Stoichiometry, Absolute Abundance, and Localization of Proteins in the Bacillus cereus Spore Coat Insoluble Fraction Determined Using a QconCAT Approach

Stelder, S.K.; Benito de Moya, C.; Hoefsloot, H.C.J.; de Koning, L.J.; Brul, S.; de Koster, C.G.

DOI
10.1021/acs.jproteome.7b00732

Publication date
2018

Document Version
Final published version

Published in
Journal of Proteome Research

License
CC BY-NC-ND

Citation for published version (APA):
Stelder, S. K., Benito de Moya, C., Hoefsloot, H. C. J., de Koning, L. J., Brul, S., & de Koster, C. G. (2018). Stoichiometry, Absolute Abundance, and Localization of Proteins in the Bacillus cereus Spore Coat Insoluble Fraction Determined Using a QconCAT Approach. Journal of Proteome Research, 17(2), 903-917. https://doi.org/10.1021/acs.jproteome.7b00732
**Stoichiometry, Absolute Abundance, and Localization of Proteins in the Bacillus cereus Spore Coat Insoluble Fraction Determined Using a QconCAT Approach**

Sacha K. Stelder,†,‡,§ Celia Benito de Moya,† Huub C. J. Hoefsloot,§ Leo J. de Koning,‡ Stanley Brul,†,⊥ and Chris G. de Koster†,⊥

†Molecular Biology & Microbial Food Safety, ‡Mass Spectrometry of Biomacromolecules, and §Biosystems Data Analysis, Swammerdam Institute for Life Sciences, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands

**ABSTRACT:** Spores of Bacillus cereus pose a threat to food safety due to their high resistance to the heat or acid treatments commonly used to make food microbiologically safe. Spores may survive these treatments and later resume growth either on foodstuffs or, after ingestion, upon entering the gut they are capable of producing toxins, which cause either vomiting or diarrhea. The outer layers of the spore, the spore coat and exosporium, consist primarily of proteins that may serve as potential biomarkers for detection. The major morphogenetic protein CotE is important for correct assembly and attachment of the outermost layer, the exosporium, and by extension retention of many proteins. However, characterization of the proteins affected by deletion of CotE has been limited to electrophoretic patterns. Here we report the effect of CotE deletion on the insoluble fraction of the spore proteome through liquid chromatography–Fourier transform tandem mass spectrometry (LC–FTMS/MS) analysis. A total of 560 proteins have been identified in both mutant and wild-type spore coat isolates. A further 163 proteins were identified exclusively in wild-type spore isolates that are dependent on CotE for their association with the spore. Several of these are newly confirmed as associated with the exosporium, namely BC_2569 (BclF), BC_3345, BC_2427, BC_2878, BC_0666, BC_2984, BC_3481, and BC_2570. A total of 153 proteins were only identified in ΔCotE spore isolates. This was observed for proteins that are known or likely to be interacting with or are encased by CotE. Crucial spore proteins were quantified using a QconCAT reference standard, the first time this was used in a biochemically heterogeneous system. This allowed us to determine the absolute abundance of 21 proteins, which spanned across three orders of magnitude and together covered 5.66% of the total spore weight. Applying the QconCAT methodology to the ΔCotE mutant allowed us to quantify 4.13% ± 0.14 of the spore total weight and revealed a reduction in abundance for most known exosporium associated proteins upon CotE deletion. In contrast, several proteins, either known or likely to be interacting with or encased by CotE (i.e., GerQ), were more abundant. The results obtained provide deeper insight into the layered spore structure such as which proteins are exposed on the outside of the spore. This information is important for developing detection methods for targeting spores in a food safety setting. Furthermore, protein stoichiometry and determination of the abundance of germination mediating enzymes provides useful information for germination and outgrowth model development.

**KEYWORDS:** Bacillus cereus, spore coat, QconCAT, CotE, quantitative proteomics

**INTRODUCTION:** Bacillus cereus presents a problem for the food industry due to its ability to survive common processing methods such as heat or acid treatment by forming (endo)spores. Although spores are metabolically dormant and damaged by the preservation strategies, after such treatments, when conditions become favorable for growth many can still germinate, repair incurred damage and grow out. Once back in a vegetative state they can start producing toxins either in the food product or after ingestion in the gut resulting in food poisoning causing vomiting or diarrhea, respectively. Proper detection of spores therefore determines an important benchmark for food safety.1 Spores consist of a layered structure (Figure 1) with a central core containing the DNA, surrounded by a peptidoglycan cortex layer, which in turn is surrounded by proteinaceous inner and outer coat layers. In the case of B. cereus, this compact spore structure is then surrounded by a loose layer composed of protein, lipids, and carbohydrates called the exosporium.2−4 As the outermost layers consist mainly of protein, there is a potential for these proteins to...
be used as spore specific biomarkers. However, to utilize spore proteins as biomarker targets for detection, it is important to know which are abundant and accessible.

Our approach for analysis of the spore proteome utilizes bead-beating to disintegrate the spore's structure and make proteins accessible to tryptic digestion. Clearly, as such this procedure eliminates any spatial information. Localization of proteins may nevertheless still be investigated by comparison of the wild type proteome to that of mutants lacking certain morphogenetic proteins. Deletion of the major morphogenetic protein CotE, which normally forms a shell around the outer coat, results in gross defects in the exosporium. Proteins normally associated with the exosporium are then likely to be affected by this perturbed assembly. When comparing the proteome of CotE knockout spores to wild type isolates, proteins that are dependent on CotE for their localization to the spore coat would either be missing or present at reduced levels in ΔCotE isolates. A similar approach has previously been used to show CotE-dependent localization of proteins in the outer coat of *Bacillus subtilis* and the exosporium of *Bacillus anthracis*. However, the approach of these previous studies relied on solubilization of spore proteins and resolution on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), which eliminates the contribution of the insoluble fraction. Ignoring this renders an incomplete picture of the effect of CotE deletion. Furthermore, the characterization of the changes observed in these previous studies was limited to electrophoretic patterns and, with the exception of BAS2377, did not include identification of the proteins in differentially observed bands. Here, we report the first proteome wide investigation into CotE dependent localization of proteins in the spores of a *B. cereus* ΔCotE mutant *B. cereus* strain previously described by Bressuire et al.

Most investigations into structural spore proteins are based on analyses of knockout mutants with which major morphogenetic proteins have been identified. SaA and CotE are examples of such proteins that are essential for the deposition of entire layers of the spore coat. While structural proteins such as these are likely to be abundant, less so are functional proteins such as enzymes or receptors present in the coat. Generally, spore proteins are known to be cross-linked together to form a coherent structure built on the framework created by these major morphogenetic proteins. Knockout strategies to identify protein functions might not always result in a clear phenotype as potential effects on spore structure may be masked by the remainder of the cross-linked network, determining quantitative protein abundance can give an indication of potential function of uncharacterized proteins.

Advances in mass spectrometry have allowed for the first quantitative analyses of the spore proteome; however, these have so far only used relative measurements that allow for comparison of different conditions but give no information on the number of proteins per spore. QconCAT approaches rely on creating a synthetic isotopically labeled protein consisting of concatenated tryp tic peptides, which are quantotypic for the targeted proteins. This protein is then spiked into the sample and digested alongside the analyte creating a set of reference peptides for all targeted proteins in a 1:1:1 ratio at a known quantity allowing for absolute quantification of several proteins at once. Being able to select 20–25 target proteins per QconCAT construct when using two selected quantotypic peptides per target protein makes it perfectly suited for investigations of this scope as the spore proteome is estimated to consist of several hundred spore specific proteins. This targeted approach has an added benefit as current methods are not able to isolate only the spore specific proteins. In the case of *B. cereus*, this is in part due to the fact that the exosporium traps proteins from the mother cell in the so-called interspace region adding a confounding factor in the case of nontargeted methods. QconCATs are normally implemented in a homogeneous system; however, as spore proteins are highly cross-linked, a large part of the spore proteome cannot be fully solubilized by standard methods such as SDS-extraction. By using a method first described by Abhyankhar et al., it is possible to identify these proteins by using trypsin to cleave peptides exposed by prior SDS-extraction and using reducing conditions that should break cross-links such as dicysteines and dityrosines. By selecting only peptides that were frequently identified in preliminary analyses, we are able to select a set of reproducibly identifiable peptides that are likely not obstructed by cross-links. By optimizing digestion time, we can ensure completed digestion, thus enabling us to quantify protein content using a QconCAT approach in a heterogeneous system.

