

## RESEARCH ARTICLE

# Early events of *Bacillus anthracis* germination identified by time-course quantitative proteomics

Pratik Jagtap<sup>1</sup>, George Michailidis<sup>2</sup>, Ryszard Zielke<sup>1</sup>, Angela K. Walker<sup>3</sup>, Nishi Patel<sup>1</sup>, John R. Strahler<sup>3</sup>, Adam Driks<sup>4</sup>, Philip C. Andrews<sup>3</sup> and Janine R. Maddock<sup>1</sup>

<sup>1</sup> Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, MI, USA

<sup>2</sup> Department of Statistics, University of Michigan, Ann Arbor, MI, USA

<sup>3</sup> Department of Biological Chemistry, University of Michigan, Ann Arbor, MI, USA

<sup>4</sup> Department of Microbiology and Immunology, Loyola University Medical Center, Maywood, IL, USA

Germination of *Bacillus anthracis* spores involves rehydration of the spore interior and rapid degradation of several of the protective layers, including the spore coat. Here, we examine the temporal changes that occur during *B. anthracis* spore germination using an isobaric tagging system. Over the course of 17 min from the onset of germination, the levels of at least 19 spore proteins significantly decrease. Included are acid-soluble proteins, several known and predicted coat proteins, and proteins of unknown function. Over half of these proteins are small (less than 100 amino acids) and would have been undetectable by conventional gel-based analysis. We also identified 20 proteins, whose levels modestly increased at the later time points when metabolism has likely resumed. Taken together, our data show that isobaric labeling of complex mixtures is particularly effective for temporal studies. Furthermore, we describe a rigorous statistical approach to define relevant changes that takes into account the nature of data obtained from multidimensional protein identification technology coupled with the use of isobaric tags. This study provides an expanded list of the proteins that may be involved in germination of the *B. anthracis* spore and their relative levels during germination.

Received: April 27, 2006

Revised: June 1, 2006

Accepted: June 4, 2006

**Keywords:**

*Bacillus anthracis* / Germination / iTRAQ / Proteome / Spore

## 1 Introduction

*Bacilli* and *Clostridia*, in response to nutrient deprivation, form spores – cells possessing a series of protective structures that encase the bacterial genome, allowing it to resist a wide range of external challenges and remain metabolically

inactive for extended periods of time. Bacterial spores are unique in nature; no other cell types possess their structure or protein composition, nor are as well adapted to dormancy.

In *Bacillus anthracis*, the outermost structure of the spore is the exosporium, a pleiomorphic shell of undefined function that surrounds spores of some species, including non-pathogenic organisms such as *B. megaterium* [1], but not others. Within this is the coat, a multilayered protein shell that protects the interior and participates in revival from dormancy (the germination process) [2–4]. The coat is present in every species that has been examined [5, 6]. Underneath the coat is a thick layer of a specialized peptidoglycan known as the cortex; the cortex directly surrounds the core, the compartment that houses the spore DNA. Germination, which initiates rapidly once nutrient returns to the environment, begins with rehydration of the spore interior and initiation of degradation of the coat, cortex and proteins in

**Correspondence:** Dr. Janine Maddock, Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, MI 48109–1048, USA

**E-mail:** maddock@umich.edu

**Fax:** +1-734-647-0884

**Abbreviations:** **DSM**, DIFCO sporulation medium; **GPR**, germination protease; **iTRAQ**<sup>™</sup>, isotope tag for relative and absolute quantitation, isotopically enriched forms of N-methylpiperazine acetic acid N-hydroxy succinimide; **SASP**, small acid soluble protein; **SCX**, strong cation exchange

the core. These steps occur in the absence of any detectable metabolic activity [7]. Because the early events in germination do not involve gene expression or protein synthesis, analysis of these events is well suited to a proteomics approach.

The best understood of the events occurring in the first few minutes of germination is the proteolysis of proteins that bind the spore chromosome, the small acid soluble proteins (SASP) [8]. These abundant small proteins are required for protection of the spore DNA, but must be degraded for successful transcription of the chromosome after germination. SASP degradation not only frees the DNA, but also provides a pool of amino acids to support the subsequent initiation of metabolism. Less well understood is the degradation of the cortex. This is accomplished by spore-associated peptidoglycan degrading enzymes, a number of which have been isolated from spores of a variety of species [9]. Less clear is the mechanism by which the coat is shed, something that must occur for the germinated cell to expand its volume as new cell growth begins. Electron microscopy shows that coat ultrastructure changes as germination proceeds [10]. Fragments of the coat are usually abundant in a germinated spore culture, frequently seeming to have cracked in half [2], suggesting that the coat harbors a protease that helps to fragment it without destroying it. Very little is known about proteolytic events during germination (see, for example, [11]); in particular, the question of whether coat degradation is targeted to a subset of polypeptide species remains open. This is an especially difficult question to address with non-proteomic approaches, due to the copious number of both candidate substrates and potential proteolytic products.

Due to their impregnable nature, decontamination of spores is a long-standing challenge that has yet to be solved in an entirely satisfactory manner. Currently used methods involve either extremely harsh approaches that are essentially impractical except on a small scale within a laboratory (such as autoclaving), or require evacuation of the contaminated space and significant amounts of equipment (such as treatment with chlorine dioxide gas). A wide range of scenarios in which neither technology is suitable is easy to imagine, particularly in light of the potential use of *B. anthracis* spores as a biological weapon. One possible route to superior decontamination strategies is study of the mechanism of spore coat degradation. If this process could be mimicked by application of a reagent such as a protease cocktail, for example, the spore would be rendered sensitive to typical bactericidal chemicals such as lysozyme. A first, critical step in understanding coat shedding, therefore, is identification of proteolytic targets, understanding of the way in which they are cleaved and identification of the proteases responsible. To address these questions, we chose to directly monitor changes in steady state coat protein levels during germination and early outgrowth, using a recently described technique that allows relative quantification of changes in protein samples, isotope tag for relative and absolute quan-

titation (iTRAQ) [12]. In two independent experiments, we labeled with isobaric tags, individually, tryptic peptides from *B. anthracis* spores collected over a 17-min germination time course. Our studies demonstrate that the isobaric tags represent an efficient and sensitive method to measure subtle changes in protein levels, with the additional advantage that offline 2-D LC-MS/MS allows thorough analysis of samples by permitting the mass spectrometer to re-query the column fractions multiple times. As expected, levels of the majority of proteins did not change over the course of this study. Significantly, however, the levels of 19 proteins decreased and the levels of 20 proteins modestly increased. In general, proteins that increased are indicative of entry into vegetative growth. The degraded proteins are of particular interest as they represent both known and novel targets of early proteolysis.

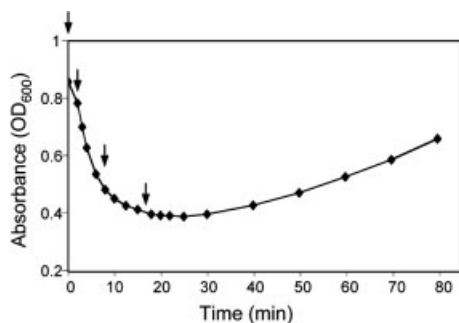
## 2 Materials and methods

### 2.1 Sporulation

*B. anthracis* Sterne (from Paul Jackson, Los Alamos National Laboratory, USA) spores were generated after nutrient exhaustion in DIFCO sporulation medium (DSM) [13] as described [14]. Briefly, *B. anthracis* cells grown on Luria broth -agar plates supplemented with 5 g/L of nutrient broth (DIFCO) were inoculated in DSM medium at an  $OD_{600}$  of 0.1. The *B. anthracis* spores were collected post-sporulation (after growth for approximately 26 h) in DSM media. The spores were separated from vegetative cells by heat treatment at 65°C for 30 min and Renografin gradient separation, washed three times with MilliQ water and centrifuged at  $10\,000 \times g$  at 4°C for 10 min. The purity of the spores was >95% as determined by their phase brightness. The spore suspension was stored at 4°C for up to a week until further processing.

### 2.2 Sample preparation

The *B. anthracis* (Sterne) spores were subjected to heat treatment (65°C for 30 min) and cooled on ice. For germination, spores were inoculated at  $OD_{600} = 1.0$  in DSM media at 37°C and the A,  $OD_{600}$ , was monitored (Fig. 1). Germinating spores were collected at 0, 2, 7 and 17 min after inoculation (Fig. 1) and germination was stopped by the addition of PMSF, 1 mM final concentration), a serine protease inhibitor, known to inhibit germination [15–17]. Spores were collected by centrifugation ( $10\,000 \times g$ , 3 min at 4°C) and washed in 3 mL ice-cold MilliQ water containing 1 mM PMSF and complete protease inhibitor mix (Roche Molecular Biochemicals). The washed pellets were immediately frozen using liquid nitrogen and stored at  $-80^\circ\text{C}$ . Proteins were obtained and analyzed from two independently obtained spore preparations.



