Available online at www.sciencedirect.com
ScienceDirect
Journal of Chromatography B, 843 (2006) 25–33
JOURNAL OF CHROMATOGRAPHY B
www.elsevier.com/locate/chromb

Small protein biomarkers of culture in *Bacillus* spores detected using capillary liquid chromatography coupled with matrix assisted laser desorption/ionization mass spectrometry

David Wunschel [a,*], Jon Wahl [a], Alan Willse [b], Nancy Valentine [a], Karen Wahl [a]

[a] *Chemical and Biological Sciences, Pacific Northwest National Laboratory, Richland, WA 99354, United States*
[b] *Statistical Sciences, Pacific Northwest National Laboratory, Richland, WA 99354, United States*

Received 20 December 2005; accepted 16 May 2006
Available online 23 June 2006

Abstract

Capillary liquid chromatography (cLC) coupled with matrix-assisted laser desorption/ionization (MALDI) time-of-flight mass spectrometry (TOF-MS) was used to compare small proteins and peptides extracted from *Bacillus subtilis* spores grown on four different media. A single, efficient protein separation, compatible with MALDI–MS analysis, was employed to reduce competitive ionization between proteins, and thus interrogate more proteins than possible using direct MALDI–MS. The MALDI–MS data files for each fraction are assembled as two-dimensional data sets of retention time and mass information. This method of visualizing small protein data required careful attention to background correction as well as mass and retention time variability. The resulting data sets were used to create comparative displays of differences in protein profiles between different spore preparations. Protein differences were found between two different solid media in both phase bright and phase dark spore phenotype. The protein differences between two different liquid media were also examined. As an extension of this method, we have demonstrated that candidate protein biomarkers can be trypsin digested to provide identifying peptide fragment information following the cLC–MALDI experiment. We have demonstrated this method on two markers and utilized acid breakdown information to identify one additional marker for this organism. The resulting method can be used to identify discriminating proteins as potential biomarkers of growth media, which might ultimately be used for source attribution.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Capillary liquid chromatography; *Bacillus*; Spores; MALDI mass spectrometry; Differential display; Small proteins

1. Introduction

Numerous mass spectrometry approaches have been developed for detection and identification of virulent bacteria, with recent focus on *Bacillus anthracis* and near neighbors. In particular, matrix-assisted laser desorption/ionization mass spectrometry (MALDI–MS) has been widely used for the analysis of bacterial proteins and peptides because it requires small sample volumes and is relatively tolerant of suppression from buffer components [1–5]. Many discriminating masses under 20 kDa have been reported, though the actual masses reported have varied between studies [6–8]. Many peaks characteristic of strains in vegetative state appear to correspond to constitutively expressed ribosomal proteins [9–11].

Less attention has been paid to MALDI–MS detection of differences in bacterial low molecular weight protein content resulting from growth conditions. Protein expression in general is known to vary with changes in nutrients [12,13]. Therefore, the effects of variable culture conditions are important to study if MALDI–MS is to be useful in detection, identification and forensic applications. A few reports have shown variability in direct MALDI–MS with growth stage, growth rate, and media type [14–16]. Similar media effects have been observed in low mass ion signatures (<6000 *m*/*z*) detected by the intact cell mass spectrometry method of Walker et al. [3]. It should be noted that these studies have focused on vegetative cells; similar studies have not been reported for bacterial spores.

The spores of *Bacillus* species have been well studied concerning their morphology, metal content, and gene regulation [17–20]. Spores have recently been described to be far more dynamic than previously thought with growth parameters such

* Corresponding author.

1570-0232/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2006.05.021

as temperature and humidity playing important roles in effecting spore characteristics [17,21,22]. *Bacillus* spores have high protein content compared to vegetative cells [20], with up to 15% of the total protein content reported to be small acid-soluble proteins (SASPs) [23]. These proteins play a critical role in protecting the spore DNA and are therefore highly conserved [19]. Several reports have detailed the small protein content of *Bacillus* spores detected by MALDI–MS for use as species-specific biomarkers [4,5,23–25]. However this has not been studied in relation to different growth conditions.

Here, we apply a capillary high performance liquid chromatography (cLC)–MALDI–MS method to visualize and identify small proteins and peptides in *Bacillus subtilis* spores that vary with culture conditions. Previous MALDI–MS experiments with vegetative *Bacillus* found a subset of fingerprint peaks to be invariant to culture conditions, whereas other peaks were discriminatory only for certain culture conditions [15]. It is important to develop methods to visualize and compare larger numbers of protein markers. Beyond validating protein markers for detection of a bacterial strain, it would be useful to compare profiles from known media types and unknown samples for clues to how the unknown was prepared for forensic purposes. Microbial forensic attribution presents a technical challenge that has received considerable attention recently [26–28]. In particular, the intentional use of *B. anthracis* spores as a weapon of terror in 2001 brought additional focus on this organism. Protein patterns indicative of specific growth conditions may prove useful in determining how a culture was grown.

An advantage of this technique is that markers detected in an initial analysis can be resolubilized for further manipulation because it is estimated that less than 70% of the protein in the sample is depleted by even an exhaustive MALDI experiment [29]. In particular, the protein can be subjected to proteolytic digestion with enzymes such as trypsin. Direct trypsin digestion on bacterial samples on the MALDI–MS probe has been demonstrated [30]. However, to our knowledge, this approach has not been applied to digestion of fractions after an initial MALDI–MS analysis. The resulting tryptic peptide masses can be used to search a database and support intact mass information for protein identification.

Spores of *B. subtilis* and similar organisms have long been used as a simulant for *B. anthracis*. Specifically, a group of pigmented strains originally called *B. globigii* later called *B. subtilis* var. *niger* and finally classified as *B. atrophaeus* has often been the simulant of choice [31,32]. Spores of a pigmented subtilis strain, *B. subtilis* 49760 were used in this study because they are readily grown on a variety of culture media types. While this proved advantageous, we also note the previously described diversity within this species and anticipate protein sequence and therefore mass variation compared to the type strain [31,33].

