form hexagonal subunits to which the hairlike structures are attached (11, 14, 15). Filaments are mostly composed of trimers of the major BclA glycoprotein (for *Bacillus* collagen-like protein of *anthracis*) that are associated with the underlying layer through their N-terminal region (16, 17). In all species of the *B. cereus* group, the BclA glycoprotein contains three domains: the N-terminal domain that generally contains 44 amino acids through which BclA is anchored to the basal layer; an extended highly polymorphic collagen-like region (CLR)[2] composed of G*XX* repeats, most of which are GPT; and the globular C-terminal domain (CTD) forming the distal end of each filament that is composed of multiple β-strands and known to promote the protein trimerization (16). One of the most remarkable features of BclA is its glycosylation pattern, which is characterized by the presence of densely packed *O*-linked glycans substituting the repeat unit of the CLR (18). Although glycans are known to play fundamental roles in the interaction of spores with surfaces and in the conformation of its surface glycoproteins (7, 19, 20), very little information is currently available regarding the molecular nature of the carbohydrate moiety of the *Bacillus* spore surface. The only exception concerns the CLR-associated oligosaccharides of *B. anthracis*. Indeed, the structures of these glycans have only been firmly established in *B. anthracis* as two major tri- and pentasaccharides with the sequences 3-*O*-Me-Rha(α1–2)Rha(α1–3)GalNAc and Ant(β1–3)Rha(α1–3)Rha-(α1–2)Rha(α1–3)GalNAc (18, 21). Although the exact se-

* This work was supported by the "Agence Nationale de la Recherche" under the "Programme National de Recherche en Alimentation et Nutrition Humaine" Project ANR-07-PNRA-009, InterSpore. The authors declare no competing financial interest.

[S] This article contains supplemental Figs. S1–S3 and Table S1.

[1] Supported by Très Grand Equipement Résonance Magnétique Nucléaire Très Hauts Champs Fr3050. To whom correspondence should be addressed. Tel.: 33-3-20-43-69-41; Fax: 33-3-20-43-55-65; E-mail: yann.guerardel@univ-lille1.fr.

[2] The abbreviations used are: CLR, collagen-like region; CTD, C-terminal domain; Rha, rhamnose; GalNAc-ol, *N*-acetylgalactosaminitol; GlcNAc, *N*-acetylglucosamine; CT, C-terminal; T, tesla; TOCSY, total correlation spectroscopy; ROESY, rotating frame nuclear Overhauser enhancement spectroscopy; HSQC, heteronuclear single quantum correlation; HMWG, high molecular weight glycan(s); LMWG, low molecular weight glycan(s); Bc, *B. cereus*; Hex, hexose; HexNAc, *N*-acetylhexosamine; HexNAc-ol, *N*-acetylhexosaminitol; HMBC, *heteronuclear multiple bond correlation; Cro,* cereose; Ba, *B. anthracis*; Ant, anthrose; Gul, gulose; dGul, deoxygulose; dHex, deoxyhexose.

quence of potential oligosaccharides in other members of the *B. cereus* group is yet unknown, composition analyses on a wide panel of spores of *Bacillus* species have established the presence of rhamnose and/or $GalNH_2$ derivatives in the exosporia of nine strains of *B. cereus*, two of *B. thuringiensis*, and of the avirulent *B. anthracis* 9131 strain, which lacks pXO1 and pXO2 plasmids. Similar residues were also observed on the spore surface of non-exosporium-producing species including *Bacillus subtilis*, *Bacillus sporothermodurans*, and *B. licheniformis*. Further composition analyses of *B. cereus* ATCC 14579 exosporia containing BclA glycoproteins truncated in their different domains have established that the glycosylation patterns of CLR and CTD differ (22). Indeed, whereas glycosylation of CLR is restricted to 3-*O*-Me-Rha, Rha, and GalNAc, CTD showed a more complex glycosylation pattern characterized by the additional presence of 2-*O*-Me-Rha and 2,4-*O*-Me-Rha.

Both CTD and CLR domains were shown to modulate the general physicochemical properties of the spore surface and interactions between spore and abiotic surfaces such as stainless steel as their deletion results in an increased resistance of adherent spores to detachment (7, 23). Conversely, CTD and CLR seem to exert antagonistic influences on the spore hydrophobicity. Thus, the exact mechanisms through which BclA influences spore adhesiveness to surfaces are yet to be established. Among these factors, surface glycosylation is known to exert a strong influence on cellular interactions either through its intrinsic physicochemical properties or its specific ligand-receptor recognition (24).

In the present study, we have investigated the domain-specific glycosylation of BclA by analyzing the structure and composition of glycan moieties of wild type (WT) BclA and BclA deleted of the CT domain. We have purified glycan fractions differentially associated with each domain using a set of *B. cereus* strains that express truncated BclA proteins and then established that they exhibit different structures and organizations. Furthermore, we have shown that not only *B. cereus* but also *B. anthracis* synthesize both CLR and CT domain-specific polysaccharide-like compounds.

Experimental Procedures

Growth of Bacteria and Generation of Mutants—The *B. cereus* ATCC 14579 wild type and related mutants and *B. anthracis* strains are listed in Table 1. Spores were produced as described previously (6). When over 95% mature spores were obtained, spores were scraped from the agar surface, washed five times with water at 4 °C, and stored at 4 °C in sterile water until use. Before each experiment, two additional washing steps were carried out to remove most of the cell debris, and potential spore aggregates were disrupted by a sonication step (bath sonicator; 1 min 30 s twice at 42 kHz). All experiments were carried out on at least two independent spore batches. The following antibiotic concentrations were used when necessary: tetracycline at 10 $\mu g\ ml^{-1}$, kanamycin at 200 $\mu g\ ml^{-1}$, and erythromycin at 10 $\mu g\ ml^{-1}$.

Purification of Glycans—Exosporia were isolated from spores as described previously (14). Briefly, spores were washed twice in water, resuspended in water, and subjected to four successive passages through a French press at 20,000 p.s.i. Spores were eliminated by centrifugation (3000 × g, 30 min, 4 °C). Insoluble fractions of exosporia were pelleted by ultracentrifugation (120,000 × g, 30 min, 4 °C). Glycan moieties were released from purified exosporia by reductive β-elimination as described previously (25). Briefly, samples were incubated for 72 h at 37 °C in 100 mM NaOH containing 1 M $NaBH_4$. The reaction was stopped by the addition of Dowex 50 × 8 (25–50 mesh, H^+ form) at 4 °C until reaching pH 6.5. After filtration and complete evaporation, boric acid was eliminated by repetitive co-distillation in the presence of methanol. The oligosaccharides alditols were submitted to cationic exchange chromatography on Dowex 50 × 2 (200–400 mesh, H^+ form) in water to separate acid and neutral oligosaccharides. Oligosaccharide alditols were desorbed in water, and pH was adjusted to 7 with ammonium hydroxide. The fractions containing sugars were desalted on a Bio-Gel® P2 column and freeze-dried for further analysis.

TABLE 1
List of wild type and related mutant strains of *B. cereus* ATCC 14579 and *B. anthracis*
aa, amino acids.

Strains and plasmids	Description	Ref.
B. cereus ATCC 14579	Wild-type strain	
ATCC 14579 Δ*bclA* Δ*exsH*	Kan^r Tet^r; *bclA*::*kan* and *exsH*::*tet* deletions	7
B. anthracis 9131	The *B. anthracis* Sterne plasmidless strain obtained by curing RP31 of pXO1 Y	54
pHT304 *bclA* ATCC 14579	Amp^r Ery^r; plasmids containing the *bclA* gene and 317 bp upstream and 138 bp downstream; used for complementation	7
pYL304	Amp^r Ery^r; plasmids expressing a BclA protein of 189 aa with the CTD deleted; *bclA*-ΔCT	7

Separation of Oligosaccharides—Released glycans were fractionated by gel filtration chromatography on an ÄKTA purifier apparatus (GE Healthcare) fitted with a ToyoPearl HW40 column (1 × 40 cm) irrigated in 0.1% acetic acid. Elution was monitored by UV spectroscopy at 206 nm, and the glycan-containing fractions identified by orcinol staining were collected and pooled in one excluded and one included fraction. Oligosaccharides from the included fraction were further submitted to HPLC separation on a C_{18} column (C18 Gemini, 250 × 4.6 mm, 5 μm; Phenomenex) using a gradient of acetonitrile, water, and 0.1% trifluoroacetic acid (TFA) and flow rate of 0.8 ml/min. Oligosaccharide alditols were detected by UV spectroscopy at 206 nm.

Mass Spectrometry Analysis—MALDI-TOF mass spectra were acquired on a voyager Elite DE-STR mass spectrometer (Perspective Biosystems, Framingham, MA) in the reflectron positive mode by delayed extraction using an acceleration mode of 20 kV, a pulse delay of 200 ns, and grid voltage of 66%. Samples were prepared by mixing on the target 1 μl of oligosaccharide solution (1–5 pmol) with 1 μl of 2,5-dihydroxybenzoic acid matrix solution (10 mg/ml in CH_3OH/H_2O, 50:50, v/v). Between 50 and 100 scans were averaged for each spectrum. Nano-electrospray ionization-tandem mass spectrometry fragmentation analyses were performed using a Q-STAR pulsar quadrupole time-flight mass spectrometer (Applied Biosystems/MDS Sciex, Toronto, Canada) fitted with a nanoelectrospray ion source (Protana, Odense, Denmark). Glycans dissolved in a solution of 50% methanol and 1% formic acid (1 pmol/μl) were sprayed from gold-coated "medium length"

borosilicate capillaries (Protana). An 800-V current was applied to the capillary tip. For the generation of MS/MS data, the precursor ion was selected by the quadrupole and subsequently fragmented in the collision cell using nitrogen at a pressure of 5.3×10^{-5} torr and appropriate collision energy. The collision-induced dissociation spectra were recorded by the orthogonal TOF analyzer over the mass range m/z 50–1000.

Composition Analysis—The monosaccharide composition of the exosporium fraction was established by GC and GC/MS as alditol acetate derivatives. Briefly, samples were hydrolyzed in 4 M TFA for 4 h at 100 °C and then reduced with sodium borohydride in 0.05 M NaOH for 4 h. Reduction was stopped by dropwise addition of acetic acid until pH 6 was reached, and borate salts were co-distilled by repetitive evaporation in dry methanol. Peracetylation was performed in acetic anhydride at 100 °C for 2 h. All monosaccharide derivatives were identified according to their specific retention times and the electronic impact-MS fragmentation patterns of individual derivatives. Individual monosaccharides were quantified by comparison with a myoinositol internal standard according to response factors established in the laboratory.

