Chapter 1

Introduction

1.1 Proteomics

1.1.1 Introduction of Proteomics

The increasing availability of fully or partially sequenced genomes for a variety of organisms (including humans) has ushered in a new era of biological system research aimed towards the understanding of life processes. Although the study of genes has generated much attention and has successfully explained many phenomena in biological systems, it is the proteins that perform most life functions and make up the majority of cellular structures. Proteomics is considered the next step in the study of biological systems after genomics. The word "proteome", a portmanteau of "*prote*in" and "gen*ome*", is the entire complement of proteins produced by an organism or system. The term "proteomics", coined in analogy with genomics, the study of the genes, is the large-scale study of proteins, particularly their structures and functions [1, 2].

Proteomics is much more challenging than genomics. While the genome of the cell is static, nearly identical for all cells of an organism, and consistent across a species, the proteome is dynamic as one organism has radically different protein expression in different parts of its body, different stages of its life cycle and different environmental conditions. Therefore, there is no fixed proteome. Another major difficulty is the

complexity of proteomics is considered much greater than that involved in the genome. It has been estimated that the average number of proteins per gene is one or two in bacteria, three in yeast and three to more than six in humans [3]. The human genome is estimated to contain 20,000–25,000 genes but it is estimated that more than 500,000 proteins are derived from these genes. The result is that direct gene expression analysis is required either at the mRNA or protein level. Although some information about protein expression levels may be obtained from the analysis of an organism's transcriptome at the mRNA level [4-7], the cellular mRNA expression may not be correlated directly to protein expression [8-10] due to post-translational modification (PTM) and protein degradation. It is therefore not surprising that proteome analysis has become a key enabling technology in the emerging science of systems biology. The advantage of proteomics lies in the ability to directly examine the biomolecules and assemblies of biomolecules that are most responsible for the function of biological systems.

The diversity and extent of proteome will not and could not be solved by a single technology. Proteomic studies over the last several years demonstrated that the most effective proteomic analysis of even the simplest biological system uses a combination of protein separation and identification techniques. The general strategy in proteomic research includes sample preparation, protein or peptide separation, their identification, and data interpretation.

1.1.2 Protein Separation and Protein Expression Mapping

To achieve the most comprehensive protein separation, old techniques have been refined and new ones introduced. For adequate representation of the proteome, only multidimensional separation techniques can provide resolving capability in proteomic

analysis and have proven to be superior to one-dimensional approaches. Until recently, two-dimensional gel electrophoresis (2-DE) is the most widely used separation technique in proteomics, in which proteins are focused according to their isoelectric point (p*I*) by isoelectric focusing (IEF) in the first dimension and then resolved in the second dimension based on their relative molecular masses (*M*r) typically by sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Using this approach, proteins are first separated by 2DE, detected by staining and quantified based on their staining intensities. Proteins of interest are then cut out of the gel, digested and identified by mass spectrometric techniques. Although this approach has been proven to be a successful means of imaging a large number of proteins within a given sample, no information for intact protein could be obtained since protein must be digested for further analysis. Also, this approach selects against low abundance proteins, membrane proteins, and proteins with extreme isoelectric points and molecular weights. In addition, 2DE is difficult to interface with mass spectromic techniques for further protein identification. These inherent limitations of 2-DE have forced researchers to look for other methods of protein separation such as capillary electrophoresis [11], protein microarrays [12], and multi-dimensional liquid chromatography [13].

Our laboratory has introduced and developed a 2-D liquid separation proteomic strategy to profile protein expression in bacterial cell lysates, cancer tissue and cell line samples [14-18]. Using this strategy, the protein extraction is first fractionated according to p*I* by chromatofocusing (CF), and further separated based on hydrophobicity by nonporous silica reversed phase HPLC (NPS-RP-HPLC). The separated proteins are then mapped to generate an image of the cellular protein content. There are several different

ways for protein expression mapping. Firstly, a 2-D UV map is produced by plotting the UV intensity against retention time during NPS-RP-HPLC separation. Secondly, the second dimensional NPS-RP-HPLC is on-line coupled with electrospray ionization time-of-flight mass spectrometer (ESI-TOF MS) for intact protein *M*r measurement, and thus generates a 2-D mass map that resembles the format of a stained 2-D gel but in a digitized format. The 2-D liquid separation based proteomic strategy has considerable advantages over 2-DE. A major issue is that separated proteins are purified in the liquid phase, which provides an automated way to collect hundreds of proteins for subsequent enzymatic digestion and identification by mass spectrometry. Furthermore, the three separation component of this strategy (CF, NPS-RP-HPLC, and ESI-TOF MS) could be integrated effectively without loss of each individual advantage. By this means, the proteins experienced a third dimensional separation based on molecular weight after being separated by p*I* and hydrophobicity using CF and RP-HPLC. In addition, the 2D mass mapping technique makes it possible for intact protein analysis. With the use of accurate molecular weight, the potential protein ID could be determined among several candidates in database searching. Another advantage of this approach is that map comparisons become much easier using this digitized format. Also, the 2-D liquid separation based proteomic strategy makes it possible for analyzing the proteins with extreme p*I*s by selecting the start and elute buffers at extreme pH in the CF separation.

1.1.3 Protein Identification by Mass Spectrometry and Database Searching

By definition, a mass spectrometer consists of an ion source that ionized the analytes, a mass analyzer that measures the mass-to-charge ratio (m/z) of the ionized analytes, and a detector that records the number of ions at each m/z value. The

development of two ionization methods – electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI), as recognized by 2002 Nobel Prize in chemistry, revolutionized protein chemistry and fundamentally changed the analysis of proteins [19, 20]. These methods solved the difficulty of generating ions from large, nonvolatile analytes such as proteins and peptides without producing significant fragmentation of the analytes. Due to the lack or minimal extent of analyte fragmentation, they are referred to as "soft" ionization techniques. In fact, they are so soft that even noncovalent interactions may be maintained during the ionization processes.

ESI creates ions by applying a potential to an analyte-containing liquid and thus causing the liquid to charge and subsequently spray. The electrospray generates very small droplets of solvent-containing analyte. The solvent is removed as the droplets enter the mass spectrometer by heat or other form of energy (e.g. energetic collisions with a gas), and multiply-charged ions are formed in the process. ESI gained immediate popularity because of the compatibility with on-line liquid separations by chromatography and capillary electrophoresis. Furthermore, the propensity of ESI to produce multiply-charged ions, arising by either proton or alkali ion attachment, allows the detection of high molecular weights using the simple quadrupole instrument and other types of mass analyzers with limited m/z range.

MALDI creates the ions by exciting the analyte molecules that are isolated from the energy of the laser by an energy absorbing matrix. The matrix, typically a low molecular weight organic compound, is believed to serve three functions: absorption of energy from the laser light, isolation of the biopolymer molecules from each other, and donation of protons. To maintain the efficient and controllable energy transfer, a molar

ratio ranging from 100:1 to 50,000:1 of matrix to analyte is optimal for ion production. The laser energy strikes the crystalline matrix to cause rapid excitation of the matrix and subsequent ejection of matrix and analyte ions into the gas phase. Irradiation of these crystals by short pulse (ns time scale) of UV or infrared light initiates desorption and ionization, where predominantly singly protonated intact molecular ions $[M+H]^+$ are produced. Given this short duration and the fact that laser beams can easily be focused to small spot sizes, ions are generated essentially at a point source in space and time. This feature makes it ideally compatible with time of flight (TOF) mass analyzer, which is robust, simple, and sensitive and has a large mass range.

There is a direct relationship between mass spectrometry data and amino acid sequence. Peptide molecular weight measurements are predictive of amino acid composition, and peptide fragmentation information relates to amino acid sequence. Protein identification is performed in two main ways, one is peptide map fingerprinting (PMF) and another is peptide sequencing using MS/MS. PMF uses masses of peptide fragments which are generated from enzymatic digestion of proteins. Using probability search algorithms such as MS-Fit and Mascot, measured masses of peptide fragments could be matched with the masses of fragments derived theoretically from a protein in a protein database such as SwissProt and NCBInr. In PMF, protein identification was based on the suggestion that at least five matched peptides with mass accuracy of 50 ppm and sequence coverage of at least 20% were available[21]. Although PMF has been widely used in protein identification, in practice, however, it is often difficult to obtain protein IDs largely due to an insufficient number of detected peptides.

Another protein identification approach is based on peptide sequencing using peptide collision-induced dissociation (CID) spectra. The major advantage of this approach is that the proteins that are difficult to identify by PMF due to an insufficient number of detected peptides could be identified using this approach if the peptide signal and its fragment signals are of sufficient intensity. In peptide sequencing, proteins are usually fragmented into peptides called parent ions by proteases prior to introduction to MS. One of the parent ions is selected and fragmented *in situ* by collision induced dissociation (CID) to generate daughter ions. Daughter ions can result in a series of consecutive fragments ions indicating the amino acid sequence of the peptide. Masses of parent and daughter ions are used for database matching to identify a protein.

1.2 Glycoproteomics

1.2.1 Introduction of Protein Glycosylation

Once the proteins are identified, the next step is characterization of their post-translational modifications (PTMs). Among more than 100 different types of PTMs, glycosylation is by far the most common and functionally important in both eukaryotes and prokaryotes [22]. Alterations in either the level of or type of glycosylation has been shown to influence cellular processes such as growth, differentiation, transformation, adhesion, metastasis and immune surveillance of tumors [23-25]. For these reasons, glycosylation increases efforts to identify protein expression patterns that correlate with disease states [26]. Many existing cancer biomarkers are glycoproteins, such as Her2/neu in breast cancer, prostate-specific antigen (PSA) in prostate cancer, CA125 in ovarian cancer, and carcinoembryonic antigen (CEA) in colorectal, bladder, breast, pancreatic and lung cancer [27]. There are four types of glycosylation: (1) the *N*-linked glycans

which are attached to asparagine residues in the consensus sequence Asn-Xxx-(Ser/Thr) via an *N*-acetylglucosamine (GlcNAc) residue, where Xxx can be any amino acid except proline; (2) the O-linked glycans where the oligosaccharide's reducing end is attached to the hydroxyl groups of serine or threonine residue within a protein through *N*-acetylgalactosamine (GalNAc); (3) glycoslphophatidylinositol anchors, which are attached to carboxyl terminus of certain membrane-associated proteins; (4) C-glycosylation, which has been found attached to tryptophan residues in certain membrane-associated and secreted proteins [28]. In rare cases, cysteine or lysine may also be glycosylated [29]. *N*-glycans are composed of the common trimannosyl chitobiose core structure with two or more antennas extending from it. Changes in *N*-glycosylation have long been associated with the development of disease. In order to address the question of disease-related glycosylation alteration, sensitive and fast strategies for the analysis of glycosylation of proteins are required.

1.2.2 Isolation of Glycoproteins by Lectin Affinity Chromatography

Generally, if necessary, glycoproteins can be purified by most conventional protein separation methods, such as ion exchange, size exclusion, partition, hydrophobic interaction, dye-ligand, and affinity chromatography. These chromatography techniques can be used to reduce the complexity of some samples [30]. Lectin affinity chromatography is the most widely used technique to purify glycoproteins with specific carbohydrate structures of the proteome and provide substantial simplification of the mixture. Lectins are a family of carbohydrate-recognizing proteins that are classified into a number of specificity groups based on the carbohydrate for which they exhibit the highest affinity [31-33]. Lectins are often not affected by other features of the glycan in

which they reside, and therefore demonstrate broad specificities towards glycans. The lectin affinity columns were prepared by adding agarose-bound lectin into a small column [34, 35], whereby the sample is loaded by a gravity-flow mode. After eliminating non-specific binding by washing, glycoproteins are eluted by displacement from the column with an elution buffer containing a haptene saccharide. Various lectins can be used to isolated glycoproteins having distinct types of carbohydrate structures, *e.g.* galectins specific for LacNAc-containing glycans found exclusively in both *N*-glycans and O-glycans [36, 37], Concanavalin A (ConA) binds with preference to oligomannosidic, hybrid, and bi-antennary *N*-glycans, either unconjugated or attached to proteins or peptides [38], *Aleuria aurentia* lectin (AAL) binds fucose linked (α-1, 6) to *N*-acetylglucosamine or (α -1, 3) to *N*-acetyllactosamine related structure [39], both *Sambucus nigra* bark lectin (SNA), *Maackia amurensis* lectin II (MAL) recognize sialic acid on the terminal branches [40], and peanut agglutinin (PNA) binds desialylated galactosyl (β-1, 3) *N*-acetylgalactosamine [41]. The potential of lectin affinity chromatography was recently used to identify serum/plasma glycoprotein that correlate with various cancers [42-45].

1.2.3 Identification of *N*-glycosylation Changes by Glycoprotein Microarrays

Protein microarrays are emerging as a powerful high throughput proteomic means to detect proteins [46], monitor their expression levels [46-48], track how they are modified [49-51], and determine how they interact with other biomolecules, such as protein or DNA [52]. In a typical protein microarray experiment, proteins are arrayed on a solid substrate as spots, the array is probed with a capture molecule, such as a monoclonal antibody or nucleic acid aptamer, and the interaction recorded through

different detection methods [53]. Since protein microarray elements can be miniaturized to contain tens of thousands of capture features they are considered a multiplexed device, thus protein microarrays have been a promising tool to perform high throughput clinical validation of putative biomarkers. Recently, lectins have been applied in arrays for rapid profiling of glycan expression patterns [54-56]. We have extended the ideas of protein microarrays and lectin arrays into a multi-lectin detection based protein microarray strategy for high throughput and multiplex profiling glycan pattern changes in glycoproteins from human serum [50]. To analyze the plasma glycosylation patterns, all fractions containing the separated intact glycoproteins are arrayed on nitrocellulose slides. Subsequently, the slides are screened to analyze the different glycan structures using five different lectins: ConA, MAL, SNA, AAL, and PNA. The utilization of these five lectins have been proved to be highly successful in covering >95% of *N*-glycan types reported and differentiating them according to their specific structures [57].

1.2.4 Statistical Analysis of Glycoprotein Microarray Data

The amount of data in glycoproteomics has increased rapidly along with the development of techniques and methods. Statistical and computational analysis of glycoproteomic data have become crucial to validate and interpret the huge amount of data [58-60]. In chapter IV and V, principal components analysis (PCA) and hierarchical clustering (HC) analysis of the normalized glycoarray data have been performed to differentiate the plasma samples in terms of their overall *N*-glycosylation patterns and to relate these patterns to clinical status. An alternative way to analyze the lectin glycoarray data is to search for signature proteins that might differentiate the plasma samples of different clinical status. This was done by calculating the z statistics of each array spot.

PCA is carried out using log-transformed and normalized array spot intensities. The leading two eigenvectors of the sample covariance matrix are used for visualization. Sample pairs falling close together in the scatter plot are more similar in terms of their overall patterns of normalized glycoform abundances. HC procedure is used without any prior knowledge of grouping to find criteria appropriate for classifying the cases according to the glycosylation pattern from glycoarrays. To do this, the normalized array spot intensities are log transformed, and the pair-wise Pearson correlations are used to carry out HC in which more closely correlated pairs of samples were joined at a lower point on the dendrogram. The scale on the dendrograms is $100 - 100 \times r$, where r is the Pearson correlation coefficient. For differential abundance analysis, Z-statistics for each protein detected by each lectin was calculated. Comparisons are made of normal versus adenoma, normal versus cancer as well as adenoma versus cancer. Based on the Bonferroni correction for two-sided testing of 36 peaks, Z values of ≥ 3.2 or ≤ -3.2 could be deemed to have significantly different glycosylation levels at a 95% significance level.