In this study, small spore protein content was compared between cultures grown on two solid and two different liquid media. Differential display of these small protein profiles was used to identify masses that differed between media types, suggestive of small protein biomarkers for growth culture. However, this paper is not intended to be a definitive description of media specific protein markers for this organism. Instead, this report describes an effective combination of methods for *visualizing and comparing* protein biomarkers using cLC and MALDI–MS.

2. Experimental

2.1. Cell cultivation and preparation

B. subtilis 49760 (ATCC, Manassas, VA, USA) was cultured in Tryptic Soy Broth (TSB) as follows. A small quantity of freezer stock was added to 3.0 mL of sterile TSB and incubated at 30 °C for ∼14 h, 150 rpm. Following preculturing, 150 μL of the vegetative cells were spread plated onto either solid media or liquid media. Two solid media types were used: Lab Lemco (LL – 23 g/L from Oxoid Scientific, Hampshire England, UK) and the related Nutrient Sporulating Media (NSM in 1 L – Bacto-agar 2 g, tryptone 3 g, Lab-Lemco Agar 23 g, yeast extract 3 g, 1 mL 1% $MnCl_2{\cdot}4\,H_2O$/L). The two liquid media types have been termed "modified schaeffer media" or MSM (rich media in 1 L – 8 g nutrient broth, 0.1% KCl, 0.012% MgCl, 1 mM $Ca(NO_3)_2$, 0.01 mM $MnCl_2$, 0.001 mM $FeSO_4$) following Schaeffer et al. [18] and "Slepecky" (minimal media in 1 L – sucrose 1 g, NaCl 1 g, KH_2PO_4 5 g, $CaCl_2$ 5 mg $(NH_4)_2HPO_4$ 1 g, $ZnSO_4$ 10 mg, $MgSO_4$ 0.2 g, $FeSO_4$ 10 mg, $MnSO_4$–H_2O 7 mg) following Slepecky and Foster [20].

The solid media samples were incubated upside down for 3–5 days in a 30 °C incubator. The cultures are checked by phase contrast microscopy for sporulation and culture purity and harvested when >95% spores are present. The spores are washed from the plates with 10 mL of sterile Milli-Q water (18 MΩ), centrifuged, decanted and repeated (four to five times) to remove the vegetative cell debris. There should be >95% spores present. The spores are stored in water at 4 °C. The spores were washed with sterile Milli-Q water every 4–5 weeks.

The concentration of spores in each preparation was determined and was typically 2 to 3×10^8 cfu/mL. The sample size used for protein extraction was 450 μL, therefore the starting concentration of spores was about 1×10^9 cfu.

The spores were characterized using phase contrast microscopy where freshly prepared spores appear "phase bright" relative to debris or vegetative cells. Cultures were only used for their initial extraction of proteins if they were at least 95% phase bright spores relative to other visible structures over the course of 10 fields of view. However, after extended storage of the spores, the spore appearance will change to phase dark. In this case, the transition occurred naturally after storage at 4 °C for a period of 6 months. A visual comparison of spore appearance is given in Fig. 1 where the phase bright (Fig. 1A) are clearly distinguished from the phase dark spores (Fig. 1B). Despite the dark appearance, the phase dark spores were still distinct from the vegetative forms in size and staining characteristics. Fig. 1C shows the typical polar staining pattern of the phase dark spores subjected to gram stain, contrasted with the uniformly gram staining vegetative cells shown in Fig. 1D.

2.2. Protein sample preparation

Protein extracts of spores were obtained by placing 450 μL of spore suspension into a glass vial with 50 μL triflouroacetic acid