**Figure 1.** Germination of *B. anthracis* spores. *B. anthracis* spores were allowed to germinate and the relative absorbance ( $OD_{600}$ ) was monitored over time, as indicated. At the times indicated by arrows (0, 2, 7 and 17 min), samples were removed, proteins were processed and resulting peptides labeled with isobaric tags (114, 115, 116 and 117, respectively) as described in the Section 2.

Spores were mechanically disrupted with 0.1 mm Zirconium beads (Biospec Products) in 200  $\mu$ L 0.1% SDS containing complete protease inhibitor mix in O-ring sealed microfuge tubes using a Minibead beater for seven cycles of 1 min, followed each time by incubation on ice. The lysates were separated from beads and unbroken cells by centrifugation (6,000 g, 10 min) and the supernatant protein fraction was stored at  $-80^{\circ}\text{C}$  until further processing.

Prior to iTRAQ labeling, protein samples were precipitated with acetone, as follows. Ice-cold acetone was added to each sample to a final concentration of 80% and the samples were incubated overnight at  $-20^{\circ}\text{C}$ . After centrifugation (14 000  $\times$  g, 10 min), the pellet was resuspended in 100% cold acetone, recentrifuged, and resolubilized in 0.05% SDS, 5 mM Tris pH 8.8. The concentration of solubilized protein was estimated using the 2-D Quant kit, according to the manufacturer's instructions (Amersham Biosciences).

### 2.3 Protein isobaric labeling with iTRAQ reagents

For four-plex isobaric labeling, separate aliquots of *B. anthracis* proteins were treated in parallel, essentially as described [12]. Stock reagents and buffer [SDS, triethylammonium bicarbonate (TEAB), Tris-2-carboxyethyl phosphine (TCEP), methylmethane thiosulfate (MMS) and four isobaric tagging reagents] were obtained in kit form (Applied Biosystems). Protein (100  $\mu$ g in 20  $\mu$ L 0.5 M TEAB containing 0.05% SDS) was reduced with 2.5 mM TCEP at  $60^{\circ}\text{C}$  for 1 h and cysteine residues blocked with 10 mM MMS at room temperature for 15 min. Protein was digested with trypsin (porcine modified, Promega; 1:20, w/w) for 20 h at  $40^{\circ}\text{C}$ . iTRAQ reagent (1 U in ethanol) was added directly to the protein digest (70% ethanol final) and the mixture was incubated at room temperature for 1 h. The reaction was quenched by addition of nine volumes 0.1% TFA in water (Optima grade, Fisher Scientific). The reaction mixtures were combined and stored at  $-20^{\circ}\text{C}$ .

### 2.4 RP HPLC separation

An aliquot of the four-plex peptide mixture (200  $\mu$ g) was applied to a sulfoethyl aspartamide strong cation exchange (SCX) spin column (SEM HIL-SCX, PolyLC, The Nest Group, Southboro, MA) equilibrated with 20%  $\text{CH}_3\text{CN}$  in 10 mM  $\text{KH}_2\text{PO}_4$  pH 4.5. For peptide adsorption to the column and subsequent washing and elution steps, centrifugal force was applied in approximately 2-s bursts such that 50  $\mu$ L of solution passed through the column over a 50–60-s interval. Excess reagent was washed from the column with 800  $\mu$ L equilibration buffer. Peptides were eluted using 50  $\mu$ L volumes of KCl in equilibration buffer in a stepwise gradient (25, 50, 75, 100, 150, 225 and 350 mM KCl). Fractions were dried in a vacuum centrifuge.

Peptides from the SCX fractions were separated by  $\text{C}_{18}$  nano LC using a 1100 Series nano HPLC equipped with  $\mu$ WPS autosampler, 2/10 microvalve, MWD UV detector (214 nm) and Micro-FC fraction collector/spotter (Agilent). Each SCX salt step was reconstituted with 43  $\mu$ L 0.1% TFA (v/v in water) and centrifuged. With the valve in LOAD position, 40  $\mu$ L sample was injected onto a  $\text{C}_{18}$  cartridge (Zorbax300SB, 5  $\mu$ m,  $5 \times 0.3$  Lmm; Agilent), desalted with solvent C ( $\text{CH}_3\text{CN}:\text{H}_2\text{O}:\text{TFA}$ , 5:95:0.1, v/v/v) at 20  $\mu$ L/min for 9 min and the effluent directed to waste. In ELUTE position, the enrichment cartridge was placed ahead of a  $\text{C}_{18}$  column (Zorbax300SB, 3.5  $\mu$ m,  $150 \times 0.1$  mm; Agilent) previously equilibrated with solvent A ( $\text{CH}_3\text{CN}:\text{H}_2\text{O}:\text{TFA}$ , 6.5:93.5:0.1, v/v/v). Peptides were eluted with a 6.5–50% gradient of solvent B ( $\text{CH}_3\text{CN}:\text{H}_2\text{O}:\text{TFA}$ , 90:10:0.1, v/v/v) over 120 min at a flow rate of 0.3  $\mu$ L/min. Column effluent was mixed with matrix [2 mg/mL CHCA in  $\text{CH}_3\text{OH}:\text{isopropanol}:\text{CH}_3\text{CN}:\text{H}_2\text{O}:\text{acetic acid}$  (12:33.3:52:36:0.7, v/v/v/v) containing 10 mM ammonium phosphate] in a mixing Tee (micro Tee, Agilent). Matrix was delivered with a PHD200 infusion pump (Harvard Apparatus) at 0.8  $\mu$ L/min. Fractions were spotted at 30-s intervals onto stainless steel MALDI targets (192 wells/plate, Applied Biosystems).

### 2.5 MS

Mass spectra were acquired on an Applied Biosystems 4700 Proteomics Analyzer (TOF/TOF). MS survey spectra were acquired from 800–3500 Da for each fraction. The eight most intense peaks in each MS spectrum above an S/N threshold of 100 were selected for MS/MS analysis. For peaks observed in consecutive fractions, only the most intense instance was selected for MS/MS analysis. Fragmentation of the labeled peptides was induced by the use of atmosphere as a collision gas with a pressure of  $\sim 6 \times 10^{-7}$  torr and collision energy of 1 kV.

Peptide identifications were obtained using GPS Explorer (v3.0, Applied Biosystems), which is a front end for the MASCOT search engine (v1.9 MatrixScience, London UK). Each MS/MS spectrum was searched against NCBI nr (July 2004, October and December 2005) filtering on firmi-

cutes. The genomes of the Sterne strain and the Ames strain were also searched in all six reading frames. Trypsin specificity with one missed cleavage was selected. S-mercapto-methylcysteine, as well as the N-terminal and lysine iTRAQ labels were selected as fixed modifications, as both the free N-terminal amino group and the free  $\alpha$ -amino group of lysine residues in proteolytic peptides are derivatized by iTRAQ reagents. Oxidized methionine was considered a variable modification. The precursor tolerance and MS/MS fragment tolerances were set to  $\pm 0.7$  and  $\pm 0.3$  Da, respectively.

The peak areas were corrected for overlapping isotope contributions from the iTRAQ tags according to the certificate of analysis provided by Applied Biosystems. The corrected peak areas represent quantitative measurements observed at 114.1, 115.1, 116.1 and 117.1  $m/z$ , which correspond to 0, 2, 7 and 17 min of germination, respectively.

## 2.6 Normalization of data

Normalization of raw peak areas was critical to quantification, as differences in actual, as opposed to theoretical, total protein labeled would result in deviations from the expected ratios. In this study, we did not anticipate many changes to occur during germination and outgrowth and, therefore, normalized the data at the level of the individual peptide peak areas.

Table 1 shows summary statistics of the peak area measurements, at the peptide level, for both biological replicate experiments. The median peak areas vary from 56 145 to 73 591 and 54 108 to 71 616 for first and second replicates, respectively (Table 1). A similar trend is apparent for peptides with lower or higher peak areas (first and third quartiles) (Table 1). These data indicate that, without normalization, the average relative ratios would deviate from 1. To correct for these differences, normalization was accomplished by matching the quantiles of the distributions of the 115, 116, and 117 measurements to the quantiles of the 114 measurements using a monotone piece-wise linear function [18]. Summary statistics indicate that, after normalization, all four measurements are well matched (Table 2).