Beyond determining the abundance, using a QconCAT also allows for determination of protein stoichiometry. Interplay between proteins is an important factor for acquisition of spore resistance mechanisms, and many spore proteins are known to interact with each other. Information about protein stoichiometry can provide insight into how proteins relate to each other in *vivo* and give a sense of spatial dimensions occupied. Perturbations in this stoichiometry give a measure for integrity of the spore coat, which may be connected to acquired spore characteristics such as heat resistance or germination efficiency.

To further expand upon the QconCAT application, we quantified the selected proteins in both wild type *B. cereus* ATCC14579 and a ΔCotE mutant *B. cereus* strain previously described by Bressuire et al. This allowed us to determine whether these proteins were partially localized in the exosporium or in fact interact with CotE and gave an indication of their accessibility.

The 21 proteins selected for quantification cover known structural proteins, enzymes, a transcription factor, and several uncharacterized proteins. These proteins are located across the various different layers of the spore though in many cases localization has not been confirmed experimentally in *B. cereus*.
Table 1. Selected Proteins and Corresponding Q-Peptide Sequences*

| gene ID | protein ID | protein description | peptide sequences |
|---------|------------|----------------------|-------------------|
| BC_0059 | SpoVT      | Stage V sporulation protein T. Forepore specific transcriptional regulator. Identified in spore coat isolates (unpublished preliminary results). | (EHK)AVNTAASFLAK(QME) |
| BC_0212 | YuaW       | Putative uncharacterized protein. Previously identified in spore coat isolates (unpublished preliminary results). | (RJR)EGDPLEFVDR(DGE) |
| BC_0389 | CotB1      | Spore coat protein B1. Exosporium basal layer protein. In B. anthracis, CotB is CotE independent, found in both coat and exosporium isolates. | (ITK)LSPILQELK(FDK) |
| BC_0390 | CotB2      | Spore protein B2. Exosporium basal layer protein. In B. anthracis, CotB is CotE independent, found in both coat and exosporium isolates. | (DLK)LNPNEFDLKD(ADY) |
| BC_0987 | BC_0987    | Putative uncharacterized protein. Previously identified in spore coat isolates (unpublished preliminary results). | (FLK)DLIGSFVR(VNR) |
| BC_1281 | CalY       | Cell envelope-bound metalloprotease. Spore associated protease detected in exosporium isolates. Potentially involved in biofilm formation. | (IVK)JEIIILAIK(HIK) |
| BC_1284 | InhA       | Immune inhibitor A. Secreted virulence factor which regulates the secretome in Bacillus anthracis and Bacillus thuringiensis. | (ISR)VJGELVSLGK(DYL) |
| BC_2064 | Alr1       | Alanine racemase 1. Quorum sensor for premature germination by l-alanine in exosporium. | (HIK)SVQSVQKVC(KCK) |
| BC_2569 | BC_2569    | Collagen triple helix repeat protein. Previously identified in spore coat isolates (unpublished preliminary results). | (ISR)TPFSLUPR(ITT) |
| BC_2752 | YpeB       | Hypothetical Membrane Spanning Protein. Cortex lysis partner to SleB. | (RSR)NTFPTQNCELVESR(TFY) |
| BC_2753 | SleB       | Spore cortex-lytic enzyme. Key protein, together with CwlJ, for cortex degradation during germination. | (KVK)FLNWNDK(QSE) |
| BC_2872 | CotX1      | Spore coat protein X. Heavily cross-linked insoluble protein, important for structural integrity (density) of outer coat. | (WDR)QSEPVTYYTTLAQLQKD(VPQ) |
| BC_2889 | IunH       | Inosine-uridine pyrophosphorylase. Involved in inosine-adenosine nucleotide hydrolase. | (PGK)AADYGADAAASGGHGNK(GPK) |
| BC_3607 | YaaH       | Spore peptidoglycan hydrolase. Assists breakdown of cortex by cleaving products of SleB/CwlJ activity. | (GLK)FEVQGQADDNASGAVR(LYR) |
| BC_3712 | BcC        | Hypothetical Membrane Spanning Protein. Frequently found in spore coat protein isolates. BLAST matches to Hypothetical proteins. Contains collagen region with GXX repeats common in exosporium proteins. | (VYK)GDDGDSVAFK(EFF) |
| BC_3770 | CotE       | Spore coat protein E. Major morphogenic protein, anchors exosporium and many outer coat proteins. | (VSR)DVNVTVPYFLR(YLA) |
| BC_4419 | YheN/ConA  | Sporulation cortex protein. Unknown function, forespore specific Sig F/G. B. subtilis mutants have hampered outgrowth. | (AGR)IPNTPSIPRT(AQL) |
| BC_4420 | SflA       | SpoVID-dependent spore coat assembly factor. Major morphogenic protein, directs inner coat proteins. | (ISQ)ARESTIPRTIPRT(LIQ) |
| BC_4640 | YtfJ/GerW  | Putative uncharacterized protein. Germination receptor, stimulus unknown. | (MR)HIVVQK(GDQ) |
| BC_5056 | BC_5056    | Collagen adhesion protein. Unknown function, frequently found in spore coat protein isolates. Adhesin-like protein indicates it may be important as a virulence factor. | (HMK)QQAGASAPP(KYQ) |
| BC_5391 | GerQ       | Spore protein GerQ, Structural protein, potentially highly cross-linked. Important for CwlJ localization. Important for germination with Ca2+. | (PIK)AADGSLVLTYSK(VSF) |

*Bracketed sequence represents the native primary sequence context, which was not included in the QconCAT.

(Figure 1). A short functional description of each target protein and the selected Q-peptides included in the QconCAT as well as accompanying notable bibliography are displayed in Table 1. The abundance of 21 spore proteins is determined using a QconCAT approach in a heterogeneous system. In this manner, 5.66% ± 0.51 of the total spore weight has been quantified in wild type B. cereus and 4.13% ± 0.14 in a ΔCotE mutant strain. Using the obtained abundances, it is possible to classify proteins by function, structural, or otherwise. Furthermore, determining protein stoichiometry provides insight into the way either known or potential interactors may relate to each other. The potential of the most abundant proteins for use as biomarkers is assessed using the immunogenic epitope prediction algorithm SVMTrilP. 22

**MATERIALS AND METHODS**

B. cereus Growth and Sporulation Conditions

B. cereus ATCC 14579 was grown in Tryptic Soy Broth medium at 30 °C. A ΔCotE knockout mutant strain previously described by Bressuire-Isard et al.29 was grown in Tryptic Soy Broth supplemented with 250 μg/mL spectinomycin. Overnight cultures were harvested and washed once and transferred to 250 mL of chemically defined growth and sporulation medium (CDGS) 25 at an optical density of 2 at 600 nm. Growth and
sporulation were allowed to continue for 4 days until a >95% spore crop was obtained. Spores were harvested by centrifugation, washed with 1.25% Tween-20 once, followed by four washes with cold mili-Q water after which spores were freeze-dried and stored at −80 °C before protein isolation.

**Spore Coat Protein Isolation**

Proteins in the spore coat insoluble fraction were isolated using a modified version of the method developed by Abhyankar et al. Lyophilized spores were resuspended in cold 10 mM Tris-HCl (pH 7.8), 0.2% SDS, 100 mM Na-EDTA, 100 mM β-ME) for 10 min at 80 °C. This also leaves the insoluble fraction of the proteome more accessible to trypsin. After centrifugation, the insoluble pellet was washed three times with mili-Q water and freeze-dried.

**Peptide Sample Preparation**

Freeze-dried protein material was resuspended in 100 mM ammonium bicarbonate and reduced with 10 mM DTT at 55 °C for 60 min and alkylated with 55 mM iodoacetamide for 45 min in the dark. Samples were centrifuged, and the pellet was resuspended in 100 mM ammonium bicarbonate, 10% acetonitrile, and 5 μg of Trypsin was added per mg of material. At this point, QconCAT was added in different quantities as indicated. The digestion was allowed to proceed for 18 h at 37 °C unless otherwise indicated. After digestion, the supernatant was collected and the pellet was washed twice with 100 mM ammonium bicarbonate. Up to 1% TFA was added to acidify the buffer, which removed ammonium bicarbonate and inactivated trypsin. Samples were freeze-dried and resuspended in 50% ACN, 0.1% TFA and stored at −80 °C until use.