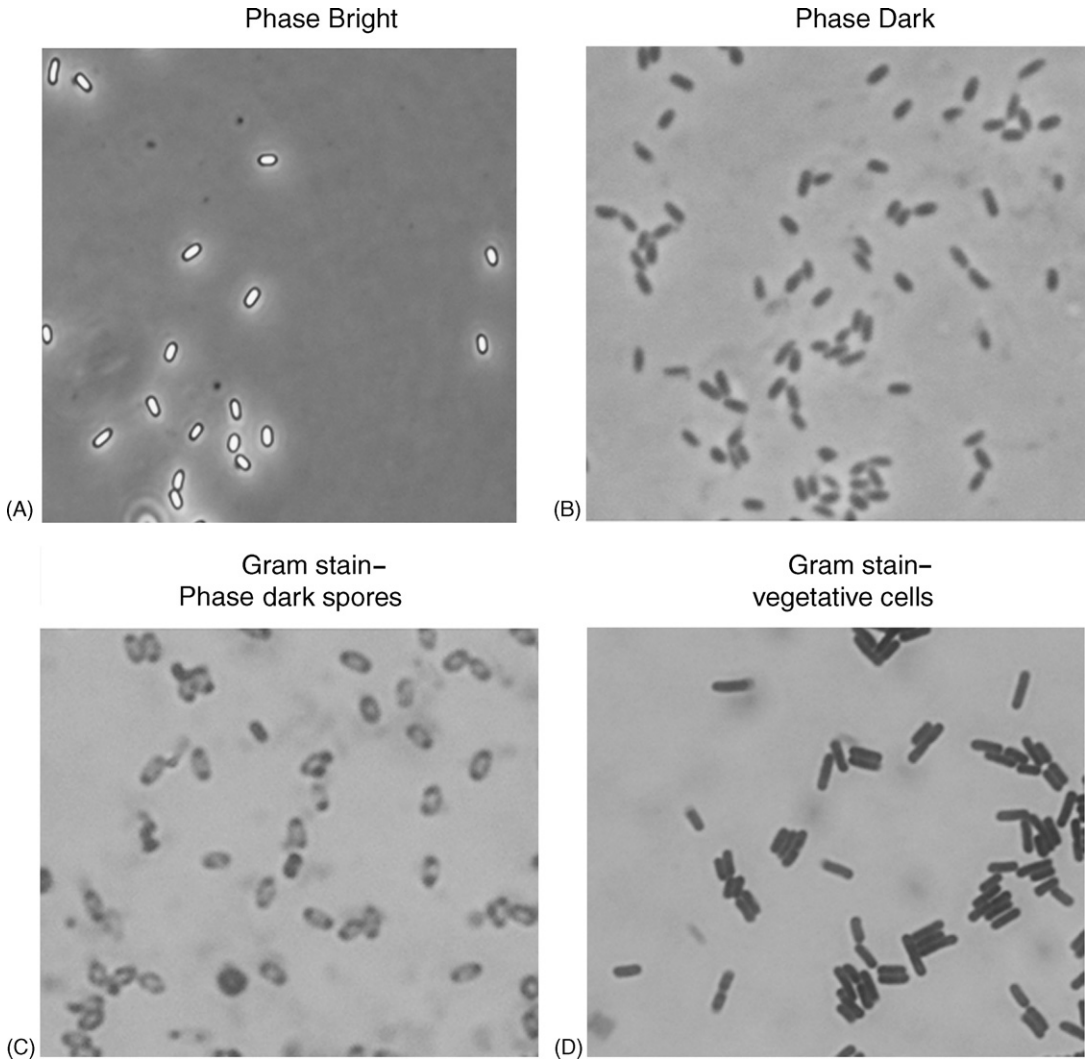

Fig. 1. The spore preparation of *B. subtilis* 49760 observed under phase contrast microscopy. Phase bright (A) or phase dark (B) spores are contrasted with the Gram stain of the phase dark spores (C). These spores demonstrated only poor polar staining, distinct in appearance from the vegetative cells (D).

(5% final concentration), 500 μL acetonitrile (50% final concentration), similar to the approach described in Hathout et al. [23]. Samples were sonicated in 2 mL glass auto sampler vials for 60 min, and then spores removed with a 2 min spin filtration on 0.2 mm filter (Millipore Corp, Billerica, MA, USA). Filtrate was dried and resuspended in 50 mL of 10% acetonitrile/0.1% TFA.

2.3. Reverse phase HPLC—fraction collection

Ten microliters injection used for each sample were made onto a 45 cm long fused silica capillary column with dimensions of 250 μm ID, 360 μm OD (Polymicro technologies, Boise ID) packed with 5 μm Jupiter C-18 (2) material. The capillary columns were packed in-house using ISCO high pressure pumps at 6000 psi. The column flow rate was 4–5 μL/min measured manually and controlled using a 1:30 flow split from gradient pumps (Jasco Inc., Tokyo, Japan). The two solvents used for reverse phase chromatography were A: 10% acetonitrile, 0.1% TFA; B: 90% acetonitrile, 0.1% TFA. A linear gradient of 10–95% B was used to elute proteins over 30 min. Ferulic acid matrix (10 mg/mL in 30% acetonitrile) was co-added to the fractions at a flow rate equal to the column flow rate. One hundred nanograms each of bovine insulin and cytochrome *c* were used as internal retention time standards. Fractions were collected using a "Probot" (LC Packings, San Francisco, CA, USA) microfraction collector to spot the gold coated well plates. Each fraction was collected using an 18 s residence time per fraction resulting in a 3–3.5 μL fraction volume.

Protein fractions selected for further analysis and identification were retrieved from the sample plate using two successive 1 μL additions of 50% acetonitrile with aspiration to resolubilize. These two fractions were combined in a 0.5 mL eppendorf tube. Further added to the reaction mixture was 1 μL of 100 mM ammonium bicarbonate, 1 μL of agarose-bound trypsin and 1 μL of water. The TPCK-trypsin agarose beads (Pierce Inc, Rockford, IL, USA) were prepared by washing 25 μL of trypsin slurry with 475 μL of 50 mM ammonium bicarbonate two times and resuspended in 50 μL of the ammonium bicarbonate solution. The reaction mixtures were incubated 2 h at 37 °C. One microliter of the digestion reaction was respotted.

Each plate was visually inspected to ensure full coverage of the spot with matrix crystals. The HPLC fractions in the early part of the gradient (time 5–15 min) were often thinly crystallized due to the low acetonitrile content and required a second addition 0.5 μL of matrix.

2.4. MALDI–TOF MS

The mass spectrometry data collection for the separated fractions utilized an Applied Biosystems DE-RP (Framingham, MA, USA) equipped with a nitrogen laser (337 nm), reflector and delayed extraction features. Data was collected in linear mode using an accelerating voltage of 25 kV with a source grid voltage at 95% (23.75 kV) and a delay time of 90 ns. Mass spectral data was collected over a 1000 to 30,000 *m/z* range. In automated mode, 256 laser shots were collected over 12 standardized locations using a preprogrammed spiral raster pattern. Data was collected for each HPLC fraction in an automated routine over 100 sample wells. The protein mixture of angiotensin I, insulin and bovine cytochrome *c* was used to externally mass calibrate at wells 12, 19, 82 and 89 of the 100 well plates to achieve an expected mass accuracy of at least 0.1% over the entire plate using these four spots.

The peptides generated from trypsin digested fractions were analyzed in linear mode as well for increased sensitivity. A neighboring position on the sample plate was used to spot a standard peptide mixture of Insulin chain b (mw 3695.5 Da) and leucine–Enkephalin (mw 555.6 Da). External mass calibration was performed for each spot with a measured average mass accuracy of 0.05% for the digest peptides. Averages of 128 laser shots were collected for the peptide mixtures.

2.5. Data analysis

Raw mass spectral data files were imported into Matlab, and each cLC–MALDI run stored in a 100 × 14,470 data matrix, with rows corresponding to MALDI mass spectra (for masses between 2 and 20 kDa) for each LC fraction. Four of the rows (corresponding to wells 12, 19, 82, and 89) were used only for mass calibration, and were not used in subsequent comparisons of growth conditions. For each cLC–MALDI run, LC fractions were pre-processed separately by baseline correcting and smoothing their individual MALDI–MS, and then detecting peaks, which putatively correspond to proteins.

Following pre-processing, two different types of comparative analyses were performed. The first is based on total mass spectral intensity values summed over all fractions. This approach is robust to retention time mismatches in the LC fractions, and it is still likely that more proteins will be compared than using direct MALDI–MS.