Nuclear Magnetic Resonance Analysis—Samples were solubilized in highly enriched deuterated water (99.96% deuterium atom; EurisoTop®, St-Aubin, France) and lyophilized; this operation was repeated twice. Experiments were recorded on 9.4-T (Plateforme Résonance Magnétique Nucléaire, Université Lille 1), 14.1-T (Institut Pasteur de Lille), and 21.4-T spectrometers (Unité de Glycobiologie Structurale et Fonctionnelle, Infrastruture de Recherche-Très Hauts Champs-Résonance Magnétique Nucléaire, CNRS) where protons resonate at 400, 600, and 900 MHz and ^{13}C resonates at 100, 151, and 250 MHz, respectively. The 9.4-T spectrometer was equipped with 5-mm triple broadband inverse (*i.e.* ^{1}H, ^{13}C, *X*) probe head with a z-gradient. The 14.1-T spectrometer was equipped with 5-mm quadruple resonance cryoprobe inverse (QCI) cryoprobe head with ^{1}H, ^{2}H, ^{19}F, ^{13}C cooled channels and a ^{15}N channel with z-gradients. The 21.4-T spectrometer was equipped with 5-mm triple resonance cryoprobe inverse (TCI) cryoprobe with ^{1}H, ^{2}H, ^{13}C cooled channels and a ^{15}N channel with a z-gradient. Moreover, the latter magnet was equipped with a sample jet robot. All samples were put in 5-mm tubes matched for D_2O. Acetone was added as an internal standard, starting from a solution of 2.5 μl of acetone in 10 ml of D_2O. All pulse sequences are taken from the Bruker library of pulse programs and then optimized for each sample. Spectral widths were 12 and 200 ppm for proton and carbon observations, respectively. TOCSY was achieved with various mixing times from 40 to 120 ms, and ROESY spectra were recorded with a 300-ms mixing time. Edited ^{1}H-^{13}C HSQC were recorded with 1536 data points for detection and 256 data points for indirect direction.

Results

CLR and CT Domains of Bc-BclA Are Differentially Glycosylated—We first investigated the saccharide composition of the glycan moiety of WT-BclA (with both CLR and CT domains). Glycan moieties were released from the protein backbones by reductive β-elimination and desalted by a combination of cation exchange and gel filtration chromatography (Bio-Gel P2) according to well established protocols (25). This method generates reduced oligo- or polysaccharides that are protected from the so-called peeling reaction as previously observed for the hydrazinolysis-released oligosaccharides from the CLR of *B. anthracis* BclA that lacked the terminal GalNAc residue (18). Composition analysis showed that the free carbohydrate fraction isolated from WT-BclA spores was composed of a mixture of 2,4-*O*-Me-Rha, 2-*O*-Me-Rha, 3-*O*-Me-Rha, Rha, and GalNAc (Fig. 1*A*). These monosaccharides are in similar proportions to those observed from acid hydrolysis of the untreated exosporia (Fig. 1*A*), which strongly suggests that β-elimination quantitatively released carbohydrate moieties from glycoproteins. The released glycans were then separated according to their relative molecular weights by gel filtration chromatography (ToyoPearl HW40) into two distinct fractions: one excluded fraction called high molecular weight glycans (HMWG) and one included fraction called low molecular weight glycans (LMWG). As seen in Fig. 1*B*, monosaccharide analysis established that LMWG and HMWG do not only differ in size but also in composition.

In similar experimental conditions, spores expressing a BclA homologue that is truncated in the CT domain (ΔCT-Bc) produced a LMWG with a composition identical to that of the WT-Bc but did not show any HMWG (Fig. 1*C*). Altogether, these data established that the CLR domain of Bc-BclA was substituted by small glycans composed of GalNAc, Rha, and 3-*O*-Me-Rha, whereas the CT domain carried large glycans with a more complex monosaccharide composition. The observation of LMWG associated with the Bc-BclA CLR domain shows similarities to the previously identified tri- and tetrasaccharides on the CLR domain of *B. anthracis* BclA (18), whereas the presence of a large rhamnosylated polysaccharide on the CT domain has, to our knowledge, yet to be documented.

B. cereus CLR Is Substituted by Original Rhamnosylated Oligosaccharides—To understand the structural basis of the differential glycosylation of both domains, the structures of LMWG and HMWG were further studied by a combination of HPLC, mass spectrometry, and NMR. MALDI-MS profiling of total glycans from LMWG showed two major signals at m/z 552.3 and 712.4 as well as a minor signal at 509.3 that were tentatively attributed to oligosaccharides (Fig. 2*A*). To establish their structures, LMWG isolated from WT-Bc were separated by HPLC on a C_{18} reverse phase column using a gradient of water and acetonitrile. As observed in Fig. 2*B*, three major peaks were observed at 11.0, 17.8, and 33.3 min. Corresponding molecules were collected and freeze-dried for further analysis. Following analysis by MS and NMR, the signals at 11.0 and 33.3 min were identified as reduced glycans, whereas the signal at 17.8 min was identified as cytidine based on its NMR parameters (supplemental Table S1 and supplemental Fig. S1). MALDI-MS spectra of compound eluting at 11.0 min showed two $[M + Na]^+$ and $[M + K]^+$ signals at m/z 552.3 and 568.2 (Fig. 2*C*) that correspond to an oligosaccharide with the tentative $HexNAc_1dHex_2Me_1$-ol composition. *Electrospray* MS/MS sequencing of the signal at 552 generated a set of Y-type ions at m/z 246 and 392 and C ions at m/z 183 and 329 that established the sequence (Me)dHex-dHex-HexNAc-ol (Fig. 2*D*). A series of homo- and heteronuclear NMR experiments were performed

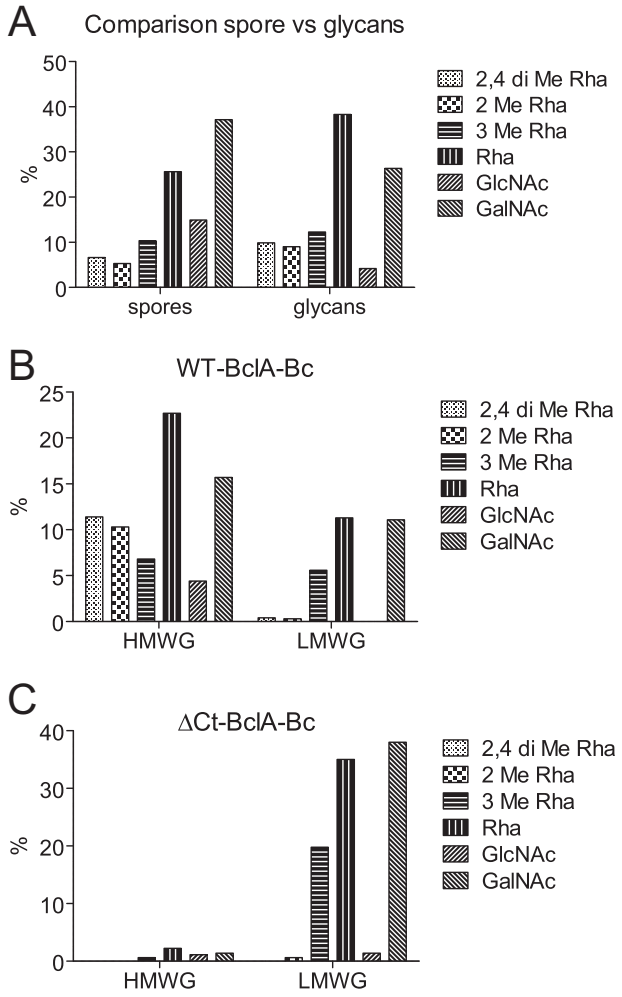


FIGURE 1. **Monosaccharide composition analysis of *B. cereus* exosporia.** Exosporia were purified from *B. cereus* strains inhibited in the synthesis of major surface glycoproteins (BclA and ExsH) in which WT or CT-truncated BclA-Bc are expressed. The glycan moieties of recombinant BclA were purified and analyzed by GC-MS. *A*, monosaccharide composition of exosporia (*left*) and total glycan moiety; *B*, monosaccharide composition of HMWG and LMWG fractions from WT-BclA-Bc; *C*, HMWG and LMWG fractions from Δ-CT-BclA-Bc.

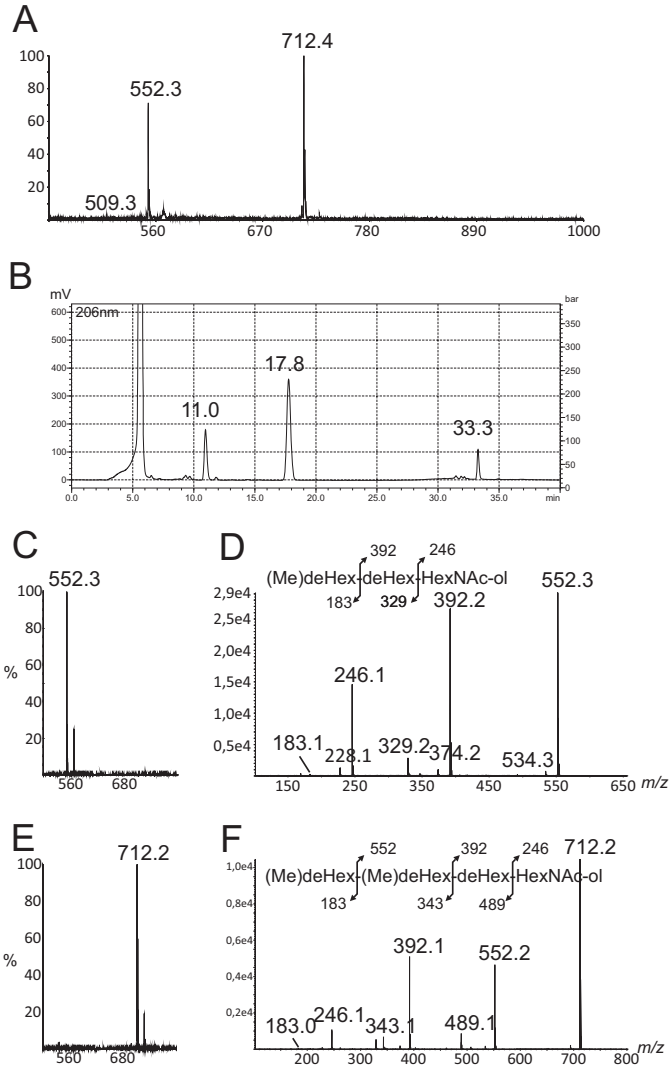


FIGURE 2. **MS sequencing of major oligosaccharides of LMWG fraction from WT-BclA-Bc.** Shown are MALDI-MS analysis (*A*) and reverse phase HPLC separation (*B*) of oligosaccharides of WT-BclA-Bc LMWG, MALDI-MS (*C*) and MALDI-MS/MS (*D*) of oligosaccharide at *m/z* 552 from reverse phase HPLC fraction at retention time 11.0 min, and MALDI-MS (*E*) and MALDI-MS/MS (*F*) of oligosaccharide at *m/z* 712.2 from reverse phase HPLC fraction at retention time 33.3 min. *deHex*, deoxyhexose.

to elucidate this sequence. Two anomeric proton/carbon couples were identified at δ 5.03/102.0 (I) and δ 5.05/103.6 (II) ppm from the ^{1}H-^{13}C HSQC spectrum (Fig. 3), indicating the presence of two monosaccharide residues. Moreover, the observation of a proton at 4.35 ppm (ol-2) carried by a carbon-bearing nitrogen at 53.1 ppm and a pseudotriplet at δ 3.86/71.4 (ol-5) typified a HexNAc-ol residue. The identification was further supported by the observation of two signals at 3.99/79.4 and 3.62/70.8 ppm attributed to H3 and H4, respectively, of the *N*-acetylgalactosaminitol (GalNAc-ol) residue. An *N*-acetyl group signal was identified at 2.05/23.3 ppm and easily attributed to a GalNAc-ol residue. Spin systems of the two identified monosaccharides and their respective coupling constants established that they had an α-*manno* configuration (Table 2). Indeed, $^{3}J_{1,2}$, $^{3}J_{2,3}$, $^{3}J_{3,4}$, and $^{3}J_{4,5}$ were small, small, large, and large, respectively (26). The α-anomery of these two units was confirmed by observation of an intraresidual strong NOE effect between H1 and H2 and the typical chemical shifts of its H5/C5 around 3.8/71 ppm. Finally, two methyl groups resonating like doublets were identified at 1.30/18.0 ppm and were correlated with H5 of α-*manno* configuration-related sugars. These observations indicated that both monosaccharides were α-rhamnosyl residues. An additional singlet with a relative intensity of 3 compared with H1 signals was observed at 3.44/57.2 ppm. It is characteristic of an *O*-methyl group. This group is attached to the O3 of the Rha II in a non-reduced terminal position as proved by its deshielded C3 at 80.6.ppm, which defined residue II as 3-*O*-Me-Rha, also called acofriose (27). Finally, GalNAc-ol and Rha I residues were shown to be substituted in positions 3 and 2, respectively, as deduced from the deshielded values of GalNAc-ol C3 at δ 79.4 and Rha I C2 at δ 79.7. Altogether, data from MS and NMR established the structure of this compound as 3-*O*-Me-Rha(α1–2)Rha(α1–3)GalNAc-ol.