## 2.7 Quantitation of differentially expressed proteins

In determining whether a protein was increased or decreased, we took into consideration the organization of the data; namely, that each protein can potentially be identified by a number of peptides and every peptide can be measured multiple times. As previously noted [19], there are cases in which the relative expression levels exhibit a relatively high degree of variability amongst multiple measurements of the same peptide for a specific protein and, even more importantly, across peptides for the same protein. A simple averaging of relative ratios as suggested in the literature [12, 19] ignores (or bypasses) this issue. Our method, however, rigorously accounts for these two sources of variability (the parameters  $\sigma^2$  and  $\tau^2$ ), which are each taken into consideration

**Table 1.** Average peak areas in the 114.1–117.1  $m/z$  range from raw data

	114	115	116	117
<b>First replicate</b>				
First quartile	25 863	29 970	33 953	30 179
Median	56 145	64 307	73 591	69 553
Third quartile	110 806	126 247	143 206	137 061
<b>Second replicate</b>				
First quartile	30 805	35 775	32 668	28 226
Median	63 040	71 616	64 213	54 108
Third quartile	125 175	137 581	123 324	106 042

**Table 2.** Average peak areas in the 114.1–117.1 range from normalized data

	114	115	116	117
<b>First replicate</b>				
First quartile	25 863	25 869	25 738	25 982
Median	56 145	56 191	55 937	56 185
Third quartile	110 806	111 561	111 497	111 523
<b>Second replicate</b>				
First quartile	30 805	30 823	30 890	30 810
Median	63 040	62 801	62 834	62 960
Third quartile	125 175	124 892	124 614	125 252

when calculating the standard error of the ratio  $R$  that was used in the subsequent hypothesis testing procedure for biological activity (see below). Thus, let  $r(i,j,k,l)$  denote the ratio of the corrected and normalized peak area of MS/MS spectrum  $k$  corresponding to identified peptide  $j$  for protein  $i$  for labeled sample  $l = 115, 116, 117$ , by the corresponding corrected and normalized peak area for 114. For ease of notation we drop from the subsequent discussion indices  $i$  and  $l$ , as the proposed model is used for each protein separately and for each (115–117) labeled sample. We model these ratios in a  $\log_2$  scale, to overcome the fact that the ratio scale is bounded from below by 0, by the following random one-way analysis of variance (ANOVA) model:  $\log_2(r(j,k)) = R + R(j) + u(j,k)$ , where  $R$  is the relative abundance of protein in the labeled samples (technically, the overall mean of the modeled ratios at the protein level),  $R(j)$  is the specific effect of peptide  $j$  on the ratio  $r(j,k)$ , and  $u(j,k)$  is an MS/MS spectrum effect assumed to be normally distributed with mean zero and constant variance  $\sigma^2$ , where  $\sigma^2$  is a random component. It is also assumed that the peptide effects follow a normal distribution with mean zero and variance  $\tau^2$ , where  $\tau^2$  is a random component. Therefore, the posited model accounts for variability of the observed MS/MS measurements at the MS/MS level and at the peptide level, allowing every observed relative peak area ratio to be accounted for by the overall protein ratio, by a peptide specific value and by an MS/MS specific value. The parameters of the random one-

way ANOVA model were estimated by restricted maximum likelihood method [20]. For a protein identified by a single peptide, the estimation of the above model simplifies significantly as, in this case,  $R$  simply corresponds to the average of the corresponding MS/MS ratios.

In terms of biological activity, we are interested in estimating and subsequently testing  $R = 0$  vs. the alternative hypothesis that  $R \neq 0$  (as we operate on a  $\log_2$  scale and  $\log_2(1) = 0$ ), where  $R > 0$  or  $R < 0$  would indicate an increase or decrease in the protein level, respectively. To incorporate biological significance, however, we modified the null hypothesis as follows:  $H_0: -.23 \leq R \leq .20$ , vs. the alternative hypothesis  $H_A: R < -.23$  or  $R > .20$  which requires, in the original scale, an increase or decrease of at least 15% to be deemed significant. Hypothesis testing was carried out basically as described [21]. Proteins reported to increase or decrease during germination met these criteria, unless otherwise indicated.

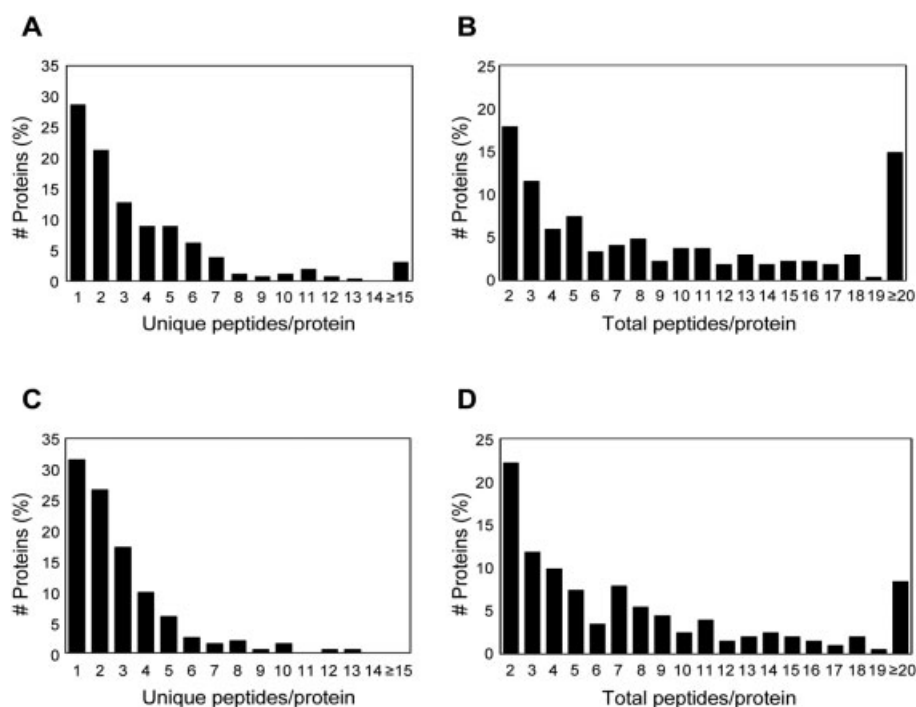
### 3 Results

#### 3.1 Proteomic analysis of germinating cells

In our initial analysis of *B. anthracis* germinating spores, we examined 4486 parent ions and identified 295 unique *B. anthracis* proteins from 3096 identified peptides ( $\geq 90\%$  confidence). When the experiment was repeated using an independent spore preparation, we obtained 2559 parent ions and identified 273 unique *B. anthracis* proteins from 1879 identified peptides ( $\geq 90\%$  confidence). It is important to

note that the parent ions were obtained using MALDI, a technique that results in significantly fewer parent ions than ESI-MS/MS. Of the proteins identified, the evidence for 37 and 70 protein identifications (first and second replicate, respectively) was marginal, based only on a single measurement of a single peptide. As such, these proteins were not subjected to further analysis with the exception of several instances of proteins fewer than 100 amino acids (BAS0043, BAS048S2A, BAS1086, BAS4311, BAS4595 and BAS4644A). The peptide assignment for these proteins was verified by manual inspection of the MS/MS spectra (Supplementary Fig. S1).

The average numbers of unique peptides/protein were 3.7 and 2.8 for proteins with at least 2 peptide measurements (first and second replicate, respectively) (Fig. 2). The distribution of peptides/protein, however, was biased, in that the majority (62 and 58%, first and second replicate, respectively) was identified by  $\leq 3$  unique peptides (Fig. 2). In many cases, multiple measurements of the same peptide were obtained, resulting in an average number of total peptides per protein of 11.7 and 8.9 and median total peptides per protein of 6 and 5 (first and second replicate, respectively) (Fig. 2). The majority of the proteins (86 and 91%, first and second replicate, respectively) had fewer than 20 total peptides identified (Fig. 2). A few abundant proteins, however, such as elongation factor Tu, had a large number of identifying peptides (22 unique and 104 total, and 13 unique and 86 total, first and second replicate, respectively) (Fig. 2). The lower average of unique and total peptides in the second replicate reflects the lower total number of identified peptides (1879 vs. 3096 in the first replicate) in this experiment.



**Figure 2.** Peptide distribution for identified proteins. Spore and germinating spore proteins were subjected to 2-D LC and identified by MALDI MS/MS as described in Section 2. Only proteins with at least two peptide measurements at  $\geq 90\%$  confidence are included. (A, C) The distribution of unique peptides/protein in the first and second replicate, respectively. (B, D) The distribution of total peptides/protein in the first and second replicate, respectively.

Of the 261 and 203 proteins identified with 2 or more independent peptide measurements (first and second replicate, respectively), a common set of 167 proteins, hereafter called the “common spore proteome”, was identified in both studies (for gene assignments and peptides used to identify the proteins, see Supplementary Table S1). Interestingly, although the theoretical 2-D representations of the predicted *B. anthracis* proteome (pI vs. molecular weight of all predicted proteins) indicate a relatively even bipolar distribution (56% of the proteins below and 44% above a pI of 7, <http://www.tigr.org/>), in the common spore proteome we observed a strong bias towards acidic proteins (72% below pI of 7; Supplementary Table S1), indicating either that acidic proteins are present at higher levels than basic proteins in the *B. anthracis* spore and in germinating cells, or that our methodology introduces a bias. Such a bias for acidic proteins has previously been reported for *B. anthracis* spores identified by MudPIT [22].