1.3 References:

[1] Anderson, N. L. and Anderson, N. G., *Electrophoresis* 1998, 19, 1853-1861.
[2] Blackstock, W. P. and Weir, M. P., *Trends Biotechnol.* 1999, 17, 121-127.
[3] Wilkins, M. R., Sanchez, J. C., Gooley, A. A., Appel, R. D*., et al.*, *Biotechnol. Genet. Eng. Rev.* 1996, 13, 19-50.
[4] Velculescu, V. E., Zhang, L., Vogelstein, B. and Kinzler, K. W., *Science* 1995, 270, 484-487.
[5] Zhang, L., Zhou, W., Velculescu, V. E., Kern, S. E*., et al.*, *Science* 1997, 276, 1268-1276.
[6] Velculescu, V. E., Zhang, L., Zhou, W., Vogelstein, J*., et al.*, *Cell* 1997, 88, 243-251.
[7] Schena M., Shalon, D., Davis, R. W. and Brown, P. O., *Science* 1995, 270, 467-470.
[8] Gygi, S. P., Rochon, Y., Franza, B. R. and Aeberold, R., *Mol. Cell. Biol.* 1999, 19, 1720-1730.
[9] Anderson, L. and Seihamer, J., *Electrophoresis* 1997, 18, 533-537.
[10] Griffin, T. J., Gygi, S. P., Iderker, T., Rist, B*., et al.*, 2002.
[11] Simpson, D. C. and Smith, R. D., *Electrophoresis* 2005, 26, 1291-1305.
[12] MacBeath, G. and Schreiber, S. L., *Science* 2000, 289, 1760-1763.
[13] Kachman, M. T., Wang, H., Schwartz, D. R., Cho, K. R. and Lubman, D. M., 2002.
[14] Zheng, S., Schneider, K. A., Barder, T. J. and Lubman, D. M., *BioTechniques* 2003, 35, 1202-1212.
[15] Wang, Y., Wu, R., Cho, K. R., Shedden, K. A*., et al.*, *Molecular and cellular proteomics* 2005.
[16] Wang, H., Kachman, M. T., Schwartz, D. R., Cho, K. R. and Lubman, D. M., *Electrophoresis* 2002, 23, 3168-3181.
[17] Qiu, Y., Kathariou, S. and Lubman, D. M., *Proteomics* 2006, 6, 5221-5233.
[18] Kreunin, P., Urquidi, V., Lubman, D. M. and Goodison, S., *Proteomics* 2004, 4, 2754-2765.
[19] Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. and Whitehouse, C. M., *Science* 1989, 246, 64-71.
[20] Karas, M. and Hillenkamp, F., *Anal. Chem.* 1988, 60, 2299-2301.
[21] Jensen, O. N., Podtelejnikov, A. V. and Mann, M., *Anal. Chem.* 1997, 69, 4741-4750.
[22] Moens, S. and Vanderleyden, J., *Arch. Microbiol.* 1997, 168, 169-175.
[23] Hakomori, S., *Adv. Exp. Med. Biol.* 2001, 491, 369-402.
[24] Hakomori, S., *Proc. Natl. Acad. Sci. USA* 2002, 99, 225-232.
[25] Choudhury, A., Moniaux, N., Ulrich, A. B., Schmied, B. M*., et al.*, *Br. J. Cancer* 2004, 90, 657-664.
[26] Morelle, W., Canis, K., Chirat, F., Faid, V. and Michalski, J., *Proteomics* 2006, 6, 3993-4015.
[27] Diamandis, E. P., *Mol. Cell. Proteomics* 2004, 3, 367-378.
[28] Hofsteenge, J., Müller , D. R., de Beer, T., Löffler, A*., et al.*, *Biochemistry* 1994, 33, 13524-13530.

[29] Rajeswari, P., Natarajan, R., Nadler, J. L., Kumar, D. and Kalra, V. K., *J. Cell Physiol.* 1991, 149, 100-109.
[30] Ottens, A. K., Kobeissy, F. H., Wolper, R. A., Haskins, W. E.*, et al.*, *Anal. Chem.* 1995, 77, 4836-4845.
[31] Sharon, N. and Lis, H., *Adv. Exp. Med. Biol.* 2001, 491, 1-16.
[32] Sharon, N. and Lis, H., *J. Agric. Food. Chem.* 2002, 50, 6586-6591.
[33] Sharon, N. and Lis, H., *-234Science* 1989, 246, 227.
[34] Cummings, R. D., *Methods Enzymol.* 1994, 230, 66-86.
[35] Taketa, K., *Electrophoresis* 1998, 19, 2595-2602.
[36] Hirabayashi, J., Satoh, M. and Kasai, K., *J. Biol. Chem.* 1992, 267, 15485-15490.
[37] Hirabayashi, J., Ubukata, T. and Kasai, K., *J. Biol. Chem.* 1996, 271, 2497-2505.
[38] Cummings, R. D. and Kornfeld, S., *J. Biol. Chem.* 1982, 257, 11230-11234.
[39] Kochibe, N. and Furukawa, K., *Biochemistry* 1980, 19, 2841-2846.
[40] Knibbs, R. N., MacCallum, D. K., Lillie, J. H. and Goldstein, I. J., *Glycobiology*, 4, 419-428.
[41] Neurohr, K. J., Young, N. M. and Mantsch, H. H., *J. Biol. Chem.* 1980, 255, 9205-9209.
[42] Zhao, J., Patwa, T. H., Qiu, W., Shedden, K.*, et al.*, *J. Proteome Res.* 2007, 6, 1864-1874.
[43] Kreunin, P., Zhao, J., Rosser, C., Urquidi, V.*, et al.*, *J. Proteome Res.* 2007, 6, 2631-2639.
[44] Yang, Z., Harris, L. E., Palmer-Toy, D. E. and Hancock, W. S., *Clin. Chem.* 2006, 52, 1897-1905.
[45] Xiong, L., Andrews, D. and Regnier, F., *J. Proteome Res.* 2003, 2, 618-625.
[46] Kattah, M. G., Alemi, G. R., Thibault, D. L., Balboni, I. and Utz, P. J., *Nat. Methods.* 2006, 3, 745-751.
[47] Sreekumar, A., Nyati, M. K., Varambally, S., Barrette, T. R.*, et al.*, *Cancer Res.* 2001, 61, 7585-7593.
[48] Qiu, J., Madoz-Gurpide, J., Misek, D. E., Kuick, R.*, et al.*, *J. Proteome Res.* 2004, 3, 261-267.
[49] Pal, M., Moffa, A., Sreekumar, A., Ethier, S. P.*, et al.*, *Anal. Chem.* 2006, 78, 702-710.
[50] Patwa, T. H., Zhao, J., Anderson, M. A., Simmeone, D. M. and Lubman, D. M., *Anal. Chem.* 2006, 6411-6421.
[51] Martin, K., Steinberg, T. H., Cooley, L. A., Gee. K.R.*, et al.*, *Proteomics* 2003, 3, 1244-1255.
[52] Oh, S. W., Moon, J. D., Lim, H. J., Park, S. Y.*, et al.*, *FASEB. J.* 2005, 19, 1335-1337.
[53] Cretich, M., Damin, F. and Pirri, G., *Biomol. Eng.* 2006, 23, 77-88.
[54] Kuno, A., Uchiyama, N., Koseki-Kuno, S., Ebe, Y.*, et al.*, *Nat. Methods.* 2005, 2, 851-856.
[55] Zheng, T., Peelen, D. and Smith, L. M., *J. Am. Chem. Soc.* 2005, 127, 9982-9983.
[56] Angeloni, S., Ridet, J. L., Kusy, N., Gao, H.*, et al.*, *Glycobiology* 2005, 15, 31-41.
[57] Patwa, T. H., Zhao, J., Anderson, M. A., Simeone, D. M. and Lubman, D. M., *Anal. Chem.* 2006, 6411-6421.
[58] Holzman. T. and Kolker, E., *Curr. Opin. Biotechnol.* 2004, 15, 52-57.

[59] Kolker, E., Higdon, R. and Hogan, J. M., *Trends Microbiol.* 2006, 14, 229-235.
[60] Cox, B., Kislinger, T. and Emili, A., *Methods* 2005, 35, 303-314.

Chapter 2

Proteomic Analysis of Cold Adaptation in a Siberian Permafrost Bacterium – *Exiguobacterium Sibiricum* 255-15 by Two Dimensional Liquid Separation Coupled with Mass Spectrometry

2.1 Introduction

Low temperature is a predominant environmental characteristic of the majority of the earth's surface. Approximately 70% of the earth surface covered by oceans with an average temperature of 4°C while over 20% is occupied by permafrost. Often in these harsh environments, the only form of life is bacteria [1-5]. These bacteria have apparently developed various adaptive mechanisms that allow them to survive in such hostile environmental conditions through long-term evolutionary processes. Among such adaptive processes, not only the bacteria themselves might be affected by environmental low temperature and induced cold adapted features, but also the production of organic molecules within them, such as enzymes and proteins that sustain their metabolism. In discussion of bacterial low temperature adaptation, specific sets of cold shock proteins (Csps) and cold acclimation proteins (Caps) have been considered to facilitate and allow cell growth at low temperature. The term "Csps" is used here for proteins that are transiently over-expressed after an abrupt shift to a low temperature, and the term "Caps" is used for proteins synthesized at a greater level during continuous growth at low

temperature as compared with high temperature. The response to cold temperature has been extensively studied in mesophilic, psychrotrophic and psychrophilic bacteria such as *Escherichia coli* (*E. coli*) [6-18], *Bacillus subtilis* (*B. subtilis*) [19-30], *Bacillus cereus* [31], *Bacillus psychrophilus* [32], *Arthrobacter globiformis* SI55 [33, 34], *Enterococcus faecalis* [35], *Listeria monocytogenes* [36-38], *Mycobacterium tuberculosis* [39], *Pseudomonas fragi* [40-42], *Salmonella typhimurium* [43-45], *Streptomyces clavuligerus* [46] and *Trichosporon pullulans* [47]. The mechanism of bacterial response to cold temperature is still poorly understood, while this problem is of great importance with regard to ecological safety and the development of ecological biotechnologies in the temperate and subpolar climatic regions.

E. sibiricum 255-15 is a non-spore forming gram-positive bacterium of family bacillaceae and was first isolated from 2-3 million year old Siberian permafrost sediment [48]. This bacterium strain was found to grow well at -2.5°C and remain physiologically active down to at least -12°C [49, 50]. This strain was chosen for sequencing based on its excellent survival after a long-term freeze, rapid growth at low temperature and the age of the permafrost sediment from which it was originally isolated. *E. sibiricum* 255-15, together with other permafrost inhabiting bacteria, may serve as a model for low temperature exobiological niches on Mars and other planets and satellites in our solar system.

Bacterial cold adaptation at the protein level has been most successful when applied to small sets of proteins isolated in specific functional contexts. However, to be biologically useful, as opposed to simply highlighting analytical methods, large-scale proteomic studies are needed to underlie all cellular processes. No method or instrument

exists that is able to identify and quantify the components of complex protein samples in a simple, single-step operation. Rather, different components for separation, identification and quantification as well as tools for data analysis must be combined together [51]. We have developed a 2-D liquid separation based proteomic technology to compare protein profiling in human cells [52-58] as well as bacterial samples [59]. The entire method is performed in the liquid phase and, consequently, could be directly coupled with ESI-TOF-MS for accurate intact protein *M*r measurement and interlysate quantification. MS-based interlysate quantification provides high reproducibility compared with 2-DE in which poor reproducibility is always an issue for comparing sample-to-sample protein expression.

In the present work, the 2-D liquid phase separation coupled with ESI-TOF MS, MALDI-TOF MS and MALDI-QIT-TOF MS has been demonstrated to be a rapid and effective means to study cold adaptation in *E. sibiricum* 255-15 at the proteome-wide scale. The accurate *M*r from ESI-TOF MS that is not available in 2-DE separation, together with p*I*, provides an essential factor for protein identification in complex protein samples. As a result, over 500 proteins were resolved in the 2-D liquid separation and 256 of which were positively identified by PMF with MALDI-TOF MS and peptide sequencing by MS/MS using MALDI-QIT-TOF MS. Among these proteins 39 Caps were preferentially or uniquely expressed at the low temperature, which are probably one of the key determinants that allow life at low temperatures.

2.2 Materials and Methods

2.2.1 Chemicals

Bovine serum albumin (BSA), dithiothreitol (DTT), *n*-octyl-β-D-glucopyranoside (OG), glycerol, phenylmethanesulfonyl fluoride (PMSF), iminodiacetic acid, trifluoroacetic acid (TFA), ammonium bicarbonate (NH_4HCO_3), urea, thiourea, tris-(hydroxymethyl) aminomethane (Tris-base), 1,3 bis[tris(hydroxymethyl methylamino] propane (bis-tris propane), α-cyanohydroxycinnamic acid (CHCA), angiotensin I, angiotensin II, human adrenocorticotropic hormone (ACTH) fragment 1-17, ACTH fragment 18-39, bradykinin fragment 1-7, P14R, sodium chloride (NaCl), hydrochloride acid (37%), HPLC grade acetonitrile (ACN) and 2-isopropanol were purchased from Sigma-Aldrich (St. Louis, MO, USA). N-tosyl-L-phenylalanine chloromethyl ketone (TPCK)-treated trypsin was supplied from Promega (Madison, WI, USA). Formic acid (98-100%) was purchased from Riedel de Haen (Seelze, Germany). Tryptic soy broth (Difco) and yeast extract (Difco) were obtained from Fisher Scientific. Polybuffer 74 and 96 were purchased from Amersham Pharmacia Biotech (Piscataway, NJ). 2,5 dihydroxy benzoic acid (DHB) was purchased from LaserBio Labs (Sophia-Antipolis, Cedex, France). The deionized water was obtained using a Milli-Q water filtration system (Millipore, Inc., Bedford, MA).

2.2.2 Cell Culture and Sample Preparation

The *E. sibiricum* 255-15 cell pellets described in this study were obtained from the Department of the Food Science at North Carolina State University. All cells were cultured in tryptic soy broth with 7% yeast extract to mid-log phase (OD = 0.7). The cells were washed with 50 mM Tris-HCl (pH 7.5) for three times before they were frozen at -80°C. The cell pellets were shipped on dry ice.

A 0.2-0.5 g cell pellets were resuspended in 5 mL lysis buffer (pH 7.3) containing 50 mM Tris-HCl, 6 M urea, 2 M thiourea, 0.1 M DTT, 2% OG and 2 mM PMSF and then vortexed for 2 min. After being sonicated for 10 min, cell walls and other insoluble debris were removed by centrifugation at 40, 000 rpm for 20 min at 4 °C (Model L-70 Ultracentrifuge, 80 Ti Rotor, Beckman-Coulter, Fullerton, CA, USA). The resulting whole cell lysates were desalted using a PD-10 Sephadex G-25 gel filtration column (Amersham Biosciences, Piscataway, NJ, USA) and the protein concentration was determined using the Bradford-based protein assay (Bio-Rad, Hercules, CA, USA) with BSA as a standard.