In a similar manner, MALDI-MS spectra of the compound eluting at 33.3 min showed two $[M + Na]^+$ and $[M + K]^+$ signals at *m/z* 712.2 and 728.2 indicative of an oligosaccharide

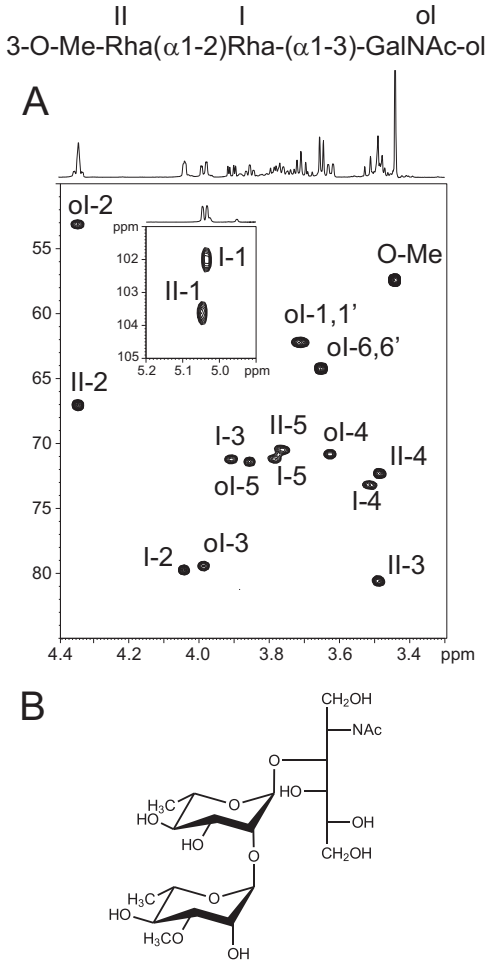


FIGURE 3. **NMR analysis of major trisaccharide of LMWG fraction from WT-BclA-Bc.** *A*, ^{1}H-^{13}C HSQC spectrum of trisaccharide from WT-BclA-Bc LMWG; *B*, deduced structure of the trisaccharide.

with the $dHex_3Me_2$-$HexNAc_1$-ol composition (Fig. 2*E*). Electrospray MS/MS sequencing of signal at *m*/*z* 712 generated a set of Y-type ions at *m*/*z* 246, 392, and 552 and C-type ions at *m*/*z* 183, 343, and 489 that typify the (Me)dHex-(Me)dHex-dHex-HexNAc-ol sequence (Fig. 2*F*). In addition to a GalNAc-ol residue, the one-dimensional and two-dimensional NMR spectra of this compound showed three ^{1}H/^{13}C anomeric signals at δ 5.03/102.0 (I), 5.05/103.3 (II), and 4.79/103.9 (III), which confirmed that this was a tetrasaccharide (Fig. 4*A* and Table 3). Residues I and II exhibited very similar spin systems to those of the trisaccharide and were identified as α-Rha and α-3-*O*-Me-Rha residues, respectively. However, the chemical shift of the C4 position of III was deshielded to 3.7/80.7 ppm, which established that residue III was further substituted in its C4 position (Fig. 4*D* and Table 3). In addition, the tetrasaccharide spectra showed four signals associated with methyl groups. Three resonated as doublets and were assigned to C6 methyl groups of dHex residues at δ 1.30/18.0 (I-C6), δ 1.37/18.4 (II-C6), and δ 1.22/16.8 (III-H6), whereas one resonated as a singlet at δ 1.36/23.5 (III-H7) (Fig. 4*B*). On the homonuclear ^{1}H-^{1}H TOCSY spectrum, the anomeric proton III-1 was connected to a doublet III-2 ($^{3}J_{1,2}$ ~ 8 Hz) but showed no other correlation. The 6-methyl group at δ 1.36/23.5 was correlated with two protons, III-5 and III-4, at 4.18 and 3.27 ppm, respectively. No additional correlation was observed, thereby suggesting that carbon 3 was not protonated. Taking account of the molecular weight and the fragmentation, this unit was identified as a methylated deoxyhexose carrying a *C*-methyl group in the C3 position. The chemical shift of this methylated group strongly suggested that it is linked to a hydroxylated carbon. Moreover, the C3 quaternary carbon at δ 74.9 can be observed through the $^{2}J_{H,C}$ starting from both H4 and *C*-methyl group on the ^{1}H-^{13}C HMBC spectrum (Fig. 4*C*). Taken together, these data strongly suggested the presence of a quaternary carbon in C3 substituted by the supplementary methyl group. The configuration of unit III has been established on the basis of $^{3}J_{H,H}$ and NOE effect. The strong vicinal coupling constant (8 Hz) between III-1 and III-2 and the strong dipolar NOE effect between III-1 and III-5 (supplemental Fig. S2) indicate that H1, H2, and H5 were in axial positions. The very small vicinal coupling constant (<1 Hz) between H4 and H5 established that H4 was in an equatorial position. The observation of a strong NOE effect between methyl group (III-7) carried by quaternary carbon (III-3) and both H2 and H4 indicated that -CH_3 was in an equatorial position. Altogether, these observations established that unit III exhibits a *β-gulo* conformation (*R*, *R*, *S*, *R*) with two methyl groups linked to C5 and C3 in place of the proton, which defines it as a 6-deoxy-3-*C*-methylgulose. This molecule is an epimeric molecule of evalose (6-deoxy-3-*C*-methyl-D-mannopyrannose) and 6-deoxy-3-*C*-methyl-D-talopyrannose or vinelose and to our knowledge is described here for the first time (28, 29). Its optical isomery, D or L, remains unresolved. This can only be unambiguously resolved by synthesizing and analyzing the two different isomers. We gave this novel sugar the trivial name cereose (Cro; 3-*C*-Me-6-dGul). The final sequence was established through ROESY and HMBC experiments. In particular, the NOE contacts (supplemental Fig. S2) observed among III-1, II-3, and II-4 clearly established that unit III is linked to unit II. Then an HMBC experiment showed a $^{3}J_{H,C}$ correlation between III-H1 and II-C4, indicating that unit III is linked to unit II in the C4 position (data not shown). Moreover, the deshielding of C4 of unit II confirmed that it is *O*-4-substitued. Data from MS and NMR established the structure of this compound as 3-*C*-Me-6-dGul(β1–4)3-*O*-Me-Rha(α1–2)Rha(α1–3)GalNAc-ol or Cro(β1–4)3-*O*-Me-Rha(α1–2)Rha(α1–3)GalNAc-ol. As a whole, content analysis of the LMWG fraction isolated from WT-BclA-Bc established that it contains two major oligosaccharides: one trisaccharide, 3-*O*-Me-Rha(α1–2)Rha(α1–3)GalNAc-ol, and one tetrasaccharide, Cro(β1–4)3-*O*-Me-Rha(α1–2)Rha(α1–3)GalNAc-ol, containing a so-far unique monosaccharide. Detailed MS/MS analysis of the total fraction prior to HPLC separation also permitted us to identify a minor trisaccharide ($[M + Na]^{+}$ at *m*/*z* 509) with the partial sequence Me-dHex-Me-dHex-dHex-ol (data not shown). The very low quantities in which this compound is observed deterred us from fine structural analysis. Based on the fine structure of the two major polysaccharides described, one can infer its structure either as Cro(β1–4)3-*O*-Me-Rha(α1–2)Rha-ol or as 3-*O*-Me-Rha(α1–2)3-*O*-Me-Rha(α1–2)Rha-ol.

A similar experimental approach used on the LMWG fraction isolated from the ΔCT-Bc showed an identical set of oligo-

TABLE 2
Proton and carbon chemical shifts of major trisaccharide 3-O-Me-Rha(α1–2)Rha(α1–3)GalNAc-ol isolated from LMWG WT-BclA-Bc (Fig. 3)
Bold values indicate substitution positions.

		1	2	3	4	5	6,6′	Add -CH3
GalNAc-ol (ol)	^{1}H	~3.71	4.35	**3.99**	3.62	3.86	~3.65	2.05[a]
	^{13}C	62.2	53.1	**79.4**	70.8	71.4	64.2	23.3
α-Rha (I)	^{1}H	5.03	**4.04**	3.91	3.51	3.78	1.29	No
	^{13}C	102.0	**79.7**	71.2	73.2	71.2	18.0	No
α-3-*O*-Me-Rha (II)	^{1}H	5.05	4.35	**3.49**	3.48	3.76	1.29	3.44[b]
	^{13}C	103.6	67.1	**80.6**	72.3	70.5	18.0	57.2

[a] Methyl group from *N*-acetyl group.
[b] Methyl group carried by an oxygen.