To provide insight into the spore proteome, we examined the functional categories, as defined by the Clusters of Orthologous Groups classification [23], for the proteins identified. The predicted *B. anthracis* proteome encodes proteins that fall into 22 functional categories (<http://www.ncbi.nlm.nih.gov/COG/>) (Supplementary Fig. S2A, Supplementary Table S1). The common spore proteome contained proteins representing 17 of these categories (Supplementary Fig. S2B). Unsurprisingly, proteins involved in defense mechanisms (V), intracellular trafficking and biosynthesis (U), and transport and catabolism of secondary metabolites (Q) were not detected. Protein categories over-represented ( $\geq 1.5$ -fold) in the spore and germinating proteomes include translation (J; 27%), protein modification and turnover (O; 6.3%), energy production and conversion (C; 7%) and nucleotide transport and metabolism (F; 4.5%). The categories of proteins that were underrepresented ( $\leq 1.5$ -fold) include: replication, recombination and repair (L; 2.5%), amino acid transport and metabolism (E; 5%), coenzyme transport and metabolism (H; 1.2%), signal transduction mechanisms (T; 1.9%), cell membrane and wall biogenesis (M; 1.9%), lipid transport and metabolism (I; 1.2%), function unknown (S; 3.1%), cell motility (N; 0.6%) and general function prediction only (R; 5.7%). These differences had been previously noted in other proteomic studies [22] and were not unexpected as the highly specialized spore is vastly different from vegetative cells [24, 25].

### 3.2 Identification of proteins not annotated in the *B. anthracis* Sterne strain

Direct queries of the translated Sterne strain genome (all six reading frames) revealed several proteins that had not previously been annotated. In all cases, the corresponding gene is present in other *B. anthracis* genomes (Ames and 2012) and/or in *B. cereus*, and typically these genes are also present in at least one other *Bacillus* species (Table 3). We have assigned these genes BAS numbers based on their genomic

location (Supplementary Table S1). Significantly, all of these proteins are small (<10 kDa), and it is not, therefore, unreasonable that several of them were identified by only one peptide measurement (Supplementary Fig. S1) and/or only in one replicate study. The levels of several of these proteins decreased during germination and outgrowth, including BAS0469A, BAS0488A and BAS4644A. We include discussion of these proteins as they may reflect novel early targets of proteolysis during germination.

### 3.3 Relative changes in protein levels during germination and outgrowth

To examine whether global changes in the relative ratios of spore coat proteins occurred during germination and outgrowth, we analyzed the total datasets for both germination experiments separately, at both the peptide and at the protein levels. The distributions of the estimated peptide (Fig. 3A and B) and protein (Fig. 3C and D) ratios (reconverted from the  $\log_2$  scale to the original scale) for the 2- and 7-min time points (115 and 116) relative to  $T_0$  (114) are centered at 1 (range 0.99–1.05) with SD of approximately 0.2 (range 0.16–0.34) (Fig. 3). This distribution indicates that there were few differences in the relative abundance of most peptides and proteins during early germination. It is worth noting that the spread of these distributions at the protein level is significantly smaller than that seen at the peptide level (compare Fig. 3A and B with 3C and D, respectively), due to the data aggregation process. Moreover, there is a high degree of reproducibility between the two biological replicates. The larger variances in the second experiment could be attributed to the smaller number of MS/MS spectra acquired (about 60% of the number obtained in the first experiment). Significantly, our SD are in agreement with the ones reported previously using a different method of data display (Supplementary Fig. S3) [12]. We, however, prefer to visualize the data using absolute ratios as this provides relationships that are more meaningful for evaluation of the biology.

In contrast to the earlier time points, the relative ratio distribution of 17 min/0 min (117/114) is centered at 1.09–1.10 with SD ranging from 0.32 to 0.46, revealing changes in the relative abundance of peptides (Fig. 3A and B) and proteins (Fig. 3C and D) in late germination compared to pre-germination. The shape of the histograms, too, indicates changes in the relative levels of peptides and proteins.

Based on quantitative data obtained from the iTRAQ analysis, proteins in the common spore proteome were classified as proteins whose levels change during germination (39 proteins) and those that do not (115 proteins). In addition, the ratios for 13 proteins were inconsistent between datasets. Of the proteins that showed significant changes at 17 min as compared to 0 min, the levels of 19 decreased and 20 increased (Supplementary Table S1). Because the number of proteins whose levels change significantly is much greater than that predicted by chance [26], these likely represent *bona fide* proteomic changes that occur during germination.

**Table 3.** Uncharacterized proteins that decrease during germination

ORF	Additional comments <sup>a)</sup>	Length (amino acids)	% Glutamine	Predicted operon structure <sup>b)</sup>	Potential regulation <sup>c)</sup>	Homologs <sup>d)</sup>	RNA expression <sup>e)</sup>	Previously identified spore protein <sup>f)</sup>
BAS0469A <sup>g)</sup>	Not annotated in 2012 or Sterne. Downstream of <i>ytaF</i> (BAS0469)	52	1.9%	BAS0469-BAS0469A (F)	$\sigma^G$	Bc, Bt	Wave 4	ORF01172
BAS0488A <sup>g)</sup>	Might be cotranscribed with <i>sspK</i> (BAS0488B; 43 bases downstream) also not annotated in Sterne	50	12%	BAS0488A(R)		Bc, Bs <sup>g)</sup> , Bt <sup>g)</sup>	Wave 5	No
BAS0770	In operon with BAS0769, a glycoxy-lase. Divergently transcribed from <i>cotJA</i> , <i>cotJB</i> and <i>cotJC</i>	259	2.7%	BAS0769-BAS0770 (F)	$\sigma^G$	Bc, Bl, Bh, Bt	Wave 4	ORF00802
BAS1086		65	13.8%	BAS1086 (R)		Bc, Bl, Bt	Wave 5	ORF00393
BAS1169		143	3.5%	BAS1169 (R)	$\sigma^K$	Bc, Bt	ND	No
BAS1302	Related to BAS1957, a permease. YqfX-like (35% identical). Not annotated in 2012	123	.08%	BAS1302 (R)	$\sigma^G$	Bc, Bt	Wave 5	ORF0106 <sup>i)</sup>
BAS1511	Related to YpzA (34% identical)	82	3.6%	BAS1514-BAS1515 (F)		Bc, Bt, Gk	ND	ORF06512
BAS2560	Forespore/cortex related. Related to YlaJ, YhcN and YrbB/CoxA. Related to BAS4834 and BAS4323. Transcribed with a hydrolase (BAS2562)	198	4.5%	BAS2560-BAS2561-BAS2562 (R)	$\sigma^{Ej}$	Bc, Bl, Bs, Bt, Gk	Wave 5	ORF05206
BAS3115		63	14.3%	BAS3115 (F)	$\sigma^{Ej}$	Bc, Bt	Wave 5	ORF04501
BAS3634	Related YmfJ and YfiH (58 and 40% identical, respectively). In large operon	82	7.3%	BAS3634 (R)	$\sigma^G$	Bc, Bh, Bl, Bcl, Bs, Bt, Gk, Oi,	Wave 5	
BAS4323	Forespore/cortex related. Related to YrbB/CoxA, YlaJ and YhcN. Related to BAS2560 and BAS4834	214	1.4%	BAS4323 (R)	$\sigma^G$	Bc, Bt	Wave 5	ORF0385 <sup>i)</sup>
BAS4595		92	10.9%	BAS4595 (F)		Bc, Bl, Bcl, Bs, Bt, Gk, Oi.	Wave 5	ORF02663
BAS4644A <sup>g)</sup>	Transcribed with BAS4645	49	16.3%	BAS4644A-BAS4645 (R)	$\sigma^G$	Bc, Bt <sup>g)</sup>	Wave 5	ORF03001
BAS4645	Transcribed with BAS4644A. Similar to YtzC (51% identical)	87	19.5%	BAS4644A-BAS4645 (R)	$\sigma^G$	Bc, Bl, Bs, Bt, Gk	Wave 5	ORF02605
BAS4870	Spore coat protein F-like. In operon with BAS4871	82	12.2%	BAS4870-BAS4871 (R)	$\sigma^G$	Bc, Bt, Cp, Ct, Oi	Wave 5	ORF02319
BAS4871	In operon with BAS4870	61	14.8%	BAS4870-BAS4871 (R)	$\sigma^G$	Bc, Bt, Cp, Ct, Oi	Wave 5	ORF02318
BAS4993	In operon with <i>spoVAC</i> (BAS4995), <i>spoVAD</i> (BAS4996) and <i>spoVAE</i> (BAS4997)	68	20.6%	BAS4993-BAS4994-BAS4995-BAS4996-BAS4997-BAS4998-BAS4999 (F)	$\sigma^{Ej}$	Bc, Bt, Gk, Oi,	Wave 5	ORF02170

a) ORF names starting with Y are in *B. subtilis*.

b) From <http://vimss.org/operons> [32]. F, forward. R, reverse.