**QconCAT Design**

The QconCAT was designed based on initial nonquantitative analyses of the spore coat insoluble fraction (unpublished preliminary results). Reproducibly observed proteins were selected for quantification if two peptides were observed, which could be considered Q-peptides as described by Brownridge et al.

To be able to include several proteins of interest, some liberties had to be taken with the strict requirements for Q-peptides proposed previously, for instance peptides starting with methionine, and pyroglutamation of N-terminal glutamic acid were allowed. The list of selected peptides and nonspecific proteins.

The soluble fraction of the spore proteome was removed by SDS extraction (50 mM Tris-HCl (pH 7.8), 0.2% SDS, 100 mM Na-EDTA, 100 mM β-ME) for 10 min at 80 °C. This also leaves the insoluble fraction of the proteome more accessible to trypsin. After centrifugation, the insoluble pellet was washed three times with mili-Q water and freeze-dried.

**Data Analysis**

Raw Fourier transform (FT)-MS/MS data were processed with the MASCOT DISTILLER program, version 2.4.3.1 (64 bit), MDRO 2.4.3.0 (MATRIX science, London, UK) including the Search toolbox and the Quantification toolbox. Peak-picking for both MS and MS/MS spectra was optimized for the mass resolution of up to 60 000. Peaks were fitted to a simulated isotope distribution with a correlation threshold of 0.7, with minimum signal-to-noise ratio of 2.

The processed data were searched with the MASCOT server program 2.3.02 (MATRIX science, London, UK) against a complete B. cereus ATCC14579 predicted proteome database obtained from UniPROT (accessed January 2016) complemented with the QconCAT protein amino acid sequence. Trypsin was used as enzyme and one missed cleavage was allowed. Carbamidomethylation of cysteine, oxidation of methionine, and pyroglutamination of N-terminal glutamic acid were included as variable modification. The peptide mass tolerance was set to 25 ppm and the peptide fragment mass tolerance was set to 0.03 Da. Where applicable, the quantification method was set to the metabolic 15N labeling method to enable MASCOT to identify both the 14N target and 15N QconCAT peptides.

For identification results, identified proteins were determined using MASCOT with MudPIT peptide identification score set to a cutoff of 20. At this cutoff and based on the number of assigned decoy peptide sequences, the significance threshold was adjusted to obtain a peptide false discovery rate of 2% for all analyses. The MASCOT protein identification reports were exported as XML and then imported in a custom-made VBA software program running in Microsoft Excel. A list of all proteins identified is represented in Supplementary Table 1.

Using the quantification toolbox, the quantification of the light target peptides relative to the corresponding heavy QconCAT peptides was determined as the heavy over light isotopic ratio.
Figure 2. continued
Figure 2. continued
using Simpson integration of the peptide MS chromatographic profiles for all detected charge states. Correlation threshold for isotopic distribution fit was set to 0.98, with a detected 15N label content of 99.6%. XIC threshold, 0.1; all charge states, on; max XIC width, 200 s. Allowed elution time shift for the heavy and light peptides was limited to 20 s.

All isotope ratios were manually validated by inspecting the MS spectral data. For target peptides, for which overall replica analyses were occasionally not MS/MS identified, the isotope ratio was manually calculated from their extracted heavy and light peptide ion chromatograms as the ratio of the areas under the curves using Bruker DataAnalysis 4.2 program (Bruker Daltonics, Bremen, Germany).

The heavy over light isotopic ratios for all target peptides over all analyses are systematically listed in Supplementary Table S2. The mass spectrometry proteomics data of the wild-type QconCAT analyses have been deposited to the ProteomeX-change Consortium via the PRIDE partner repository with the data set identifier PXD003569.

**Method Verification**

QconCAT concentration was determined by PolyQuant GmbH using a Bradford protein assay.

QconCAT approaches are usually performed in a homogeneous setting requiring complete digestion of both sample and QconCAT. The insoluble nature of a large part of the spore proteome therefore presents an issue in translating this method to spore coat analyses. Though insoluble residual material remains after performing digestion with trypsin, completed digestion of spore material was assured by optimizing digestion time until no change was observed in isotope ratios. Triplicate digests of an insoluble fraction isolate were made with digestion time varying from 16 to 22 h. Isotope ratios were plotted versus digestion time and minimal change was observed for most peptides over the selected time frame. In the ones that did change, this was mostly due to data points conceived through manual integration of the chromatogram. Overall 18 h of digestion time was selected as optimal. Data are made available in Supporting Information S2. This validates the QconCAT approach insofar that the obtained data quantifies the entirety of the proteins of interest that can be accessed by trypsin. Selective inaccessibility (i.e., cross-linking of one of the Q-peptides) is assessed by using two Q-peptides per protein and comparing the difference between them (see below). There remains a possibility that a fraction of the population of a protein of interest is completely blocked off; however, using mechanical and chemical disruption as well as enzymatic cleavage makes this possibility remote.

**Statistical Approach**

Acquired isotope ratios were multiplied by the amount of QconCAT used in each experiment to convert the unit to pmol/mg. Data points were lumped together for each protein across all analyses: three biological replicates with three technical replicates each, using four different QconCAT amounts for each or only one concentration of QconCAT (37.5 pmol/mg of spore material) for the ΔCotE analyses. The average value and 95% confidence interval were determined using a bootstrap approach. The obtained data set was resampled using a vector length equal in size to the number of acquired data points separately for each protein and bootstrapped to a value of n = 1000 using Microsoft Excel.

This statistical approach relies on extensive experimental sampling and allows all types of variance to be indicated under one denominator. This is enabled by quantifying across a range of QconCAT reference concentrations. To appreciate the different
sources of variance, the isotope ratios for each peptide at each concentration were plotted and linearity on a double log scale was verified across the range. As the total amount of peptide material released from the sample should be the same across all technical replicates, the sum of the quantified peptides in each digestion should be the same. We can correct for any mixing/weighting error by normalizing to the average yield across each biological replicate. The resultant plots are shown in Figure 2.

**Table 2. Identified Proteins Known To Be Related to Spores Both in Presence and Absence of CotE**

| Protein Description                        | ID     |
|--------------------------------------------|--------|
| 1-pyrroline-5-carboxylic acid dehydrogenase | rocA   |
| Alanine racemase 1                       | alr1   |
| Catalase                                   | BC_1155|
| Cell envelope-bound metalloprotease (Camelysin) | BC_1281|
| Cell surface protein                      | BC_3547|
| Cell surface protein *                    | BC_2639|
| Cell wall hydrolase cwj                   | BC_5390|
| Cell wall-associated hydrolase            | BC_2849|
| Collagen adhesion protein                 | BC_5056|
| CotJA protein                             | BC_0823|
| CotJB protein                             | BC_0822|
| CotJC protein                             | BC_0821|
| D-α-alanyl-d-alanine carboxypeptidase      | BC_4075|
| Enolase                                    | eno    |
| Ferredoxin-nitrite reductase              | BC_1424|
| Germination protease                      | gpr    |
| Hypothetical Cytoplasmic Protein          | BC_3986|
| Hypothetical Cytoplasmic Protein *        | BC_3195|
| Hypothetical Exported Protein             | BC_1334|
| Hypothetical Membrane Spanning Protein    | BC_3712|
| Hypothetical Membrane Spanning Protein *  | BC_2752|
| Hypothetical Membrane Spanning Protein    | BC_1221|
| IG hypothetical 17193                     | BC_3534|
| IG hypothetical 18063                     | BC_1029|
| Immune inhibitor A                        | BC_1284|
| Inosine-uridine preferring nucleoside hydrolase | BC_3552|
| Inosine-uridine preferring nucleoside hydrolase | BC_2889|
| N-acetylmuramoyl-l-alanine amidase         | BC_2823|
| N-acetylmuramoyl-l-alanine amidase *       | BC_1494|
| Nonheme chloroperoxidase                  | BC_4774|
| Oligopeptide-binding protein oppA *       | BC_3586|
| Oligopeptide-binding protein oppA          | BC_3585|
| Oligopeptide-binding protein oppA *       | BC_3584|
| Oligopeptide-binding protein oppA *       | BC_2026|
| Protein translocase subunit YajC *        | BC_4410|
| Putative hydrolase                        | BC_3133|
| Putative stage IV sporulation protein     | BC_4303|
| Putative uncharacterized protein          | BC_5391|
| Putative uncharacterized protein *        | BC_4640|
| Putative uncharacterized protein          | BC_4419|
| Putative uncharacterized protein          | BC_4387|
| Putative uncharacterized protein          | BC_4047|
| Putative uncharacterized protein          | BC_3992|
| Putative uncharacterized protein          | BC_3582|
| Putative uncharacterized protein          | BC_3090|
| Putative uncharacterized protein          | BC_2969|
| Putative uncharacterized protein *        | BC_2745|