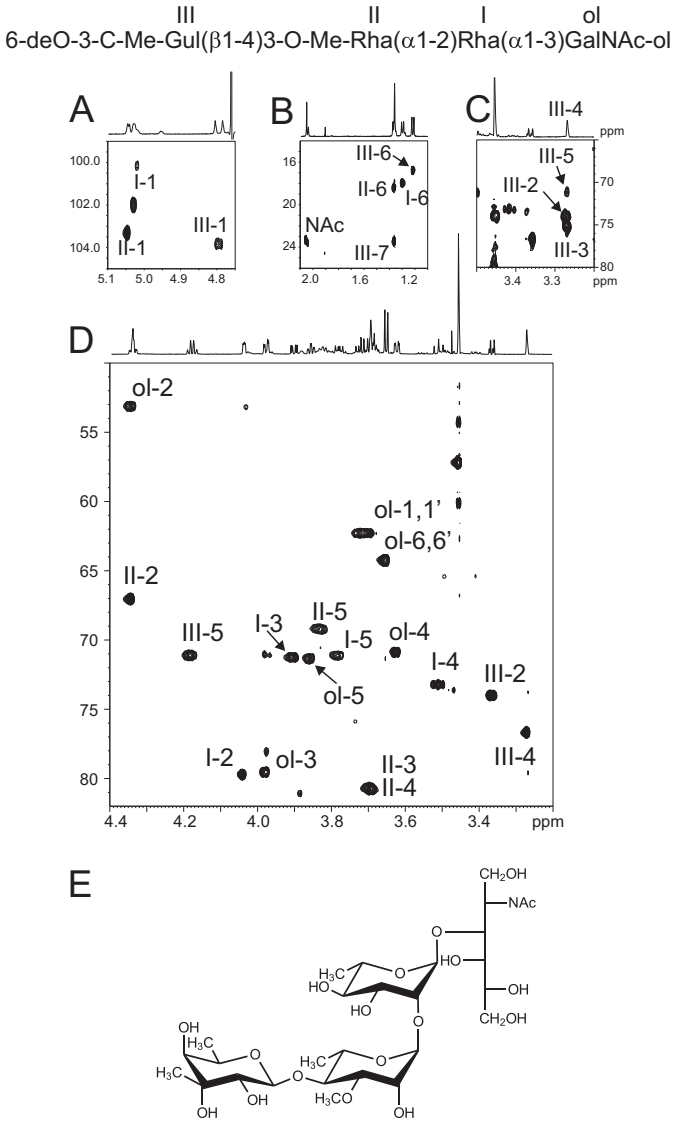


FIGURE 4. **NMR analysis of major tetrasaccharide of LMWG fraction from WT-BclA-Bc.** Shown is the ^{1}H-^{13}C HSQC spectrum of tetrasaccharide from WT-BclA-Bc LMWG. *A*, zoom of anomeric region of ^{1}H-^{13}C HSQC; *B*, zoom of methyl group chemical shift region; *C*, zoom of ^{1}H-^{13}C HMBC showing quaternary carbon of unit III; *D*, zoom of ^{1}H-^{13}C HSQC spectrum where skeleton proton and *O*-methyl groups resonate; *E*, deduced structure of the trisaccharide.

saccharides characterized by the presence of two major tri- and tetrasaccharides with $[M + Na]^+$ at *m*/*z* 552 and 712 as well as a minor trisaccharide with $[M + Na]^+$ at *m*/*z* 509 (data not shown), establishing that BclA isolated from WT-Bc and ΔCT-Bc share similar oligosaccharides. This was further confirmed by HPLC separation of these oligosaccharides, which presented identical patterns to WT-Bc (data not shown).

B. cereus CTD Is Substituted by a Heterogeneous Polysaccharide—Composition analysis established that the large molecular weight glycan associated with the C-terminal domain of Bc-BclA is made up of a complex mixture of monosaccharides, including several derivatives of Rha and methylated Me-Rha and of *N*-acetylated monosaccharides. As expected, we could not obtain any spectrum by MALDI-MS and electrospray MS analysis due to the heavy molecular weight of HMWG and were thus obliged to exclusively rely on NMR analysis. ^{1}H-^{1}H COSY, TOCSY, and ROESY and ^{1}H-^{13}C HSQC and HMBC analyses confirmed that HMWG was an extremely heterogeneous polysaccharide constituted by Rha and HexNAc derivatives. Indeed, at least 28 anomer signals with ^{1}H values ranging from 4.49 to 5.30 ppm and ^{13}C values from 69.0 to 107.4 ppm were observed (supplemental Fig. S3). Of these, eight monosaccharides were tentatively identified as β-anomers, and 20 were tentatively identified as α-anomers based on the ^{1}H chemical shift and $^{1}J_{H\text{-}C}$ correlation constant values of their anomer signal. The observation of multiple -NH-CO-CH_3 ^{1}H/^{13}C signals resonating between 2.018 and 2.110/23.4 ppm as well as -CH- signals at 3.97/52.8 confirmed the presence of *N*-acetylated monosaccharides. At least seven ^{1}H/^{13}C signals in the 1.216–1.385/16.7–23.2 range were identified as methyl groups from Rha residues based on the observation on the ^{1}H/^{1}H COSY spectrum of $^{3}J_{H\text{-}H}$ correlations with Rha H5 signals around 3.8 ppm. Finally, a set of at least five intense signals at 3.442–3.566/57.2–61.3 ppm established the presence of a complex pattern of *O*-methylation. $^{3}J_{H\text{-}C}$ correlations observed on HMBC spectra with signals at δ 3.295/82.7, 3.674/80.5, 3.78/80.9, and 3.825/80.6 strongly supported the presence of methyl residues substituting the C2, C3, and C4 positions of rhamnose residues (30, 31, 32). However, despite these data, we could not obtain further sequence information that would allow the presence of a polysaccharide repetition unit to be established. A mild acid hydrolysis degradation procedure of the native polysaccharide failed to generate oligosaccharides, exclusively generating monosaccharides, which further hampered the sequencing efforts. As such, this compound appears to be a high molecular weight non-repetitive polysaccharide whose structure is too heterogeneous to allow straightforward structural elucidation. Considering the novelty of the structural data, we assessed whether the presence of a CTD-associated polysaccharide was specific to *B. cereus* or could be extended to other species of exosporium-producing *Bacillus* species. Doing this, we hoped to uncover a species in which the CTD-associated polysaccha-

TABLE 3
Proton and carbon chemical shifts of major tetrasaccharide Cro(β1–4)3-O-Me-Rha(α1–2)Rha(α1–3)GalNAc-ol isolated from LMWG WT-BclA-Bc (Fig. 4)
Bold values indicate substitution positions.

		1	2	3	4	5	6,6′	Add -CH3
GalNAc-ol (ol)	^{1}H	~3.7	4.34	**3.98**	3.62	3.86	~3.65	2.05[a]
	^{13}C	62.3	53.1	**79.5**	70.9	71.3	64.2	23.4
α-Rha (I)	^{1}H	5.03	**4.04**	3.90	3.51	3.78	1.30	No
	^{13}C	102.0	**79.7**	71.2	73.2	71.1	18.0	No
α-3-*O*-Me-Rha (II)	^{1}H	5.05	4.35	**~3.7**	**~3.7**	3.83	1.37	3.45[b]
	^{13}C	103.3	67.0	**~80.7**	**~80.7**	69.2	18.4	57.2
β-3-*C*-Me-6-dGul (III)	^{1}H	4.79	3.37	No	3.27	4.18	1.22	1.36[c]
	^{13}C	103.9	74.0	74.9[d]	76.7	71.1	16.8	23.5

[a] Methyl group from *N*-acetyl group.
[b] Methyl group carried by an oxygen.
[c] Methyl group carried by a carbon.
[d] Value obtained on HMBC experiment through $^{3}J_{H,C}$.

rides were simpler, allowing a validation of our hypothesis of a CTD-associated polysaccharide in the *B. cereus* group.

The Glycosylation of Ba-BclA Is Also Domain-specific—A similar experimental approach was applied to *B. anthracis*, previously shown to produce an exosporium and to synthesize glycosylated BclA (14). Previous studies concerning the glycosylation profiles of Ba-BclA established the presence of rhamnose-containing short oligosaccharides whose structures, although different, show similarities to those identified in the present report from *B. cereus*. In particular, oligosaccharides from *B. anthracis* are characterized by the presence of an unusual *N*-acylated monosaccharide in the terminal non-reducing position, named anthrose, which was not observed in *B. cereus*. As for *B. cereus*, reduced oligosaccharides released by reductive β-elimination from the exosporium of *B. anthracis* could be separated by gel filtration chromatography into two distinct HMWG and LMWG fractions that contain equivalent quantities of total monosaccharides (Fig. 5). HMWG contains 2,4-*O*-Me-Rha, 3-*O*-Me-Rha, Rha, GlcNAc, and GalNAc, whereas LMWG contains 3-*O*-Me-Rha, Rha, GlcNAc, and GalNAc in different proportions, again suggesting the presence of two differently glycosylated domains in Ba-BclA. Contrarily to *B. cereus*, 2-*O*-Me-Rha was not identified in the HMWG fraction.

MS and MS/MS analysis of LMWG fraction revealed the presence of two major oligosaccharides at *m*/*z* 552.3 and 943.5 plus a set of minor oligosaccharides at *m*/*z* 797.3, 957.4, and 1202.5. Each oligosaccharide was enriched by HPLC in conditions identical to those used for *B. cereus* and then sequenced by MS/MS. The collision-induced dissociation-MS fragmentation of the $[M + Na]^+$ signal at *m*/*z* 552 produced a pattern identical to that observed for the similar ion from *B. cereus*, which strongly suggests that its sequence is similarly 3-*O*-Me-Rha(α1–2)Rha(α1–3)GalNAc-ol (Fig. 6*A*). Furthermore, this sequence is in total agreement with the previously identified trisaccharide in *B. anthracis* (21). The fragmentation of signal at *m*/*z* 943.5 is characterized by the loss of a monosaccharide in the terminal non-reducing position with a mass of 259 mass units that was assigned to anthrose (Fig. 6*B*) in accordance with its previous identification in *B. anthracis*. The recurrent Y- and C-type fragmentation established the sequence of the oligosaccharide as Ant-dHex-dHex-dHex-HexNAc-ol, which typified it on the basis of its previous identification as Ant(β1–3)Rha(α1–3)Rha(α1–2)Rha(α1–3)GalNAc-ol (Fig. 6*B*) (18, 21).

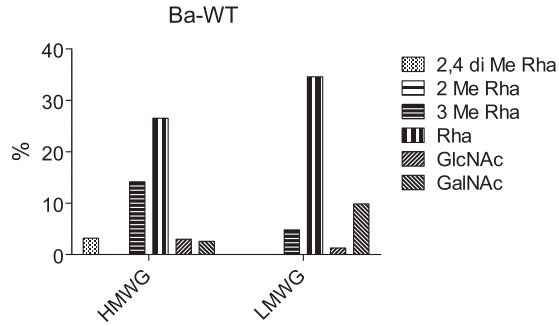


FIGURE 5. **Monosaccharide composition analysis of HMWG and LMWG from *B. anthracis*.**

The NMR analysis in D_2O confirmed this structure assignment by comparison with previous data (data not shown) (18). Indeed, in addition to the three internal Rha and terminal GalNAc-ol residues, the presence of β-anthrose was easily typified on the one-dimensional ^{1}H and ^{1}H/^{1}H TOCSY spectra by the observation of its 2-*O*-methylated C2 position as a H2 triplet ($^{3}J_{H1\text{-}H2} = {}^{3}J_{H2\text{-}H3} = 8.5$ Hz) at 3.453 ppm. Furthermore, Ant exhibited a C6 CH_3 group that resonated at δ 0.992/18.53 compared with 1.066–1.099/17.9–21.3 ppm for internal Rha residues and an intense $(CH_3)_2$-COH- signal at δ 1.068/29.4 ppm. In addition to the two major oligosaccharides, the sequences of the three minor compounds at *m*/*z* 797.3, 957.4, and 1202.5 were established by MS/MS fragmentation, and their structures were assumed by structural similarity with the two major compounds. Indeed, all three oligosaccharides were shown to be variations of the major tri- and pentasaccharides. The MS/MS fragmentation pattern of the $[M + Na]^+$ signal at *m*/*z* 797.3 established that this glycan was constituted by a stretch of two dHex, one GalNAc-ol, and one Ant with an Ant-dHex-dHex-dHex-HexNAc-ol sequence (Fig. 6*C*). Considering its similarity to the compound at *m*/*z* 943, we hypothesized that its structure is Ant(β1–3)Rha(α1–2/3)Rha(α1–3)GalNAc-ol. Then the 14-mass unit difference between the compound at *m*/*z* 957 and the previously identified pentasaccharide at *m*/*z* 943 established that both molecules only differ by an extra methyl (Fig. 6*D*). MS/MS fragmentation established that the Rha residue in the third position from the end was changed into a 3-*O*-Me-Rha residue in agreement with the monosaccharide composition of the LMWG fraction. Thus, this oligosaccharide was identified as Ant(β1–3)Rha(α1–2/4)-3-*O*-Me-Rha(α1–2)Rha(α1–3)GalNAc-ol. However, the linkage position of the 3-*O*-Me-Rha can-

