ELSEVIER

Journal of Chromatography B, 748 (2000) 167–177

JOURNAL OF CHROMATOGRAPHY B

www.elsevier.com/locate/chromb

# Use of non-porous reversed-phase high-performance liquid chromatography for protein profiling and isolation of proteins induced by temperature variations for Siberian permafrost bacteria with identification by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and capillary electrophoresis–electrospray ionization mass spectrometry

Bathsheba E. Chong[a], Jeongkwon Kim[a], David M. Lubman[a,*], James M. Tiedje[b], Sohpia Kathariou[c]

[a]*Department of Chemistry, The University of Michigan, 930 N. University Avenue, Ann Arbor, MI 48109-1055, USA*
[b]*Center for Microbial Ecology, Michigan State University, East Lansing, MI 48824-1325, USA*
[c]*Department of Microbiology, University of Hawaii at Manoa, Honolulu, HI 96822, USA*

**Abstract**

Non-porous reversed-phase high-performance liquid chromatography (NP-RP-HPLC) has been used to separate and isolate proteins from whole cell lysates of ED 7-3, a bacterium from the buried Siberian permafrost sediment. The proteins collected from the liquid eluent of this separation were then analyzed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) and capillary electrophoresis–electrospray ionization mass spectrometry (CE–ESI-MS). In order to study the differences in expression of cold-shock proteins (CSPs) at different growth temperatures, cultures of the ED 7-3 strain were prepared at 4°C and 25°C. The goals of this work were twofold: firstly, to identify the presence of CSPs and other proteins that are highly expressed at 4°C but not at 25°C; and secondly, to isolate these proteins for MALDI-TOF-MS and CE–ESI-MS identification. In this initial work, distinct protein profiles were observed for these cultures as a function of temperature. Fraction collection from the eluent of NP-RP-HPLC of some of the highly expressed proteins was performed and the proteins were mass analyzed for molecular mass. Peptide maps of the proteins were generated by tryptic digestion and were analyzed by CE–ESI-MS and MALDI-TOF-MS for database identification of the expressed proteins. 

*Keywords:* Proteins

## 1. Introduction

An area of current interest is the use of non-porous reversed-phase high-performance liquid chromatography (NP-RP-HPLC) for rapid profiling of the protein content of cellular extracts. The use of NP-RP-HPLC has been shown to provide rapid, high-resolution separations of proteins compared to porous-based separations. Separations from whole cell lysates of proteins using NP columns have been

*Corresponding author. Tel.: +1-734-764-8776; fax: +1-734-764-1669.
*E-mail address:* dmlubman@umich.edu (D.M. Lubman).

0378-4347/00/$ – see front matter 
PII: S0378-4347(00)00288-7

achieved in typically 15–30 min. The use of $C_{18}$ coated, silica-based NP packing materials eliminates the problem of proteins sticking inside the porous materials resulting in improved resolution and protein recovery [1–4]. NP-RP-HPLC has been used for separations of proteins from whole cell lysates of human erythroleukemia (HEL) cells [5], bacterial cells [6], and human breast cancer cells [7,8]. Other applications of this method include separation of proteins and peptides [9–12], vitamins [13,14], drug products [15] and various phenols [16].

The mechanism of survival of extremophiles is an area of current interest. In this study, we have chosen to examine bacteria strain ED 7-3, which has been isolated from the buried Siberian permafrost sediment that has been continuously frozen for thousands of years [17]. ED 7-3 is a gram-positive bacterium that belongs to the genus *Exiguobacterium*. The growth of ED 7-3 involves a complex series of integrated chemical reactions that are affected directly by external temperature. The extremely cold weather in Siberia affects the most susceptible cell components, namely the proteins. Cellular proteins will require adjustments to cope with the extreme temperature in the environment to survive as well as to maintain balanced growth. Thus, cold temperature has a profound impact on the physiology of bacterial cells and specific sets of cold-shock proteins (CSPs) are generated to protect the cells from freezing. CSPs may affect the formation of ice crystals by their ability to adsorb to ice and thus inhibit ice crystal growth, even though the actual mechanism of this process is still unclear. Several studies [18–26] have shown that in bacteria these unique proteins help to retard freezing or lessen the damage incurred upon freezing and thawing.

In this work, tandem-column [27] NP-RP-HPLC is demonstrated as a rapid method for separation and isolation of proteins in the liquid phase from bacterial whole cell lysates for further analysis by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) and capillary electrophoresis–electrospray ionization mass spectrometry (CE–ESI-MS). Bacterium ED 7-3, obtained from the Siberian permafrost, was studied in order to examine the cold shock response and resulting induction of CSPs at 4°C and room temperature. The whole cell lysates were rapidly profiled and separated using a tandem-column NP-RP-HPLC method where distinct protein profiles were obtained from each culture. Proteins that were differentially expressed at 4°C were collected and sized for molecular mass by MALDI-TOF-MS. The proteins were digested with trypsin to generate peptide maps for identification by MALDI-TOF-MS analysis and further confirmed by CE–ESI-MS using database searching procedures [28–33]. As a result, several CSPs and other cold-shock related proteins were identified.