**RESULTS AND DISCUSSION**

**Identities**

A total of 713 proteins were identified in ΔCotE isolates, having a Mascot identification in at least two out of three replicates. Beyond these, 92 were only identified in one replicate in ΔCotE isolates and are disregarded for discussions below. A total of 560 proteins overlapped between the mutant and wild type cells. Of those proteins previously identified in wild type isolates, 163 proteins were not detected in ΔCotE isolates indicating they were CotE dependent for their association with the spore. A further 153 proteins were identified exclusively in the ΔCotE...
isolates, which indicates that their accessibility is higher upon removal of CotE. Proteins of interest are displayed in the tables below. All identifications across ΔCotE isolates compared to the collected identifications from wild type isolates are displayed in Supplementary Table S1.

CotE Independent Proteins

The majority of protein identifications were not dependent on CotE. Identified proteins known to be spore associated are listed in Table 2. Interestingly, these included proteins originating from all layers of the spore. Examples are SpoIVA, SafA, CotX, and InhA, which are located at the cortex boundary, inner coat, outer coat, and exosporium, respectively. This suggests that the CotE deletion does not completely remove any of the spore layers.

CotE Dependent Localization

Comparing proteins isolated from wild type spores to those isolated from CotE knockout spores with perturbed exosporia can reveal proteins that are dependent on CotE for their association with the spore, and by extension are likely localized in the exosporium. Important to note is that removing the exosporium or damaging its integrity also eliminates the interspace region of the spore. Proteins indicated to depend on CotE for their association with the spore are therefore not necessarily spore specific and could merely be part of the mother cell cytosol trapped in the interspace. Interpretation then requires background knowledge of the function of identified proteins to determine whether they are likely spore specific. Identified proteins with a known function in sporulation or spore biogenesis, showing dependence on CotE are displayed in Table 3. The full list of identified proteins is available in Supplementary Table S1.

An alternate strategy was previously reported by Abhyankar et al. who identified proteins in purified exosporium fractions of B. cereus as well as combined spore coat and exosporium isolates. By using French press to apply a shear stress, it is possible to remove parts of the exosporium; however, this method is not capable of completely removing the exosporium. Likely some fragments remain at the cap region where the exosporium is supposedly anchored to the coat. This was reflected in their results as the number of proteins missing from spore coat isolates where a French press step was applied was quite limited. Only three proteins, InhA, CalY, and a Putative uncharacterized protein BC_1591, were specific to exosporium containing isolates. InhA and CalY are quantified using QconCAT and discussed below. BC_1591 was identified in only one replicate analysis making biological interpretation of this particular identification difficult. The results obtained in this study complement the results obtained previously as several proteins were both detected in exosporium isolates and are missing from ΔCotE spore isolates as displayed in Table 3. This strongly suggests that BC_2569 (BclF), BC_3345, BC_2427, and BC_2878 are exosporium proteins. Beyond that, several proteins can be localized to the exosporium based on homology to known exosporium proteins. For instance, Immune inhibitor A (BC_1284) is a known exosporium component, its homologues BC_0666 and BC_2984 are thus likely to be exosporium associated. Furthermore, BC_3481 and BC_2570 both contain collagen-like triple helix repeat domains frequently found in exosporium proteins are therefore likely exosporium proteins. Other proteins will require further evidence to be appointed to the exosporium.

CotE Dependent Accessibility

Apart from proteins specifically being identified in the wild type, there are also a number of proteins that were exclusively identified in the ΔCotE isolates. As CotE is a major structural protein that interacts with many different proteins directly, lacking CotE might result in interacting proteins to become less tightly bound (i.e., lacking cross-links) or no longer being encased in CotE or CotE dependent proteins. Therefore, they can become more accessible to tryptic digestion and become newly identified. The resultant new identifications specific to the ΔCotE mutant isolates are displayed in Table 4. The full list of identified proteins is available in Supplementary Table S1.

Table 3. Identified Proteins Known To Be Related to Spores Exclusively Identified in Wild Type Isolates

| protein description | ID       |
|---------------------|----------|
| Collagen triple helix repeat protein | BC_3481 |
| Collagen triple helix repeat protein | BC_2570 |
| Collagen triple helix repeat protein | BC_2569 (BclF) |
| Collagen-like triple helix repeat protein | BC_3345 |
| Hypothetical Glycosyltransferase | BC_3515 |
| Immune inhibitor A | BC_0666 |
| Immune inhibitor A | BC_2984 |
| N-acetylmuramoyl-L-alanine amidase | BC_2822 |
| N-acetylmuramoyl-L-alanine amidase | BC_3677 |
| Peptidoglycan N-acetylglucosamine deacetylase | BC_1960 |
| Putative uncharacterized protein | BC_2427 |
| Putative uncharacterized protein | BC_2878 |
| Putative uncharacterized protein | BC_0663 |
| Putative uncharacterized protein | BC_2375 |
| Putative uncharacterized protein | BC_p002 |
| Small acid-soluble protein gamma type | sspE |
| Spore coat protein E | BC_1984 |
| Stage IV sporulation protein B | BC_3770 (CotE) |
| Stage IV sporulation protein B | BC_4172 |

“Indicates proteins previously identified by Abhyankar et al. 14

“Indicates proteins previously identified by Abhyankar et al. 14

“Indicates a protein that is removed during creation of the CotE knockout strain.

Table 4. Identified Proteins Known To Be Related to Spores Exclusively Identified in ΔCotE Isolates

| protein description | ID       |
|---------------------|----------|
| Antisigma F factor | spoIIAB |
| Antisigma factor antagonist | BC_4074 |
| Oligopeptide-binding protein oppA | BC_2848 |
| Prespore specific transcriptional activator mtfA | BC_3922 |
| Putative uncharacterized protein | BC_0623 |
| Putative uncharacterized protein | BC_1708 |
| SpoIIA like protein | BC_2436 |
| Stage III sporulation protein D | BC_5282 |
| Stage V sporulation protein S | BC_2142 |
| Stage VI sporulation protein D | BC_4467 |

“Indicates proteins previously identified by Abhyankar et al. 14

SpoVID is involved in spore coat deposition; specifically it is at the interface between the cortex and coat layers where it anchors the coat layer and is known to interact with CotE directly. Identification of this protein corroborates the notion that coat integrity is compromised allowing trypsin to access this protein. Other spore related proteins specific to ΔCotE isolates are mostly regulatory proteins important early during sporulation. 5

DOI: 10.1021/acs.jproto.7b00732
J. Proteome Res. 2018, 17, 901-917
To understand how much of the spore is covered by these analyses, we made a weight comparison for each protein by multiplying copy numbers with protein masses (Table 5). This indicated 5.66% ± 0.51 of the total spore weight is covered by these proteins with InhA taking up 2.33% ± 0.27 as the largest contributor. The total amount of protein per spore which is estimated at around 70 wt % indicating around 8% of the entire spore protein weight is quantified using only the 21 proteins included in this QconCAT analysis. As the total number of proteins we identify from spore samples ranges in the hundreds (700 in a single LC–MS/MS analysis), this signifies a substantial part was covered using this limited set of proteins.

The rather unique properties of some spore proteins do not lend them suitable for use within a QconCAT strategy due to them containing either a large number of cysteines, repeat sequences prone to missed cleavage, regions very rich in lysine, etc. This made it impossible to include several proteins frequently identified in initial analyses such as CotX2, CotY1, CotY2, CwJL, and CotAlpha a.o. To complete the quantitative assessment of the spore coat, other methods such as AQUA will need to be employed.