The second approach used the full two-dimensional information (retention time and mass). A challenge of this approach is that proteins might elute at slightly different retention times (fractions) from run to run, so there will not be a one-to-one correspondence between fractions. Because only 100 fractions were obtained, nonlinear retention time alignment algorithms commonly used in gas chromatography are not applicable. The internal standard provides a single landmark (eluting at about fraction 57) that was used to coarsely align fractions between runs. This shift correction will not usually perfectly align fractions, so to make comparisons between samples we first applied single linkage cluster analysis to two-dimensional peak locations across all samples to identify well-separated regions of interest, within which both qualitative and quantitative comparisons were made between media types.

Three separate comparative experiments were performed. The first experiment compared phase bright spores grown in two different agar media (LL and NSM). The second compared LL and NSM cultured phase dark spores. The third compared phase bright spores grown in two different broth media (MSM and Slepecky). Because the three experiments were performed at different times, we analyzed each experiment separately to avoid the influence of artifacts due to instrument drift or other subtle laboratory changes.

Three replicate extractions were used for each media type unless otherwise noted.

2.6. Identification using MALDI–TOF data

Intact protein masses were used to generate pools of candidate proteins for identification of a few markers. The mass information was submitted to a web-based Tagident tool using a 1% mass error searching either *Bacillus* or *B. subtilis* as the taxonomy search tools (http://us.expasy.org/cgi-bin/tagident0.pl; [34]). No restrictions on protein isoelectric point were used.

Peptide mass information used for identification was entered into the Mascot peptide mass fingerprinting tool (http://www.matrixscience.com; [35]). The data was compared to the Gram-positive sequences in the general mascot database. The search parameters included two missed cleavages for tryptic peptides, variable methionine oxidation, and average peptide masses with a 1.5 Da error (assuming singly charged ions). The top 20 candidates were collected and reduced to preferentially include *Bacillus* proteins with predicted masses within 10% of the measured protein mass.

3. Results and discussion

3.1. Processing of HPLC–MALDI–MS data

One of the major advantages of coupling a cLC separation to MALDI–MS is that it greatly expands the number of components that can be detected from a sample. It does this by increasing the amount of sample that can be loaded onto a column versus spotted on a plate, while also reducing the competitive ionization between proteins. This method increases the number of masses detected over direct MALDI–MS analysis. For example, in a separate analysis not described here, direct MALDI analysis of spores detected 51 peaks across replicate analysis where cLC–MALDI detected 631 total peaks for the same sample.

Fig. 2 shows a two-dimensional display of a typical cLC–MALDI chromatogram following pre-processing, so that only intensity values above noise levels are visible. Rows corre-

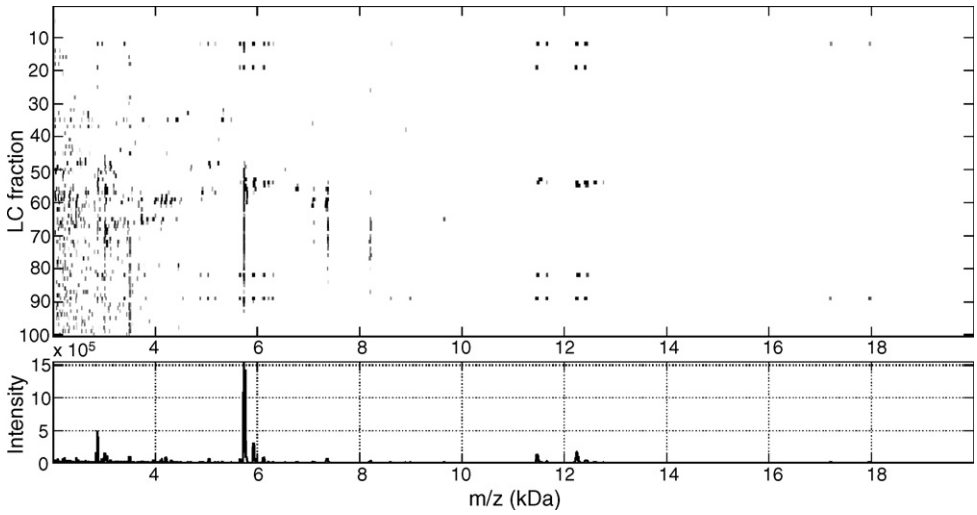

Fig. 2. Two-dimensional display of cLC–MALDI chromatogram for *B. subtilis* spores prepared on LL media (bright phase) following pre-processing. Only intensity values above noise levels are visible. Rows correspond to LC fraction, columns correspond to *m/z* values. Fractions 12, 19, 82, and 89 were used for external mass calibration. The insulin internal standard (mass 5735 Da) elutes at about fraction 54. The spectrum just below the image is the sum of MALDI mass spectra across all fractions.

spond to MALDI spectra for each fraction. The insulin internal standard (mass 5735 Da) elutes at about fraction 54. The spectrum just below the image is the sum of MALDI mass spectra across all fractions. Excluding the four external standard fractions, we detected 737 peaks, some of which correspond to the same peptide or protein (e.g., proteins that elute at consecutive fractions). Because peak detection algorithms are not perfectly accurate, we used these detected peak locations merely as guides for where quantitative comparisons should be made, either at their 2D locations (allowing for retention time shifts), or summed over all fractions.