## 2. Experimental

### *2.1. NP-RP-HPLC analysis*

A Beckman (Fullerton, CA, USA) System Gold HPLC system was utilized for this work. The pump (Model 128) has a gradient solvent delivery module with a built-in system controller. The detector was a programmable detector module (Model 166) with an analytical flow cell. All separations in this work were monitored at 214 nm. The columns were from Eichrom Technologies (Darien, IL, USA) and contained non-porous silica bead packing material of a 1.5 μm particle diameter coated with $C_{18}$. The dimension of the first column was 33×4.6 mm I.D. (ODS3E) while the second column was 14×4.6 mm I.D. (ODS1). These wide-bore columns enabled the collection of sufficient material for further analysis and the relatively short length allowed for rapid separations on the order of 20 min. The tandem-column NP-RP-HPLC separations of proteins were performed via gradient elution of two solvents [solvent A: Milli-Q water with 0.1% trifluoroacetic acid (TFA); solvent B: acetonitrile with 0.1% TFA] with a flow-rate of 1 ml/min. The temperature of the first column was maintained at 60°C in a Timberline (Boulder, CO, USA) column heater while the second column was at room temperature (25°C). The gradient profile used for solvent B was generally as follows: 0% for 1.5 min; 0 to 30% in 15 min; 30 to 50% in 8 min; 50 to 100% in 1 min; 100% for 3 min; 100 to 0% in 1 min. In order to obtain a reproducible separation profile, the sample was ''conditioned'' to the column environment by mixing the sample with an equivalent amount of water (0.1% TFA), or a 1:1

ratio. This acidifying step was performed prior to sample injection. Each injection contained an average of 50 to 150 μg of protein. Using a Speed-Vac, the volume of the fractions collected was reduced to dryness. Depending on the subsequent analysis, the proteins were re-solubilized in either 60% acetonitrile with 1% TFA (for MALDI mass analysis) or 50 m*M* of ammonium hydrogencarbonate (for tryptic digestion).

### 2.2. Chemicals

The chemicals involved in this study were used without prior purification. Acetone (HPLC grade) was obtained from Fisher (Fair Lawn, NJ, USA). Acetonitrile, guanidine hydrochloride (gu-HCl), α-cyano-4-hydroxycinnamic acid (α-CHCA), TFA and octyl glucopyranoside (OCG) were from Aldrich (Milwaukee, WI, USA). Bovine cytochrome *c*, lysozyme, bovine serum albumin (BSA), leucine-enkaphalin (leu-enk), angiotensin III, bradykinin and neurotensin were purchased from Sigma (St. Louis, MO, USA). Trypsin was acquired from Promega (Madison, WI, USA). Distilled and deionized water was obtained from a Milli-Q reagent grade purification system from Millipore (Bedford, MA, USA). The nitrocellulose (NC), Immobilon-NC Pure, was also from Millipore.

### 2.3. Bacterial cell cultures

The ED 7-3 strain was grown in liquid medium on a shaker at either 4°C (7-3-4) for 6 days or 25°C (7-3-25) for 24 h. The liquid medium was 1/2 strength Trypticase Soy Broth with 5% glycerol. Cells were harvested at an optical density range of 0.8 to 1.0 by centrifugation, then resuspended in fresh medium and kept on wet ice. Cells were subsequently stored at −80°C. Details of isolation of ED 7-3 and other strains from Siberian permafrost frozen sediments were discussed in previous work by Rivkina et al. [17].

### 2.4. Matrix, substrate and sample preparation

α-CHCA was the matrix used in these experiments. For MALDI analysis, the matrix was prepared as a saturated solution in a 3:2 ratio of acetonitrile to Milli-Q water containing 1% TFA. Before applying the sample solution to the steel probe tip, 2 μl of nitrocellulose (NC; ~10 mg/ml in acetone) was applied to the tip and air dried followed by 3–5 μl of sample solution, which was also air dried. The final step involved the addition of 3 μl of matrix solution onto the probe tip that was also air-dried. External calibrations were performed throughout these experiments using bovine cytochrome *c*, lysozyme and BSA for protein analysis, while leu-enk, angiotensin III, bradykinin and neurotensin were used for peptide analysis. Calibrations were determined at the beginning of the analysis and after each spectrum with the sample. The mass accuracy for protein analysis (up to 70 kDa) was approximately ±5 Da, while for peptide analysis (500 to 3000 Da) it was about ±0.5 Da.

The bacterial cell pellets were lysed in 200 μl of 50 m*M* Tris buffer (~pH 7), 6 *M* gu-HCl and 5–10 m*M* OCG. This mixture was vortex-mixed for 30 s and stored overnight at −20°C. Subsequently, the lysate was thawed on ice and fractionated by centrifugation in an Eppendorf centrifuge at 16 500 rpm for 20 min. In order to remove as much particulate and cell debris as possible, the soluble portion (supernatant) was centrifuged at 16 500 rpm for another 10 min. The most reproducible results were observed when the supernatant was analyzed within 24 h.

To remove the acetonitrile and TFA, the total volume of the NP-RP-HPLC collected fractions was reduced to near dryness with a Speed-Vac. Then 20 μl of 50 m*M* ammonium hydrogencarbonate was added to each fraction and vortex-mixed. Finally, 2 μl of the sequencing-grade-modified trypsin (0.5 μg/μl) was added to the fractions. Again the mixtures were vortex-mixed and incubated at 37°C for 24 h.

After tryptic digestion, the resulting peptides were subjected to pulsed delayed extraction (PDE) MALDI analysis to obtain peptide maps. In the PDE–MALDI analysis, one part of the saturated matrix solution was diluted 10-fold, i.e., 9 μl of 60% acetonitrile (1% TFA) was added to 1 μl of the saturated matrix solution. A 2-μl volume of this diluted matrix solution was then added directly to the probe tip. To improve the PDE–MALDI mass spectra, the tryptic digestions were dried down using

the Speed-Vac and re-acidified by adding 6 μl of 60% acetonitrile (1% TFA). A 3-μl volume of this acidified digest solution was then mixed with 2 μl of the diluted matrix solution. This mixture was vortex-mixed and applied to the dried matrix layer. The other 3 μl was dried down and reconstituted in water before being subjected to CE–ESI-MS analysis.