2.2.3 Chromatofocusing

A Beckman Gold HPLC equipped with model 127S pump and model 166 detector (Beckman Coulter, Fullerton, CA, USA) was used in the CF separation. The separation was performed on an HPCF-1D column (250 × 2.1 mm) (Eprogen, Inc., Darien, IL) on which a linear pH gradient was generated using a start buffer (SB) and elute buffer (EB). The upper limit of the pH gradient was determined by the SB (25 mM bis-tris propane, pH 8.5) and the lower limit was determined by the EB (3% v/v poly-buffer 74, 7% v/v poly-buffer 96, pH 4.0). Both buffers were prepared in 6 M urea and 0.1% OG with the pH adjusted using iminodiacetic acid. A total of 5 mg extracted proteins were loaded on the CF column that was equilibrated at 0.2 mL/min with the SB for two hours followed by elution with the EB at the same flow rate. After the pH gradient was finished, the column was washed with 1 M NaCl solution to elute the proteins with the p*I* values lower than 4.0. The pH was detected online by an off-column pH electrode (Lazar Research Laboratories, Inc. Los Angeles, CA) and the separation was monitored at 280 nm.

Effluent from the CF separation was collected from pH 8.5 to 4.0 every 0.3 pH unit intervals. All fractions were then stored at -80°C for NPS-RP-HPLC separation.

2.2.4 NPS-RP-HPLC Online Coupled with ESI-TOF MS

Using the same HPLC system, the NPS-RP-HPLC separation was performed at a flow rate of 0.5 mL/min on a 33 × 4.6 mm ODS III column packed with 1.5 μm nonporous silica beads (Eprogen, Inc.). The column temperature was maintained at 60°C using a Model 7971 column heater (Jones Chromatography, Resolution Systems, Holland, MI, USA) to improve the resolution and speed of the separation. The separation was performed using a water (solvent A) and acetonitrile (solvent B) gradient both of which contained 0.1% v/v TFA. The gradient profile used was as follows: (1) 5 to 26% B in 1 min; (2) 26 to 35% B in 3.5min; (3) 35 to 40% B in 9.5 min; (4) 40 to 50% B in 13 min; (5) 50 to 58% B in 4 min; (6) 58 to 75% B in 1 min; (7) 75-100% B in 1 min; (8) 100 to 5% B in 1 min.

A post-column splitter was used after NPS-RP-HPLC so that half of the effluent was collected for further MS based identification and the other half was directly online monitored by the ESI-TOF MS (LCT, Micromass, Manchester, U.K.). To improve ionization efficiency, a 10% formic acid solution was continuously infused into the splitter at 5 μL/min by a syringe pump to be mixed with the effluent before being delivered into the ionization source. The LCT parameters were set as follows: capillary voltage at 3200 V, sample cone at 45 V, extraction cone at 3 V, reflection lens voltage at 750 V, desolvation temperature at 250°C, source temperature at 120°C, desolvation gas at 600 L/h and nebulizer gas flow at the maximum. ESI-TOF MS was externally calibrated by directly infusing NaI-CsI standard solution and internally normalized by adding 1μg

of insulin as a standard for the quantification of each protein based on its peak area. During the separation, one spectrum was acquired *per* second. The intact *M*r was obtained by deconvoluting the multiple-charged-ion umbrella of the ESI-TOF MS spectra into a MaxEnt spectrum on a real-mass scale by MassLynx 4.0 software (Micromass). Deconvolution was performed using a target mass range of 4-95 kDa, 1 Da resolution, 0.75 Da peak width and 65% peak height value.

2.2.5 Tryptic Digestion

The collected NPS-RP-HPLC fractions were first concentrated to 80 μL by a SpeedVac concentrator (Centrivap Concentrator, Labconco, Kansas City, Missouri, USA) to remove acetonitrile and TFA followed by the addition of 10 μL of 1 M NH_4HCO_3 and 10 μL of 100 mM DTT to neutralize the remaining TFA and to denature the protein. The mixtures were incubated at 37°C for 20 min, and then 0.5 μg TPCK-treated trypsin was added to start the digestion. The digestion was maintained at 37°C for 24 hours and terminated by adding 2 μL TFA. Before MS-based identification, tryptic digests were desalted and pre-concentrated in 5 μL of 60% acetonitrile with 0.1% TFA using 2 μm C18 Zip-Tips (Millipore, Inc.).

2.2.6 MALDI-TOF MS for Peptide Map Fingerprinting

In preparation for MALDI-TOF MS analysis, a saturated CHCA matrix solution in 60% acetonitrile with 0.1% TFA was freshly prepared. Angiotensin I ($[M+H]^+$ 1296.69), ACTH 1-17 ($[M+H]^+$ 2093.09) and ACTH 18-39 ($[M+H]^+$ 2465.20) were added as internal standards into 1:4 diluted matrix solution and 0.5 μL of the resulting

solution was spotted on a 96-spot MALDI plate (Micromass) followed by layering 0.5 μL peptide sample on top of the spot.

Peptide mass was measured on a Micromass TofSpec2E system (Micromass/Waters, Milford, MA, USA) equipped with a 337 nm nitrogen laser source and delayed extraction. The MALDI-TOF MS was operated in positive ion reflector mode and the parameters were set as follow: reflectron voltage at 25 kV, operating voltage at 20 kV, pulse voltage at 2 kV, delay time at 520 ns, reflectron ratio at 1:3, suppression setting at 500 V, and sampling rate at 2 GHz. The final MALDI spectrum was summed over the average of 100-150 spectra and calibrated with the three internal standards within 50 ppm. The peptide masses were obtained using MassLynx 4.0 over the range of 800-4000 Da and then submitted to MS-Fit (http://prospector.ucsf.edu/ucsfhtml4.0/msfit.htm) for protein identification. The NCBInr database was searched with a mass tolerance of 50 ppm and one missed cleavage set as fixed parameters.

2.2.7 MALDI-QIT-TOF MS/MS Peptide Sequencing

Mass spectrometric peptide fragmentation and sequencing was performed on an Axima MALDI-QIT-TOF mass spectrometer (Shimadzu Corporation, Kyoto, Japan and Kratos Analytical, Manchester, UK) equipped with a 337nm nitrogen laser source. MALDI-TOF MS/MS analysis was performed in the positive ion mode using an external calibration with a mixture of bradykinin fragment 1-7 ($[M+H]^+$ 757.40), angiotensin II ($[M+H]^+$ 1046.54), P14R ($[M+H]^+$ 1533.86) and ACTH fragment 18-39 ($[M+H]^+$ 2465.20). The matrix used in this case was DHB solution with a concentration of 10 mg/mL. Each MALDI-QIT-TOF profile resulted from the accumulation of five laser shots. The signal was further processed and analyzed by Lanchpad software (Kratos

Analytical Ltd., Manchester, UK). The parent ion mass and the resulting fragment ion masses were searched against the NCBInr database using Mascot 1.8 (Matrix Science, London, UK) setting a peptide tolerance of 1.2 Da, a MS/MS tolerance of 0.6 Da and one missed cleavage site as fixed parameters.

2.3. Results and Discussion

2.3.1 Protein Purification Using 2-D Liquid Phase Separation

Using the 2-D liquid phase separation method followed by ESI-TOF MS intact protein molecular mass measurement, 2-D mass maps as shown in Figure 2.1 were generated to visualize the protein profiling of *E. sibiricum* 255-15 at different growth temperatures. 2-D mass maps are analogous to 2-DE images but in a more advanced digitized format. The 18 lanes in the mass map represented 18 p*I* fractions from the first dimensional CF separation. The first three lanes were the NaCl wash fractions and the other lanes were the pH gradient fractions starting with pH 4.0-4.3 at lane 4 and ending up with pH 8.2-8.5 at lane 18. Approximately 500 proteins were detected by ESI-TOF MS with the majority of proteins eluted at the p*I* range of 4.0-6.1. The vertical axis showed the *M*r values of proteins in range from 5 kDa up to 95 kDa based on ESI-TOF MS analysis. A differential mass map was created by point-by-point subtraction to compare protein expression of different samples. The actual resolution, relative quantification and mass accuracy of specific protein bands were recovered from the original ESI-TOF MS data. Most proteins were similarly expressed at the two temperatures, which were referred to as housekeeping proteins. They are essential for the proper function of the bacterial cells during the cold adaptation process. In this study, we were most interested in the cold acclimation proteins that were preferentially or uniquely

expressed at the low temperature. Using a relative abundance threshold of two, differentially expressed proteins were then identified by subsequent peptide mapping using MALDI-TOF MS and peptide sequencing with MALDI-QIT-TOF MS.

2.3.2 Proteins Identification by PMF Using MALDI-TOF MS

Proteins eluted from HPLC were digested by trypsin and peptide mass maps were obtained using MALDI-TOF MS. The peptide masses were submitted to NCBInr by MS-Fit software for protein identification. The identification was based on the suggestion that at least five matched peptides with mass accuracy of 50 ppm and sequence coverage of at least 20% [60]. Although this approach has been proven to be a successful means in protein identification, in practice, however, it is often difficult to obtain protein identification largely due to an insufficient number of detected peptides. Peptides were often prevented from being observed by weak response of lysine-terminated peptides, internal fragmentation, non-specific and incomplete enzymatic digestion and modification. In table 1, the first 31 listed proteins were identified using the PMF method and had more than five theoretical tryptic peptides in the "working range" of MALDI-TOF MS with uracil phosphoribosyltransferase having the highest sequencing coverage of 83%. All identities were confirmed by intact *M*r values from online NPS-RP-HPLC/ESI-TOF MS experiments. In contrast, the last 11 proteins in the table were not identified with ≤ 4 peptides detected, which did not meet the requirement for protein identification by the PMF method and thus need to be sequenced by MS/MS for identification. Although these proteins could not be identified using the PMF method, peptide maps from MALDI-TOF MS provided very important information of potential interesting peptides for further sequencing using MALDI-QIT-TOF MS.

The intact *M*r is extremely important in terms of protein identification using peptide mapping and peptide sequencing. With the use of an accurate *M*r one can find the potential protein ID among several possibilities in protein database searching. Also, there are often isoforms and truncated forms of proteins and the *M*r provides confidence as to which form is being studied. In addition, there may be several proteins that are of similar *M*r and difficult to distinguish; however, with the use of accurate *M*r measurements one can readily identify the presence of these proteins [57]. The discrepancy between the experimental and the theoretical *M*r values in Table 1 may result from as yet unidentified post-translational modifications or protein truncation. There is a *M*r associated with each protein identified in this work.

2.3.3 Protein Identification by Peptide Sequencing Using MALDI-QIT-TOF MS

Another protein identification method is based on peptide sequencing using peptide collision-induced dissociation (CID) spectra. CID amide bond fragmentation along the length of peptide generates b- and y- ions, which reveal the peptide sequence. Protein identification is obtained by searching the protein database with the detected fragment ions in the experiment. The MALDI-QIT-TOF MS used in the work provides high sensitivity and resolution as well as the advantage of sharing the same sample preparation procedure with MALDI-TOF MS. Proteins analyzed by MALDI-TOF MS can be reanalyzed by MALDI-QIT-TOF MS after placing a new sample spot on the MALDI-QIT-TOF plate. The quality of MS/MS data and the effectiveness of using MALDI-QIT-TOF MS in protein identification have been discussed [61-63].

Protein identification by peptide sequencing is more reliable than that achieved by PMF and may allow the identification of proteins based on a single peptide fragmentation

sequence [51]. The major advantage of this method is that the small proteins that are difficult to identify by PMF due to an insufficient number of detected peptides may be identified using peptide sequencing if the peptide signal and its fragment signals are of sufficient intensity. In the case of 7.150 kDa cold shock protein, only one peptide with *M*r of 1098 was detected by MALDI-TOF MS. The peptide was then subjected to MALDI-QIT-TOF sequencing and the resulting CID spectrum in Figure 2.2 showed b- and y-ions with rich sequence information. When the parent ion and product ions were used to search against NCBInr using MASCOT software, 7.150 kDa cold shock protein was identified with the corresponding peptide sequence comprising amino acid 40-56. Its identity was confirmed by the accurate *M*r measurement in Figure 2.3 in which two of the multiple-charged-ion umbrella of ESI-TOF spectra were as shown in Figure 2.3 (A) and (B) and the intact *M*r values obtained by deconvoluting Figure 2.3 (A) and 3 (B) into a MaxEnt spectrum on a real mass scale were shown in Figure 2.3(C) and 3 (D). In this case, proteins of 7.150 kDa and 7.414 kDa co-eluted at both 4ºC and 25ºC while a protein of 7.444 kDa was only expressed at 4ºC. In order to identify these proteins, the tryptic digests of the RPLC fraction that contain all these proteins were peptide mapped by MALDI-TOF MS (Figure 2.4), and each of the detected peptides was then subjected to MALDI-QIT sequencing. Two possible homologous cold shock proteins with the same molecular mass of 7.409 kDa but slightly different pH values of 4.55 and 4.41 were detected by MALDI-QIT-TOF sequencing when peptide 1937 (SLDEGQEVSFEVEEGQR) was submitted for CID fragmentation as show in Figure 2.5. The identities of both proteins were confirmed with the fragmentation of peptide 2098 (ESGDDVFVHFSAIQTDGFK) (Figure 2.6) and 2125 (ENGDDVFVHFSAIQ TDGFK)

(Figure 2.7), which were also reconfirmed by the accurate *M*r measurement shown in Figure 2.3. Using peptide sequencing by MALDI-QIT-TOF MS, the proteins not identified by PMF in table 1 were all identified.

2.3.4 Homologous Cold Shock Proteins (Csps)

The three cold shock proteins discussed above have similar *M*r, p*I* and amino acid sequence with 65.15, 66.67 and 59.09% overlap with the sequence of major cold shock protein A (CspA*)* in *E. coli* and over 74% when compared to *CspB*, *CspC* and *CspD* in *B. subtilis* (Table 2), which suggests that these three cold shock proteins are the major homologous Csps in *E. sibiricum* 255-15. An interesting finding in this study is that, unlike in *E. coli* and *B. subtilis*, the three Csp family homologous proteins were found similarly expressed at 25°C and 4°C based on ESI-TOF MS analysis. In fact, the three Csps represent about 10% of the total soluble proteins in cells grown at both 4°Cand 25°C. This result suggests that the genes for these proteins are turned on continuously to produce proteins to protect the cells from cold damage. Such behavior has been observed in other extremophiles such as psychrobacter where it has been shown in gene expression array studies that certain genes are always expressed. Apparently, these organisms which survive for long periods of time under cold conditions have adapted such continuous expression as a means of survival [64].