3.2. Comparison of growth conditions

Visible differences are apparent in many of the MALDI mass spectra. For example, Fig. 3 shows differences in MALDI spectra at fraction 60 between LL and NSM media as phase bright spores. It is important to make sure that these apparent differences are not simply due to mismatched fractions. The image plots at the right of Fig. 4 show MALDI intensity values for the same mass range and between fractions 50 and 70, making it clear that these differences are not due to misaligned fractions (the retention time variation in this time range is at most two

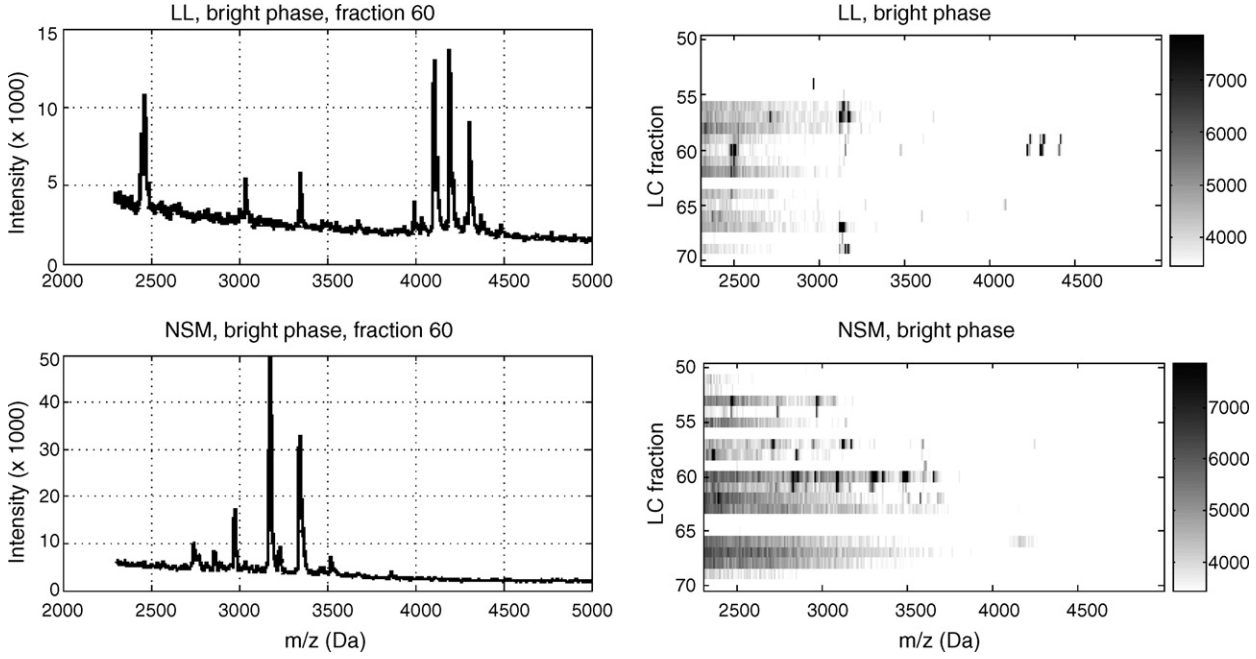

Fig. 3. Comparison of mass spectra for fraction 60 between LL and NSM bright phase samples. Adjacent plots show *m/z* 2500–5000 portions of fractions 50–70 to demonstrate differences appear from differential expression and not due to peak shifting. The relative intensity in the plots are given on the intensity scale at the left and right of the mass spectra and two-dimensional plot, respectively.

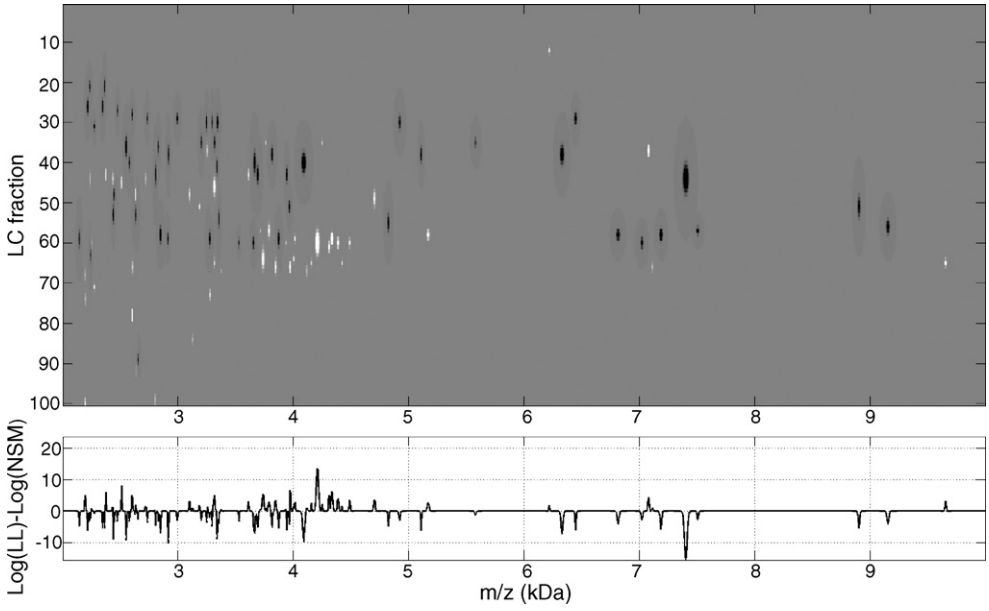

Fig. 4. Differential two-dimensional display comparing LL and NSM bright phase samples. White spots indicate peak regions where peaks were detected in all LL samples and no NSM samples, with corresponding large intensity differences. Dark spots indicate peak regions where peaks were detected in all NSM samples and no LL samples, with corresponding large intensity differences.

fractions). Although not shown, these differences are visible in all LL and NSM phase bright samples. Automated statistical analysis helps to facilitate the detection of these differences in replicate samples.

Given the small sample size, we focus here on detection of large differences, and do not attempt to comprehensively catalogue all potential biomarkers. In many cases, the large differences happen to correspond to proteins that are present in one condition and absent in the other (or in such low abundance that they are not detected).

Several media differences were found even in the total MALDI–MS comparisons. For this analysis, comparisons were made only at *m*/*z* locations corresponding to an apex in the total filtered second derivative values, averaged over all runs. For the three experiments, between 127 and 154 peak comparisons were made in the total MALDI spectra. To preferentially select large differences, differences were ranked using a modified or 'shrunken' 2-sample *t*-test, where the standard deviation used in the denominator is equal to the standard deviation of (log-transformed) intensity values at the *m*/*z* location *plus* the 90th percentile of the standard deviations computed for all selected *m*/*z* locations. This prevents selection of baseline-level proteins that happen to have very small within-class variance, perhaps an artifact of the baseline-correction algorithm.