### 2.5. MALDI analysis

The TOF mass spectrometer employed in these studies was a modified Wiley–McLaren design with a four-plate acceleration stage [34]. It was capable of high-voltage acceleration up to ±20 kV (R.M. Jordan, Grass Valley, CA, USA). The laser source used to produce MALDI was a Minilite 10 Hz Nd–YAG laser system (Continuum, Santa Clara, CA, USA). All mass spectra were obtained using 355 nm radiation. In this work, between 15 and 30 laser shots were averaged from 5 to 10 spots on the probe tip. The laser power density was estimated at $\sim 5 \cdot 10^6$ to $1 \cdot 10^7$ W/cm$^2$. The detector was a triple microchannel plate (MCP) detector (R.M. Jordan), which adapted a CuBe conversion dynode with post-acceleration (PA) capability up to ±12 kV [35] in front of the MCP. The total ion acceleration across the TOF device may thus be >30 kV. The PA stage enhances the detection of heavy species, but at the expense of resolution. In addition, PDE as described in previous studies [36–39] could be used to enhance the resolution for the analysis of the tryptic digests. The 1-m long flight tube was pumped to a base pressure of $8 \cdot 10^{-7}$ to $1 \cdot 10^{-6}$ Torr by a diffusion pump (Varian, Lexington, MA, USA) (1 Torr=133.322 Pa). Data was recorded using a LeCroy 9310AM (400 MHz) digital oscilloscope (LeCroy, Chestnut Ridge, NY, USA) and was processed on a Gateway 586 computer.

### 2.6. CE–ESI-MS analysis

CE analysis was performed using fused-silica capillaries of 110 μm O.D.×40 μm I.D. (Polymicro Technologies, Phoenix, AZ, USA). The total length of the capillaries was 50 cm. The capillary was coated with polybrene according to the procedure given by Li et al. [40]. 100 m*M* formic acid and ammonia were used to prepare pH 3 buffer solution. 100 m*M* formic acid was prepared first and titrated to the point of pH 3 with ammonia. CE was performed using a 30 kV power supply (Model CZE 1000R, Spellman High Voltage Electronics, Plainview, NY, USA). Electrokinetic injection method was used to inject the sample. In the beginning both the electrode (at −12 kV) and the capillary end were in the buffer solution. Both were then placed in the sample solution, where the voltage applied to the electrode was between −2 and −8 kV for 10 to 15 s, causing the sample to enter the capillary. Next, the sample vial was replaced by the buffer vial and −15 kV was applied to the electrode. After 1 min, the voltage decreased to −12 kV and 3 kV was simultaneously applied to the electrospray needle for electrospray ionization. As a result, the potential difference across the capillary was 15 kV.

An ion trap storage reflectron time-of-flight mass spectrometry (IT-reTOF-MS) system was used as a detector for the CE–MS analysis as described in previous work [41]. It consists of a quadrupole ion trap storage device (Model C-1251, R.M. Jordan) interfaced to a reflectron TOF mass analyzer (Model D-850, R.M. Jordan). Detection was performed with a 40 mm triple-microchannel plate detector (Model C-2501, R.M. Jordan).

### 2.7. Database searching procedure for protein identification

The MS-Fit sequence database located in the Protein Prospector program was used for protein identification by entering the peptide masses generated by tryptic digestion. The program is available on the Internet at http://prospector.ucsf.edu. Subsequently, other relevant parameters such as protein species, molecular mass and isoelectric point (p*I*) range are also entered in order to narrow down the search. In this case, *Bacillus subtilis* was chosen as the species database because it is the species phylogenetically closest to ED 7-3. Since these proteins were obtained from HPLC, no p*I* information was available. Thus, the p*I* range was set between 3 and 10. The range of molecular mass values for each search was determined by MALDI-TOF-MS analysis. The tolerance for the search of peptides against the database was set at 3 Da (CE–ESI-MS) or 5 Da (MALDI-TOF-MS). The reasons for using the 3–5

Da mass range were: to compare the peptide maps from both techniques in obtaining possible protein identification; to accommodate mutations and other modifications such as phosphorylation; and since a protein database for ED 7-3 is not currently available, the database of *Bacillus subtilis* was used with a wider mass range to accommodate small differences.

## 3. Results and discussion

Fig. 1a and b show the protein profiles of whole cell lysates of the ED 7-3 using the tandem-column NP-RP-HPLC method. Employing a relatively short gradient program (about 25 min), distinct profiles were obtained for each growth temperature. For the cold culture, 7-3-4, the majority of proteins elute between 3 and 17 min while for 7-3-25 most of the proteins elute after 15 min into the gradient. Fig. 1a shows the protein profile for the culture grown at 4°C where a large number of new proteins have been expressed relative to the 7-3-25 culture. Some of these peaks have been identified and are tabulated in Table 1 where several CSPs such as CSPC, CSI4B, CSI15 and CSI5 are included. Other proteins that are highly expressed belong to the general stress family (GSP170), translation and transcription family (IF-1, SENS and ABRB), degradation regulation enzyme family (DEGQ), and the ribosomal family (S6). Trisephosphate isomerase (TIM) and RNA polymerase sigma-B factor (Sigma-37) are also highly expressed. This RNA-polymerase is also known as general stress protein 84 or GSP84. Many of these proteins are housekeeping proteins and have been previously identified in *B. subtilis* and other bacteria [18–26]. They are essential for the proper function of the bacterial cells during the cold-shock process [19–22].