2.3.5 Cold Acclimation Proteins (Caps)

2.3.5.1 Caps Characterized as Csps in *E. coli*

From peptide mapping and peptide sequencing analysis, 39 identified proteins with *M*r ranging from 7 to 95 kDa presented an increased level of synthesis at the lower

temperature and were considered to be Caps, 16 of which were not detected at 25°C. Some of these Caps were characterized as Csps in *E. coli*, such as trigger factor (TF) and pyruvate dehydrogenase. TF is a molecular chaperone with prolyl-isomerase activity that may be a bottleneck in the folding of some polypeptides at low temperature. TF can be cross-linked to the nascent polypeptide chain on the ribosome and binds to the GroEL chaperone. This TF enhances the affinity of GroEL toward unfolding proteins and activates the degradation of some polypeptides [65-67]. Unlike the trigger factor, the role of pyruvate dehydrogenase has not yet been well understood. Presumably, they are involved in the intensification of glycolysis and the suppression of the tricarboxylic acid cycle, *i.e.* in the processes that are observed upon the retardation of cell growth and the adaptation of cells to stresses [12, 68, 69].

2.3.5.2 Caps Characterized as Other Stress-induced Proteins

Like in *E. coli*, the over-expression of heat shock protein 70 (Hsp70) molecular chaperons was also observed in *E. sibiricum* 255-15 during the cold adaptation process. Most heat shock proteins may function as molecular chaperones that play an important role in protein folding. It seems that folding of exiguobacterial proteins at near-freezing temperatures is also jeopardized and, therefore, Hsp70 and other heat shock proteins are required. Thus, these so-called “heat shock proteins” are not simply heat shock-specific proteins. They should more appropriately be called “temperature-stress proteins”. Hsp molecular chaperones were found to be actively synthesized in response to heat, cold and chemical stress[70]. Based on the 2-D mass map analysis, phage shock protein A (PspA) was the most differentially expressed protein at different growth temperatures whose expression ratio was over 70. Presently, the exact function of PspA remains elusive.

Various stress conditions, including severe heat shock, cold shock, osmotic shock and exposure to ethanol [71-73], which is involved in protein translation across the cytoplasmic membrane, have been reported to induce expression of the Psp operon. These stress conditions might all lead to the dissipation of the proton-motive force, and it has been demonstrated that expression of the Psp operon, and more specifically of PspA, helps the cells to maintain the proton-motive force under such stress conditions [74]. Penicillin tolerance protein was also found greatly over-expressed at 4°C. These facts suggest that a single stress could induce other stress-induced proteins that are organized in a complex and highly sophisticated adaptation network.

2.3.5.3 Enzymes

E. sibiricum 255-15 is able to grow efficiently at near freezing temperatures. Clearly, this organism has found mechanisms of temperature compensation in order to cope with the reduction of chemical reaction rates induced by low temperatures. Synthesizing more enzymes and synthesizing cold-efficient enzymes are two possible means for adaptation [75]. Adjustment in enzyme concentration has been reported during cold adaptation. For instance, increased expression of polynucleotide phosphorylase has been detected in *E. coli* at low temperatures [76]. Another mechanism for survival is to possess enzymes with temperature-independent reaction rates. This is the case of perfectly evolved enzymes, however such enzymes are relatively rare: typical examples are carbonic anhydrase, acetylcholinesterase and triosephosphate isomerase. In this study, the higher level of triosephosphate isomerase has been detected in a cold adapted population of *E. sibiricum* 255-15. Perfectly evolved enzymes could be extremely useful to probe the various hypotheses related to enzyme adaptation because, apparently, they

do not need to be adapted to low temperatures from a kinetic point of view. In this study, 28 out of 39 identified Caps are enzymes. Although the roles of these enzymes concerning how they interact together to maintain the adequate metabolic fluxes need to be further studied, it is possible that these enzymes maintain the bacterial metabolism enabling the cells to adapt to cold temperatures.

2.3.5.4 Possible Post-translational Modification

From ESI-TOF MS measurement in Figure 2.3, a protein of 7.442 kDa co-eluted with the three homologous Csps at 4°C, which suggests that this 7.442 kDa protein might play an important role in the physiology of the cells during cold adaptation. From the peptide map by MALDI-TOF MS analysis in Figure 2.4, there were no detected peptides that could be assigned to a protein other than the three co-eluted cold shock proteins. Based on the mass difference of this protein and the three identified cold shock proteins, this protein may be another homologous cold shock protein with the above three or be a post-translationally modified cold shock protein that was induced during the cold adaptation process.

2.4 Concluding Remarks

In proteomic level, 2-D liquid separation coupled with ESI-TOF MS, MALDI-TOF MS and MALDI-QIT-TOF MS provides rapid, accurate and reproducible protein fractionation and identification in the study of cold adaptation in *E. sibiricum* 255-15. The accurate *M*r of protein from ESI-TOF MS, together with the p*I* from CF, is essential for protein identification and characterization. The results from this study indicated that the adaptive nature of *E. sibiricum* 255-15 at near freezing temperatures could be regulated by cellular physiological processes through the regulation of certain cellular

proteins. Although the cold adaptation is still far from being properly understood, it is possible that new and increasingly synthesized proteins at the low temperature may support temperature homeostasis and enable the cells to adapt to the near or below freezing temperatures. Chaperone proteins could be involved in protection of proteins from denaturation and damage when *E. sibiricum* 255-15 cultures were exposed to cold temperature. Moreover, the unique nature of these proteins, differing from previously reported Csps, also warrants further enquiry.

Table 2.1 Caps identified in *E. sibiricum* 255-15.

Protein ID (Access. No.)	Experimental *Mr*	Experimental. *pI*	MOWSE Score	Masses Matched	% Coverage	Theoretical *Mr*/*pI*	Expression Ratio
Enzyme related to GTP cyclohydrolase I (45531353)	19096	6.1-5.8	6.88E+04	8	57	19095/5.2	3.803
2'-5' RNA ligase (46113530)	19752	5.8-5.5	1.76E+06	10	74	19743/5.4	2.986
F0F1-type ATP synthase, subunit b (45531034)	20237	4.3-4.0	4.38E+05	11	47	20367/4.7	N/A
Uncharacterized conserved protein (46113851)	21810	6.7-6.4	7.95E+04	8	47	21941/6.4	48.83
Uracil phosphoribosyltransferase (53771257)	22910	5.8-5.5	7.76E+09	17	83	23043/5.5	6.064
Hypothetical protein (45532138)	23572	5.8-5.5	1.53E+05	6	48	23572/5.4	2.503
Phage shock protein A (IM30), suppresss sigma 54-dependent transcription (45532491)	24464	4.9-4.6	3.87E+06	18	53	24460/4.9	74.09
Sporulation control protein (45530919)	28742	5.5-5.2	2.20E+04	6	30	28739/5.0	N/A
Hydroxymethylpyrimidine/phosphomethylpyrimidine kinase (46113160)	28843	6.1-5.8	6.42E+03	6	29	28970/5.6	N/A
Translation elongation factor Ts (45532587)	31893	5.5-5.2	8.62E+05	8	39	31890/5.0	N/A
ABC-type uncharacterized transport system, ATPase component (45530831)	33396	5.8-5.5	7.95E+05	12	52	33397/5.6	5.336
NADPH:quinone reductase and related Zn-dependent oxidoreductases (53771726)	33826	5.2-4.6	1.60E+04	8	40	33826/4.8	N/A
Fructose-1,6-bisphosphatase/sedoheptulose 1,7-bisphosphatase and related proteins (53771404)	33886	5.5-5.2	9.66E+08	15	61	33886/5.0	7.187
Predicted HD-superfamily hydrolase (53771602)	34626	6.1-5.8	1.45E+06	8	32	34621/5.9	25.68
Penicillin tolerance protein (53771372)	34706	6.1-5.8	1.23E+06	7	38	34703/5.7	31.35
Protoheme ferro-lyase (ferrochelatase) (53771592)	34788	5.2-4.9	5.26E+06	9	35	34789/4.9	N/A
Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase (46113193)	35971	5.8-5.2	4.72E+07	13	33	36105/5.4	3.126
Transcriptional regulators (46112980)	37045	5.8-5.2	5.31E+04	8	21	37047/5.4	5.082
Cellulase M and related proteins (46113630)	39295	5.8-5.2	4.70E+04	8	36	39298/5.3	28.79
Alanine dehydrogenase (53771421)	39670	5.8-5.2	2.62E+05	6	26	39672/5.2	N/A
Pyruvate/2-oxoglutarate dehydrogenase complex, dehydrogenase (E1) component, eukaryotic type, alpha subunit (46113512)	40141	5.2-4.6	8.08E+05	10	31	40137/5.1	2.704
3-oxoacyl-(acyl-carrier-protein) synthase (46113125)	44108	5.8-5.5	6.47E+07	9	37	44111/5.3	N/A
Uncharacterized conserved protein (46113892)	45624	5.8-5.5	5.11E+08	13	44	45756/5.4	N/A
Glycine/serine hydroxymethyltransferase (46113662)	45723	6.1-5.8	8.23E+07	16	41	45716/5.7	2.279
Glutamate-1-semialdehyde aminotransferase (46113490)	46903	5.8-5.5	2.84E+05	8	33	47038/5.4	N/A
Predicted GTPase (46113059)	47472	5.2-4.9	3.61E+04	11	37	47488/5.2	N/A
FKBP-type peptidyl-prolyl cis-trans isomerase (trigger factor) (46113050)	48029	4.3-4.0	3.98E+06	9	29	48026/4.3	4.094
Thiamine pyrophosphate-requiring enzymes [acetolactate synthase, pyruvate dehydrogenase (cytochrome), glyoxylate carboligase, phosphonopyruvate decarboxylase] (46113116)	61994	5.2-4.9	1.11E+10	19	46	61988/5.2	3.082
Hsp70 Molecular chaperone (45531750)	64781	< 4.0	1.20E+04	7	25	64904/4.6	2.815
ATPases with chaperone activity, ATP-binding subunit (46113174)	90678	5.8-5.5	9.70E+09	18	34	90773/5.8	2.384
Uncharacterized protein conserved in bacteria (46112941)	13364	5.8-5.5	8.55E+03	3	54	13332/5.3	2.941
Chromosome segregation ATPase (45532021)	19304	5.5-5.2	6.23E+02	3	24	19301/5.0	N/A

DNA-directed RNA polymerase specialized sigma subunit, sigma24 homolog (46113445)	21365	5.5-5.2	7.26E+02	4	36	21363/5.4	N/A
Dehydrogenases with different specificities (46113222)	23028	5.8-5.5	1.39E+02	4	17	23028/5.3	23.72
Alpha-acetolactate decarboxylase (45531272)	26705	5.2-4.9	1.49E+03	4	31	26704/5.0	28.02
Triosephosphate isomerase (46113191)	26825	5.2-4.9	1.49E+03	4	25	26827/4.9	3.171
Lysophospholipase (53771292)	27078	5.8-5.5	1.79E+03	4	26	27079/5.9	N/A
Predicted sugar phosphate isomerase (46114194)	31763	5.5-5.2	6.65E+02	4	20	31767/5.2	N/A
Uncharacterized conserved protein (46114036)	33644	4.9-4.6	3.88E+02	4	20	33640/4.8	N/A

a), Protein quantification was based on its peak area from ESI-TOF MS analysis. Expression ratios were obtained by dividing the amount of protein expressed at 4°C with the amount of the same protein expressed at 25°C.
b), The expression ratios of the proteins that were exclusively expressed at 4°C were not available (N/A) in the table.

Table 2.2 Sequence alignment of the homologous Csps in *E. sibiricum* 255-15.

CspA* in *E. coli	MTQGTVKWFN SEKGFGFISS ETGTDVFAHF SEIKVDGFKT LEEGQKVTFD IQDGQRGPQA TNINLVK	Identity (%)
7.409 kDa Csp (pH 4.55)	MEQGKVKWFN AEKGFGFIER ESGDDVFVHF SAIQTDGFKS LDEGQEVSFE VEEGQRGPQA TNVTKL	65.15
7.409 kDa Csp (pH 4.41)	MEQGTVKWFN AEKGFGFIER ENGDDVFVHF SAIQTDGFKS LDEGQEVSFE VEEGQRGPQA TNVTKL	66.67
7.150 kDa Csp (pH 4.54)	MNTGKVKWFN AEKGFGFIEV EGGEDVFVHF SAITGEGFKS LDEGQEVEFE ITEGARGAQA ANVVKL	59.09
CspB* in *B. subtilis	MLEGKVKWFN SEKGFGFIEV EGQDDVFVHF SAIQGEGFKT LEEGQAVSFE IVEGNRGPQA ANVTKEA	Identity (%)
7.409 kDa Csp (pH 4.55)	MEQGKVKWFN AEKGFGFIER ESGDDVFVHF SAIQTDGFKS LDEGQEVSFE VEEGQRGPQA TNVTKL	75.76
7.409 kDa Csp (pH 4.41)	MEQGTVKWFN AEKGFGFIER ENGDDVFVHF SAIQTDGFKS LDEGQEVSFE VEEGQRGPQA TNVTKL	74.24
7.150 kDa Csp (pH 4.54)	MNTGKVKWFN AEKGFGFIEV EGGEDVFVHF SAITGEGFKS LDEGQEVEFE ITEGARGAQA ANVVKL	77.27
CspC in B. subtilis	MEQGTVKWFN AEKGFGFIER ENGDDVFVHF SAIQSDGFKS LDEGQKVSFD VEQGARGAQA ANVQKA	Identity (%)
7.409 kDa Csp (pH 4.55)	MEQGKVKWFN AEKGFGFIER ESGDDVFVHF SAIQTDGFKS LDEGQEVSFE VEEGQRGPQA TNVTKL	83.33
7.409 kDa Csp (pH 4.41)	MEQGTVKWFN AEKGFGFIER ENGDDVFVHF SAIQTDGFKS LDEGQEVSFE VEEGQRGPQA TNVTKL	86.38
7.150 kDa Csp (pH 4.54)	MNTGKVKWFN AEKGFGFIEV EGGEDVFVHF SAITGEGFKS LDEGQEVEFE ITEGARGAQA ANVVKL	74.24
CspD in B.subtilis	MQNGKVKWFN NEKGFGFIEV EGGDDVFVHF TAIEGDGYKS LEEGQEVSFE IVEGNRGPQA SNVVKL	Identity (%)
7.409 kDa Csp (pH 4.55)	MEQGKVKWFN AEKGFGFIER ESGDDVFVHF SAIQTDGFKS LDEGQEVSFE VEEGQRGPQA TNVTKL	78.79
7.409 kDa Csp (pH 4.41)	MEQGTVKWFN AEKGFGFIER ENGDDVFVHF SAIQTDGFKS LDEGQEVSFE VEEGQRGPQA TNVTKL	77.27
7.150 kDa Csp (pH 4.54)	MNTGKVKWFN AEKGFGFIEV EGGEDVFVHF SAITGEGFKS LDEGQEVEFE ITEGARGAQA ANVVKL	78.79

a), *CspA* in *E.coli*, *CspB,CspC and CspD* in *B. subtilis* are the reference molecules for this sequence alignment. b), The identical residues are underlied.