In the first experiment, two LL extractions were compared with three NSM extractions, both for phase bright spores. The 10 masses that most differentiate the two media types for each of the three comparisons are given in Table 1. In the second experiment, we compared protein profiles of the same spores after their appearance had changed to phase dark. Although the protein profiles could still be differentiated, it was a different set of ions that differed between the phase bright spores from the LL and NSM media types (Table 2). In the third experiment, two liquid media types were compared, one with a rich nutrient source, MSM, and the other with minimal media, termed Slepecky. The replicate analysis of spores grown on MSM was easily differentiated from the minimal media grown Slepecky spores.

Note that even by aggregating LC fractions, we detect two to three times more peaks in the total MALDI spectra than would be obtained using direct MALDI–MS which would reflect the expected increased ionization suppression in direct MALDI. Even more peaks can be detected and compared by using the full two-dimensional map provided by cLC–MALDI.

Fig. 4 displays a 2D difference map between LL and NSM phase bright spores, which can be used as a first step toward identifying and validating protein differences between two media types. The figure was constructed using two LL samples and three NSM samples. White spots indicate peak regions where peaks were detected in all LL samples and no NSM sam-

Table 1
Top 10 MALDI intensity differences based on total MALDI mass spectra ranked in order of greatest to least significant

	LL vs. NSM (bright)	LL vs. NSM (dark)	MSM vs. SLEPECKY (bright)
1	4207	2700	2774
2	4311	7748	4925
3	4117	2151	2642
4	3009	5332	2677
5	6915	2669	7107
6	2765	2769	5521
7	2975	3775	3054
8	2348	9885	4684
9	2737	4080	2399
10	3179	7369	5564

Table 2
Database entries with similar parent protein and tryptic peptide mass fragment data to digested cLC fraction

Protein name and accession number	Position	Predicted [M + H]	Missed	Sequence
SASP-2 *B. subtilis* SAS2_BACSU	1–26	**2701.89**	0	AQNSQNGNSSNQLLVPGAAQAIDQMK–oxidation (M)
(M_r): 7328 score: 66 expect: 0.049	27–44	**1929.06**	0	FEIASEFGVNLGAETTSR
70 aa, seq. coverage: 100%	56–70	**1757.93**	1	RLVSFAQQNMSGQQF–oxidation (M)
	57–70	**1601.75**	0	LVSFAQQNMSGQQF–oxidation (M)
	45–56	**1189.29**	1	ANGSVGGEITKR
SASP-1 *B. subtilis* SAS1_BACSU	1–25	*2572.77*		PNQSGSNSSNQLLVPGAAQAIDQMK–oxidation (M)
(M_r): 7068, score: 34, expect 78	26–43	*1929.06*	0	FEIASEFGVNLGAETTSR
69 aa, seq. coverage: 79%	44–55	*1189.29*	1	ANGSVGGEITKR
SASP C-5 *B. megaterium* C24543	44–68	2573.85	2	ANGSVGGEITKRLVQMAEQQLGGGR oxidation (M)
(M_r): 7675; score: 44, expect: 7.8	6 25	1928.10	0	SSNELAVHGAQQAIDQMK oxidation (M)
73 aa, seq. coverage: 61%	55–70	1759.01	2	RLVQMAEQQLGGGRSK
	56–70	1602.82	1	LVQMAEQQLGGGRSK
	44–55	1189.29	1	ANGSVGGEITKR
***B. subtilis* strain 49760 unknown digest**		Measured [M + H]		Putative identity
(M_r): *7107*, **7373**		**2703.1**		**SASP-2**
		2573.4		*SASP-1*
		1929.8		**SASP-2**; *SASP-1*
		1758.9		**SASP-2**
		1602.6		**SASP-2**
		1298.1		
		1190.0		**SASP-2**; *SASP-1*
		1141.5		

Predicted and measured peptide masses given as protonated [M + H] masses. Trypsin fragment mass, position, number of missed cleavages, mascot score and sequence for top candidates matching peptide mass fingerprinting data. The measured mass [M + H] for the proteins and the resulting tryptic peptides from Fig. 5 are listed with masses corresponding to each protein given in italics or bold.

ples, with corresponding large intensity differences. Dark spots indicate peak regions where peaks were detected in all NSM samples and no LL samples, with corresponding large intensity differences.

3.3. Protein digestion for identification of select proteins

Protein masses alone can be used to differentiate between spores grown on different types of media; however, it is desirable to also have information on the identity of these markers. Mass alone provides a limited amount of identifying information, so additional information needs to be obtained to support marker identification. Therefore, we have investigated methods to provide more identifying information.

An important characteristic of MALDI–MS is that only a small portion of the small MALDI sample is interrogated with the laser. The remaining sample could be utilized for further analysis, including trypsin digestion. To confirm the identity of a few protein markers, we selected two fractions of abundant protein markers for trypsin digestion and identification. Following the initial MALDI–MS analysis, selected spots were resolubilized using 1 μL of 50% acetonitrile and 0.1% TFA. Sufficient residual ferulic acid remained in the digestion mixture to allow for detection of resulting peptides. It was also possible to recrystallize the digested sample by adding an additional 1 μL of ferulic acid matrix solution. Because ferulic acid served well as a matrix for detecting peptides after digestion, no attempt was made to exchange it out for an alternative matrix such as alpha-cyano-hydroxy cinnamic acid (ACHC). Attempts to overlay the reaction mixture with ACHC were unsuccessful.