One of the advantages of using the tandem-column separation method compared to a single column is the enhanced loadability of sample material. Employed on its own, the first column (ODS3E, 33×4.6 mm I.D.) can separate up to 50 μg of protein sample. However, by adding a second shorter column (ODS1, 17×4.6 mm I.D.) up to 150 μg of sample can be loaded. The second column also provides enhanced focusing of each band, thus improving peak resolution for peaks not well resolved by the first column. The use of two different temperatures for the two columns speeds up the separation (under 25 min), and further improves the quality of the separation by improving the selectivity for a particular sample. As a result, large numbers of protein from a complex mixture are separated in a short time with sufficient resolution for further analysis by MS. Since enhanced amounts of material can be loaded and separated using this tandem-column technique, fractions from one separation alone contain sufficient amounts of protein for tryptic digestion for analysis by MALDI-TOF-MS and CE–ESI-MS. Even though protein recovery from the non-porous column is protein-dependent, for a complex mixture up to 80% of the proteins are recovered or eluted off the column [5–8].

This tandem-column NP-RP-HPLC method has several advantages for separation of proteins from whole cell lysates in these studies compared to more traditional one- (1-D) or two-dimensional (2-D) sodium dodecyl sulfate–polyacrylamide gel electrophoresis. A major advantage of the NP-RP-HPLC method is the separation and isolation of proteins in the liquid phase. The purified proteins in the liquid phase are readily amenable to analysis by MALDI-TOF-MS or CE–ESI-MS as compared to gel methods which require time consuming extraction and purification procedures for analysis of the proteins embedded in the gels. In addition, even after these purification procedures the percent coverage of the protein digests of proteins isolated by NP-RP-HPLC is generally much improved compared to that observed for proteins excised from gels. A second advantage of the NP-RP-HPLC method in these studies is the speed of analysis where a typical tandem-column separation of a cell lysate may only take 30 min. In comparison, an equivalent 1-D gel separation might require more than 10 h to obtain reasonable separation, while a 2-D gel separation would require 2–3 days. Further, the liquid separation can be totally automated and the data is obtained in digitized form. The gel methods are not readily automated and the gel image must be digitized. Furthermore, the NP-RP-HPLC method provides improved separations over 1-D gels [5–8]. Although 2-D gels provide higher resolution than the NP-RP-HPLC method, the resolution obtained is

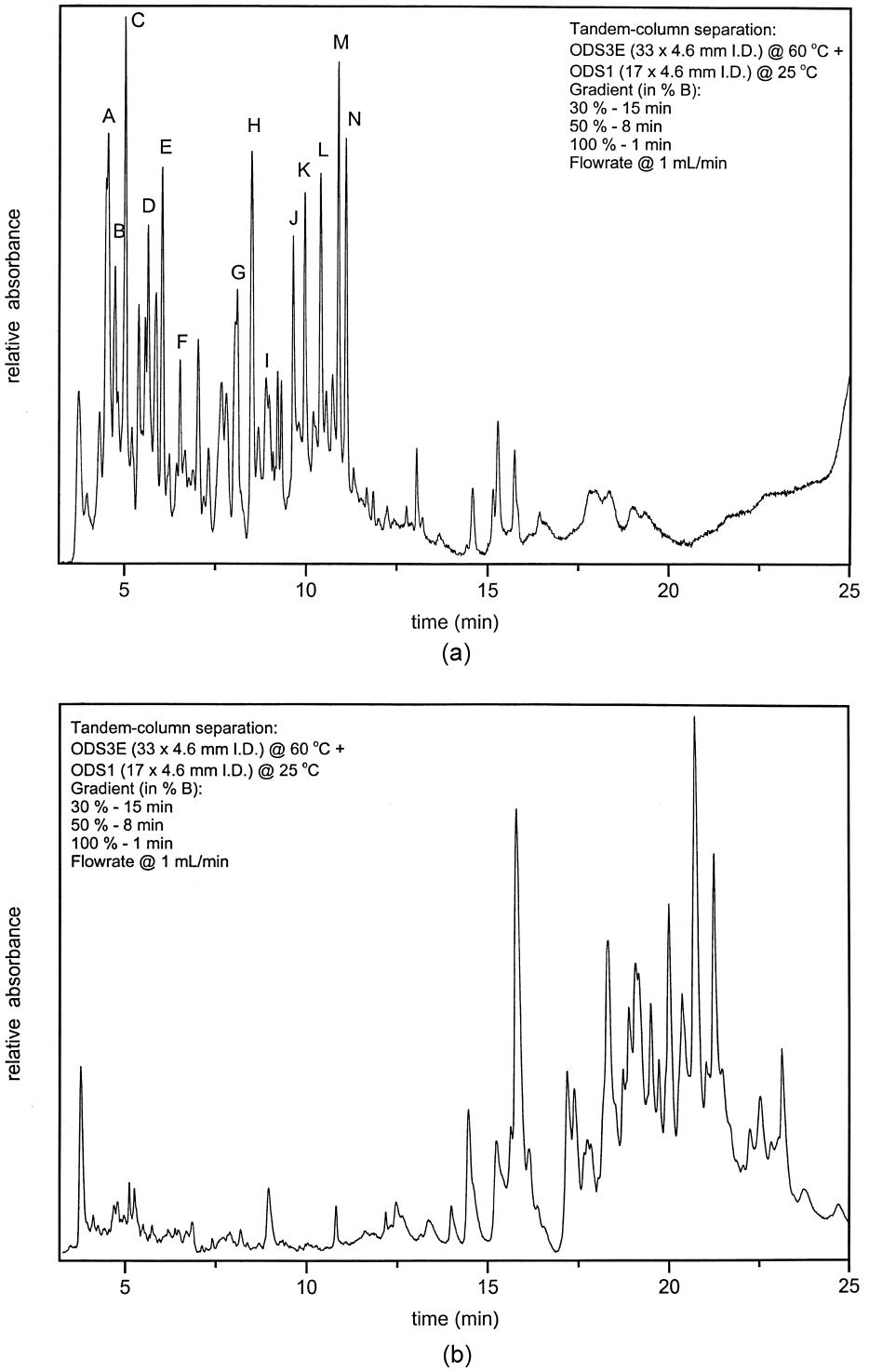


Fig. 1. (a) NP-RP-HPLC protein profile of the whole cell lysate of the 7-3-4 strain. (b) NP-RP-HPLC protein profile of the whole cell lysate of the 7-3-25 strain.