**Quantitative Dependence**

We used the Qstel01 synthetic QconCAT protein to quantify 21 proteins from across all layers of the spore. In this manner, we can determine CotE dependence or interaction beyond the yes/no answer of identification as described above. Many proteins were significantly different in abundance compared to wild type, as displayed in Figure 3. Though CotE and BclF light peptides were not identified in the ΔCotE mutant strain, the method for quantification searches for light peptides matching the heavy peptides introduced with the QconCAT. This results in CotE being quantified at an extremely low level corresponding to noise levels and BclF at severely reduced levels.

Fourteen out of 20 proteins quantified (excluding CotE itself) are significantly affected by the ΔCotE mutation. The proteins

---

**Table 5. Protein Abundance in Molecules Per Spore and Percentage of Total Weight**

| proteins | wild type | ΔCotE | % weight | wild type | ΔCotE |
|----------|-----------|-------|----------|-----------|-------|
| CotB1    | 1.9 × 10⁴ ± 1.7 × 10⁴ | 1.4 × 10⁴ ± 7.6 × 10² | 0.60 ± 5.6 × 10⁻² | 0.44 ± 2.4 × 10⁻² |
| InhA     | 1.6 × 10⁴ ± 1.9 × 10⁴ | 7.9 × 10² ± 1.6 × 10⁴ | 2.3 ± 0.27 | 1.1 ± 2.3 × 10⁻² |
| CotB2    | 1.5 × 10⁴ ± 1.1 × 10⁴ | 2.3 × 10³ ± 3.3 × 10³ | 0.42 ± 3.0 × 10⁻² | 0.64 ± 9.2 × 10⁻³ |
| CotX     | 1.5 × 10⁴ ± 7.0 × 10⁴ | 1.3 × 10³ ± 3.0 × 10³ | 0.40 ± 1.9 × 10⁻² | 0.35 ± 8.1 × 10⁻³ |
| SasS     | 1.0 × 10⁴ ± 4.7 × 10⁴ | 3.1 × 10² ± 8.8 × 10² | 0.25 ± 1.1 × 10⁻² | 7.5 × 10⁻⁴ ± 2.1 × 10⁻³ |
| GerQ     | 8.9 × 10⁴ ± 45 × 10³ | 1.7 × 10¹ ± 2.6 × 10¹ | 0.24 ± 1.2 × 10⁻² | 0.45 ± 6.9 × 10⁻³ |
| CjY      | 5.6 × 10⁴ ± 86 × 10³ | 2.6 × 10¹ ± 1.8 × 10¹ | 0.20 ± 3.1 × 10⁻² | 9.4 × 10⁻³ ± 6.7 × 10⁻³ |
| CotE     | 3.5 × 10⁴ ± 16 × 10³ | 0.12 ± 5.5 × 10⁻³ | 0.12 ± 5.5 × 10⁻³ |
| BclC     | 2.7 × 10⁴ ± 2.9 × 10⁴ | 2.5 × 10² ± 2.6 × 10² | 0.34 ± 3.6 × 10⁻² | 0.32 ± 3.3 × 10⁻² |
| SaA      | 2.6 × 10⁴ ± 9.0 × 10³ | 1.5 × 10² ± 2.6 × 10² | 0.29 ± 9.8 × 10⁻³ | 0.16 ± 2.8 × 10⁻³ |
| YhcN     | 1.9 × 10⁴ ± 1.3 × 10³ | 2.1 × 10¹ ± 1.9 × 10¹ | 7.9 × 10⁻² ± 5.3 × 10⁻³ | 8.6 × 10⁻² ± 7.8 × 10⁻³ |
| BC_5056  | 1.5 × 10⁴ ± 1.5 × 10³ | 9.4 × 10¹ ± 1.1 × 10¹ | 8.9 × 10⁻² ± 8.7 × 10⁻³ | 5.4 × 10⁻² ± 6.2 × 10⁻³ |
| YsaH     | 1.1 × 10⁴ ± 4.2 × 10² | 1.4 × 10⁰ ± 4.5 × 10⁰ | 9.1 × 10⁻² ± 3.3 × 10⁻³ | 0.11 ± 3.6 × 10⁻³ |
| AlrI     | 1.1 × 10⁴ ± 3.3 × 10² | 1.0 × 10⁰ ± 1.9 × 10⁰ | 8.1 × 10⁻² ± 2.4 × 10⁻³ | 7.5 × 10⁻² ± 1.4 × 10⁻³ |
| YusW     | 7.2 × 10⁴ ± 1.8 × 10³ | 7.9 × 10⁰ ± 3.6 × 10⁰ | 2.1 × 10⁻¹ ± 1.1 × 10⁻¹ | 2.3 × 10⁻¹ ± 1.1 × 10⁻¹ |
| IunH     | 5.6 × 10⁴ ± 66 × 10² | 5.5 × 10⁰ ± 5.1 × 10⁰ | 3.4 × 10⁻³ ± 4.0 × 10⁻³ | 3.3 × 10⁻³ ± 3.1 × 10⁻³ |
| SpoVT    | 5.1 × 10⁴ ± 3.7 × 10³ | 6.8 × 10⁰ ± 2.9 × 10⁰ | 1.7 × 10⁻² ± 1.2 × 10⁻² | 2.2 × 10⁻² ± 9.6 × 10⁻³ |
| Yij      | 4.0 × 10⁴ ± 1.4 × 10³ | 5.0 × 10⁰ ± 1.3 × 10⁰ | 9.4 × 10⁻³ ± 3.2 × 10⁻³ | 1.2 × 10⁻² ± 2.9 × 10⁻³ |
| BclF     | 3.6 × 10⁴ ± 1.9 × 10³ | 8.8 × 10⁰ ± 6.6 × 10⁰ | 3.2 × 10⁻³ ± 1.7 × 10⁻³ | 7.8 × 10⁻³ ± 5.9 × 10⁻³ |
| YpeB     | 3.0 × 10⁴ ± 2.9 × 10³ | 3.2 × 10⁰ ± 1.1 × 10⁰ | 2.5 × 10⁻³ ± 2.9 × 10⁻³ | 2.7 × 10⁻³ ± 8.9 × 10⁻³ |
| SleB     | 2.3 × 10⁴ ± 7.4 × 10³ | 3.0 × 10⁰ ± 2.7 × 10⁰ | 1.1 × 10⁻³ ± 3.5 × 10⁻⁴ | 1.4 × 10⁻³ ± 1.3 × 10⁻³ |

*Margins indicate 95% confidence intervals of the bootstrapped average.*

Expression of CotE or lack thereof does not affect this early regulation once again indicating it is increased accessibility which is causing these proteins to be identified. Altogether, this suggests these proteins are encased by CotE yet it is apparently not the (sole) reason why they are retained in the spore.

**Protein Abundance**

Quantitative analyses were performed by combining the spore coat insoluble fraction isolates with different concentrations of QconCAT (18.75, 37.5, 75, 187.5 pmol per milligram dry weight of insoluble coat material). Isotope ratios for each peptide at each concentration were plotted and linearity on a double log scale was verified across the range (Figure 2A), these also show the variance due to experimental errors to be small. Variance between two Q-peptides of the same protein also becomes apparent and, though small for most proteins, differed significantly for CotB1, SpoVT, IunH, BclC, YhcN, BC_5056 (Figure 2B). Ratios of both peptides were used for calculating the quantified value for the protein with the exception of CotB2, SaA, and InhA where only one peptide was reliably measured (Figure 2C). Using the higher of the two quantified peptides in these cases is worthy of consideration if the observed variance is biological in nature as discussed below. Having confirmed the linear fit to the isotope ratios, we calculated the copy number of proteins per milligram insoluble fraction by multiplying the isotope ratios with the amount of QconCAT added. To convert these obtained quantities of proteins per milligram of spore dry weight to copy numbers per spore, we need to know the dry weight of a single spore which in the case of B. cereus is estimated to be around 1 picogram. As spores can differ in size, this may have a larger variance than the data obtained here. Assuming the insoluble fraction represents the majority of the spore’s dry weight the unit (pmol/mg) of the numbers obtained above can then be converted to molecules per spore by multiplying with a conversion factor of 600 (pmol/mg = 6 × 10²⁵ molecules/10²⁵ spores) (Table 5).
most severely affected include known exosporium proteins (CotB1, InhA, CotB2, CalY, BclF). This corresponds to the visual phenotype of spores having defects in their exosporium. Likewise, levels of candidate exosporium proteins SasS and BC_5056 are indeed reduced in ΔCotE isolates. This indicates that though their incorporation in the spore is not wholly dependent on CotE, they are severely affected by its absence. It is likely that the reduced amounts recovered from ΔCotE isolates are maintained in the small fragments of exosporium which still form despite the absence of CotE.