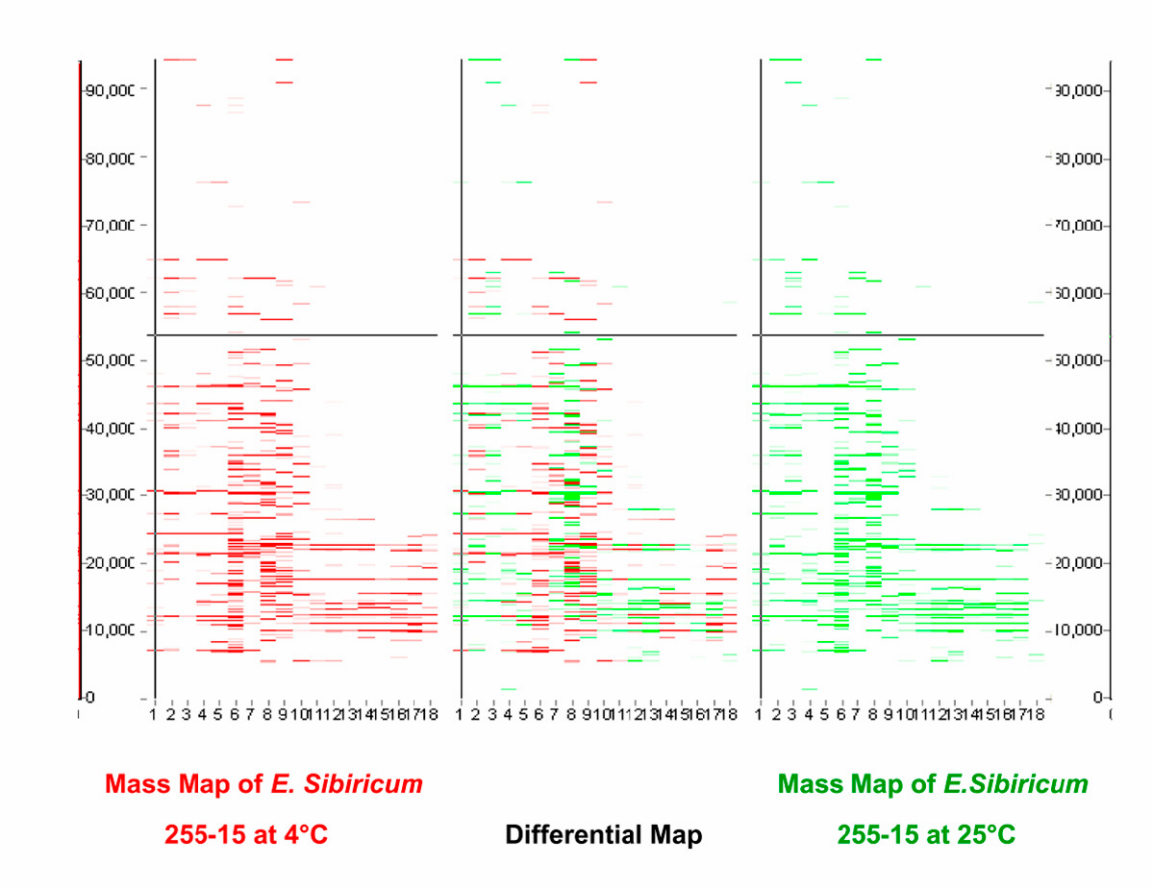

Figure 2.1 Comparison of 2-D mass maps between *E. sibiricum* 255-15 grown at 4ºC and 25ºC. 2-D mass map of *E. sibiricum* 255-15 grown at 4ºC is shown in shades of red on the left while protein expression of *E. sibiricum* 255-15 grown at 25ºC is shown in shades of green on the right. The vertical axis shows the *M*r values of proteins. The 18 lanes represent the 18 p*I* fractions from CF separation. A differential mass map is shown in the middle in which the difference in protein expression between two growth temperatures is quantitatively plotted in shades of green and red.

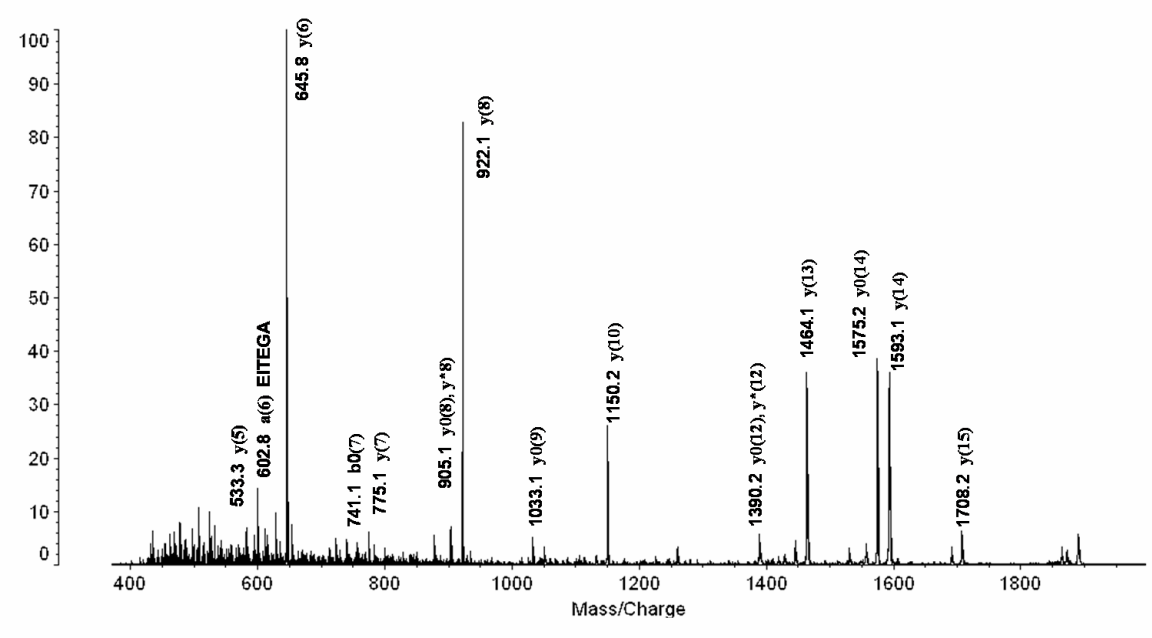

Figure 2.2 MALDI-QIT-TOF peptide sequencing of 1098.7 (SLDEGQEVEFEITEGAR) from Csp (7.150 kDa, pH 4.54). In this case, y*(i)=y(i)–NH_3, y0(i)=y(i)–H_2O and b0(i)=b(i)–H_2O.

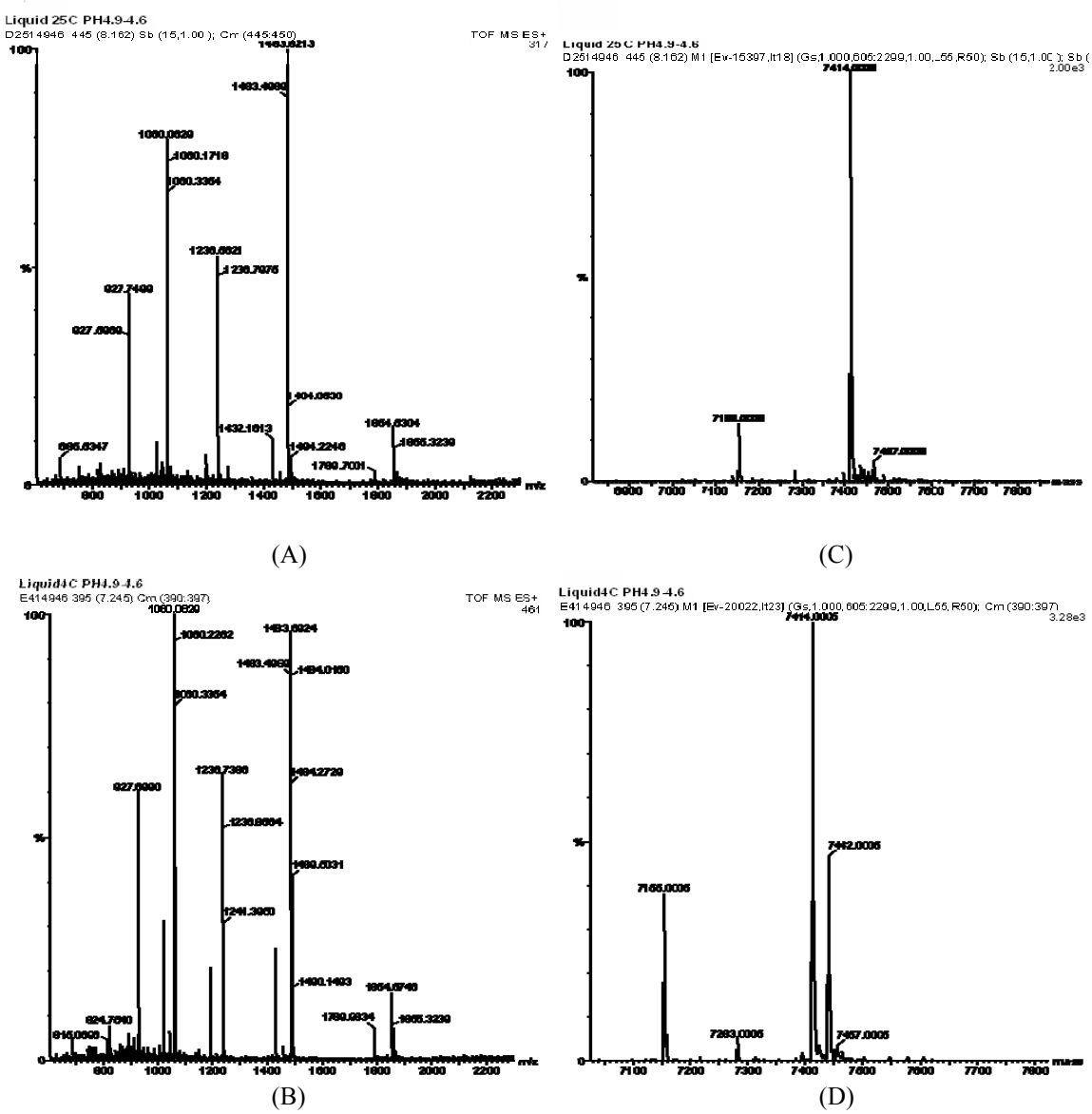

Figure 2.3 ESI-TOF spectra of NPS-RP-HPLC fraction containing 7.150 kDa and 7.414 kDa Csps at 25°C (A) and its corresponding HPLC fraction at 4°C (B). Their deconvoluted spectra are shown in (C) and (D). A protein of 7.442 kDa was exclusively expressed at 4°C as shown in (D).

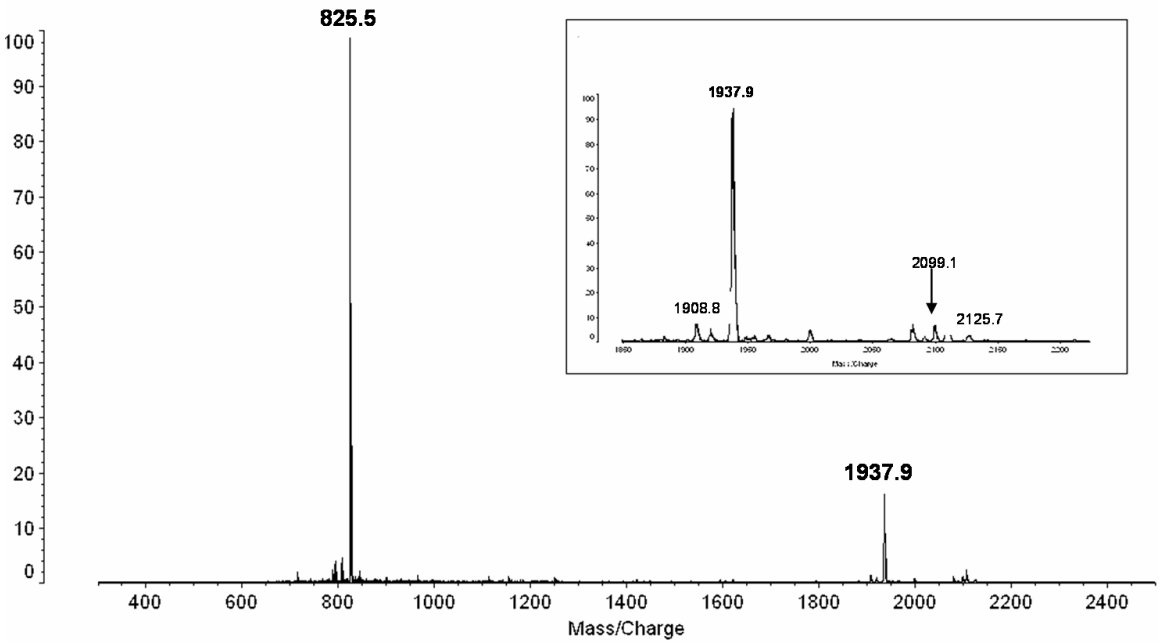

Figure 2.4 MALDI-TOF peptide mapping of NPS-RP-HPLC fraction containing proteins of 7.150 kDa, 7.414 kDa and 7.444 kDa.

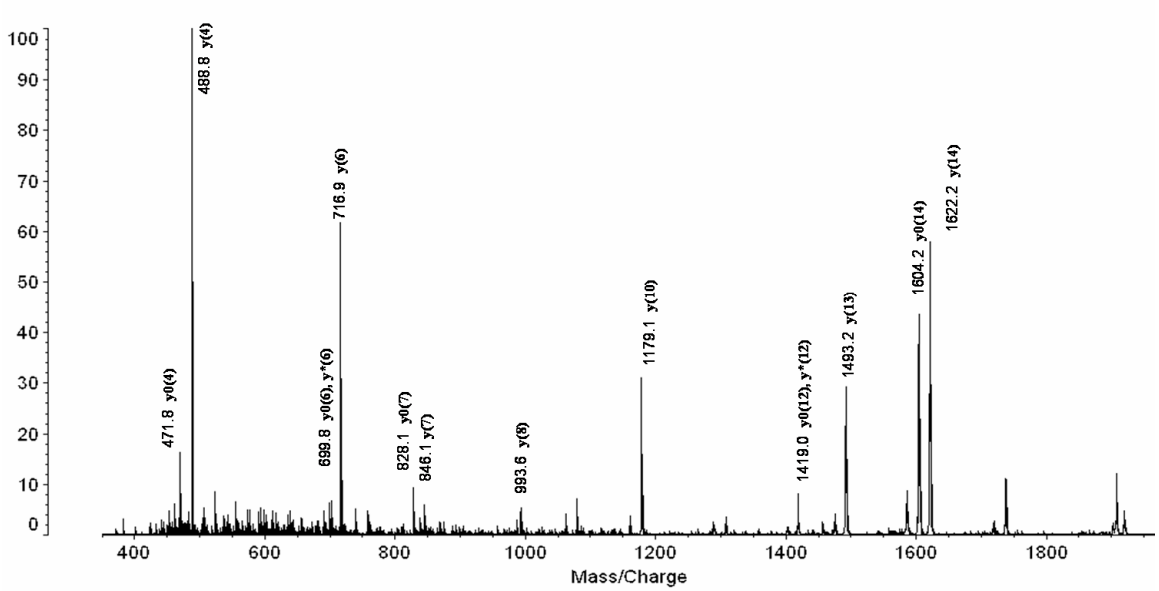

Figure 2.5 MALDI-QIT-TOF peptide sequencing of 1937.9 (SLDEGQEVSFEVEEGQR) from homologous Csps of 7.409 kDa (pH 4.55) and 7.409kDa (pH 4.41).