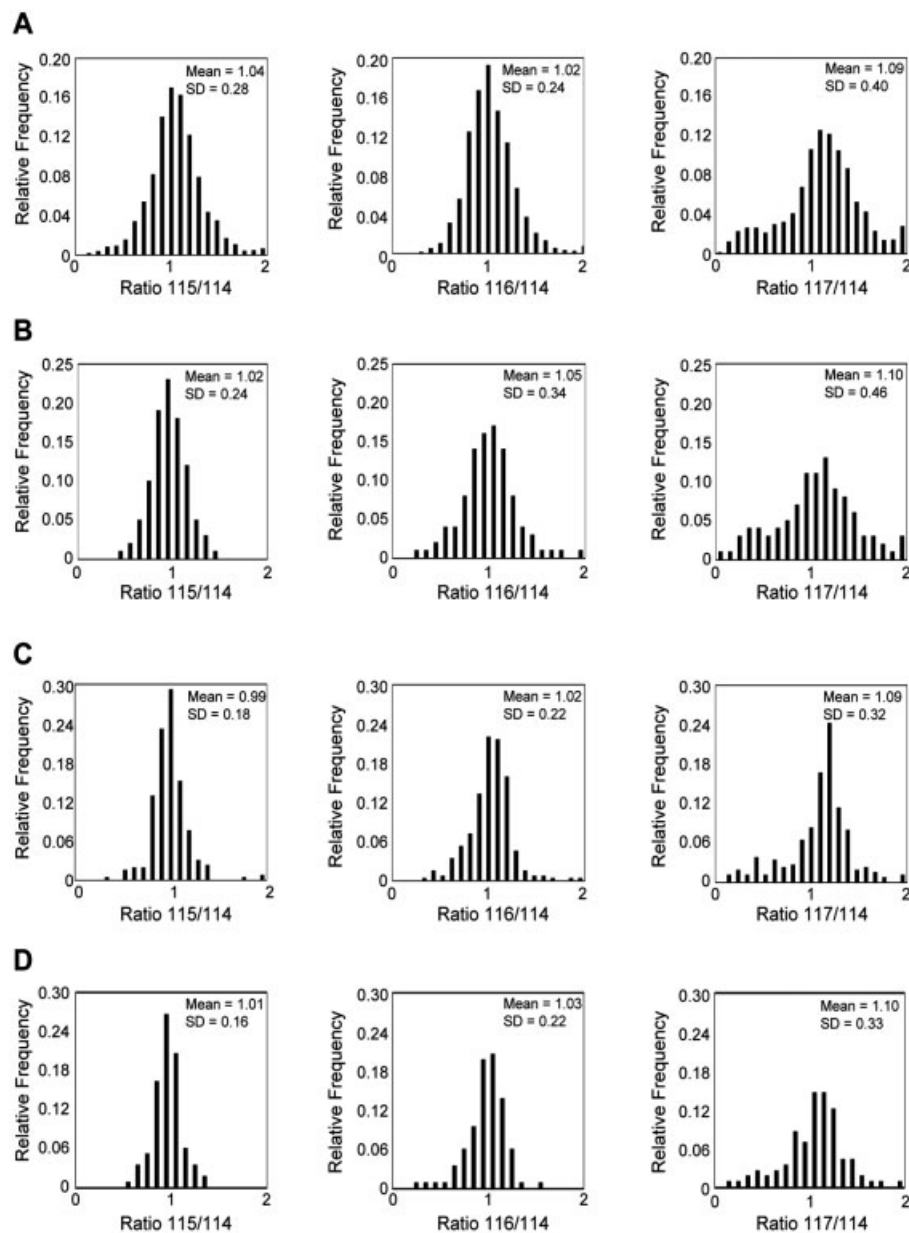
c) Upstream region possesses the following sequences:  $\sigma^E$ , Ata(16–18)cATAcanT;  $\sigma^K$ , AC(16–17)CATAnnnT;  $\sigma^G$ , gnaTr 15 cAtnnTA.

d) Protein with  $\geq 45\%$  identity also found in the indicated bacteria. Bc, *B. cereus*; Bcl, *B. clausii*; Bh, *B. halodurans*; Bl, *B. licheniformis*; Bs, *B. subtilis*; Bt, *B. thuringiensis*; Cp, *Clostridium perfringens*; Ct, *C. tetani*; Gk, *Geobacillus kaustophilus*; Oi, *Oceanobacillus iheyensis*.

e) Expression wave pattern as determined by Liu *et al.* [22].

f) Corresponding ORF number previously found in *B. anthracis* Ames spore proteome [22]. ND- no data. Protein not annotated in *B. anthracis* Ames strain.

g) Not annotated in the genome.



**Figure 3.** Relative distribution of peptides and proteins quantified with isobaric tags. Shown are the distributions of the peptide (A and B; first and second replicate, respectively) and protein (C and D; first and second replicate, respectively) ratios for 2 min/0 min (115/114), 7 min/0 min (116/114) and 17 min/0 min (117/114). For each plot, the mean and SD are given. The data in panel (A) are plotted as described [35] in Supplementary Fig. S3.

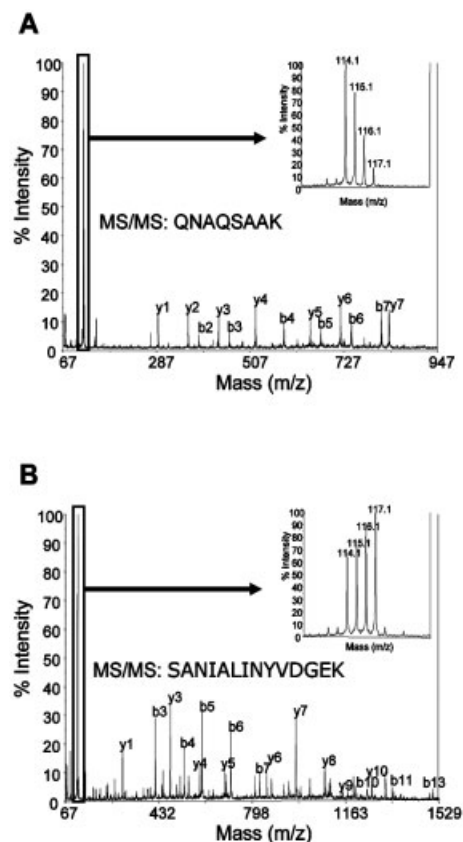
Typical quantitative data profiles for a protein that decreases (BAS0493, an SASP  $\gamma$ -type protein) and one that increases (RplB, a ribosomal protein) are shown in Fig. 4. Average ratios for all proteins and time points are in Supplementary Table S1.

### 3.4 Proteins that decrease during germination

We identified 19 proteins that decreased in abundance during germination in both replicates (Supplementary Table S1), including SASP, known coat proteins and hypothetical proteins. An additional 12 proteins may also decrease, but were only identified by one peptide or in only one

experiment. We attribute the reduction in protein levels to protein degradation resulting from the germination process. Consistent with this idea, the majority (17) of these proteins were identified in a study of the spore proteome [22], whereas only 3 were identified in the vegetative cell proteome [27], indicating that most, if not all, of these proteins are *bona fide* spore proteins. Not surprisingly, the vast majority of these proteins are not assigned a functional category (Supplementary Fig. S2D), indicating that they are either not widely conserved or are uncharacterized. In addition to proteins in the common spore proteome, we also include a discussion of decreased proteins that were found in only one study (Supplementary Table S1).





**Figure 4.** MS/MS spectra and iTRAQ data for selected proteins that increase or decrease during germination. Shown are the MS/MS spectra for one peptide whose level (A) decreases (BAS0493, a  $\gamma$ -type SASP) and (B) increases (BAS0113, ribosomal protein RplB). The b and y fragment ions from the MS/MS spectra are indicated. Inset shows the reporter ion region for each peptide. Peptide sequences are indicated.