Previously, CotB was indicated to be CotE independent in B. anthracis. B. cereus CotB shows the same behavior qualitatively; however, quantitative differences exist. Interestingly, unlike its orthologue CotB1, CotB2 is actually quantified at a higher level than in wild type. Following the general theme of proteins detected exclusively in ΔCotE isolates being more accessible to trypsin, the same thing appears to be happening quantitatively to CotB2. What is striking is that CotB1 is actually less abundant in ΔCotE isolates indicating that although they are orthologues, their interactions appear different.

Structural proteins SafA and GerQ are components of the inner coat layer and were significantly affected by deletion of CotE. Interestingly, where SafA was less abundant in ΔCotE isolates, GerQ was actually more abundant. As previously mentioned, the likely explanation is related to the nature of the interaction between proteins. GerQ has been indicated to be cross-linked by transglutaminase, an enzyme located across all coat layers, resulting in glutamyl-lysine isopeptide cross-links in B. subtilis. Although the specific N-terminal lysines indicated to be cross-linked are not present in the GerQ orthologue used by B. cereus, 19 out the first N-terminal 50 amino acids are glutamines indicating this function is likely maintained. CotE could then either be cross-linked to GerQ directly, or it is responsible for retention of proteins normally cross-linked to GerQ. In both scenarios, knocking out CotE would directly or indirectly make GerQ more accessible to trypsin. This generic effect of higher abundance in ΔCotE isolates is also seen among several other quantified proteins associated with the inner coat layers such as SleB, YtfJ, and YaaH but is most prevalent in the case of GerQ. Contrary to GerQ’s dense structure, SafA has an extended conformation in which the N-terminus spans the inner coat to interact with SpoVID at the spore cortex boundary. Potentially part of the SafA present is not directly attached to SpoVID and relies on outer coat integrity for its retention to the coat. Though the ΔCotE spores still maintain an outer coat layer,
as mentioned above, the integrity of this layer is affected if this abundant protein is removed.

Of the proteins that show no significant difference in their abundance, IunH deserves special mention. IunH is an enzyme involved in modulating germination in response to inosine by hydrolyzing the germinant and reducing the signal triggering germination.\textsuperscript{33,36} This allows the spore to prevent premature germination in an environment where the environment is not suitable for growth. Bressuire-Isoard et al.\textsuperscript{6} report inosine triggered germination is impaired in the ΔCotE knockout strain. As IunH abundance appears not to be affected by deletion of CotE it seems that this phenotype of impaired germination is likely correlated to the germination receptors rather than the signal of the germinant itself.

**Functional Classification by Protein Abundance**

Protein abundance is determined across three orders of magnitude showing significant differences depending on the role of the protein (Figure 3). Structural proteins (CotB, CotX, CotE, SaFA, GerQ) are generally more abundant than proteins with enzymatic functions (YaaH, SleB, Alr1, IunH) as is to be expected. Proteins located in the exosporium layer are also generally more abundant, which correlates to the size of the exosporium compared to the coat layers.

The uncharacterized protein BC_0987 was identified in many spore coat protein isolates indicating it may be spore associated and was therefore included in the QconCAT. It is determined to be as highly abundant as known structural proteins indicating, for lack of conserved (enzymatic) domains, it may be a structural spore protein. BLAST analysis reveals a homologue of this protein can be found across the B. cereus group. Strangely, there is a gene coding for a (nearly) identical protein (Accession number CKG89570.1) found in a singular isolate classified as *Streptococcus pneumoniae*\textsuperscript{55} casting doubt onto whether or not BC_0987 is an actual spore protein as *S.pneumoniae* is incapable of producing spores. Interestingly, other verified spore proteins, such as CotX and CotB, also have nearly identical homologues (CJGK03595.1 and CJA37784.1) in this particular isolate of *S. pneumoniae* indicating it perhaps has several residual nonfunctional spore protein homologues or this isolate is simply misclassified. Though the function of BC_0987 needs to be further confirmed via traditional knockout analysis, based on the data shown here we propose to rename this protein to structural abundant spore protein SasS.

BC_2569 contains collagen-like regions, which relates it to one of the best characterized exosporium proteins BclA. BclA is glycosylated protein, rich in GXX repeats known to be a major structural component of the exosporium in *Bacillus anthracis*.\textsuperscript{57} BclA is targeted to the exosporium by an N-terminal sequence, which is similar to an N-terminal sequence found in BC_2569 (LLGPTLPAPPP vs consensus L_{1-2}\textsubscript{GGP}T_{1-2}P_{3-4}P_{5-6}).\textsuperscript{38} Applying the same reasoning for classification as a structural protein to BC_2569 suggests it is in fact not a structural protein, or at least not a major one like BclA. Nonetheless, its structure can be classified as belonging to the group of Bcl proteins and Abhyankar et al.\textsuperscript{14} previously also identified it in *B. cereus* exosporium isolates. Here we propose to rename it BclF.

Alr1 and IunH are enzymes which modify germinant molecules, L-alanine and inosine, respectively, to prevent the initiation of germination prematurely.\textsuperscript{51,56} As such they are of crucial importance for the survival of the bacterium. Kinetics of triggering germination in response to stimuli with germinants rely on the conversion rate, which depends on the enzyme’s kinetic parameters and its abundance. The obtained abundances may therefore be useful to reinforce models for protein germination.

**Stoichiometry: Interactors Have Similar Abundance**

Proteins known to interact or function in a similar process appeared to be present at near equimolar ratios. For instance, paralogs CotB1 and CotB2 are roughly equally abundant (5:4) and, though structurally different, are thought to fulfill a similar role as a structural component of the exosporium.\textsuperscript{37} Their equal abundance suggests they may interact directly and have equally important roles in ensuring spore structure.

CotE and SaFA are two of the major morphological proteins of the spore coat, each being required for deposition of an entire coat layer.\textsuperscript{56,57} In *Bacillus subtilis*, the two proteins have been shown to be anchored to the developing spore through the protein SpoVID though not at the same binding site.\textsuperscript{59,60} They appear to be present in a roughly 4:3 ratio CotE:SaFA numerically suggesting the nonequimolar stoichiometry in interaction. This means either a fraction of the CotE present in spores does not interact with SpoVID or there is space for more copies of CotE to interact compared to SaFA. Considering their respective protein masses, we can get an indication of the relative space occupied by the respective proteins. Multiplying the observed numeric ratio with respective protein masses, the proteins show a 2:5 ratio CotE:SaFA. This indicates it is likely that there is indeed space for multiple copies of CotE to interact with one copy of SpoVID. To get a full picture of protein stoichiometry of the morphogenetic proteins in the spore coat layers and update the proposed structural model, quantification of SpoVID and SpoIVA is pending. Unfortunately, in our hands SpoVID and SpoIVA were not reproducibly detected when selecting proteins to include in the QconCAT and were therefore not included in the quantitative analyses. Hence, determining full stoichiometry of these morphogenetic proteins awaits complementary approaches.

YpeB and SleB are known to interact\textsuperscript{34} and appear to be present in a 5:4 ratio. YaaH is known to break down products of SleB and CwlJ\textsuperscript{59} and is more abundant. It would be interesting to see if the combined copy numbers of SleB and CwlJ equal that of YaaH or if these enzymes are kinetically balanced instead. Unfortunately, CwlJ is structurally incompatible with QconCAT analysis as possible tryptic peptides are either N-terminal, contain cysteines, or end in RR which are prone to missed cleavage and was therefore not included.