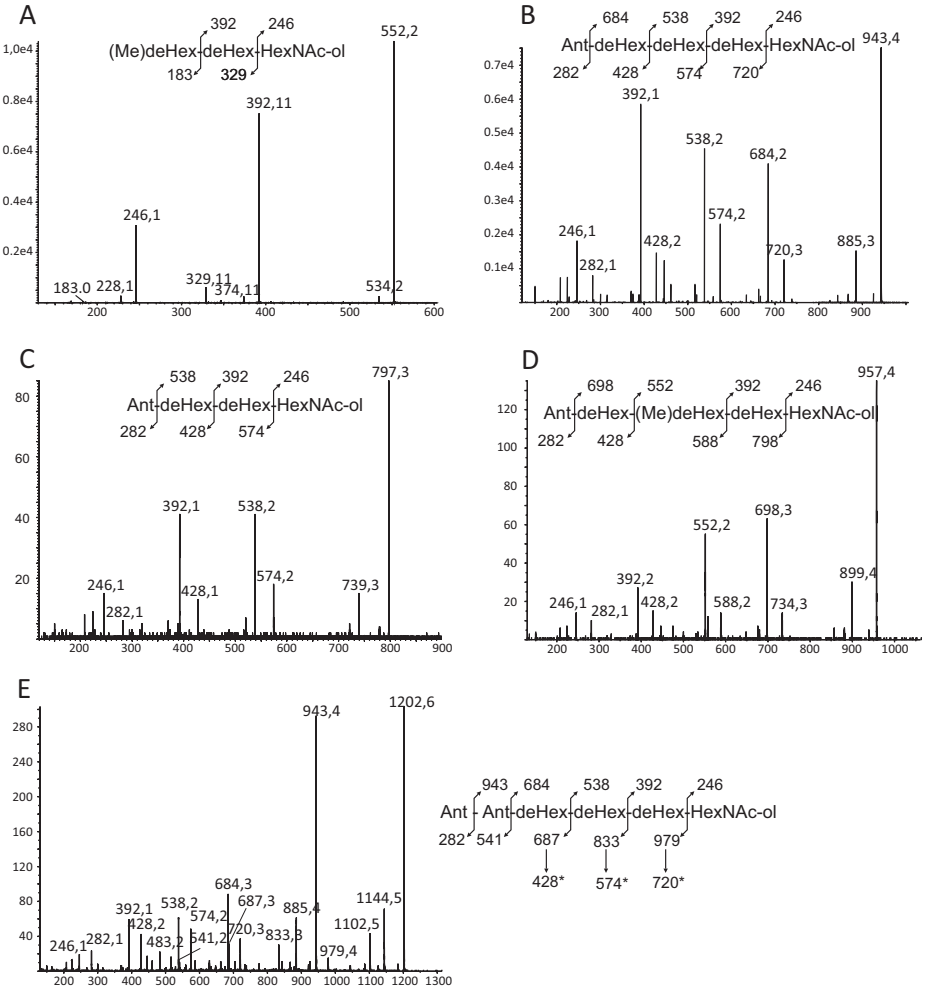


FIGURE 6. **MS/MS sequencing of major oligosaccharides of LMWG fraction from WT-BclA-Ba (*A–E*).** *deHex*, deoxyhexose.

not be definitely inferred from these data. Finally, the oligosaccharide at *m*/*z* 1202 was clearly identified as an equivalent of the anthrose-containing pentasaccharide substituted by an additional Ant residue in the terminal non-reducing position. The presence of a terminal Ant-Ant motif was unambiguously established by the observation of a complete set of C-type ions at *m*/*z* 541, 687, 833, and 979 that were further decomposed by secondary cleavage of the terminal Ant into ions at *m*/*z* 282, 428, 574, and 720 (Fig. 6*E*).

B. anthracis CT Is Substituted by a Homogeneous O-Methylated Rhamnan—As was observed in *B. cereus*, reductive β-elimination of BclA isolated from *B. anthracis* released an HMWG fraction that was separated from the small oligosaccharides by gel filtration. MALDI-TOF MS analysis of this fraction revealed a complex but homogenous pattern of signals observed in the *m*/*z* 2200–5400 range (Fig. 7*A*). Two distinct series of about 20 signals each separated by *m*/*z* 160 increments can be observed: one ranging from *m*/*z* 2445.2 to *m*/*z* 5006.4 and the other ranging from *m*/*z* 2488.1 to *m*/*z* 5207.8. In accordance with composition analysis and calculation of *m*/*z* values, the first was tentatively attributed to a series of $[M + Na]^+$ adducts of $(MeRha)_n$-MeRha-ol polysaccharides with 14 ($m/z = 2445.2$) $< n <$ 30 ($m/z = 5006.4$), whereas the second one was attributed to $[M + Na]^+$ adducts of $(MeRha)_p$-HexNAc-ol with 14 ($m/z = 2488.1$) $< p <$ 31 ($m/z = 5207.8$). For both rhamnan series, maximum signal intensities were observed for $n, p = 21$ and 23. The nature of the rhamnosylated polysaccharide was then confirmed by NMR. As seen in Fig. 7, *B–D*, ^{1}H-^{1}H TOCSY and ^{1}H-^{13}C HSQC spectra of the HMWG fraction from *B. anthracis* established the presence of a single major residue identified as a 3-*O*-methylated rhamnose residue based on its spin system and the strong deshielding of C3 at δ 3.588/79.3 (Table 4). Then the strong deshielding of C2 at 4.344/73.5 established that the rhamnan was polymerized through its C2 position in accordance with data from methylated rhamnosylated LPS from *Mesorhizobium* and *Xanthomonas* strains (32, 33). The absence of HexNAc-associated signals in the spectra probably originates from the very low ratio of HexNAc in the HMWG fraction (Me-Rha/HexNAc around 50:1). Also, NMR and MS analyses on HMWG did not show any anthrose signal, which established that this monosaccharide is a specific component of LMWG associated with the CLR domain of Ba-BclA but absent from the HMWG associated with CTD.

As a whole, the analysis of HMWG isolated from *B. anthracis* exosporium revealed the presence of a methylated rhamnan polysaccharide of a far greater homogeneity than that observed in *B. cereus*. Mass spectrometry analysis strongly suggests that this may be linked to the protein backbone either through a

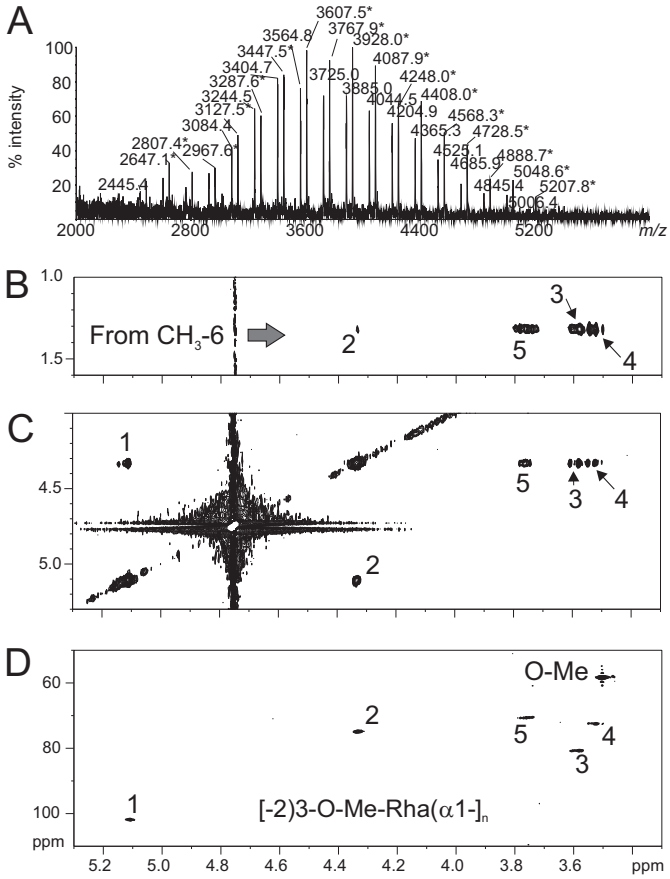


FIGURE 7. **Structural analysis of HMWG fraction from WT-BclA-Ba.** *A*, MALDI-MS spectrum of HMWG. All values are $[M + Na]^+$ adducts of two polysaccharide series: *no star*, $(Me\text{-}dHex)_n Me\text{-}dHex\text{-}ol$; with *star*, $(Me\text{-}dHex)_n$-HexNAc-ol. Shown is ^1H-^1H TOCSY analysis from C6 methyl group of Me-dHex (*B*) and H1 and H2 of Me-dHex (*C*). *D*, ^1H-^{13}C HSQC spectrum.

3-*O*-Me-Rha residue or through a HexNAc residue. In the latter case, it was not possible to unambiguously establish the nature of this residue, but one may speculate that the polysaccharide is an elongated version of the GalNAc-attached oligosaccharides identified on the CLR region and thus would be $[2)3\text{-}O\text{-}Me\text{-}Rha(\alpha1]_n\text{-}2(3\text{-}O\text{-}Me)Rha(\alpha1\text{–}3)GalNAc$.

BclA CT Domain Contains Potential O-Glycosylation Sites—Sequences of BclA proteins are well conserved within the *B. cereus* group with about 95% identity for the N-terminal domain and 88% identity for the CTD. In particular, CTD of BclA from *B. cereus* ATCC 14579 and *B. anthracis* (strains BF1 and EJY92615) show 90% identity over two 134-amino acid sequences. They include many hydroxylated Ser and Thr amino acids, 10 Ser and 15 Thr for *B. cereus* ATCC 14579, and 13 Ser and 15 Thr for *B. anthracis*. According to the crystal structure of the C-terminal domain of *B. anthracis* BclA (Protein Data Bank code 1WCK), 15 serine/threonine residues of the 28 are present at the surface of the CTD trimer and thus present the highest potential for being glycosylated as shown in Fig. 8 (34). Similarly, 14 serine/threonine residues of 25 are potentially *O*-glycosylated in *B. cereus* according to their localization (Fig. 8, *A* and *B*). Of those 16 putative *O*-glycosylation sites, 13 are common to both strains of which 6 are localized at the apical side of the predicted protein structure. Because of their ideal position at the very top of the predicted macromolecule, these 6 residues in *B. anthracis* (Thr-142, Ser-146, Ser-169, Thr-194, and Ser-199) are primary targets for future work dedicated to precisely localizing the *O*-glycosylation sites on BclA CTD. Altogether, structural analysis of glycans from WT and CTD-deleted *Bacillus* strains enabled us to propose a BclA model in which CLR is substituted by short oligosaccharides, whereas the C-terminal domain is covered by a polysaccharide-like material (Fig. 8*C*).