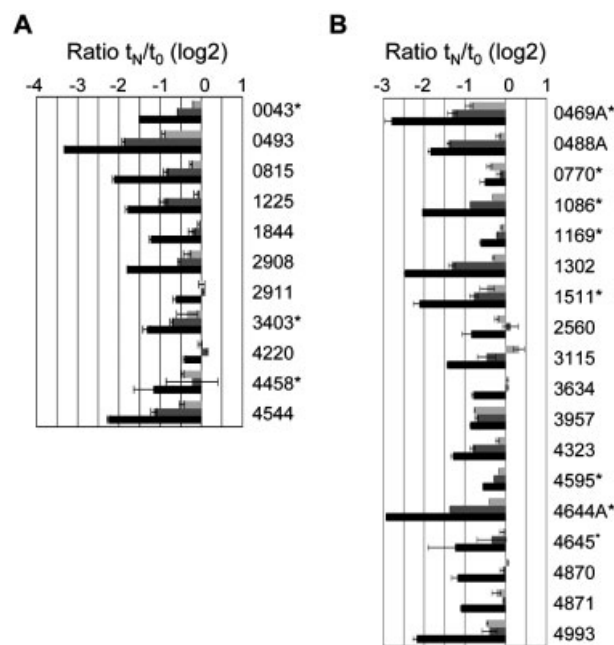
### 3.5 SASPs

SASPs are a predominant class of proteins in the spore core comprising approximately 15% of the total protein content in sporulating bacteria such as *Bacillus*, *Clostridium* and *Thermoactinomyces* [28]. The Sterne genome encodes 12 annotated SASPs and an additional unannotated protein (BAS3059A) that belongs to the  $\alpha/\beta$  family of SASPs. The  $\alpha/\beta$  type of SASPs bind to the spore core double-stranded DNA, causing a change in DNA from the B-form to the A-form [29]. These SASPs protect the DNA from chemical, enzymatic and UV damage [30–32]. *B. anthracis* Sterne also encodes one  $\gamma$ -type SASP (BAS0043). This abundant SASP species does not bind to DNA and might have a role in osmoregulation in the spore core [33] and in providing amino acids for protein metabolism in the initial phase after germination [34]. The  $\alpha/\beta$  and  $\gamma$  SASPs are known to be rapidly degraded by the germination-specific protease, GPR, during the first few minutes of germination [35]. GPR is inactive in the dehydrated spore but rapidly activates upon

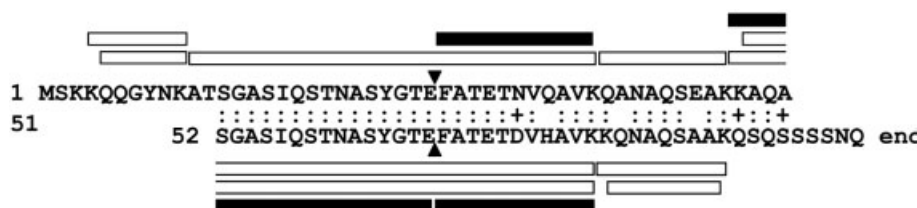
spore rehydration during germination. The roles of the remaining minor SASPs (SspH, SspI, SspO and Tlp) are unknown but, in *B. subtilis*, these proteins are not critical for spore generation or germination [36].

In this study, 7 of the 13 *B. anthracis* SASPs were identified in both replicates and three additional SASPs were found only in the first replicate (Supplementary Fig. S1). In total, we identified 6/7 of the  $\alpha/\beta$  SASPs (BAS0815, BAS1125, BAS1844, BAS2908, BAS2911 and BAS4544), the  $\gamma$ -SASP (BAS0493), Tlp (BAS3403) and two minor SASP proteins (BAS0043 and BAS4458). All 6 of the identified  $\alpha/\beta$ -type SASPs and the  $\gamma$ -type SASP have recognizable consensus sites for GPR cleavage (data not shown) [28, 37]. As expected, all 10 of these SASPs were significantly degraded during germination (Fig. 5A, Supplementary Table S1). Furthermore, during outgrowth, levels of the protease GPR were also decreased, albeit at a modest level (Fig. 5A), consistent with the role of proteolysis in activation of GPR (see, for example, [16]).

We noted that BAS0493, the  $\gamma$ -type SASP, possessed two GPR consensus sequences (Fig. 6). The *B. subtilis*  $\gamma$ -type SASP, SspE, possesses only one GPR cleavage site. Interestingly, it appears that in *B. anthracis* and other closely related Bacilli, the  $\gamma$ -type SASP resulted from a duplication event (amino acids 13–51 and 52–90 are 87% identical) (Fig. 6). To determine whether we could provide experimental evidence



**Figure 5.** Quantitation of proteins that decrease during spore germination. The relative ratios (2 min/0 min, light gray; 7 min/0 min dark gray; 17 min/0 min black) of the indicated proteins are shown. (A) SASP and GPR proteins. (B) Unknown proteins. Data are plotted on a  $\log_2$  scale. Error bars are SEM. Proteins for which there was only one measurement do not have error bars. An asterisk indicates that the protein was only detected in one germination experiment.



**Figure 6.** Identification of peptides and GPR cleavage sites for a  $\gamma$ -type SASP. The sequence of BAS0493 is shown such that the duplicated regions (amino acids 13–51 and 52–90) overlay. Amino acid identities and similarities are indicated by : and +, respectively. The GPR protease cleavage sites are indicated by arrowheads, the positions of the identified tryptic peptides indicated by open boxes, and the proposed GPR protease cleavage peptides by shaded boxes.

for *in vivo* cleavage at the predicted GPR cleavage sites, we queried the MS/MS peak lists for semi-tryptic cleavages (peptides having one tryptic and one non-tryptic cleavage site). We detected peptides consistent with GPR cleavage for two SASPs, BAS0493 and BAS4544 (Fig. 6 and data not shown), direct indication that the GPR protease is involved in turnover of these proteins.

### 3.6 Other decreased proteins

Whereas few coat proteins have been characterized in *B. anthracis*, many have been identified based on sequence similarity to *B. subtilis* coat proteins [14, 38]. In this study, we isolated relatively few of these predicted coat proteins, an expected result, as our extraction protocol was relatively gentle as compared to our previous method (0.1% SDS and mechanical shearing compared to 2% SDS and boiling) [14]. Regardless, we identified several proteins that were clearly spore coat proteins: BAS3957 (Cot $\alpha$ ), BAS3619 (CotE), BAS3402 (YaaH), BAS0340, BAS0341, BAS2008, BAS2377, BAS3619, BAS4870, cortex related proteins BAS2560 and BAS4323, and exosporium protein BAS1898 (Supplementary Table S1) [14]. Not surprisingly, a number of these proteins (BAS2560, BAS3957, BAS4323 and BAS4870) are degraded during germination and outgrowth (Supplementary Table S1, Fig. 5B). Of note, not all spore coat proteins are degraded during germination, as CotE, YaaH and the protein encoded by BAS3619 (in both germination studies) were stable during germination (Supplementary Table S1). This differential stability suggests that not all spore coat proteins are early targets of proteolysis and that germination does not involve indiscriminate degradation of all coat proteins. This observation is consistent with the observed spore fragments present in germination media [10] and, together with the ultrastructural observations of spore cracking [2], suggests that selective proteolysis and zonal distribution of spore coat proteins may be responsible.

The vast majority of the uncharacterized decreased proteins are found exclusively in sporulating bacteria and have been previously shown to be expressed late in sporulation [22] (Table 3), consistent with a spore-specific cellular role. In many cases, however, these unknown proteins are only found in the closely related bacterium, *B. cereus*, and reitera-

tive PSI-BLAST searches, in general, did not reveal sequence features that suggested a function. Particularly striking is that 12/17 of these unknown proteins are less than 100 amino acids in length. Included are 3 proteins whose genes were previously unannotated in the Sterne genome (Table 3). One of these genes, BAS0488A is also unannotated in the *B. subtilis* genome. We also note that several of these proteins contain significant levels of the amino acid glutamine ( $\geq 10\%$ , Table 3).

To provide clues as to the role of the unknown degraded proteins, we examined the promoter regions for consensus motifs based on the *B. subtilis* sigma factor binding sites [39]. Spore formation is driven by a cascade of temporally and spatially controlled phases of transcription that results in the appearance of specific proteins in the forespore (such as the SASP) and others in the mother cell (such as the coat proteins) [25, 40–43]. These phases of expression are controlled by a series of sigma factors that become active sequentially and in specific compartments.  $\sigma^F$  and  $\sigma^G$  are both active in the forespore and are responsible for synthesis of spore-specific proteins including the SASPs, and certain proteins involved in cortex biosynthesis (such as PbpF and PbpG) [44]. Gene expression in the mother cell is directed by  $\sigma^E$  and  $\sigma^K$ . Their action results in the synthesis of proteins involved in cortex production and all the coat proteins. DNA array analysis indicates that, in *B. subtilis*,  $\sigma^E$  directs the expression of 262 genes and that  $\sigma^K$  activates an additional 121 genes [40–42].