Table 1
Proteins induced during cold shock response in ED 7-3 at 4°C (from Fig. 2)[a]

| Peak | Protein name | kDa/p*I* | AC |
|---|---|---|---|
| A | [Still to be determined] | – | – |
| B | [Still to be determined] | – | – |
| C | DNA-binding protein II (HB) (HU) | 9884.3/8.96 | P08821 |
| C | Stage V Sporulation protein M | 3017.8/10.57 | P37817 |
| D | Cold shock protein CSI4B (9 KD cold shock protein) | 1924.3/4.52 | P81094 |
| D | General stress protein 170 (GSP170) | 2158.4/8.80 | P80242 |
| E | Cold-shock induced protein 15 (CSI15) | 2376.8/10.83 | P54616 |
| E | Transcriptional regulatory protein SENS | 7912.5/10.00 | P21344 |
| F | Translation initiation factor IF-1 | 8213.6/6.83 | P20458 |
| G | [Still to be determined] | – | – |
| H | Cold shock proteinC (CSPC) | 7255.0/4.72 | P39158 |
| H | Cold shock protein CSI5 (11 KD cold shock protein) | 1359.7/11.17 | P81095 |
| I | Transcription state regulatory protein ABRB | 10 772.7/6.29 | P08874 |
| J | Degradation enzyme regulation protein DEGQ | 5546.5/6.16 | Q99039 |
| K | 30S Ribosomal protein S6 (BS9) | 11 124.6/5.2 | P21468 |
| L | [Still to be determined] | – | – |
| M | Trisephosphate isomerase (TIM) | 27 029.8/5.00 | P27876 |
| N | RNA polymerase sigma-B factor (Sigma-37) (general stress protein 84) (GSP84) | 29 901.4/5.55 | P06574 |

[a] Tryptic peptide maps used for MS-Fit database search were obtained from both CE–ESI-MS and MALDI-TOF-MS.

sufficient for these studies, when combined with the mass spectrometry capabilities.

The use of MALDI-TOF-MS can provide molecular mass information of the proteins separated in HPLC, while MALDI-TOF-MS and CE–ESI-MS can provide mass analysis of peptides generated by enzymatic mapping. By using this information with a database search, we were able to identify many of the bacterial proteins separated by NP-RP-HPLC. Only proteins that appeared in both search results were used as possible candidates for identification. Table 2 shows an example of the results obtained from CE–ESI-MS and MALDI-TOF-MS analysis of one particular protein fraction. Database search

Table 2
An example of MS-Fit database search results for fraction H using (a) CE–ESI-MS data and (b) MALDI-TOF-MS data[a]

| *m/z* submitted | MH+ matched | Delta Da | Start | End | Peptide sequence | Modifications |
|---|---|---|---|---|---|---|
| *(a) Results from CE–ESI-MS data*[b] | | | | | | |
| 779.2900 | 776.8243 | 2.4657 | 40 | 46 | (K)SLDEGQK(V) | |
| 825.2500 | 825.9459 | −0.6959 | 14 | 20 | (K)GFGFIER(E) | |
| 1557.5900 | 1559.6176 | −2.0276 | 1 | 13 | (−)EQGTVKWFNAEK(G) | 1PO4 Acet N |
| *(b) Results from MALDI-TOF-MS data*[c] | | | | | | |
| 796.2100 | 792.9329 | 3.2771 | 1 | 7 | (−)MEQGTVK(W) | |
| 796.2100 | 794.8883 | 1.3217 | 8 | 13 | (K)EFNAEK(G) | |
| 954.6450 | 958.0660 | −3.4210 | 57 | 66 | (R)GAQAANVQKA(−) | |
| 1568.9770 | 1568.7980 | 0.1790 | 1 | 13 | (−)MEQGTVKWFNAEK(G) | |
| 1651.9540 | 1648.7777 | 3.1763 | 1 | 13 | (−)MEQGTVKWFNAEK(G) | 1PO4 |
| 2109.3410 | 2113.2549 | −3.9139 | 21 | 39 | (R)ENGDDVFVHFSAIQSDGFK(S) | |

[a] Search parameters used: mass range=7000–8000 Da; p*I*=3–10; peptide masses=average; digest used=trypsin; database used=Swiss Protein; species=*Bacillus subtilis*; considered modifications=peptide N-terminal Gln to pyroGlu, phosphorylation of S, T and Y, oxidation of M, protein N-terminus acetylated.

[b] 7255.0 Da, p*I*=4.72, Acc. P39158. Cold shock protein C (CSPC). The matched peptides cover 40% (27/66 amino acids) of the protein.

[c] 7255.0 Da, p*I*=4.72, Acc. P39158. Cold shock protein C (CSPC). The matched peptides cover 62% (41/66 amino acids) of the protein.

results from the tryptic maps generated by MALDI-TOF-MS and CE–ESI-MS were able to confirm the presence of cold shock protein C (CSPC, 7255.0 Da) in fraction H (see Fig. 1a). The matched peptides cover 62% of the protein for MALDI-TOF-MS data and 40% for the CE–ESI-MS data. Fig. 2 shows the PDE–MALDI-TOF-MS of the tryptic digest of fraction H. It should be noted that many of the collected fractions from NP-RP-HPLC were shown to contain two or more proteins when sized by MALDI-TOF-MS. For these fractions, the MS-Fit database search on the tryptic peptides was performed twice (or more) using the corresponding mass range for each protein. When digesting these fractions, a blank tryptic digest was always performed and then analyzed on MALDI-TOF-MS for trypsin autolysis peaks. When sufficient protein is present for digestion, these autolysis peaks are seldom detected. A continuation of this work will involve performing another digest, possibly CNBr followed by trypsin, to further ascertain the identity of these proteins as well as tandem MS analysis on peptides of interest.