TABLE 4
Proton and carbon chemical shifts of total HMWG fraction from WT-BclA-Ba
Bold values indicate substitution positions.

	1	2	3	4	5	6	*O*-Me
^1H	5.117	**4.344**	3.588	3.526	3.762	1.315	3.503
^{13}C	100.4	**73.5**	79.3	71.0	69.3	16.7	56.8

Discussion

Spores of *Bacillus* species are surrounded by a carbohydrate-rich external layer called an exosporium (35). The majority of these carbohydrates are linked to the BclA protein that is the main constituent of the hairlike surface of the exosporium in *B. cereus* and *B. anthracis* (1, 14, 17). Composition analyses of exosporia purified from a large panel of *Bacillus* species strongly suggest that the surface carbohydrate moiety exhibits a high level of species or even strain specificity (14, 36). Indeed, although most of the *Bacillus* species express rhamnose derivatives, including rhamnose methylated in different positions, the exact nature and ratios of these monosaccharides may vary from one species to another. Despite the fundamental role that glycans play in the interaction of spores with biotic and abiotic surfaces and in the conformation of its surface glycoproteins (7, 19, 20), very little information is currently available as to the molecular nature of the carbohydrate moiety of the *Bacillus* spore surface with the exception of the CLR-associated oligosaccharides of *B. anthracis* (18). In the present work, we have first established that *B. cereus* biosynthesizes rhamnose-containing oligosaccharides associated with the CLR domain of BclA. Detailed structural analysis of the oligosaccharide fraction from *B. cereus* identified two major glycans, $3\text{-}O\text{-}Me\text{-}Rha(\alpha1\text{–}2)Rha(\alpha1\text{–}3)GalNAc\text{-}ol$ and $Cro(\beta1\text{–}4)3\text{-}O\text{-}Me\text{-}Rha(\alpha1\text{–}2)Rha(\alpha1\text{–}3)GalNAc\text{-}ol$. To the best of our knowledge, this particular monosaccharide has never been described before. In *B. anthracis*, we identified the same two major glycans as observed previously (18, 21). Along with the major oligosaccharides, we observed a number of minor compounds, in particular other anthrose-containing oligosaccharides, including a hexasaccharide substituted by two Ant residues in the terminal non-reducing position. It is noteworthy that *B. anthracis* and *B. cereus* strains had the trisaccharide $3\text{-}O\text{-}Me\text{-}Rha(\alpha1\text{–}2)Rha(\alpha1\text{–}3)GalNAc\text{-}ol$ in common but none of the larger oligosaccharides that therefore appear to be species-specific.

Consensus models put the CTD at the tip of the long BclA fibrils that are at the forefront of cellular interaction, which underlines the importance of knowing its glycosylation status (16). In the present report, we have not only confirmed that BclA-CT domains from both *B. anthracis* and *B. cereus* are glycosylated but also that they bear a different type of glycosylation

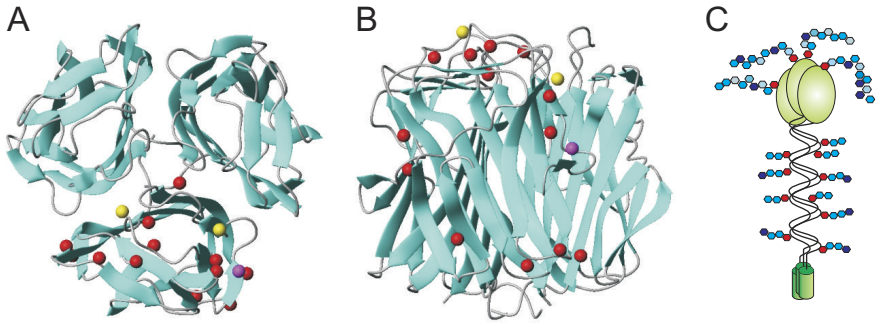


FIGURE 8. **Prediction of the CTD glycosylation sites.** Putative sites of *O*-glycosylation on the C-terminal domain of *B. anthracis* and *B. cereus* BclA were inferred from the crystal structure of *B. anthracis* BclA (34). The best resolution crystal structure of the C-terminal domain of *B. anthracis* BclA (Protein Data Bank code 1WCK) was used to illuminate 16 *O*-glycosylable (Ser and Thr) residues supposed to be on the surface of BclA CTD from the top (*A*) and side (*B*). For clarity, only *β*-carbons are represented as *spheres* on one monomer of the trimer: *red* indicates surface-exposed Ser/Thr residues common to *B. anthracis* and *B. cereus* (92/96/109/119/142/146/152/169/183/184/185/194/199 in *B. anthracis* numbering); *yellow* and *magenta* indicate those present in only *B. anthracis* (144/166) and *B. cereus* (159), respectively. The end of the collagen-like domain anchored in the basal layer is at the *bottom* of the image. The image was made by MOLMOL software (53). *C*, hypothetical representation of the domain-specific glycosylation of BclA.

than BclA-CLR. Indeed, gel filtration of total released glycans enabled us to purify HMWG. These are only observed in the glycans released from *B. cereus* WT-BclA but not in those released from ΔCT-BclA, which strongly suggests that they are associated with BclA-CT domains. Furthermore, HMWG isolated from *B. cereus* exhibited a similar monosaccharide composition to the glycan fraction released from ΔCLR-BclA, thereby confirming its association with BclA-CT (22). In contrast, the LMWG isolated from *B. cereus* ΔCT-BclA strain was strictly identical to that isolated from WT-BclA strain. Further structural analyses established that the nature of HMWG was also species-specific. Indeed, although HMWG from *B. cereus* appears to be made up of a highly heterogeneous polysaccharide containing multiple *O*-methylated rhamnose residues, that from *B. anthracis* is made up of a major homogenous 3-*O*-Me-Rha polymer, probably linked to the protein backbone through a 3-substituted GalNAc residue. Despite our efforts, we could not establish the fine structure of HMWG from *B. cereus* because it is composed of a non-repetitive polysaccharide whose structure is too heterogeneous to allow straightforward structural elucidation. Further efforts are being made to establish its structure by generating *B. cereus* mutant strains in which putative rhamnosyl methyltransferases are inactivated. Indeed, the absence of methylation will drastically reduce the overall structural heterogeneity of this compound, thus allowing its sequencing. We believe that this approach will not only permit us to establish the definitive structure of *B. cereus* HMWG but also to identify the substrate specificity of the different *O*-methyltransferase toward the rhamnose positions. Preliminary screening has revealed that the genome of *B. cereus* ATCC 14579 possesses 71 annotated or putative methyltransferases. Of these, 35 appear not to be associated with any biosynthesis pathway, and seven of the 35 have no homologous gene in the genome of *B. anthracis* Ames strain. These genes may thus be involved in species- or strain-specific glycosylation of *B. cereus*. The genome organization around the *bclA* gene in both *B. anthracis* and *B. cereus* strains is well conserved. Several genes are predicted to be methyltransferases or involved in rhamnolipid modifications. This locus may be a good start place to search for genes involved in the biosynthesis of the glycosylated BclA proteins. Furthermore, the analysis of the proteins expressed during the sporulation of *B. anthracis* may give us further clues. Indeed, the *bclA* gene in *B. anthracis* is expressed in wave 4 of five of gene expression (37). One can expect that genes involved in post-translational glycosylation of BclA are expressed in wave 4 or later. Nine putative glycosyltransferases are expressed in waves 4 and 5 in *B. anthracis*. Homologous genes of these nine are also present in the *B. cereus* genome. These putative glycosyltransferases that may be involved in the glycosylation of exosporium or at least in the overall glycosylation of the spore will be targeted for further work on the structural and functional analysis of *B. cereus* rhamnan.

The presence of the unusual monosaccharide anthrose in *B. anthracis* was a major incentive in developing species-specific diagnoses and targeted therapeutic intervention (38–42). Assays conducted on monoclonal antibodies that target *B. anthracis* glycan (38, 43, 44) have established the immunodominant character of anthrose (38, 45, 46) potentially specific for *B. anthracis* spores (47). Similarly, spore surface carbohydrates have showed the potential of eliciting the anti-spore IgG antibodies in mice, which opens the possibility of inserting carbohydrate haptens in vaccine formulations (48). However, despite these successes, the usage of anthrose-specific reagents for treatment and diagnostic purposes presents several limitations, including (*a*) the more widespread presence of anthrose in *B. cereus* and *B. thuringiensis* groups than initially expected (49, 50), (*b*) the identification of *B. anthracis* strains devoid of anthrose (51), and (*c*) the presence of anthrose analogues in the capsular polysaccharide of a *Shewanella* sp. strain and on the flagella of *Pseudomonas syringae* that cross-react with anti-*B. anthracis* spore sera (52). In the present report, we have described a new family of spore surface polysaccharidic components that may open new possibilities to generate or complement carbohydrate-based species-specific immunological tools. Indeed, HMWG exhibits several interesting features that could serve as immunological targets. First, as it is ideally localized at the tip of BclA and thus at the very surface of the spore exosporium, it is easily accessible for antibody interaction. Second, its polymeric nature may further increase the accessibility of the antigenic epitope by providing a flexible spacer. Third, the structures of HMWG isolated from *B. anthracis* and

B. cereus are different, which suggests that it is species-specific. However, a thorough investigation of a large number of strains and species is now required to assess the extent of HMWG specificity.

Author Contributions—E. M. performed the NMR experiments. F. K. and E. G. purified the glycans and performed GC experiments. B. C. performed the MS experiments. X. T. analyzed proteins sequences. Y. L. and A. R. produced the spores. C. F. and Y. G. conceived and coordinated the study. Y. G. wrote the paper. All authors analyzed the results and approved the final version of the manuscript.

Acknowledgments—The 900-MHz NMR spectrometer was funded by Région Nord-Pas de Calais, European Union (Fonds Européen de Développement Économique et Régional), Ministère Français de la Recherche, Université Lille 1, and CNRS.