Transcriptional profiling of *B. anthracis* cells during sporulation reveals a coordinated temporal expression of spore mRNA [22], largely consistent with that observed in *B. subtilis* [40–43]. Thus, although experimental data for the role of specific mother and forespore sigma factors has not been reported in *B. anthracis*, it is likely that the expression of key core protein genes will be under the control of  $\sigma^F$  and  $\sigma^G$ , whereas coat protein encoding genes will be controlled by  $\sigma^E$  and  $\sigma^K$ . Using the consensus sequences Ata(16–18)cATAcAnT and AC(16–17)CATAnnnT for  $\sigma^E$  and  $\sigma^K$ , respectively, we found perfect  $\sigma^K$  binding sites in the promoter region of BAS1169 and perfect  $\sigma^E$  binding sites in the promoter regions of BAS2560, BAS3115 and BAS4993 (Table 3). Therefore, these genes may encode coat and/or germination proteins expressed from mother cell-specific promoters. BAS4993 is in an operon with several genes (*spoVAC*, *spo*

VAD, *spoVAE*) whose products, in *B. subtilis*, are important for dipicolinic acid movement across spore membranes during sporulation and germination [45]. The majority of the remaining proteins that decreased during germination (13) has consensus sequences for  $\sigma^G$  (Table 3) and, therefore, may be expressed in the forespore. None of the genes had a consensus sequence for a  $\sigma^F$  promoter.

### 3.7 Proteins that increase during germination

Twenty proteins significantly increased, albeit at relatively modest levels, by 17 min of outgrowth (Supplementary Table S1). We envision two possible reasons for their increase: differential extraction of the protein due to changes in the spore and *bona fide* increase in gene expression. It is well documented that germinating spores are transcriptionally inert [8]. At 17 min, however, the spores have germinated, considerable degradation of spore-specific proteins has occurred, and are reorganizing their genetic program to initiate vegetative cell growth (Fig. 1). Not surprisingly then, we observed that the majority of proteins whose levels increase are involved in the cellular reprogramming necessary for cell growth (Supplementary Fig. S2C) and, therefore, increases in protein level may represent new protein synthesis. Proteins involved in translation and carbohydrate transport and metabolism (categories J and G) account for over 50% of the proteins that increase late in germination. Notably, 15 proteins that increased are found in the vegetative cell proteome [27], consistent with a transition from spore to vegetative cell. For the most part, the increased proteins correlate well with the consequences of an increase in cell metabolism.

## 4 Discussion

Bacterial spores are dormant, highly resistant and, as far as can be discerned, metabolically inert. Nonetheless, in response to the presence of nutrient, they are able to germinate in minutes and resume cellular activity. This process initiates without any input of energy or detectable metabolism, and very quickly leads to cell growth and division. The early events in this process are largely undefined. In particular, with the exception of the SASPs, we know very little about changes in spore composition during germination. A better understanding of these changes would provide significant insight into the ways in which spores control the deconstruction of their protective layers and the methods by which they reinitiate metabolism in the most efficient manner. In this study, we have begun to address these questions using quantitative proteomics to describe the changes in protein levels during the first few minutes of *B. anthracis* spore germination.

Two major conclusions can be drawn from our work. First, we have shown that within the first 17 min, which encompasses the full period of loss of spore refractility and,

therefore, most of the rehydration process, a specific subset of proteins are significantly reduced in abundance. The degradation is rapid and is specific to some, but not all coat proteins, the SASPs and a set of heretofore-uncharacterized proteins. We argue, therefore, that the initial degradation events target specific spore proteins. Secondly, we found that 20 proteins modestly increase in abundance. The majority of these proteins are involved in key metabolic processes, as would be expected for a cell initiating active growth. Although we cannot definitively rule out that differential solubilization of these proteins occurred, we favor a model whereby metabolism has resumed in at least a subpopulation of the spores in our experiment. It is reasonable to assume that there is significant evolutionary pressure for a very rapid resumption of metabolism as soon as nutrient enters the environment. Given that spores will usually be present in an environment with many other organisms (e.g. in the soil), the reintroduction of a food source will spark a period of intense competition. Cells that most efficiently reestablish rapid growth will be most successful in repopulating their niches. Thus, there is a significant advantage to rapid and efficient dismantling of the spore outer layers and a mechanism of control of protein synthesis that allows the resumption of vegetative growth to occur as rapidly as possible.

Of high priority is the determination of the cellular function of the hypothetical proteins that are degraded during germination. Of the 19 proteins that decrease in abundance identified in both replicates and the 12 found in a single study, only 1 has been assigned to a COG family (Supplementary Fig. S2D) indicating that, in general, these proteins are not distributed broadly and/or that they are uncharacterized. Not surprisingly, 10 of the decreased proteins are SASPs whose rapid degradation during germination is expected, as they play a central role in protection of spore DNA and presumably interfere with transcription [28, 32, 36, 46, 47]. Several of the decreased proteins are likely coat proteins (BAS2560, BAS3957, BAS4323 and BAS4870). Several others (BAS1169, BAS2560, BAS3115 and BAS4993) may be controlled by  $\sigma^K$  and  $\sigma^E$  and, therefore, could be additional coat proteins.

Intriguingly, more than half of the unknown decreased proteins identified in this study were quite small (less than 100 amino acids; Table 3). Moreover, 3/12 of these small proteins were previously unannotated in the *B. anthracis* Sterne genome (Table 3), presumably due to their small size. These proteins would have certainly been missed using a standard 2-D gel-based system and, therefore, provide further rationale for using multiple approaches to study the spore proteome. Many of these small proteins (9) also possess biased amino acid compositions, containing more than 10% glutamine residues (Table 3). In *B. subtilis*, a number of the SASPs, including the  $\gamma$ -type SASP, SspE, and several of the minor SASPs (SspH, SspO and SspI) also possess many glutamine residues (>10%), raising the possibility that some of these proteins may be heretofore unknown SASPs. Inter-

estingly, based on its predicted operon structure BAS0488A is likely co-expressed with *sspK*. Another possibility, however, is that some of these proteins are spore coat proteins. This is likely the case for BAS4870, a CotF-like protein, and may be true for proteins whose expression may be under the control of mother cell sigma factors (Table 3).

Although most proteins showed the same trends in both independent replicates, 13 proteins displayed appreciable degradation in one dataset but not the other (Supplementary Table S1). For a new quantitative technology, a false positive rate of 8% (13/167) would be tolerable. It may be, however, that the difference in the datasets for these specific proteins reflects experimental differences and not variance in the iTRAQ technology. Whereas most *bona fide* decreased proteins had greater reductions in their iTRAQ ratios with longer germination times (Fig. 6), nine of these anomalous proteins showed similar low iTRAQ ratios, regardless of the time examined. Furthermore, we note that 5 of these anomalous proteins are spore coat and/or surface structure proteins, including BAS0340, BAS0341, BAS1898, BAS2008 and BAS2377). One possibility, therefore, is that differential extraction, perhaps due to protein cross-linking, was responsible for this anomalous behavior. We hypothesize that these and other proteins with conflicting iTRAQ ratios may also have solubilized differently in the independently germinated and processed replicate experiments.

In this study, we obtained direct *in vivo* evidence for specific degradation events through the identification of peptides from cleavage at consensus sites for the GPR protease (Fig. 6). The effects of degradation by the endogenous GPR protease might be expected to be less obvious in the early stages of germination because limited proteolysis will introduce only a few non-tryptic cleavages. The observation of reductions in protein levels at the earliest time point, however, indicates that the germinating spore is fully competent for protein degradation early in the germination process.

One characteristic of *Bacillus* spore germination is that the early stages are pre-transcriptional and apparently pre-translational, limiting direct study to the level of proteins. This characteristic is shared with other systems undergoing extreme dormancy where transcriptional activity is minimal or non-existent. An analysis of early protein changes, however, would be difficult to achieve using classical gel-based proteomics methods as many of the spore proteins are quite small and/or are difficult to solubilize. Previous studies of spore proteins using 2-DE, for instance, identified only 3 degraded proteins that fell within the observable pI and size range of 2-D gels [48]. Trypsinization of the spore coat proteins, however, results in soluble peptides that are amenable to separation by chromatography as previously shown in a non-quantitative study of *B. anthracis* spore proteins [22].

This study took advantage of the recently developed isobaric tag technology [12] to examine the time dependency of protein levels in the germinating spore. The data obtained from time courses provide more detailed information regarding changes in protein levels and a higher degree of

certainty than binary experiments, as proteins can be expected to exhibit consistent trajectories within the time frame of this study. A temporal examination of spore germination was feasible because this isotope enrichment method allows four-plex experiments and quantification to be performed in the second stage of a tandem mass spectrometer with the attendant advantages of that approach. The iTRAQ technology employed in this study used a standard MudPIT approach [49], with the collection of capillary HPLC fractions directly on a MALDI sample plate. The relatively high mass accuracy and resolution of the TOFTOF mass spectrometer contributes to the accuracy of the protein identifications. Because the capillary HPLC fractions were collected on a MALDI sample plate with inline addition of MALDI matrix, the MALDI TOFTOF could be used for these studies, with the inherent benefits of being able to re-query the samples spotted on the MALDI plate. Moreover, the higher collision energy of this instrument is well suited for reproducible analysis of iTRAQ peptides.