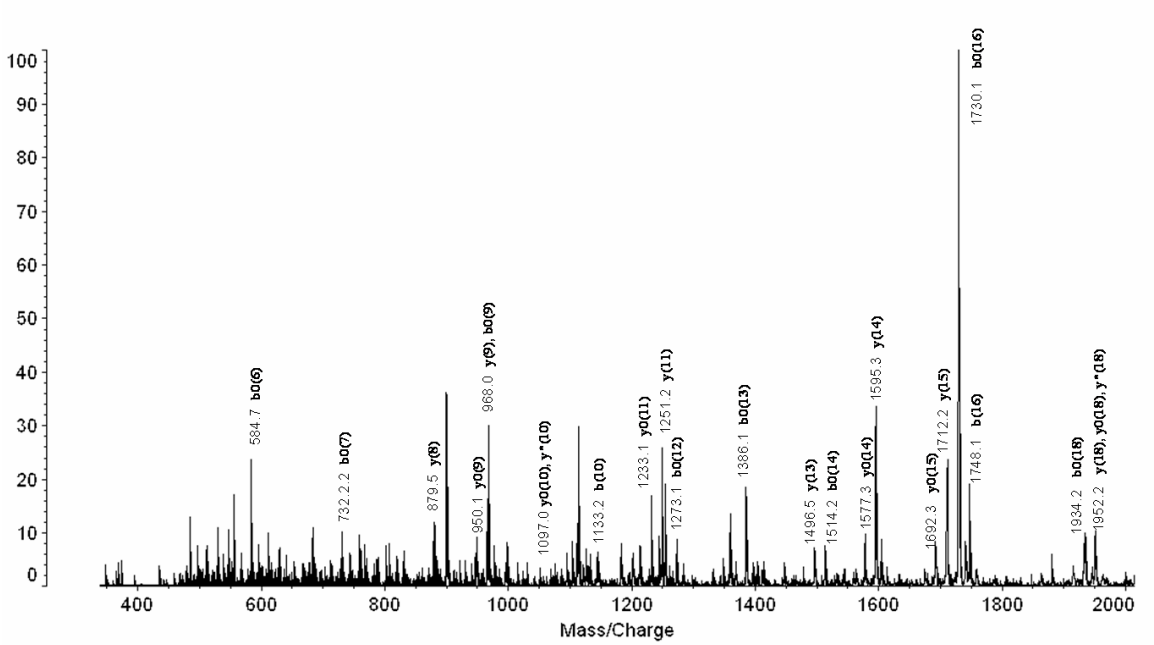

Figure 2.6 MALDI-QIT-TOF peptide sequencing of 2098.0 (ESGDDVFVHFSAIQTDGFK) from Csp (7.409 kDa, pH 4.55).

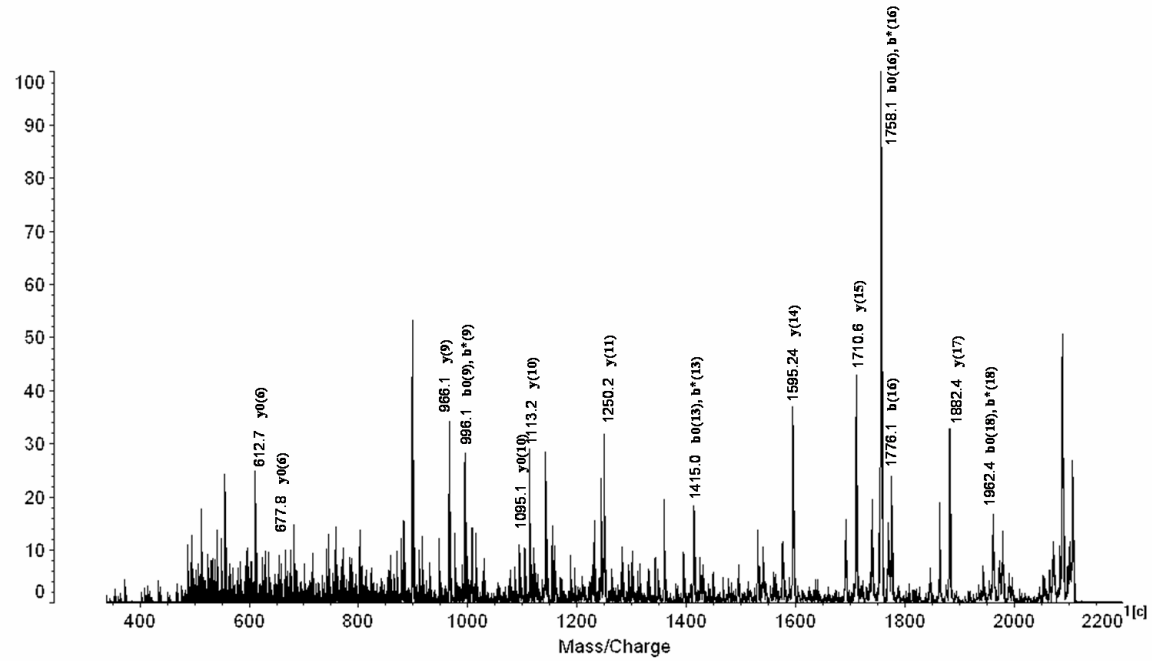

Figure 2.7 MALDI-QIT-TOF peptide sequencing of 2125.1 (ENGDDVFVHFSAIQTDGFK) from Csp (7.409 kDa, pH 4.41). In this case, b*(i)=b(i)–NH_3.

2.5 References:

[1] Gilichinsky, D., in: Bitton, G. (Ed.), *Encyclopedia of Environmental Microbiology*, Wiley, New York 2002, pp. 932-956.

[2] Abyzov, S., in: Friedmann, E. I. (Ed.), *Antarctic Microbiology*, Wiley, New York 1993, pp. 265-297.

[3] Cano, R. and Borucki, M., *Science* 1995, 268, 1060-1064.

[4] Lipman, C., *Science* 1928, 68, 272-273.

[5] Vreeland, R., Rosenzweig, W. and Powers, D., *Nature* 2000, 407, 897-900.

[6] Jones, P. G., VanBogelen, R. A. and Neidhardt, F. C., *J. Bacteriol.* 1987, 169, 2092-2095.

[7] Goldstein, J., Pollitt, N. S. and Inouye, M., *Proc. Natl. Acad. Sci USA* 1990, 87, 283-287.

[8] Lee, S. J., Xie., A., Jiang, W., Etchegaray, J. P*., et al.*, *Mol. Microbiol.* 1994, 11, 833-839.

[9] Nakashima, K., Kanamaru, K., Mizuno, T. and Horikoshi, K., *J. Bacteriol.* 1996, 178, 2994-2997.

[10] Wang, N., Yamanaka, K. and Inouye, M., *J. Bacteriol.* 1999, 181, 1603-1609.

[11] Lelivelt, M. J. and Kawula, T. H., *J.Bacteriol.* 1995, 177, 4900-4907.

[12] Graumann, P. and Marahiel, M. A., *Arch. Microbiol.* 1996, 166, 293-300.

[13] Jones, P. G., Krah, R., Tafuri, S. R. and Wolffe, A. P., *J. Bacteriol.* 1992, 174, 5798-5802.

[14] Yamanaka, K., Fang, L. and Inouye, M., *Mol. Microbiol.* 1998, 27, 247-255.

[15] Jiang, W., Hou, Y. and Inouye, M., *J. Biol. Chem.* 1997, 272, 196-202.

[16] Kandror, O. and Goldberg, A. L., *Proc. Natl. Acad. Sci USA* 1997, 94, 4978-4981.

[17] Brandi, A., Pietroni, P., Gualerzi, C. O. and Pon, C. L., *Mol. Microbiol.* 1996, 19, 231-240.

[18] Goldberg, D., Azar, I. and Oppenheim, A. B., *Mol. Microbiol.* 1996, 19, 241-248.

[19] Aguilar, P. S., Cronan, J. E. and de Mendoza, D., *J. Bacteriol.* 1998, 180, 2194-2200.

[20] Graumann, P. L. and Marahiel, M. A., *Arch. Microbiol.* 1999, 171, 135-138.

[21] Schroder, K., Zuber, P., Willimsky, G., Wagner, B. and Marahiel, M. A., *Gene* 1993, 136, 277-280.

[22] Willimsky, G., Bang, H., Fischer, G. and Marahiel, M. A., *J. Bacteriol.* 1992, 174, 6326-6335.

[23] Graumann, P., Schroder, K., Schmid, R. and Marahiel, M. A., *J. Bacteriol.* 1996, 178, 4611-4619.

[24] Graumann, P. and Marahiel, M. A., *FEBS Lett.* 1994, 338, 157-160.

[25] Schnuchel A., Wiltscheck R., Czisch M., Herrler M*., et al.*, *Nature* 1993, 364, 169-171.

[26] Schindelin H., Marahiel M.A. and Heinemann U., *Nature* 1993, 364, 164-168.

[27] Schindler T., Perl D., Graumann, P., Sieber V*., et al.*, *Proteins* 1998, 30, 401-406.

[28] Graumann P., Wendrich T.M., Weber M.H., Schroder K. and Marahiel M.A., *Mol. Microbiol.* 1997, 25, 741-756.

[29] Graumann P. and Marahiel M.A., *Mol.Gen Genet* 1997, 253, 745-752.

[30] Schindler T., Graumann P.L., Perl D., Ma S*., et al.*, *J. Biol. Chem.* 1999, 274, 3407-3413.

[31] Mayr, B., kaplan, T., Lechner, S. and Scherer, S., *J. Bacteriol.* 1996, 178, 2916-2925.

[32] Whyte, L. G. and Inniss, W. E., *Can. J. Microbiol.* 1992, 38, 1281-1285.

[33] Berger, F., Normand, P. and Potier, P., *J. Bacteriol.* 1997, 179, 5670-5676.

[34] Berger, F., Morellet, N., Menu, F. and Potier, P., *J. Bacteriol.* 1996, 178, 2999-3007.

[35] Panoff, J. M., Corroler, D., Thammavongs, B. and Boutibonnes, P., *J. Bacteriol.* 1997, 179, 4451-4454.

[36] Baylee D.O., Annous B.A. and Wilkinson B.J., *Appl.Environ. Microbiol.* 1996, 62, 1116-1119.

[37] Hebraud M. and Guzzo J., *FEMS Mcirobiol. Lett* 2000, 190, 29-34.

[38] Wemekamp-Kamphuis H.H., Karatzas A.K., Wouters J.A. and Abee T., *Appl.Environ. Microbiol.* 2002, 62, 456-463.

[39] Manganelli, R., Dubanu, E., Tyagi, S., Kramer, F. R. and Smith, I., *Mol. Microbiol.* 1999, 31, 715-724.

[40] Hebraud, M., Dubois, E., Potier, P. and Labadie, J., *J. Bacteriol.* 1994, 176, 4017-4024.

[41] Michel, V., Lethoux, I., Depret, G., Anglade, P*., et al.*, *J. Bacteriol.* 1997, 179, 7331-7342.

[42] Michel, V., Labadie, J. and Hebraud, M., *Curr. Microbiol.* 1996, 33, 16-25.

[43] Smith, C. M., Koch, W. H., Franklin, S. B., Foster, P. L*., et al.*, *J. Bacteriol.* 1990, 172, 4964-4978

[44] Horton, A. J., Hak, K. M., Steffan, R. J., Foster, J. W. and Bej, A. K., *Antonie Van Leeuwenkoek* 2000, 77, 13-20.

[45] Kim, B. H., Bang, I. S., Lee, S. Y., Hong, S. K., *et al.*, *J. Bacteriol.* 2001, 183, 5580-5588.

[46] Av-Gay, Y., Aharonowitz, Y. and Cohen, G., *Nucleic. Acids Res.* 1992, 20, 5478.

[47] Julseth, C. R. and Inniss, W. E., *Can, J. Microbiol.* 1990, 36, 519-524.

[48] Vishnivetskaya, T., Kathariou, S., McGrath, J., Gilichinsky, D. and Tiedje, J. M., *Extremophiles* 2000, 4, 165-173.

[49] Vishnivetskaya, T. A. and Kathariou, S., *Appl. Environ. Microbiol.* 2005, 71, 6954-6962.

[50] Ponder, M. A., Gilmour, S. J., Bergholz, P. W., Mindock, C. A., *et al.*, *FEMS Micriobiol. Ecol.* 2005, 53, 103-115.

[51] Aebersold, R. and Mann, M., *Nature* 2003, 422, 198-207.

[52] Kachman, M. T., Wang, H., Schwartz, D. R., Cho, K. R. and Lubman, D. M., *Anal. Chem.* 2002, 74, 1779-1791.

[53] Kreunin, P., Urquidi, V., Lubman, D. M. and Goodison, S., *Proteomics* 2004, 4, 2754-2765.

[54] Wall, D. B., Parus, S. J. and Lubman, D. M., *J. Chromatogr. B Analyt.Technol. Biomed.Life Sci.* 2002, 774, 53-58.

[55] Wang, Y., Wu, R., Cho, K. R., Shedden, K. A., *et al.*, *Mol. Cell. Proteomics* 2006, 5, 43-52.

[56] Yan, F., Subramanian, B., Nakeff, A., Barder, T. J., *et al.*, *Anal. Chem.* 2003, 75, 2299-2308.

[57] Zhu, K., Miller, F. R., Barder, T. J. and Lubman, D. M., *J. Mass Spectrom.* 2004, 39, 770-780.

[58] Zhu, K., Kim, J. K., Yoo, C., Miller, F. R. and Lubman, D. M., *Anal. Chem.* 2003, 75, 6209-6217.

[59] Zheng, S., Schneider, K. A., Barder, T. J. and Lubman, D. M., *BioTechniques* 2003, 35, 1202-1212.

[60] Jensen, O. N., Podtelejnikov, A. V. and Mann, M., *Anal. Chem.* 1997, 69, 4741-4750.

[61] Ojima, N., Masuda, K., Tanaka, K. and Nishimura, O., *J. Mass Spectrom.* 2005, 40, 380-388.

[62] Nakanishi, T., Ohtsu, I., Furuta, M., Ando, E. and Nishimura, O., *J.Proteome Res.* 2005, 4, 743-747.

[63] Koy, C., Mikkat, S., Raptakis, E., Sutton, C., *et al.*, *Proteomics* 2003, 3, 851-858.

[64] Ponder, M. A., *Center for Genomic and Evolutionary Studies on Microbial Life at Low Temperatures, Department of Microbiology and Molecular Genetics*, Michigan State University, East Lansing 2006.

[65] Hesterkamp, T. and Bukau, B., *FEBS Lett.* 1996, 385, 67-71.

[66] Valent, Q. A., Kendall, D. A., High, S., Kusters, R.*, et al.*, *EMBO J.* 1995, 14, 5494-5505.

[67] Lill, R., Crook, E., Guthrie, B. and Wickner, W., *Cell* 1988, 54, 1013-1018.

[68] Jones, P. G., VanBogelen, R. A. and Neidhardt, F. C., *J. Bacteriol.* 1987, 169, 2092-2095.

[69] Jones, P. G. and Inouye, M., *Mol. Microbiol.* 1994, 11, 811-818.

[70] Salotra, P., Singh, D. K., Seal, K. P., Krishna, N.*, et al.*, *FEMS Mcirobiol. Lett.* 1995, 131, 57-62.