References

1. Henriques, A. O., and Moran, C. P. (2007) Structure, assembly, and function of the spore surface layers. *Annu. Rev. Microbiol.* **61,** 555–588
2. Plomp, M., Leighton, T. J., Wheeler, K. E., and Malkin, A. J. (2005) Architecture and high-resolution structure of *Bacillus thuringiensis* and *Bacillus cereus* spore coat surfaces. *Langmuir* **21,** 7892–7898
3. Sylvestre, P., Couture-Tosi, E., and Mock, M. (2003) Polymorphism in the collagen-like region of the *Bacillus anthracis* BclA protein leads to variation in exosporium filament length. *J. Bacteriol.* **185,** 1555–1563
4. Weaver, J., Kang, T. J., Raines, K. W., Cao, G.-L., Hibbs, S., Tsai, P., Baillie, L., Rosen, G. M., and Cross, A. S. (2007) Protective role of *Bacillus anthracis* exosporium in macrophage-mediated killing by nitric oxide. *Infect. Immun.* **75,** 3894–3901
5. Oliva, C. R., Swiecki, M. K., Griguer, C. E., Lisanby, M. W., Bullard, D. C., Turnbough, C. L., Jr., and Kearney, J. F. (2008) The integrin Mac-1 (CR3) mediates internalization and directs *Bacillus anthracis* spores into professional phagocytes. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 1261–1266
6. Faille, C., Tauveron, G., Le Gentil-Lelièvre, C., and Slomianny, C. (2007) Occurrence of *Bacillus cereus* spores with a damaged exosporium: consequences on the spore adhesion on surfaces of food processing lines. *J. Food Prot.* **70,** 2346–2353
7. Lequette, Y., Garénaux, E., Tauveron, G., Dumez, S., Perchat, S., Slomianny, C., Lereclus, D., Guérardel, 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. *Appl. Environ. Microbiol.* **77,** 4905–4911
8. Giorno, R., Bozue, J., Cote, C., Wenzel, T., Moody, K.-S., Mallozzi, M., Ryan, M., Wang, R., Zielke, R., Maddock, J. R., Friedlander, A., Welkos, S., and Driks, A. (2007) Morphogenesis of the *Bacillus anthracis* spore. *J. Bacteriol.* **189,** 691–705
9. Kang, T. J., Fenton, M. J., Weiner, M. A., Hibbs, S., Basu, S., Baillie, L., and Cross, A. S. (2005) Murine macrophages kill the vegetative form of *Bacillus anthracis*. *Infect. Immun.* **73,** 7495–7501
10. Gerhardt, P., and Ribi, E. (1964) Ultrastructure of the exosporium enveloping spores of *Bacillus cereus*. *J. Bacteriol.* **88,** 1774–1789
11. Kailas, L., Terry, C., Abbott, N., Taylor, R., Mullin, N., Tzokov, S. B., Todd, S. J., Wallace, B. A., Hobbs, J. K., Moir, A., and Bullough, P. A. (2011) Surface architecture of endospores of the *Bacillus cereus*/*anthracis*/*thuringiensis* family at the subnanometer scale. *Proc. Natl. Acad. Sci.* **108,** 16014–16019
12. Steichen, C. T., Kearney, J. F., and Turnbough, C. L. (2005) Characterization of the exosporium basal layer protein BxpB of *Bacillus anthracis*. *J. Bacteriol.* **187,** 5868–5876
13. 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
14. Faille, C., Lequette, Y., Ronse, A., Slomianny, C., Garénaux, E., and Guerardel, Y. (2010) Morphology and physico-chemical properties of *Bacillus* spores surrounded or not with an exosporium: consequences on their ability to adhere to stainless steel. *Int. J. Food Microbiol.* **143,** 125–135
15. Rodenburg, C. M., McPherson, S. A., Turnbough, C. L., Jr., and Dokland, T. (2014) Cryo-EM analysis of the organization of BclA and BxpB in the *Bacillus anthracis* exosporium. *J. Struct. Biol.* **186,** 181–187
16. Boydston, J. A., Chen, P., Steichen, C. T., and Turnbough, C. L. (2005) Orientation within the exosporium and structural stability of the collagen-like glycoprotein BclA of *Bacillus anthracis*. *J. Bacteriol.* **187,** 5310–5317
17. Sylvestre, P., Couture-Tosi, E., and Mock, M. (2002) A collagen-like surface glycoprotein is a structural component of the *Bacillus anthracis* exosporium. *Mol. Microbiol.* **45,** 169–178
18. 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. *J. Biol. Chem.* **279,** 30945–30953
19. 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. Pathog.* **38,** 1–12
20. Oliva, C., Turnbough, C. L., Jr., and Kearney, J. F. (2009) CD14-Mac-1 interactions in *Bacillus anthracis* spore internalization by macrophages. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 13957–13962
21. Dong, S., Chesnokova, O. N., Turnbough, C. L., Jr., and Pritchard, D. G. (2009) Identification of the UDP-N-acetylglucosamine 4-epimerase involved in exosporium protein glycosylation in *Bacillus anthracis*. *J. Bacteriol.* **191,** 7094–7101
22. Lequette, Y., Garénaux, 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
23. Chen, G., Driks, A., Tawfiq, K., Mallozzi, M., and Patil, S. (2010) *Bacillus anthracis* and *Bacillus subtilis* spore surface properties and transport. *Colloids Surf. B Biointerfaces* **76,** 512–518
24. Pieters, R. J. (2011) Carbohydrate mediated bacterial adhesion. *Adv. Exp. Med. Biol.* **715,** 227–240
25. Coddeville, B., Maes, E., Ferrier-Pages, C., and Guerardel, Y. (2011) Glycan profiling of gel forming mucus layer from the scleractinian symbiotic coral *Oculina arbuscula*. *Biomacromolecules* **12,** 2064–2073
26. Koerner, T. A., Prestegard, J. H., and Yu, R. K. (1987) Oligosaccharide structure by two-dimensional proton nuclear magnetic resonance spectroscopy. *Methods Enzymol.* **138,** 38–59
27. Zdorovenko, E. L., Ovod, V. V., Zatonsky, G. V., Shashkov, A. S., Kocharova, N. A., and Knirel, Y. A. (2001) Location of the O-methyl groups in the O polysaccharide of *Pseudomonas syringae* pv. *phaseolicola*. *Carbohydr. Res.* **330,** 505–510
28. Giuliano, R. M., and Kasperowicz, S. (1988) Synthesis of branched-chain sugars: a stereoselective route to sibirosamine, kansosamine, and vinelose from a common precursor. *Carbohydr. Res.* **183,** 277–285
29. Tóth, A., Reményik, J., Bajza, I., and Lipták, A. (2003) Synthesis of the methyl ethers of methyl 6-deoxy-3-C-methyl-α-L-talopyranoside and -α-L-mannopyranoside. Examination of the conformation and chromatographic properties of the compounds. *ARKIVOC* **2003**, 28–45
30. Gilleron, M., Venisse, A., Rivière, M., Servin, P., and Puzo, G. (1990) Carbohydrate epitope structural elucidation by ^{1}H-NMR spectroscopy of a new *Mycobacterium kansasii* phenolic glycolipid antigen. *Eur. J. Biochem.* **193,** 449–457
31. Zdorovenko, E. L., Valueva, O. A., Kachala, V. V., Shashkov, A. S., Knirel, Y. A., Komaniecka, I., and Choma, A. (2012) Structure of the O-polysaccharide of *Azorhizobium caulinodans* HAMBI 216; identification of 3-C-methyl-D-rhamnose as a component of bacterial polysaccharides. *Carbohydr. Res.* **358,** 106–109
32. Senchenkova, S. N., Huang, X., Laux, P., Knirel, Y. A., Shashkov, A. S., and Rudolph, K. (2002) Structures of the O-polysaccharide chains of the lipopolysaccharides of *Xanthomonas campestris* pv. *phaseoli* var. *fuscans* GSPB 271 and *X. campestris* pv. *malvacearum* GSPB 1386 and GSPB 2388.

Carbohydr. Res. **337,** 1723–1728
33. Zdorovenko, E. L., Valueva, O. A., Kachala, V. V., Shashkov, A. S., Kocharova, N. A., Knirel, Y. A., Kutkowska, J., Turska-Szewczuk, A., Urbanik-Sypniewska, T., Choma, A., and Russa, R. (2009) Structure of the O-polysaccharides of the lipopolysaccharides of *Mesorhizobium loti* HAMBI 1148 and *Mesorhizobium amorphae* ATCC 19655 containing two O-methylated monosaccharides. *Carbohydr. Res.* **344,** 2519–2527
34. Réty, S., Salamitou, S., Garcia-Verdugo, I., Hulmes, D. J., Le Hégarat, F., Chaby, R., and Lewit-Bentley, A. (2005) The crystal structure of the *Bacillus anthracis* spore surface protein BclA shows remarkable similarity to mammalian proteins. *J. Biol. Chem.* **280,** 43073–43078
35. Fox, A., Stewart, G. C., Waller, L. N., Fox, K. F., Harley, W. M., and Price, R. L. (2003) Carbohydrates and glycoproteins of Bacillus anthracis and related bacilli: targets for biodetection. *J. Microbiol. Methods* **54,** 143–152
36. Fox, A., Black, G. E., Fox, K., and Rostovtseva, S. (1993) Determination of carbohydrate profiles of *Bacillus anthracis* and *Bacillus cereus* including identification of O-methyl methylpentoses by using gas chromatography-mass spectrometry. *J. Clin. Microbiol.* **31,** 887–894
37. Liu, H., Bergman, N. H., Thomason, B., Shallom, S., Hazen, A., Crossno, J., Rasko, D. A., Ravel, J., Read, T. D., Peterson, S. N., Yates, J., 3rd, and Hanna, P. C. (2004) Formation and composition of the *Bacillus anthracis* endospore. *J. Bacteriol.* **186,** 164–178
38. Mehta, A. S., Saile, E., Zhong, W., Buskas, T., Carlson, R., Kannenberg, E., Reed, Y., Quinn, C. P., and Boons, G.-J. (2006) Synthesis and antigenic analysis of the BclA glycoprotein oligosaccharide from the *Bacillus anthracis* exosporium. *Chemistry* **12,** 9136–9149
39. Crich, D., and Vinogradova, O. (2007) Synthesis of the antigenic tetrasaccharide side chain from the major glycoprotein of *Bacillus anthracis* exosporium. *J. Org. Chem.* **72,** 6513–6520
40. Milhomme, O., Dhénin, S. G., Djedaïni-Pilard, F., Moreau, V., and Grandjean, C. (2012) Synthetic studies toward the anthrax tetrasaccharide: alternative synthesis of this antigen. *Carbohydr. Res.* **356,** 115–131
41. Guo, H., and O'Doherty, G. A. (2007) *De novo* asymmetric synthesis of the anthrax tetrasaccharide by a palladium-catalyzed glycosylation reaction. *Angew. Chem. Int. Ed. Engl.* **46,** 5206–5208
42. Dhénin, S. G., Moreau, V., Morel, N., Nevers, M.-C., Volland, H., Créminon, C., and Djedaïni-Pilard, F. (2008) Synthesis of an anthrose derivative and production of polyclonal antibodies for the detection of anthrax spores. *Carbohydr. Res.* **343,** 2101–2110
43. Tamborrini, M., Werz, D. B., Frey, J., Pluschke, G., and Seeberger, P. H. (2006) Anti-carbohydrate antibodies for the detection of anthrax spores. *Angew. Chem. Int. Ed. Engl.* **45,** 6581–6582
44. Kuehn, A., Kovác, P., Saksena, R., Bannert, N., Klee, S. R., Ranisch, H., and Grunow, R. (2009) Development of antibodies against anthrose tetrasaccharide for specific detection of *Bacillus anthracis* spores. *Clin. Vaccine Immunol.* **16,** 1728–1737
45. Wang, D., Carroll, G. T., Turro, N. J., Koberstein, J. T., Kovác, P., Saksena, R., Adamo, R., Herzenberg, L. A., Herzenberg, L. A., and Steinman, L. (2007) Photogenerated glycan arrays identify immunogenic sugar moieties of *Bacillus anthracis* exosporium. *Proteomics* **7,** 180–184
46. Oberli, M. A., Tamborrini, M., Tsai, Y.-H., Werz, D. B., Horlacher, T., Adibekian, A., Gauss, D., Möller, H. M., Pluschke, G., and Seeberger, P. H. (2010) Molecular analysis of carbohydrate-antibody interactions: case study using a *Bacillus anthracis* tetrasaccharide. *J. Am. Chem. Soc.* **132,** 10239–10241
47. Tamborrini, M., Holzer, M., Seeberger, P. H., Schürch, N., and Pluschke, G. (2010) Anthrax spore detection by a Luminex assay based on monoclonal antibodies that recognize anthrose-containing oligosaccharides. *Clin. Vaccine Immunol.* **17,** 1446–1451
48. Milhomme, O., Köhler, S. M., Ropartz, D., Lesur, D., Pilard, S., Djedaïni-Pilard, F., Beyer, W., and Grandjean, C. (2012) Synthesis and immunochemical evaluation of a non-methylated disaccharide analogue of the anthrax tetrasaccharide. *Org. Biomol. Chem.* **10,** 8524–8532
49. Dong, S., McPherson, S. A., Tan, L., Chesnokova, O. N., Turnbough, C. L., Jr., and Pritchard, D. G. (2008) Anthrose biosynthetic operon of *Bacillus anthracis*. *J. Bacteriol.* **190,** 2350–2359
50. Tamborrini, M., Oberli, M. A., Werz, D. B., Schürch, N., Frey, J., Seeberger, P. H., and Pluschke, G. (2009) Immuno-detection of anthrose containing tetrasaccharide in the exosporium of *Bacillus anthracis* and *Bacillus cereus* strains. *J. Appl. Microbiol.* **106,** 1618–1628
51. Tamborrini, M., Bauer, M., Bolz, M., Maho, A., Oberli, M. A., Werz, D. B., Schelling, E., Zinsstag, J., Seeberger, P. H., Frey, J., and Pluschke, G. (2011) Identification of an African *Bacillus anthracis* lineage that lacks expression of the spore surface-associated anthrose-containing oligosaccharide. *J. Bacteriol.* **193,** 3506–3511
52. Kubler-Kielb, J., Vinogradov, E., Hu, H., Leppla, S. H., Robbins, J. B., and Schneerson, R. (2008) Saccharides cross-reactive with *Bacillus anthracis* spore glycoprotein as an anthrax vaccine component. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 8709–8712
53. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14,** 51–55, 29–32
54. Pezard, C., Duflot, E., and Mock, M. (1993) Construction of *Bacillus anthracis* mutant strains producing a single toxin component. *J. Gen. Microbiol.* **139,** 2459–2463