It should be noted that while the iTRAQ data presented in this manuscript exhibit well-defined precision, the authors have noted in preliminary studies (Maddock and Andrews, unpublished) that iTRAQ data collected on the model 4700 appears to be slightly dampened. Initial observations attribute this effect to the large timed ion selection window (ca. 8 Da) on this instrument, that allows significant chemical noise in the form of low level iTRAQ-tagged peptides to be included with the parent ion (unpublished results). The magnitude of this dampening effect may be as large as twofold for some ions based on preliminary studies with narrower ion selection windows on other models. This does not significantly affect the precision of the measurements, but does affect the accuracy with which the change can be measured. The changes reported in this study should therefore be considered to be minimal values.

We also describe a robust statistical method for determining the relative changes that occur. Our method takes into account the complex nature of the data for each protein, which includes both measurements of unique peptides for each protein and multiple measurements of each unique peptide. This redundancy of measurements can greatly increase the certainty with which small but significant changes can be identified. We anticipate that the precision of this method and the ability to easily acquire up to four time points will allow more detailed studies of protein expression levels and more effective correlation with mRNA expression levels than other proteomic approaches.

The degree of detail obtained from this quantitative time-course study has significantly increased our understanding of the specific events involved in *B. anthracis* spore germination. Elucidation of the selective degradation of spore proteins, verification of two substrates for GPR protease, and implication of newly identified ORF in spore germination were enhanced by the ability to perform time-course proteomics with good precision. We anticipate that this approach could be extended to provide more detailed information

about the relative kinetics of the identified substrates for the spore GPR protease and the early translation process during spore germination.

*This work was supported by grants AI053360 (J.R.M.), GM53989 (A.D) and AI53365 (A.D) and NCRR P41–18627 (P.C.A) from the National Institutes of Health. The Michigan Proteome Consortium is acknowledged for infrastructure support. We are extremely grateful to Sue Sullivan for critical editing of this manuscript.*

## 5 References

- [1] Vary, P. S., *Microbiology* 1994, **140**, 1001–1013.
- [2] Moir, A., Corfe, B. M., Behravan, J., *Cell. Mol. Life Sci.* 2002, **59**, 403–409.
- [3] Henriques, A. O., Costa, T. V., Martins, L. O., Zilhao, R., in: Ricca, R. E., Henriques, A. O., Cutting, S. M., (Eds.), *Bacterial Spore Formers: Probiotics and Emerging Applications*, Horizon Biosciences, Norfolk 2004, pp. 65–86.
- [4] Driks, A., in: Sonenshein, A. L., Hoch, J. A., Losick, R., (Eds.), *Bacillus Subtilis and its Closest Relatives*, American Society for Microbiology, Washington 2002, pp. 527–536.
- [5] Aronson, A. I., Fitz-James, P., *Bacteriol. Rev.* 1976, **40**, 360–402.
- [6] Holt, S. C., Leadbetter, E. R., *Bacteriol. Rev.* 1969, **33**, 346–378.
- [7] Setlow, P., *Curr. Opin. Microbiol.* 2003, **6**, 550–556.
- [8] Paidhungat, M., Setlow, P., in: Sonenshein, A. L., Hoch, J. A., Losick, R., (Eds.), *Bacillus Subtilis and its Closest Relatives*, American Society for Microbiology, Washington 2002, pp. 527–536.
- [9] Makino, S., Moriyama, R., *Med. Sci. Monit.* 2002, **8**, RA119–127.
- [10] Santo, L. Y., Doi, R. H., *J. Bacteriol.* 1974, **120**, 475–481.
- [11] Chada, V. G., Sanstad, E. A., Wang, R., Driks, A., *J. Bacteriol.* 2003, **185**, 6255–6261.
- [12] Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B. et al., *Mol. Cell. Proteomics* 2004, **3**, 1154–1169.
- [13] Cutting, S. M., Vander Horn, P. B., *Molecular Biological Methods for Bacillus*, John Wiley & Sons Ltd., Chichester 1990, pp. 27–74.
- [14] Lai, E.-M., Phadke, N. D., Kachman, M. T., Giorno, R. et al., Driks, A., *J. Bacteriol.* 2003, **185**, 1443–1453.
- [15] Cortezzo, D. E., Setlow, B., Setlow, P., *J. Appl. Microbiol.* 2004, **96**, 725–741.
- [16] Nessi, C., Jedrzejewski, M. J., Setlow, P., *J. Bacteriol.* 1998, **180**, 5077–5084.
- [17] Venkatasubramanian, P., Johnstone, K., *J. Gen. Microbiol.* 1989, **135**, 2723–2733.
- [18] Michailidis, G., Shedden, K., *J. Comput. Biol.* 2003, **10**, 689–698.
- [19] Aggarwal, K., Choe, L. H., Lee, K. H., *Proteomics* 2005, **5**, 2297–2308.
- [20] Snijders, T., Bosker, R., *Sage Publications Ltd.*, London 1999.
- [21] Hodges, J. L., Lehmann, E. L., *J. Royal Stat. Soc. Series B* 1954, **16**, 261–268.
- [22] Liu, H., Bergman, N. H., Thomason, B., Shallom, S. et al., *J. Bacteriol.* 2004, **186**, 164–178.
- [23] Tatusov, R. L., Galperin, M. Y., Natale, D. A., Koonin, E. V., *Nucleic Acids Res.* 2000, **28**, 33–36.
- [24] Driks, A., *Microbiol. Mol. Biol. Rev.* 1999, **63**, 1–20.
- [25] Piggot, P. J., Hilbert, D. W., *Curr. Opin. Microbiol.* 2004, **7**, 579–586.
- [26] Benjamini, Y., Hochberg, Y., *J. Royal Stat. Soc. Series B, Methodological* 1995, **57**, 289–300.
- [27] Francis, A. W., Ruggiero, C. E., Koppisch, A. T., Dong, J. et al., *Biochim. Biophys. Acta* 2005, **1748**, 191–200.
- [28] Setlow, P., *Annu. Rev. Microbiol.* 1988, **42**, 319–338.
- [29] Mohr, S. C., Sokolov, N. V., C. M., H., Setlow, P., *Proc. Natl. Acad. Sci. USA* 1991, **88**, 77–81.
- [30] Sohail, A., Hayes, C. S., Divvela, P., Setlow, P., Bhagwat, A. S., *Biochem.* 2002, **41**, 11325–11330.
- [31] del Carmen Huesca Espita, L., Caley, C., Bagyan, I., Setlow, P., *Mutat. Res.* 2002, **503**, 77–84.
- [32] Setlow, P., *Ann. Rev. Microbiol.* 1995, **49**, 29–54.
- [33] Cucchi, A., Sanchez de Rivas, C., *Curr. Microbiol.* 1995, **31**, 228–233.
- [34] Hackett, R. H., Setlow, P., *J. Bacteriol.* 1988, **170**, 1403–1404.
- [35] Loshon, C. A., Swerdlow, B. M., Setlow, P., *J. Biol. Chem.* 1982, **257**, 10838–10845.
- [36] Cabrera-Hernandez, A., Sanchez, S., J. L., Paidhungat, M., Setlow, P., *Gene* 1999, **232**, 1–10.
- [37] Carrillo-Martinez, Y., Setlow, P., *J. Bacteriol.* 1994, **176**, 5357–5363.
- [38] Driks, A., *Trends Microbiol.* 2002, **10**, 251–154.
- [39] Helmann, J. D., Morgan, C. P. J., in: Sonenshein, A. L., Hoch, J. A., Losick, R., (Eds.), *Bacillus Subtilis and its Closest Relatives*, American Society for Microbiology, Washington 2002, pp. 289–312.
- [40] Eichenberger, P., Fujita, M., Jensen, S. T., Conlon, E. M. et al., *PLoS Biol.* 2004, **2**, e328.
- [41] Eichenberger, P., Jensen, S. T., Conlon, E. M., VanOoij, C. et al., *J. Mol. Biol.* 2003, **327**, 945–972.
- [42] Steil, L., Serrano, M., Henriques, A. O., Volker, U., *Microbiology* 2005, **151**, 399–420.
- [43] Feucht, A., Evans, L., Errington, J., *Microbiology* 2003, **149**, 3023–3034.
- [44] McPherson, D. C., Driks, A., Popham, D. L., *J. Bacteriol.* 2001, **183**, 6046–6053.
- [45] Vepachedu, V. R., Setlow, P., *Microbiol. Lett.* 2004, **239**, 71–77.
- [46] Cabrera-Hernandez, A., Setlow, P., *Gene* 2000, **248**, 169–181.
- [47] Bagyan, I., Setlow, B., Setlow, P., *J. Bacteriol.* 1998, **180**, 6704–6712.
- [48] Huang, C. M., Foster, K. W., DeSilva, T. S., Van Kampen, K. R. et al., *Proteomics* 2004, **4**, 2653–2661.
- [49] Washburn, M. P., Wolters, D., Yates, J. R. Jr., *Nat. Biotechnol.* 2001, **19**, 242–247.