Both MALDI-TOF-MS and CE–ESI-MS were used together to confirm the validity of protein identification. From Table 2 it is shown that CE–ESI-MS detects mainly the lower mass range (up to 1500 Da) of the tryptic peptide fragments while MALDI-TOF-MS is more efficient in detecting peptides between 1000 and 3000 Da. Due to interference from MALDI matrix, peptides of 500 Da and under are generally suppressed. The two methods complement each other in terms of total peptide fragments detected for each tryptic digest fraction and result in an improved percent coverage of the total protein map as shown in Table 3. By combining the peptide maps from both techniques, the total

Table 3
A summary of Table 2 where the total peptide coverage of cold shock protein C (CSPC, 7255.0 Da) was increased to 100% by combining the peptide maps from MALDI-TOF-MS and CE–ESI-MS (the peptide fragments detected are underlined)[a]

| | | |
|---|---|---|
| 1 | MEQGTVKWFNAEKGFGIERENGDDVFHF | 30 |
| 31 | SAIQSDGFKSLDEGQK**VSFDVEQGAR**GAQA | 60 |
| 61 | ANVQKA | 66 |

[a] Fragment 47–56 (**VSFDVEQGA**) has a mass of 1108.2001 Da but MALDI-TOF-MS detected a fragment at 1072.17 Da (see Fig. 2). This corresponds to the $MH+-2H_2O$ ion of this fragment. As a result, the peptide fragments cover 100% of CSPC. Without this fragment, the underlined peptides cover 85% (56/66) of the protein.

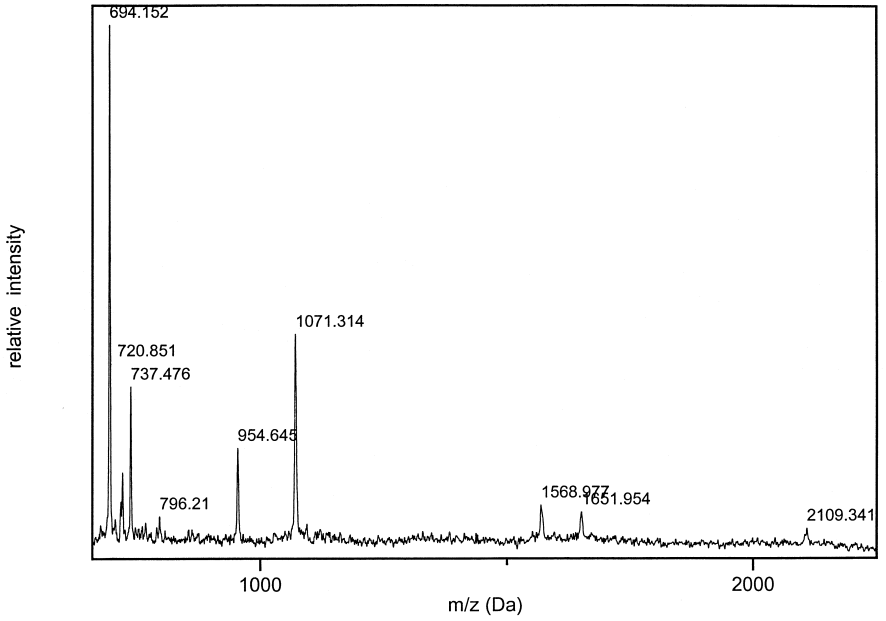


Fig. 2. PDE–MALDI-TOF mass spectrum of the tryptic digest of HPLC fraction H (Fig. 1a).

peptide coverage was increased up to 85% (55/66 amino acids) for CPSC. Table 4 shows an example of the database search results for fraction E (Fig. 1a) using both MS techniques. In this case, the matched peptides cover 61% of transcriptional regulatory protein SENS with the CE–ESI-MS data and 67% with the MALDI-TOF-MS data. This results in an overall improved coverage of up to 80% for the protein as shown in Table 5. Thus, both sets of data (CE–ESI-MS and MALDI-TOF-MS) allow confirmation of protein identity using the MS-Fit database searching procedure, especially when several reasonable protein-matches might be obtained with either technique alone.

Table 5
A summary of Table 4 where the total peptide coverage of transcriptional regulatory protein SENS (7912.4 Da) was increased to 80% by combining the peptide maps from MALDI-TOF-MS and CE–ESI-MS (the peptide fragments detected are underlined)[a]

| | | |
|---|---|---|
| 1 | MGVKKEKGRKRFRKRKTYGNQILPLELLIE | 30 |
| 31 | KNKREIINSAELMEEIYMKIDEKHTQCVTK | 60 |
| 61 | YKKTR | 65 |

[a] The matched peptides covered 80% (52/65) of the protein.