[71] Brissette, J. L., Russel, M., Weiner, L. and Model, P., *Proc. Natl. Acad. Sci. USA* 1990, 87, 862-866.

[72] Model, P., Jovanovic, G. and Dworkin, J., *Mol. Microbiol.* 1997, 24, 255-261.

[73] Kleerebezem, M. and Tommassen, J., *Mol. Microbiol.* 1993, 7, 947-956.

[74] Kleerebezem, M., Crielaard, W. and Tommassen, J., *EMBO J.* 1996, 15, 162-171.

[75] Feller, G. and Gerday, C., *Cell. mol. life sci.* 1997, 53, 830-841.

[76] Mathy, N., Jarrige, A.-C., Robert-Le Meur, M. and Portier, C., *J. Bacteriol.* 2001, 183, 3848-3854.

Chapter 3

Effect of Growth Temperature and Culture Medium on the Cryotolerance of Permafrost *Exiguobacterium Sibiricum* 255-15 by Proteome-wide Mass Mapping

3.1 Introduction

There has been growing interests in the survival mechanism of psychroactive bacteria at repeated freeze-thaw challenge largely due to the fact that the processes of freezing and thawing are common processes in nature and 80% of the earth's surface is cold. Bacteria in environments which experience seasonal temperature fluctuations are expected to be adapted to repetitive freeze-thaw cycles. On the other hand, bacteria in stable subfreezing environments are expected to be adapted to low temperature although these bacteria do not experience any repetitive freeze-thaw cycles in their native environment. The example of a stable subfreezing environment is permafrost. Permafrost, which is defined as a subsurface frozen layer, primarily soil or rock that remains frozen for more than two years, makes up more than 20% of the land surface of the earth, including 82% of Alaska, 50% of Russia and Canada, 20% of China, and most of the surface of Antarctica [1-3]. Within the buried Siberian permafrost soils, high numbers of viable microorganisms have been discovered [4-8]. The presence of the microorganisms is surprising, not only because of the constant subzero temperature of soils, averaging from -10°C to -12°C, but also because of the length of time the soils have been frozen, which ranges from a few

thousand years up to 2-3 million years. These organisms may well be the only living cells that have survived for a geologically significant period of time. These bacteria in the permafrost may be viewed as the result of a continuous process of selection for those capable of withstanding prolonged exposure to subzero temperatures. Even more significant is the fact that before the bacteria became trapped within the permafrost they were the outcome of cyclic freeze-thaw pre-selection in the original "active" tundra layers where they were exposed to warmer growth periods in summer and colder surface temperatures in winter. Cryotolerance, which has been defined as the ability of cells to recover their activity after freezing, has been studied using serial freeze-thaw treatments of five *Exiguobacterium* strains including *E. sibiricum* 255-15 [9], which has been isolated from a 2-3 million year old Siberian permafrost sediment [10]. *Exiguobacterium* spp. are low G+C Gram-positive non-spore-forming bacteria. Psychrotrophic *Exiguobacterium* strains were found to grow well in a temperature range from -6°C to 40°C [9, 11]. Our previous study [9] has shown that cryotolerance of psychrotrophic *Exiguobacterium* spp. is significantly influenced by low temperature (4 and -6^{o}C) and by surface- (agar) –associated growth. Molecular mechanisms underlying such improved cryotolerance of *Exiguobacterium* spp. remain unexplored.

The objective of this study was to compare proteomic profiles of *E. sibiricum* 255-15 grown at the different temperatures and media, and to identify proteins differentially expressed at conditions under which the bacteria developed improved cryotolerance, using a 2-D liquid separation and mass spectrometry-based mass mapping technique. The application of this technique has recently been evaluated during studying the low temperature adaptations in *E. sibiricum* 255-15 [12]. PMF by MALDI-TOF MS and

peptide sequencing using MALDI-QIT-TOF MS/MS have proved to be a powerful combination of techniques in protein identification. In the present work, with the aid of MALDI-QIT-TOF MS/MS, the percentage of proteins identified using both PMF and peptide sequencing has been improved to over 60%. In addition, the recently sequenced genome of *E. sibiricum* 255-15 (http://genome.jgi-psf.org/mic_asmb.html, accession number NZ_AADW00000000) will greatly facilitate the characterization of proteomic profiles. The current study both confirmed and significantly extended our previous studies. The results of this study would be applicable in many industrial, medical, agricultural and food technologies for improving cryoprotective characteristics in useful microorganisms and for controlling cold-resistant pathogen bacteria. An understanding of mechanisms contributing to cryotolerance would be also helpful in the search for life beyond earth, particularly because seven of the eight planets of our solar system, as well as their satellites, comets and asteroids have a cryogenic nature, *i.e.* the permafrost is a common occurrence in space.

3.2 Materials and Methods

3.2.1 Chemicals

Tryptic soy broth (TSB), tryptic soy agar (TSA), and yeast extract (YE) were obtained from Difco (BD Diagnostics Systems, Franklin Lakes, NJ). Tris-(hydroxymethyl) aminomethane (Tris-base), hydrochloride acid (37%), urea, thiourea, dithiothreitol (DTT), *n*-octyl-β-D-glucopyranoside (OG), phenylmethanesulfonyl fluoride (PMSF), bovine serum albumin (BSA), 3 bis[tris(hydroxymethyl methylamino] propane (bis-tris propane), iminodiacetic acid, trifluoroacetic acid (TFA), HPLC grade acetonitrile (ACN), 2-isopropanol, sodium chloride (NaCl), ammonium bicarbonate (NH_4HCO_3),

insulin, α- cyanohydroxycinnamic acid (CHCA), angiotensin I, angiotensin II, human adrenocorticotropic hormone (ACTH) fragment 1-17, ACTH fragment 18-39, bradykinin fragment 1-7, and P14R were purchased from Sigma-Aldrich (St. Louis, MO). Polybuffer 74 and 96 were obtained from Amersham Pharmacia Biotech (Piscataway, NJ). Formic acid (98-100%) was purchased from Riedel de Haen (Seelze, Germany). N-tosyl-L-phenylalanine chloromethyl ketone (TPCK)-treated trypsin was supplied from Promega (Madison, WI). 2,5 dihydroxy benzoic acid (DHB) was purchased from LaserBio Labs (Sophia-Antipolis, Cedex, France) and the water was purified using a Milli-Q water filtration system (Millipore, Inc., Bedford, MA).

3.2.2 *E. sibiricum* 255-15 Cell Culture

E. sibiricum 255-15 cell pellets were obtained from the Department of the Food Science at North Carolina State University. All cells were cultured in tryptic soy broth with 7% yeast extract (TSB-YE) for liquid broth growth and tryptic soy agar (TSA-YE) consisting of TSB-YE with 1.2% agar for agar surface growth. For liquid broth growth, bacteria were grown in TSB-YE to the mid-log phase (O.D. = 0.7) at 4°C for 7 days and at 25°C overnight. For growth on agar surface, bacteria were grown on TSA-YE at 4°C for 14 days and at 25°C overnight. Cells grown on agar medium were transferred with a sterile swab in TSB-YE to approximately 5 x 10^8 cell/ml (O.D. = 0.7). Cells grown at 4°C were harvested at 4°C, while cells grown at 25°C were harvested at room temperature.

3.2.3 Freeze-thawing

The cells grown on agar medium were collected in TSB-YE and frozen at -20°C while the cells grown in TSB-YE were frozen in the growth medium at -20°C. For the

freeze-thaw experiment, every two days the cells were thawed completely in a water bath at the room temperature and then refrozen at -20°C immediately. The CFU (colony forming units) of viable cells before and after 2, 6, 9, 12, 16, 20 cycles of freeze-thaw were determined on 1/10 TSA. The plates were incubated at 24°C for 48 h and the resulting colonies were counted using Protos Plus Colony Counter (Synoptics Ltd., Cambridge, UK). The results are the mean values of two independent replicate assays done in triplicate.

3.2.4 Cell Lysis

Proteins were extracted by sonicating the bacterial cells with 5 mL (pH 7.3) lysis buffer containing 50 mM Tris-HCl, 6 M urea, 2 M thiourea, 0.1 M DTT, 2% OG and 2 mM PMSF for 10 min. The resulting mixture was centrifuged at 40, 000 rpm for 20 min (Model L-70 Ultracentrifuge, 80Ti Rotor, Beckman-Coulter, Fullerton, CA) to remove cell walls and other insoluble debris. The supernatants from whole cell lysates were then desalted by a PD-10 Sephadex G-25 gel filtration column (Amersham Biosciences, Piscataway, NJ) and the protein concentrations were determined by the Bradford-based protein assay using a commercial kit (Bio-Rad, Hercules, CA). The samples were quantified in triplicate at 595 nm with BSA as a standard. The protein concentrations of the samples were between 1.0-2.0 mg/mL.

3.2.5 2-D Liquid Phase Separation

A 2-D liquid phase mass mapping method has been developed in our laboratory to profile protein expression in *Exiguobacterium sibiricum* 255-15 (Figure 3.2). Proteins extracted from *E. sibiricum* 255-15 were first fractionated by CF using a Beckman Gold

HPLC equipped with model 127S pump and model 166 detector (Beckman Coulter, Fullerton, CA) on an HPCF-1D column (250 × 2.1 mm) (Eprogen, Inc., Darien, IL). The pH gradient was generated using a start buffer (25 mM bis-tris propane, pH 8.5) and an elute buffer (3% v/v poly-buffer 74, 7% v/v poly-buffer 96, pH 4.0). Both buffers were prepared in 6 M urea and 0.1% OG with pH adjusted by adding iminodiacetic acid. A total of 5 mg of protein from each sample was loaded on the CF column which was equilibrated with the start buffer for two hours at 0.2 mL/min followed by elution at the same flow rate using the elute buffer. The pH was monitored online by a pH electrode (Lazar Research Laboratories, Inc. Los Angeles, CA) and the separation was detected at 280 nm. Effluent from the CF separation was collected from pH 8.5 to 4.0 every 0.3 pH unit intervals. After running the pH gradient, the column was washed with 1 M NaCl solution to elute proteins with p*I* lower than 4.0.

NPS-RP-HPLC separation was performed on a 33 × 4.6 mm ODS III column packed with 1.5 μm nonporous silica beads (Eprogen) at a flow rate of 0.5 mL/min using the same HPLC system as used in the CF fractionation. To improve the resolution and speed of the separation, the column temperature was maintained at 60°C using a Model 7971 column heater (Jones Chromatography, Resolution Systems, Holland, MI). The NPS-RP-HPLC separation was performed using gradient elution with a water (solvent A) and acetonitrile (solvent B) gradient, both of which were prepared in 0.1% TFA. The gradient profile used was as follows: (1) 5 to 26% B in 1 min; (2) 26 to 35% B in 3.5min; (3) 35 to 40% B in 9.5 min; (4) 40 to 50% B in 13 min; (5) 50 to 58% B in 4 min; (6) 58 to 75% B in 1 min; (7) 75-100% B in 1 min; (8) 100 to 5% B in 1 min.

3.2.6 Protein *M*r Measurement and Interlysate Quantification by ESI-TOF MS

Half of the effluent from NPS-RP-HPLC was collected for further MS-based identification and the other half was directly injected into ESI-TOF MS for quantification (LCT, Micromass, Manchester, U.K.). ESI-TOF MS was externally calibrated by directly infusing NaI-CsI standard solution and normalized internally by the peak area of 1μg of insulin added as an internal standard. The intact *M*r values were obtained from the deconvolution of the combined ESI-TOF spectra by MaxEnt I software (Micromass) using a target mass range of 4-95 kDa, resolution of 1 Da, peak width of 0.75 Da, and peak height value of 65% as fixed parameters.

3.2.7 Tryptic Digestion

Fractions for MS based identification were first reduced to a volume of 80μL using a SpeedVac concentrator (Centrivap Concentrator, Labconco, Kansas City, Missouri) to remove acetonitrile and TFA. The remaining TFA in each fraction was neutralized by adding 10 μL of 1 M NH_4HCO_3 while the proteins in the fraction were denatured by adding 10 μL of 100 mM DTT. The resulting mixtures were incubated at 37°C for 20 min, and then 0.5 μg TPCK-treated trypsin was added to start the digestion. After being maintained at 37°C for 24 hours, digestion was terminated by adding 2 μL TFA. Tryptic digests were then desalted and pre-concentrated in 5 μL of 60% acetonitrile with 0.1% TFA by C18 Zip-Tips (Millipore, Inc.) before MS-based identification.

3.2.8 Protein Identification by PMF and Peptide Sequencing

For PMF by MALDI-TOF MS, peptide mass was measured on a Micromass TofSpec2E system (Micromass/Waters, Milford, MA) equipped with a 337 nm nitrogen laser source and delayed extraction. The MALDI-TOF MS was operated in positive ion

reflector mode and the final MALDI spectrum was an average of 100-150 spectra and calibrated with an internal standard mixture of Angiotensin I ($[M+H]^+$ 1296.69), ACTH 1-17 ($[M+H]^+$ 2093.09), and ACTH 18-39 ($[M+H]^+$ 2465.20) within 50 ppm. The matrix used is α-CHCA. The peptide masses were analyzed using MassLynx 4.0 over the range of 800- 4000 Da and then submitted to MS-Fit to search against the NCBInr database (released on Feb 26, 2006) with a mass tolerance of 50 ppm and one missed cleavage as fixed parameters.

Mass spectrometric peptide fragmentation and sequencing was performed on MALDI-QIT-TOF MS (Shimadzu Corporation, Kyoto, Japan and Kratos Analytical, Manchester, UK) in the positive ion mode using an external calibration with a mixture of bradykinin fragment 1-7 ($[M+H]^+$ 757.40), angiotensin II ($[M+H]^+$ 1046.54), P14R ($[M+H]^+$ 1533.86) and ACTH fragment 18-39 ($[M+H]^+$ 2465.20). The matrix used in this case was DHB solution with a concentration of 10 mg/mL. Data acquisition and processing were controlled by Kompact software (Kratos Analytical Ltd., Manchester, UK). The parent ion mass and the resulting fragment ion masses were searched against the NCBInr database using Mascot 1.8 (Matrix Science, London, UK) setting a peptide tolerance of 1.2 Da, MS/MS tolerance of 0.6 Da, and one missed cleavage site as fixed parameters.

3.3 Results

3.3.1 Freeze-thawing Tolerance

Cell survival was monitored after repeated cycles of freeze-thaw treatments (Figure 3.1). For liquid broth growth at 25^{o}C, the cell viability was lost by more than 50% after 9 cycles and about 80% after 20 cycles. While for liquid broth growth at 4^{o}C or agar

surface growth at either 25°C or 4°C, cells showed high viability after 20 cycles of freeze-thaw treatments. Bacteria grown in liquid medium at 4°C tolerate freeze-thawing much better than those grown at 25°C. However, when grown on agar, they tolerate freeze-thawing equally well regardless of the growth temperature.