SUPPLEMENTARY MATERIAL

Supplementary material contains the structural analysis of cytidine isolated from exosporium of WT *B. cereus* by ^{1}H-NMR, the ROESY ^{1}H-^{1}H NMR analysis of the major tetrasaccharide of LMWG fraction from WT-BclA-Bc and the NMR analysis of HMWG fraction from WT-BclA-Bc.

Figure S1 - ^{1}H-NMR analysis of cytidine isolated from LMWG WT-BclA-Bc in RP-HPLC at 17.8 min (Fig. 2).

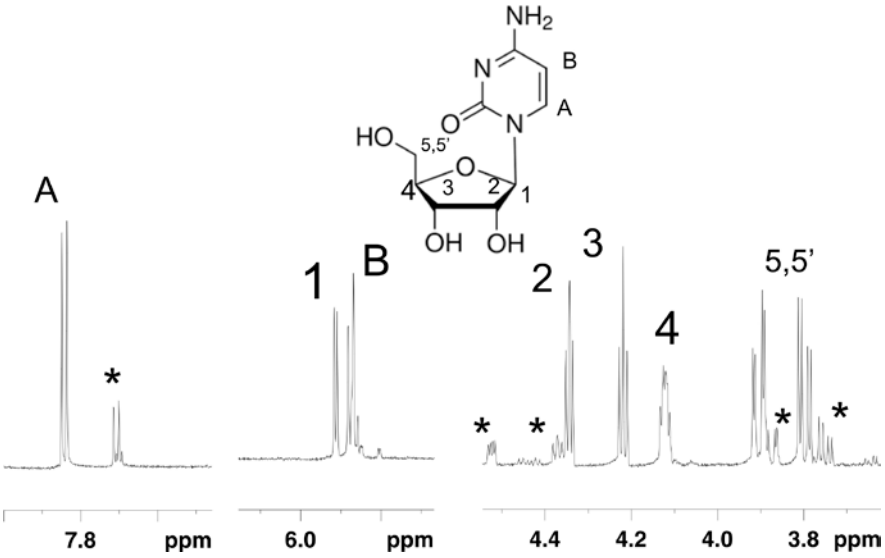


Table S1: Proton chemical shifts of cytidine isolated from LMWG WT-BclA-Bc in RP-HPLC at 17.8 min (Fig. 2).

	Protons chemical shifts (ppm)						
	1	2	3	4	5	A	B
^{1}H	5.91	4,34	4.22	4.12	3.90/3.80	7.84	5.87

Figure S2 - ROESY NMR analysis of major tetrasaccharide of LMWG fraction from WT-BclA-Bc. ROESY (in black) and TOCSY (in red) signals are superimposed and established the sequence III-II-I-ol.

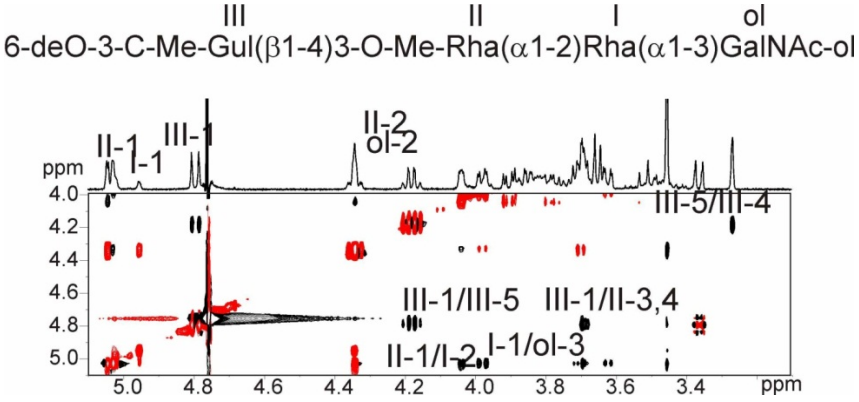
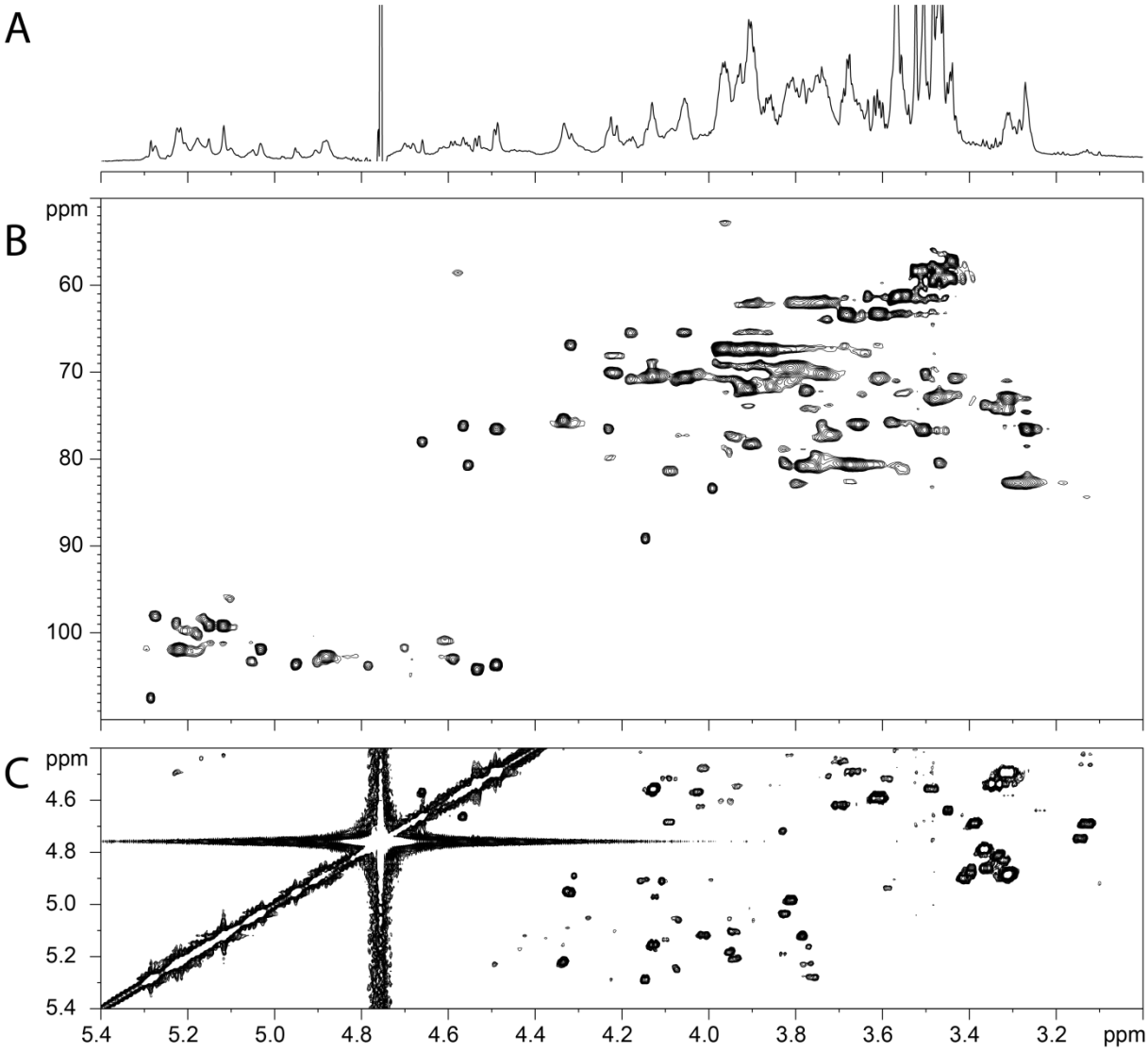


Figure S3 - NMR analysis of HMWG fraction from WT-BclA-Ba. (A) ^{1}H-NMR spectrum; (B) ^{1}H-^{13}C HSQC spectrum and (C) ^{1}H-^{1}H TOCSY spectrum.



Glycosylation of BclA Glycoprotein from *Bacillus cereus* and *Bacillus anthracis* Exosporium Is Domain-specific

Emmanuel Maes, Frederic Krzewinski, Estelle Garenaux, Yannick Lequette, Bernadette Coddeville, Xavier Trivelli, Annette Ronse, Christine Faille and Yann Guerardel

J. Biol. Chem. 2016, 291:9666-9677.
doi: 10.1074/jbc.M116.718171 originally published online February 26, 2016