It should be noted that a database for ED 7-3 is currently not available for MS-Fit searches. In these studies, *B. subtilis* was used to conduct the MS-Fit searches. *B. subtilis* is the species phylogenetically

Table 4
An example of MS-Fit database search results for fraction E using (a) CE–ESI-MS data and (b) MALDI-TOF-MS data[a]

| *m/z* submitted | MH+ matched | Delta Da | Start | End | Peptide sequence | Modifications |
|---|---|---|---|---|---|---|
| *(a) Results from CE–ESI-MS data*[b] | | | | | | |
| 147.0800 | 147.1980 | −0.1180 | 5 | 5 | (K)K(E) | |
| 147.0800 | 147.1980 | −0.1180 | 10 | 10 | (R)K(R) | |
| 147.0800 | 147.1980 | −0.1180 | 14 | 14 | (R)K(R) | |
| 147.0800 | 147.1980 | −0.1180 | 16 | 16 | (R)K(T) | |
| 147.0800 | 147.1980 | −0.1180 | 63 | 63 | (K)K(T) | |
| 258.8500 | 261.3022 | −2.4522 | 32 | 33 | (K)NK(R) | |
| 387.7700 | 390.3543 | −2.5822 | 61 | 62 | (K)YK(K) | 1PO4 |
| 439.2400 | 438.5493 | 0.6907 | 61 | 63 | (K)YKK(T) | |
| 439.7300 | 438.5493 | 1.1807 | 61 | 63 | (K)YKK(T) | |
| 486.4900 | 484.4715 | 2.0185 | 63 | 65 | (K)KTR(−) | 1PO4 |
| 504.8200 | 504.5626 | 0.2574 | 50 | 53 | (K)IDEK(H) | |
| 577.3400 | 578.7547 | −1.4147 | 1 | 5 | (−)MGVKK(E) | 1Met-ox |
| 814.8400 | 816.9582 | −2.1182 | 54 | 60 | (K)HTQCVTK(Y) | |
| 817.9000 | 816.9582 | 0.9418 | 54 | 60 | (K)HTQCVTK(Y) | |
| 1303.8400 | 1302.4975 | 1.3425 | 50 | 60 | (K)IDEKHTQCVTK(Y) | |
| 1823.8600 | 1825.0560 | −1.1960 | 17 | 31 | (K)TYGNQILPLELLIEK(N) | 1PO4 |
| 1955.2700 | 1953.2308 | 2.0392 | 16 | 31 | (R)KTYGNQILPLELLIEK(N) | 1PO4 |
| *(b) Results from MALDI-TOF-MS data*[c] | | | | | | |
| 504.6950 | 504.5626 | 0.1324 | 50 | 53 | (K)IDEK(H) | |
| 1466.7890 | 1462.4569 | 4.3321 | 50 | 60 | (K)IDEKHTQCVTK(Y) | 2PO4 |
| 1988.7900 | 1987.3552 | 1.4348 | 17 | 33 | (K)TYGNQILPLELLIEKNK(R) | |
| 1988.7900 | 1986.3219 | 2.4681 | 34 | 49 | (K)REIINSAELMEEIYMK(I) | 1Met-ox |
| 1988.7900 | 1990.0931 | −1.3031 | 35 | 49 | (R)EIINSAELMEEIYMK(I) | 2PO4 1Met-ox |
| 2381.6140 | 2379.6534 | 1.9606 | 35 | 53 | (R)EIINSAELMEEIYMKIDEK(H) | 1PO4 |

[a] Search parameters used: mass range=7000–8000 Da, p*I*=3–10; peptide masses=average; digest used=trypsin; database used=Swiss Protein; species=*Bacillus subtilis*; considered modifications=peptide N-terminal Gln to pyroGlu, phosphorylation of S, T and Y, oxidation of M, protein N-terminus acetylated.

[b] 7912.5 Da, p*I*=10.00, Acc. P21344. Transcriptional regulatory protein SENS. The matched peptides cover 61% (40/65 amino acids) of the protein.

[c] 7912.5 Da, p*I*=10.00, Acc. P21344. Transcriptional regulatory protein SENS. The matched peptides cover 67% (44/65 amino acids) of the protein.

closest to ED 7-3, for which the genome has been completely sequenced and annotated. Even though strain ED 7-3 (*Exiguobacterium* sp.) and *B. subtilis* are distinct phylogenetic units, and may not be very closely related [42], the mass spectral tryptic mapping procedure from CE–ESI-MS and MALDI-TOF-MS provided reasonably close matches to known proteins of *B. subtilis*. This may suggest that in the cold shock response the proteins are conserved between these two gram-positive microorganisms [19–22]. Strong conservation in cold shock proteins has indeed been noted across diverse microbial lineages [43].

## 4. Conclusions

NP-RP-HPLC separation and isolation of proteins coupled with MALDI-TOF-MS and CE–ESI-MS as shown here is not only promising, but is becoming an important and versatile tool in protein profiling and identification even with complex mixtures. The tryptic peptide maps from MALDI-TOF-MS and CE–ESI-MS together provide improved coverage of the total protein, thus enhancing the validity of protein identity and the ability to verify the sequence against known proteins. As a result, several of the CSPs and other highly expressed proteins were identified for ED 7-3 grown at 4°C. The long-term goals of this work will be to create a searchable protein database for ED 7-3 and other permafrost bacteria as well as to enhance our understanding of the mechanisms of cold-shock induction and the role of the CSPs in the survival of bacterial cells in less than ideal environmental conditions. Further work will entail the use of other enzymatic digestions and intensive tandem-MS analysis to confirm protein identification and protein sequences compared to those in the *Bacillus* database. Apart from ED 7-3, other strains of the permafrost bacteria will also be studied.

## Acknowledgements

We gratefully acknowledge partial support of this work by the National Science Foundation under the LExEn Grant DEB 9978271 and Grant DEB 9120006 through the Center for Microbial Ecology at Michigan State University, by the PG Research Foundation, Inc., and by the National Institutes of Health Grant R01 GM49500. We thank Dr. Tim Barder and Mike Fern of Eichrom Technologies, Inc. for generously providing the non-porous RP-HPLC columns used in this work. We also thank Dr. Klaus Nusslein for growing the bacterial cultures used in this study.