The abundant markers at *m/z* 7372 and 7107 were selected for digestion and identification (Fig. 5). Using mass and species information, there are a limited number of proteins that are predicted within the *B. subtilis* 168 genome within 1% of the measured mass for each using the Tagident web-based tool

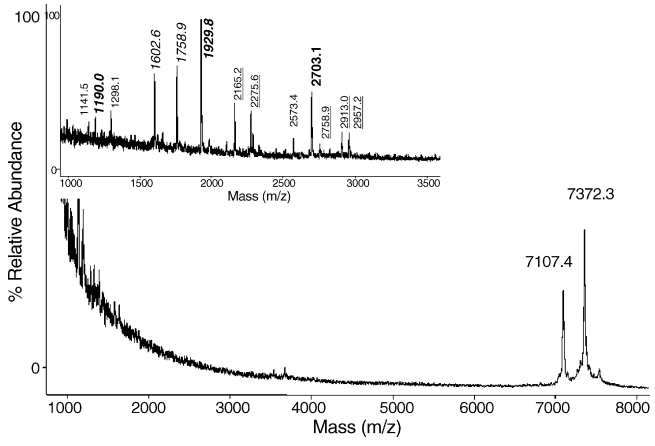

Fig. 5. Mass spectrum of two abundant masses detected in fractions 60–65 for LL media bright spores. Following resolubilization and trypsin digestion, peptide masses were detected from the proteins (inset). Tryptic fragments show highest similarity to SASP 1 (masses in italics) and to SASP 2 (masses in bold) database entries from *B. subtilis* globigii detailed in Table 2. Trypsin autolysis fragments underlined.

[34]. Twelve entries were found for 7370 Da and 10 entries for 7107 Da with no restrictions on p*I*. Among the entries were small acid soluble proteins previously described for *B. subtilis* globigii where SASP-1 (7067.79 Da) and SASP-2 (7332.03 Da) that were sequenced by mass spectrometric methods [32]. A similar entry for SASP-A (7070.89 Da) from *B. subtilis* 168 was also observed.

The peptides from the digest of the same fraction were detected (Fig. 5 inset) and compared to genomic sequence databases using the search tool Mascot [35]. A general database was searched using the web interface and restrictions were put on the taxonomy (Firmicutes), number of missed cleavages (2) and peptide mass tolerance of 1.5 Da. The possibility of protein modification (methionine oxidation) during the acid treatment indicated that protein identification needed to take this into account. The proteins matched within 10% of the intact protein mass were various *Bacillus* SASP protein sequences with the most confident match made to SASP-2 *B. subtilis* globigii. Table 2 gives a list of the detected peptide masses for both proteins as well as the candidate proteins matched with some similarity. While SASP sequences from several *Bacillus* species have predicted fragments in common with the unknown, 100% of the SASP-2 sequence can be accounted for from the experimental data. Taken together with the intact mass data, the likely matches are with the SASP-1 and SASP-2 described for *B. subtilis* globigii [32]. Using this data, the peptide masses shared between the proteins or corresponding to one particular protein are labeled accordingly in Fig. 5 inset.

3.4. Acid breakdown de novo sequence for identification

Another commonly observed marker at *m/z* 5044 was selected for further examination. Only eight masses were found to be within 2% error of this mass within the sequences for *Bacillus* using Tagident. Among them were two SASP N protein (5120.9 Da) from *B. halodurans* and SASP G (5139.3 Da), SAPS J (5030.6 Da) along with hypothetical proteins ypt A (5031.5 Da) and yhe F (4989.8 Da) for *B. subtilis*. For further identification, *m/z* 5044 had apparent degradation products eluting prior to and including the HPLC fractions where it was most abundantly observed. This is likely to be a result of acid hydrolysis, but a similar obvious mass ladder was not observed for other markers in the data set. The mass differences observed were 71.3, 147.7, 147.5, 116.7, 153.2, 115.1 and 127.0. Correlating these differences to the closest amino acid mass values gives a putative sequence read of AFFDRD(Q/K). While the mass agreement is not exact, this sequence bears a degree of similarity to the N terminal sequence (without methionine) for SASP J (*Ssp J*) from *B. subtilis* 168 (Q7WY58): GFFNKDK unlike MLNSEHF for the next closest mass ypt A. This is an even closer agreement for SASP J considering the conservative substitution of R for K.

4. Conclusion

A combination of capillary liquid chromatography and MALDI–MS was used to examine the proteins extracted from the same strain of *Bacillus* spores grown with different types of media. This method uses a (1) simple protein extraction, (2) low volume liquid separation, and (3) method for visualizing and comparing two-dimensional datasets. Approximately 10 times more peaks were detected by cHPLC–MALDI than by direct MALDI–MS. The advantage of employing a separation to increase the peaks observed was also described earlier by Dunalp and Li [36].

The spores grown on the different media appear to have unique protein markers that differentiate spores grown on each media type from one another. A more comprehensive catalog of biomarkers of growth culture will require an expanded set of experiments with (1) increased replication to allow detection of small to moderate quantitative differences in protein abundances, and (2) a larger variety of growth media. For this reason, this method serves as an excellent tool to support and expand the information gained by direct MALDI–MS.

The handling of spore samples in our lab has led us to perform microscopic examination of the spores prior to analysis. While storage of *B. subtilis* spores at 4 °C in pure water generally preserves the phase bright nature, after extended storage the appearance becomes phase dark. These phenotypic changes resulted in changes to the masses detected; however, the samples could still be differentiated with different protein markers.

We also describe a method for obtaining further information on the separated fractions after the initial data collection by resolubilization and trypsin digestion. The analysis of peptides was performed in the same well as originally collected using linear mode to maximize sensitivity. However, linear mode also afforded reduced resolution and mass accuracy compared to reflector mode. Performing the peptide digest analysis on standard plates with near neighbor external calibration will greatly improve the quality of the data. Nonetheless, identification of top candidate proteins was still possible by this technique.

Multiple pieces of information were used to identify two abundant protein markers including taxonomic similarity of source organism, intact protein mass and peptide mass fingerprint information. Two of the markers were found to be two previously described abundant proteins, SASP-1 and SASP-2 described for *B. globigii* [32]. This data also matched to highly conserved major SASP peptides from other *Bacillus* species but with lower overall similarity. There was likely oxidation of both proteins which would account for the mass difference from the previous report and is reflected in the peptide mass information. It is not entirely surprising that this data matched SASP database sequences for *B. globigii* (now *B. atrophaeus* [31]), given that *B. subtilis* 49760 is also a pigment producing strain.