**Biomarker Targets: Immunogenicity Prediction**

The obtained data are of interest for practical applications for spore recognition by proposing potential protein biomarkers. As a ubiquitous food pathogen, proper detection of *B. cereus* is important for food safety. The golden standard for detection (ISO 7932) is still enumeration of colonies which is a time-consuming approach.\textsuperscript{1} Furthermore, spores show heterogeneity in their germination response, and so-called persistors and superdormant spores\textsuperscript{61-63} are difficult to assess with this approach. A way to circumvent these issues as well as potentially increasing sensitivity and assay speed would be to use biomarker detection instead. As the majority of the outside of the spore consists of proteins, these would be a prime target. To assess the potential biomarker targets, we used the SVMTrIP algorithm\textsuperscript{22} to predict immunogenic epitopes of the most abundant proteins (CotB1, CotB2, InhA, CotX, and BC_0987), the results of which are shown in Table 6. Luckily, these proteins are all located on the outermost layers of the spore indicating they would likely be
Table 6. Epitope Predictions
table 6

| protein | location | epitope | score |
|---------|----------|---------|-------|
| CotB1   | 54–73    | KHLKSTNKANEGSSDCEW | 1.000 |
|         | 100–119  | GPEKVEGLDQVSCDFVTLL | 0.548 |
| CotB2   | 5–24     | LCCDIQKCLGVLGETVKNLRG | 1.000 |
|         | 123–142  | GDEFYIVYKSHISVQVVK | 0.932 |
|         | 46–65    | HGELVYQYKHKSLKKVKK | 0.702 |
| InhA    | 436–455  | DYEKLNKGILATYLDQSVT | 1.000 |
|         | 692–711  | AIVHTNLKPGTVSSTRFQ | 0.903 |
|         | 513–532  | FKSLLYIEEAIDFLLEIHAVT | 0.846 |
|         | 613–632  | TPQKLDGFAVSNGETKSSH | 0.832 |
|         | 643–662  | GSDNALKFARGPVYAGMV | 0.766 |
|         | 260–279  | DALKAADVSGDLSEFQFQD | 0.744 |
|         | 18–37    | TQSAYAETPANTKTSIPVD | 0.708 |
|         | 555–574  | TNGKIDKSYDLSQFKGKVK | 0.697 |
|         | 189–208  | LFGNEPFTLDGSKIEFFKQ | 0.671 |
|         | 90–109   | LTKEASDFLKKVAKADTK | 0.662 |
| CotX1   | 36–55    | Q5KTYQISEEITIVSDAD | 1.000 |
|         | 106–125  | QNKQETVIRNRSNVYTTYT | 0.534 |
| BC_0987 | none predicted |

“Immunogenic regions with location within their respective protein and score based on the SVMtTriP algorithm. Epitopes recommended by the algorithm are indicated in bold. These were submitted to BLAST and shown to be Bacillus cereus group specific with one exception in S. pneumoniae (see Results and Discussion).”

The Q-conCAT approach for quantification of proteins was applied in the heterogeneous system of the spore coat insoluble fraction of wild-type B. cereus ATCC14579 and a ΔCotE knockout strain. In this manner, 21 proteins from across all layers of the spore were quantified covering 5.66% ± 0.51 of the total spore weight in wild type and 4.13% ± 0.14 in the ΔCotE mutant strain. The protein abundance spanned three orders of magnitude and suggested functional properties of proteins correlate with their abundance. The most abundant proteins are generally known structural components, which enabled us to functionally describe the previously uncharacterized protein BC_0987, now renamed SasS. The obtained information on stoichiometry of both putative and known interactors showed them to be approaching equimolar quantities, which is of importance to developing quantitative models for spore macromolecular structures.

In line with the effects observed qualitatively in the ΔCotE mutant the 20 proteins (excluding CotE) quantified using the Q-conCAT approach, known exosporium proteins were generally less abundant in the CotE knockout whereas inner coat associated proteins were generally more abundant.

The five most abundant proteins are considered as biomarkers for detection of B. cereus group spores and immunogenic epitopes have been predicted to be present in CotB1, CotB2, CotX and InhA. These proteins are indicated to be present in the outermost layers of the spore making them ideal candidate biomarkers.

These results further contribute to our understanding of the spore structure, in particular of the outer layers of the spore. This may be applied in a food safety scenario where the makeup of the outside of the spore is of crucial importance for development of detection techniques.

- ASSOCIATED CONTENT

- Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.7b00732.

Q-conCAT design and expression report by PolyQuant (PDF)
Graphs of time required for completed digestion of each Q-peptide (XLSX)
Protein identifications in B. cereus and ΔCotE spore isolates (XLSX)
Isotope ratios (L/H) of Q-peptides in B. cereus and ΔCotE (XLSX)
SI table of contents (PDF)

- AUTHOR INFORMATION

Corresponding Author
E-mail: s.k.stelder@uva.nl.

ORCID
Sacha K. Stelder: 0000-0002-7348-4849

Author Contributions
*These authors contributed equally.

Notes
The authors declare no competing financial interest.

The mass spectrometry proteomics data of the wild-type Q-conCAT analyses have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the data set identifier PXD003569.

- ACKNOWLEDGMENTS

This project is supported by NanoNextNL, a micro and nanotechnology consortium of the Government of The Netherlands.
Netherlands and 130 partners from academia and industry. The authors would like to thank Christelle Bressure-Isard, Véronique Broussolle, and Frederic Carlin for kindly providing us with the CotE knockout strain.

REFERENCES

(1) EFSA. Opinion of the Scientific Panel on Biological Hazards on Bacillus cereus and other Bacillus spp in foodstuffs. EFSA J. 2005, 175, 1–48.

(2) Aronson, A. I.; Fitz-James, P. Structure and morphogenesis of the bacterial spore coat. *Bacteriol. Rev.* 1976, 40 (2), 360–402.

(3) Henriques, A. O.; Moran, C. P., Jr Structure, assembly, and function of the spore surface layers. *Annu. Rev. Microbiol.* 2007, 61, 555–588.

(4) Henriques, A. O.; Moran, C. P.; et al. Structure and assembly of the bacterial endospore coat. *Methods 2000*, 20 (1), 95–110.

(5) Giorno, R.; Bozue, J.; Cote, C.; Wenzel, T.; Moody, K.-S.; Mallozzi, M.; Ryan, M.; Wang, R.; Zielke, R.; Maddock, J. R.; et al. Morphogenesis of the Bacillus anthracis Spore. *J. Bacteriol.* 2007, 189 (3), 691–705.

(6) Bressure-Isard, C.; Bornard, I.; Henriques, A. O.; Carlin, F.; Broussolle, V. Sporulation temperature reveals a requirement for CotE in the assembly of both the coat and exosporium layers of *Bacillus cereus* spores. *Appl. Environ. Microbiol.* 2016, 82, 232.

(7) Zheng, L. B.; Donovan, W. P.; Fitz-James, P. C.; Losick, R. Gene encoding a morphogenetic protein required in the assembly of the outer coat of the Bacillus subtilis endospore. *Genes Dev. 1988*, 2 (8), 1047–1054.

(8) Ozin, A. J.; Henriques, A. O.; Yi, H.; Moran, C. P. Morphogenetic proteins SpoVID and SaFa form a complex during assembly of the Bacillus subtilis spore coat. *J. Bacteriol.* 2000, 182 (7), 1828–1833.

(9) Imamura, D.; Kuwana, R.; Takamatsu, H.; Watabe, K. Proteins Involved in Formation of the Outermost Layer of Bacillus subtilis Spores. *J. Bacteriol.* 2011, 193 (16), 4075–4080.

(10) McKenney, P. T.; Driks, A.; Eichenberger, P. The Bacillus subtilis endospore: assembly and functions of the multilayered coat. *Nat. Rev. Microbiol.* 2012, 11 (1), 33–44.

(11) Stewart, G. C. The Exosporium Layer of Bacterial Spores: a Connection to the Environment and the Infected Host. *Microbiol. Mol. Biol. Rev.* 2015, 79 (4), 437–457.

(12) Abhyankar, W.; Pandey, R.; Ter Beek, A.; Brul, S.; de Koning, L. J.; de Koster, C. G. Reinforcement of Bacillus subtilis spores by cross-linking of outer coat proteins during maturation. *Food Microbiol.* 2015, 45, 54–62.

(13) Brownridge, P.; Holman, S. W.; Gaskell, S. J.; Grant, C. M.; Harman, V. M.; Hubbard, S. J.; Lanthaler, K.; Lawless, C.; O’culain, R.; Sims, P.; et al. Global absolute quantification of a proteome: Challenges in the deployment of a QcomCAT strategy. *Proteomics* 2011, 11 (15), 2957–2970.