## References

[1] T.J. Barder, P.J. Wohlman, C. Thrall, P.D. DuBois, LC·GC Mag. 15 (1997) 918.
[2] D.R. Jenke, J. Chromatogr. Sci. 34 (1996) 362.
[3] K. Kalghatgi, C. Horvath, Anal. Biotechnol. 434 (1990) 162.
[4] H.J. Wirth, K.K. Unger, M.T.W. Hearn, J. Chromatogr. 550 (1990) 383.
[5] Y. Chen, D. Wall, D.M. Lubman, Rapid Commun. Mass Spectrom. 12 (1998) 1994.
[6] D.B. Wall, D.M. Lubman, S.J. Flynn, Anal. Chem. 71 (1999) 3894.
[7] B.E. Chong, D.M. Lubman, F.R. Miller, A.J. Rosenspire, Rapid Commun. Mass Spectrom. 13 (1999) 1808.
[8] B.E. Chong, D.M. Lubman, A.J. Rosenspire, F.R. Miller, Rapid Commun. Mass Spectrom. 12 (1998) 1986.
[9] J.F. Banks, E.E. Gulcicek, Anal. Chem. 69 (1997) 3973.
[10] T. Issaeva, A. Kourganov, K. Unger, J. Chromatogr. A 846 (1999) 13.
[11] R. Ohmacht, B. Boros, K. Kovacs, I. Kiss, Magy. Kem. Foly. 105 (1999) 465.
[12] J.E. MacNair, G.J. Opiteck, J.W. Jorgenson, M.A. Moseley III, Rapid Commun. Mass Spectrom. 11 (1997) 1279.
[13] R. Ohmacht, B. Boros, I. Kiss, J. Jelinek, Chromatographia 50 (1999) 75.
[14] A. Makarov, J. Szpunar, J. Anal. Atom. Spectrom. 14 (1999) 1323.
[15] X. Xu, J.T. Stewart, J. Liq. Chromatogr. Rel. Technol. 22 (1999) 1857.
[16] A.P. Kohne, T. Welsch, J. Chromatogr. A 845 (1999) 463.
[17] E. Rivkina, D. Gilichinsky, S. Wagener, J. Tiedje, J. McGrath, Geomicrobiology 15 (1998) 187.
[18] F. Berger, N. Morellet, F. Menu, P. Potier, J. Bacteriol. 178 (1996) 2999.
[19] P. Graumann, K. Schroder, R. Schmid, M.A. Marahiel, J. Bacteriol. 178 (1996) 4611.
[20] P. Graumann, M.A. Marahiel, Arch. Microbiol. 171 (1999) 135.
[21] P. Graumann, T.M. Wendrich, M.H.W. Weber, K. Schroder, M.A. Marahiel, Mol. Microbiol. 25 (1997) 741.
[22] T. Kaan, B. Jurgen, T. Schweder, Mol. Gen. Genet. 262 (1999) 351.
[23] W. Bae, P.G. Jones, M. Inouye, J. Bacteriol. 179 (1997) 7081.

[24] J.M. Panoff, D. Corroler, B. Thammavongs, P. Boutibonnes, J. Bacteriol. 179 (1997) 4451.
[25] D.O. Bayles, B.A. Annous, B.J. Wilkenson, Appl. Environ. Microbiol. 62 (1996) 1116.
[26] P.G. Jones, M. Inouye, Mol. Microbiol. 11 (1994) 811.
[27] ''Liquid chromatography: tandem system rivals 2-D gels'', High Tech Sep. News, 12 (1999) 3.
[28] Y.J. Chen, X.Y. Jin, D. Misek, R. Hinderer, S.H. Hanash, D.M. Lubman, Rapid Commun. Mass Spectrom. 13 (1999) 1907.
[29] J.R. Yates, Electrophoresis 19 (1998) 893.
[30] A. Ducret, I. Van Oostveen, J.K. Eng, J.R. Yates, R. Aebersold, Protein Sci. 7 (1998) 706.
[31] D. Arnott, W.J. Henzel, J.T. Stults, Electrophoresis 19 (1998) 968.
[32] G.J. Feistner, K.F. Faull, D.F. Barofsky, P. Roepstorff, J. Mass Spectrom. 30 (1995) 519.
[33] E. Mortz, O. Vorm, M. Mann, P. Roepstorff, Biol. Mass Spectrom. 23 (1994) 249.
[34] R.M. Whittal, L. Li, Anal. Chem. 67 (1995) 1950.
[35] J. Bai, Y.H. Liu, T.C. Cain, D.M. Lubman, Anal. Chem. 66 (1994) 3423.
[36] Y.D. Zhu, L. He, J. Srinivasan, D.M. Lubman, Rapid Commun. Mass Spectrom. 11 (1997) 987.
[37] R.M. Whittal, L.M. Russon, S.R. Weinberger, L. Li, Anal. Chem. 69 (1997) 2147.
[38] S.M. Colby, J.P. Reilly, Anal. Chem. 68 (1996) 1419.
[39] R.S. Brown, J.J. Lennon, Anal. Chem. 67 (1995) 1998.
[40] M.X. Li, L. Liu, J.T. Wu, D.M. Lubman, Anal. Chem. 69 (1997) 2451.
[41] X.Y. Jin, J.K. Kim, S. Parus, D.M. Lubman, Anal. Chem. 71 (1999) 3591.
[42] J.A.E. Farrow, S. Wallbanks, M.D. Collins, Int. J. Syst. Bacteriol. 44 (1994) 74.
[43] P. Graumann, M.A. Marahiel, Arch. Microbiol. 166 (1996) 293.