One drawback is that peptide mass fingerprinting requires fairly pure fractions for identification. More specific information obtained through fragmentation of the individual peptides by tandem mass spectrometry would potentially increase confidence of the identification and be applicable to complex fractions. Extension of this approach will be to utilize tandem mass spectrometry in conjunction with peptide mass fingerprinting for identification of spore proteins.

Acknowledgements

This work was supported by Laboratory Directed Research and Development funding under the Pacific Northwest National Laboratory's Homeland Security Initiative. Pacific Northwest National Laboratory is operated by Battelle Memorial Institute for the US Department of Energy through contract DE-AC-06-76 RLO 1830.

References

[1] J.J. Bright, M. Claydon, M. Soufin, D. Gordon, J. Microbiol. Methods 48 (2002) 127.
[2] K.L. Wahl, S.C. Wunschel, K.H. Jarman, N.B. Valentine, C.E. Petersen, M.T. Kingsley, K.A. Zartolas, A.J. Saenz, Anal. Chem. 74 (2002) 6191.
[3] J. Walker, A. Fox, V. Edwards-Jones, D. Gordon, J. Microbiol. Methods 48 (2002) 117.
[4] E. Elhanany, R. Barak, M. Fisher, D. Kobiler, Z. Altboum, Rapid Commun. Mass Spectrom. 15 (2001) 2110.
[5] V. Ryzhov, Y. Hathout, C. Fenselau, Appl. Environ. Microbiol. 66 (2000) 3828.
[6] J.O. Lay Jr., Mass Spectrom. Rev. 20 (2001) 172.
[7] C. Fenselau, P. Demirev, Mass Spectrom. Rev. 20 (2002) 157.
[8] K.H. Jarman, S.T. Cebula, A.J. Saenz, C.E. Petersen, N.B. Valentine, M.T. Kingsley, K.L. Wahl, Anal. Chem. 72 (2000) 1217.
[9] R.J. Arnold, J. Reilly, Rapid Commun. Mass Spectrom. 12 (1998) 630.
[10] R.J. Arnold, J.P. Reilly, Anal. Biochem. 269 (1999) 105.
[11] M. Stump, J. Jones, R. Fleming, J. Lay jr., C. Wilkins, J. Am. Soc. Mass Spectrom. 14 (2003) 1306.
[12] H. Antelmann, S. Engelmann, R. Schmid, A. Sorokin, A. Lapidus, M. Hecker. 179 (1997) 7251.
[13] H. Antelmann, C. Scharf, M. Heckler, J. Bacteriol. 182 (2000) 4478.
[14] R.J. Arnold, K.A. Karty, A.D. Ellington, J.P. Reilly, Anal. Chem. 71 (1999) 1990.
[15] N. Valentine, S. Wunschel, D. Wunschel, C. Petersen, K. Wahl, Appl. Environ. Microbiol. 71 (2005) 58.
[16] D.S. Wunschel, E.A. Hill, J.S. McLean, K.H. Jarman, Y.A. Gorby, N. Valentine, K.L. Wahl, J. Microbiol. Methods 62 (2005) 259.
[17] A. Driks, Proc. Natl. Acad. Sci. 100 (2003) 3007.
[18] P. Schaeffer, J. Millet, J.P. Aubert, Proc. Natl. Acad. Sci. 54 (1965) 704.
[19] B. Setlow, K.A. Mcginnis, K. ragkousi, P. Setlow, J. Bacteriol. 182 (2000) 6906.
[20] R. Slepecky, J.W. Foster, J. Bacteriol. 78 (1959) 117.
[21] A. Westphal, P. Price, T. Leighton, K. Wheeler, Proc. Natl. Acad. Sci. 100 (2003) 3461.
[22] E. Melly, P.C. Geneset, M.E. Gilmore, S. Little, D.L. Popham, A. Driks, Pl Setlow, J. Appl. Microbiol. 92 (2002) 1105.
[23] Y. Hathout, B. Setlow, R. Cabrera-Martinez, C. Fenselau, P. Setlow, Appl. Environ. Microbiol. 69 (2003) 1100.
[24] D.N. Dickenson, M.T. La Duc, W.E. Haskins, I. Gormushkin, J.D. Winefordner, D.H. Powell, K. Venkateswaran, 70 (2004) 475.
[25] B. Warscheid, C. Fenselau, Anal. Chem. 75 (2003) 5618.
[26] B. Budowle, S.E. Schutzer, A. Einseln, L.C. Kelly, A.C. Walsh, J.A.L. Smith, B.L. Marrone, J. Robertson, J. Campos, Science 301 (2003) 1852.
[27] R.S. Murch, Biosec. Bioterror. Biodef. Strat. Prac. Sci. 1 (2003) 117.
[28] H.W. Kreuzer-Martin, M.J. Lott, J. Dorigan, J.R. Ehlringer, Proc. Natl. Acad. Sci. 100 (2003) 815.
[29] J.S. Page, J.V. Sweedler, Anal. Chem. 74 (2002) 6200.
[30] W.A. Harris, J.P. Reilly, Anal. Chem. 74 (2002) 4410.
[31] Y.J. Shaver, M.L. Nagpal, R. Rudner, L.K. Nakamura, K.F. Fox, A. Fox, J. Microbiol. Methods 50 (2002) 215.
[32] J.R. Whiteaker, B. Warscheid, P. Pribil, Y. Hathout, C. Fenselau, J. Mass. Spectrom. 39 (2004) 1113.
[33] D. Wunschel, K.E. Fox, G.E. Black, A. Fox, Syst. Appl. Microbiol. 17 (1994) 625.
[34] E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel, A. Bairoch Nuc, Acids Res. 31 (2003) 3784.
[35] D.N. Perkins, D.J. Pappin, D.M. Creasy, J.S. Cottrell, Electrophoresis 20 (1999) 3551.
[36] K.Y. Dunlap, L. Li, J. Chromatogr. A 925 (2001) 123.