(14) Abhyankar, W.; Hossain, A. H.; Djasaputra, A.; Permpoonpattana, P.; Ter Beek, A.; Dekker, H. L.; Cutting, S. M.; Brul, S.; de Koning, L. J.; de Koster, C. G. In Pursuit of Protein Targets: Proteomic Characterization of Bacterial Spore Outer Layers. *J. Proteome Res.* 2013, 12 (10), 4507–4521.

(15) Brownridge, P.; Beynon, R. J. The importance of the digest: Proteolysis and absolute quantification in proteomics. *Methods 2011*, 54 (1), 351–360.

(16) Abhyankar, W.; Beek, A. T.; Dekker, H.; Kort, R.; Brul, S.; de Koster, C. G. Gel-free proteomic identification of the Bacillus subtilis insoluble spore coat protein fraction. *Proteomics* 2011, 11 (23), 4541–4550.

(17) Kuwana, R.; Okuda, N.; Takamatsu, H.; Watabe, K. Modification of GtrQ reveals a functional relationship between Tgl and YabG in the coat of Bacillus subtilis spores. *J. Biochem.* 2006, 139 (5), 887–901.

(18) Fernandes, C. G.; Plácido, D.; Lousa, D.; Brito, J. A.; Isidro, A.; Soares, C. M.; Polil, J.; Carrondo, M. A.; Archer, M.; Henriques, A. O. Structural and Functional Characterization of an Ancient Bacterial Transglutaminase Sheds Light on the Minimal Requirements for Protein Cross-Linking. *Biochemistry 2015*, 54 (37), 5723–5734.
Application in Bacillus anthracis Detection and Fingerprinting. *Appl. Environ. Microbiol.* **2009**, *75* (22), 7163–7172.

(38) Bagyan, I.; Noback, M.; Bron, S.; Paidhungat, M.; Setlow, P. Characterization of yehC, a new forespore-specific gene of Bacillus subtilis. *Gene* **1998**, *212* (2), 179–188.

(39) Takamatsu, H.; Kodama, T.; Nakayama, T.; Watabe, K. Characterization of the yrbA Gene of Bacillus subtilis, Involved in Resistance and Germination of Spores. *J. Bacteriol.* **1999**, *181* (16), 4986–4994.

(40) Cruz-Mora, J.; Pérez-Valdespino, A.; Gupta, S.; Withange, N.; Kuwana, R.; Takamatsu, H.; Christie, G.; Setlow, P. The GerW Protein Is Not Involved in the Germination of Spores of Bacillus Species. *PLoS One* **2015**, *10* (3), e0119125.

(41) Ragkousi, K.; Setlow, P. Transglutaminase-Mediated Cross-Linking of GerQ in the Coats of Bacillus subtilis Spores. *J. Bacteriol.* **2004**, *186* (17), 5567–5575.

(42) Ragkousi, K.; Eichenberger, P.; van Ooiij, C.; Setlow, P. Identification of a new gene essential for germination of Bacillus subtilis spores with Ca2+-dipicolinate. *J. Bacteriol.* **2003**, *185* (7), 2315–2329.

(43) Terry, C.; Shepherd, A.; Radford, D. S.; Moir, A.; Bullough, P. A. YwdL in Bacillus cereus: Its Role in Germination and Exosporium Structure. *PLoS One* **2011**, *6* (8), e23801.

(44) Jao, S.-C.; Ma, K.; Talafous, J.; Orlando, R.; Zagorski, M. G. Trifluoroacetic acid pretreatment reproducibly disaggregates the amyloid β-peptide. *Amyloid* **1997**, *4* (4), 240–252.

(45) Vizcaíno, J. A.; Deutsch, E. W.; Wang, R.; Csordas, A.; Reisinger, F.; Rios, D.; Diñes, J. A.; Sun, Z.; Farrah, T.; Bandeira, N.; et al. ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* **2014**, *32* (3), 223–226.

(46) Efron, B. Bootstrap Methods: Another Look at the Jackknife. *Ann. Stat.* **1979**, *7* (1), 1–26.

(47) Davison, A. C.; Hinkley, D. V. *Bootstrap Methods and Their Application (Cambridge Series in Statistical and Probabilistic Mathematics)*; Cambridge University Press: Cambridge, 1997.

(48) Steichen, C. T.; Kearney, J. F.; Turnbough, C. L. Non-uniform assembly of the Bacillus anthracis exosporium and a bottle cap model for spore germination and outgrowth. *Mol. Microbiol.* **2007**, *64* (2), 359–367.

(49) de Francesco, M.; Jacobs, J. Z.; Nunes, F.; Serrano, M.; McKenney, P. T.; Chua, M.-H.; Henriques, A. O.; Eichenberger, P. Physical Interaction between Coat Morphogenetic Proteins SpoVID and CotE Is Necessary for Spore Encasement in Bacillus subtilis. *J. Bacteriol.* **2012**, *194* (18), 4941–4950.

(50) Costa, T.; Serrano, M.; Steil, L.; Volker, U.; Moran, C. P.; Henriques, A. O. The Timing of cotE Expression Affects Bacillus subtilis Spore Coat Morphology but Not Lysozyme Resistance. *J. Bacteriol.* **2007**, *189* (6), 2401–2410.

(51) Carrera, M.; Zandoneni, R. O.; Sagripanti, J.-L. Wet and dry density of *Bacillus anthracis* and other *Bacillus* species. *J. Appl. Microbiol.* **2008**, *105* (1), 68–77.

(52) Carrera, M.; Zandoneni, R. O.; Fitzgibbon, J.; Sagripanti, J.-L. Difference between the spore sizes of Bacillus anthracis and other Bacillus species. *J. Appl. Microbiol.* **2007**, *102* (2), 303–312.

(53) Gould, G. W.; Hurst, A. *The Bacterial Spore*; Academic Press, 1969.

(54) Gerber, S. A.; Rush, J.; Stemman, O.; Kirschner, M. W.; Gygi, S. P. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (12), 6940–6945.

(55) Monroe, A.; Setlow, P. Localization of the Transglutaminase Cross-Linking Sites in the Bacillus subtilis Spore Coat Protein GerQ. *J. Bacteriol.* **2006**, *188* (21), 7609–7616.

(56) Chewapreecha, C.; Harris, S. R.; Croucher, N. J.; Turner, C.; Martinson, P.; Cheng, L.; Pessia, A.; Aanensen, D. M.; Mather, A. E.; Page, A. J.; et al. Dense genomic sampling identifies highways of pneumococcal recombination. *Nat. Genet.* **2014**, *46* (3), 305–309.

(57) Sylvester, P.; Couture-Tosi, E.; Mock, M. A collagen-like surface glycoprotein is a structural component of the Bacillus anthracis exosporium. *Mol. Microbiol.* **2002**, *45* (1), 169–178.

(58) Thompson, B. M.; Stewart, G. C. Targeting of the BcIA and BcIB proteins to the *Bacillus anthracis* spore surface. *Mol. Microbiol.* **2008**, *70* (2), 421–434.

(59) Steichen, C. T.; Kearney, J. F.; Turnbough, C. L. Characterization of the Exosporium Basal Layer Protein BxpB of Bacillus anthracis. *J. Bacteriol.* **2005**, *187* (17), 5868–5876.

(60) Mullerova, D.; Krajickova, D.; Barak, I. Interactions between *Bacillus subtilis* early spore coat morphogenetic proteins. *FEBS Microbiol. Lett.* **2009**, *299* (1), 74–85.

(61) Gould, G. W. Symposium on bacterial spores: IV. Germination and the problem of dormancy. *J. Appl. Bacteriol.* **1970**, *33* (1), 34–49.

(62) Ghosh, S.; Setlow, P. Isolation and Characterization of Superdormant Spores of Bacillus Species. *J. Bacteriol.* **2009**, *191* (6), 1787–1797.

(63) Pandey, R.; Ter Beek, A.; Vischer, N. O. E.; Smelt, J. P. P. M.; Brul, S.; Manders, E. M. M. Live Cell Imaging of Germination and Outgrowth of Individual Bacillus subtilis Spores; The Effect of Heat Stress Quantitatively Analyzed with SporeTracker. *PLoS One* **2013**, *8* (3), e58972.