3.3.2 Protein Separation and Comparison of 2-D Mass Maps of Protein Expression in Cells Grown Under Different Conditions

Protein extracts from *E. sibiricum* 255-15 were first fractionated by CF according to p*I* in the first dimension and each p*I* fraction was subsequently separated by NPS-RP-HPLC based upon the hydrophobicity before MS-based protein identification. The CF separation was achieved in approximately one hour at a flow rate of 0.2 mL/min. During the pH gradient, the proteins were collected from pH 8.5 to 4.0 at 0.3 pH unit change. For each p*I* fraction, the corresponding chromatogram of NPS-RP-HPLC separation was achieved within 35 minutes, resulting in resolution of ca.10-100 protein bands.

A mass map was generated by integrating *M*r, p*I* and protein abundance from all the 18 CF fractions (15 pH fractions from pH 8.5 to 4.0, and three NaCl wash fractions) into one single image to compare the protein expression of *E. sibiricum* 255-15 at the different growth temperatures and media. In this study, we focused on the proteins expressed to protect against the freeze-thaw treatments and thus to increase the bacterial cryotolerance. One such map is shown in Figure 3.3, which represents the comparison of 2-D mass maps between cells grown in liquid broth at 25° (left) and on agar surface at 25°C (right). The differential map shown in the middle was obtained by point-by-point subtraction. The 18 lanes in the mass map represent all the p*I* fractions from the CF separation, while the vertical axis indicates the *M*r of intact proteins in each p*I* fraction.

The mass accuracy in 2-D mass map was determined by the quality of the protein umbrella spectra from ESI-TOF MS analysis and is usually less than 100-200 ppm, which is 200-1000 times better than the 5-10% typically achieved in 2-DE separation. The accurate *M*r values of intact proteins, together with p*I*, provide essential information for protein identification and characterization. The reproducibility of protein profiling by mass mapping is displayed in Figure 3.4, which includes two differential maps of all the 18 CF fractions from two duplicate experiments.

3.3.3 Proteins Identified by PMF and Peptide Sequencing

Proteins eluting from HPLC were digested by trypsin and then identified by peptide mass fingerprinting (PMF) and peptide sequencing using MALDI-TOF MS and MALDI-QIT-TOF MS/MS respectively. The identification from PMF was obtained by searching for the best match between the experimentally determined intact masses of the peptides in the peptide map and those calculated by theoretical cleavage of the proteins in a sequence database. The PMF based identification was based on the indication that at least five peptides matched with mass accuracy within 50 ppm and sequence coverage of at least 20% [13] , which was further confirmed by protein *M*r and p*I*. In this work, around 42% of the proteins detected by ESI-TOF MS were identified by PMF using MALDI-TOF MS.

MS/MS based peptide sequencing provides a more powerful proteomic technique with higher sensitivity and accuracy in protein identification and was used to confirm these IDs. Protein identification by peptide sequencing is successful when the search score is higher than, or equal to, the homology or identity threshold scores in each search. In this case, multiple peptides are usually found and often all of their fragment spectra are used to correlate to a protein. In both PMF and peptide sequencing, the larger the number

of peptides identified the greater the confidence in the protein identification. Small proteins that are difficult to identify by PMF due to an insufficient number of detected peptides (<5) may be identified using peptide sequencing if the peptide signal and its fragment signals are of sufficient intensity. Proteins were identified with confidence when at least two MS/MS sequenced peptides matched a protein above the MASCOT significance level indicating identity or extensive homology. Figure 3.5 displays two MS/MS spectra of peptides from the digest of hypothetical protein (18.900 kDa, p*I* 5.4). The matched peptide sequences, DDATDETSGASWIDQVK and FIGIFHDESSLHQK, confirmed the protein identification from MALDI-TOF MS based PMF experiment. The discrepancy of 575 Da between the experimental and theoretical *M*r values of this protein resulted from the loss of the first five amino acids (MVLTM).

3.3.4 Cold Shock Proteins in Inducing Cryotolerance

No consistent differences in the expression of CSPs in the bacteria grown under different conditions were observed (Table 1), suggesting that there is no direct correlation between cryotolerance and the CSPs expression prior to freezing. The putative CSP of 7.154 kDa and two CSPs of 7.409 kDa with p*I* 4.5 and p*I* 4.4 respectively were found similarly expressed in the same medium regardless of the temperature. Interestingly they were up-regulated in liquid media compared with solid media.

3.3.5 Variation of Cellular Proteins According to Different Growth Conditions

The identification of about 330 soluble proteins, which represented 60% of the proteins resolved by 2-D mass mapping, was used as a starting point for physiological studies of *E. sibiricum* cells. Observation of the 2-D mass maps of Siberian permafrost *E. sibiricum* 255-15 revealed some changes in response to the different growth temperatures

and media. However, the number of proteins whose relative expression level varied more than two-fold under different growth conditions was low and represented ca. 15% of the total cellular proteins detected by 2-D liquid separation with MS technology.

3.3.5.1 Down-regulation of Proteins Associated with the Improved Cryotolerance

When compared with the protein expression in the cells grown in liquid at 25°C, 2-D mass maps revealed down-regulation of 20 identified proteins in liquid growth at 4°C as well as on agar growth at either temperature (Table 2). We observed a significant suppression of several glycolytic enzymes (phosphoglycerate kinase, enolase, fructose-1, 6-bisphosphate aldolase, pyruvate kinase, glucose-6-phosphate isomerase, pyruvate dehydrogenase) in the cells grown in liquid broth at 4°C and on agar surface at either 25°C or 4°C. Similar suppression was observed with other energy metabolic enzymes such as acetate kinase, NAD+ synthase, ATP synthase F1 beta subunit, iron-containing alcohol dehydrogenase, NADH flavin oxidoreductase and FAD dependent oxidoreductase. Enzymes in metabolism of lipids, amino acids, nucleotides and nucleic acids were also suppressed. Growing the cells on agar enhanced suppression of some enzymes. Glucose-6-phosphate isomerase and FAD dependent oxidoreductase were not even detected when the cultures were grown on the solid medium. Iron-containing alcohol dehydrogenase was drastically reduced while other proteins, including acetate kinase, were only marginally affected by surface cultivation. Purine nucleoside phosphorylase I was suppressed more during growth at 4°C. The low incubation temperature seemed to have a greater impact when coupled with agar growth for the majority of the 20 listed proteins. Surprisingly, a few enzymes exhibited the greatest suppression when the cells were grown at 25°C on agar (Table 2), such as enolase,

pyruvate kinase, phsphodiesterase, and MECDP synthase. Pyruvate kinase was the only enzyme not detected in any of the cultures grown under conditions that were associated with the increased cryotolerance, and was instead only expressed in cells grown at 25°C in liquid. Porphobilinogen synthase and certain other proteins (enolase, pyruvate dehydrogenase and hypothetical protein 68055276) were similarly expressed in the cells grown in 4°C liquid broth and on agar surface at either 25°C or 4°C (Table 2).

3.3.5.2 Up-regulation of Proteins Associated with the Improved Cryotolerance

At the same time, the consistent up-regulation of eight other proteins was observed (Table 3). The expression of the putative 2'-5' RNA ligase, hypoxanthine phosphoribosyl transferase, FeS assembly ATPase SufC, and uncharacterized conservative protein (680550340) was especially induced in surface grown bacteria of *E. sibiricum* 255-15 (Table 3). The putative hypoxanthine phosphoribosyl transferase and thioredoxin reductase were especially induced on agar in the cold (Table 3). A hypothetical protein of 24460 kDa, which was identified as phage shock protein A (IM30) using the database released in 2005, was also found consistently over-expressed in the cells associated with improved cryotolerance. Comparison of the protein sequence of this hypothetical protein to proteins in the GenBank, RefSeq, PDB, SwissProt, PIR, and PRF databases using the BLAST algorithm identified two highly similar protein sequences (Figure 3.6). PspA (IM30).

3.4 Discussion

The experiments with *E. sibiricum* 255-15 showed significant impact of growth temperature and complex (structured) medium on cryotolerance [9] . Hence, we assume

that a cryotolerance mechanism expressed during liquid broth growth in the cold was also expressed during surface growth regardless of temperature. Low temperature growth may increase cryotolerance by altering the composition of the cytoplasmic membrane, inducing ice nucleation activity or phenotypically changing the cell formation [14, 15]. On other hand, solid (agar)-medium, which is defined as a complex structured habitat, may develop multiple physical and chemical gradients (stresses) and may promote diverse adaptations including a freezing tolerance [9].

A correlation between cold shock treatment and cryotolerance was first established in mesophilic bacterium *Bacillus subtilis* [16]. When a *B. subtilis* culture grown at 37°C was cold shocked at 10°C for 2 h and frozen at -80°C, the cell viability was 96%, whereas in a culture frozen without the cold shock treatment viability was only 27%. Likewise, an *E. coli* culture that was cold shocked at 10°C for 6 h yielded a significantly higher cell viability following freeze-thawing compared to a culture frozen without the cold shock treatment [17]. Cold shock-induced cryotolerance was also displayed in *Lactococcus lactis* [18-21]. In *L. lacis* subsp. LL41-1, without cold shock treatment, the viability of cells following freezing for 1 day was 34%, 14 days 32%, 182 days 7%, and 364 days 0.2%. However, with the cold shock treatment at 10^{o}C for 5h, it was 83%, 82%, 12%, and 0.8% respectively [20]. It therefore appears that cold shock treatments prior to freezing greatly improved the cryotolerance in mesophilic bacteria. Although cryotolerance has been investigated in *E. coli*, *B. subtilis*, and *L. lactis*, little is known about the mechanisms that mediate survival following freezing. These data suggest that CSPs play a role in cryotolerance induced by cold in *L. lactis*.

Apart from the improved tolerance to freeze-thawing in response to cold shock, the overexpression of cold shock proteins (CSPs) has also been shown to occur following the cold shock treatment. Cold shock proteins are defined as proteins that are transiently up-regulated after an abrupt shift to low temperature. Do CSPs play a role in cryotolerance? Is there a direct cryoprotective role for CSPs? Currently, only limited studies that linked CSP level with the protection against freezing have been performed, mainly in *L. lactis*. It is reported that an increased survival rate (~100 fold) upon freezing of cold adapted *L. lactis* coincided with an increase of CSP expression. Also an up-regulation after freeze-thaw cycles of CspD, CspE, and CspB seemed to increase the survival to freezing 2-10 fold compared to control cells [21, 22]. These data suggest that CSPs play a role in cryotolerance induced by cold in *L. lactis*.

The above studies provide evidence that the CSPs expression level after freezing is correlated with the cold-induced cryotolerance in *L. lactis*, but is there a correlation between CSPs expression prior to freezing and cryotolerance? Do CSPs play a role to induce the improved cryotolerance? In this study, we investigated the putative CSPs expressed at different growth temperatures (25°C and 4°C) and media (liquid broth and agar surface). The major homologous CSPs of *E. sibiricum* 255-15 were similarly expressed in the same medium regardless of the temperature, and they were up-regulated in liquid medium compared with solid one. These results suggest that genes for CSPs are turned on continuously to produce "shock" proteins to protect cells from damage during abrupt changes in environmental conditions. However, cold-shock stressed (5h at 4^{o}C) cells of *E. sibiricum* 255-15 grown in liquid broth at 24^{o}C did not show significantly improved cryotolerance [9]. Therefore, our findings suggest that cold growth

(acclimation), rather than cold shock, is required for the enhanced protection to repetitive freeze-thawing stresses of psychrotrophic *Exiguobacterium* ssp. A direct role of CSPs to induce the improved cryotolerance in *E. sibiricum* 255-15 could not be ruled out yet.

Down-regulation of 20 proteins associated with the improved cryotolerance was observed at different growth temperatures and media. Using this information, there is an opportunity to analyze the regulation of major metabolic pathways such as glycolysis, electron transport, ATP synthesis, *etc*. The cells grown at conditions those enhance their cryotolerance display a significant suppression of several glycolytic and energy metabolic enzymes as well as enzymes involved in metabolism of lipids, amino acids, nucleotides and nucleic acids. The temperature dependent changes in metabolic pathways could explain the temperature specific carbon source utilization observed in *E. sibiricum* 255-15 [11]. Contrary to our study low-temperature stress induces the glycolytic activity in lactic acid bacterium *L. lactis* and psychrotrophic *Rhizobium* strains [23, 24]. However, the temperature impact on the regulation of the similar proteins could be various in bacteria from the different genera. Thus, the inhibition of ATPase activity of *E. coli* F1F0 ATP synthase was observed at 15^{o}C compare to 37^{o}C [25], while F1F0 ATP synthase b subunit was up-regulated by low temperature growth in permafrost bacterium *Psychrobacter cryohalolentis* K5 [26]. The effect of low temperature or agar growth on the suppression of several proteins was not as clearly defined as when the low incubation temperature was coupled with agar growth resulting in a greater impact for the majority of the proteins described. A few studies showed that the populations in the structured habitat (agar growth) are more likely to express a phenotype to a much greater extent than cells from broth cultures [27, 28]. Interesting, that pyruvate kinase was expressed in

cells grown at 25°C in liquid but not in any of the cells grown under conditions that were associated with the increased cryotolerance. Though, the increase of resistance to freezing and to frozen storage at -20^{o}C and oversynthesis of 4 cold induced proteins including pyruvate kinase was observed in cold stressed *Lactobacillus acidophilus* RD758[29]. Several proteins known as cold inducible proteins in other bacteria were down-regulated in *E. sibiricum* 255-15 at low temperature, nevertheless the cells were resistant to repetitive freeze-thawing. These findings suggest that *E. sibiricum* 255-15 may have evolved unknown protective mechanisms against freezing.

In the course of this study we observed up-regulation of 8 proteins at low temperature or solid (agar) medium. These results suggest these proteins may significantly enhance cryotolerance. What do we know about a role of these proteins in cold adaptation and freezing tolerance? Our previous proteomic analysis of cold adaptation in *E. sibiricum* 255-15 reveals 39 proteins that were preferentially or uniquely expressed at 4^{o}C [12]. The latter and current studies confirmed that enzyme 2'-5' RNA ligase was up-regulated at 4^{o}C as well as agar surface growth regardless of the temperature. Overexpression of the RNA ligase protein in *E. coli* led to slower growth rates and a temperature-sensitive phenotype in both wild-type and RNA ligase knockout strains. The RNA ligase reaction was studied *in vitro* using purified enzyme and was found to be reversible, indicating that this enzyme may perform cleavage or ligation *in vivo*. The direction of the equilibrium *in vivo* depends on the effective concentrations of substrates available for each reaction. If the true *in vivo* substrate is tRNA, and the ligated tRNA products are utilized for translation and thereby removed from the pool of substrates, the ligation reaction will be favored [30]. Another study has shown the involvement of DNA ligase reaction in

recovery of *E. coli* from cold shock [31]. There are no direct indications that proteins hypoxanthine phosphoribosyl transferase (HPRT), FeS assembly ATPase SufC, and thioredoxin reductase play any role in cold adaptation and freezing tolerance in bacteria, while their expression may increase upon other stresses. HPRT is a key enzyme for purine salvage [32, 33]. In living organisms, purine nucleotides needed for cellular metabolism or used as the precursors of DNA and RNA are synthesized either via the salvage mechanisms that recycle pre-formed purine bases, or via *de novo* pathway starting from amino acid, carbon dioxide and ammonia. Data reported to GenBank indicated that the HPRT of distantly related organisms share extensive primary sequence homology. For example, there is 41% identity in the deduced amino acid sequence between the human HPRT and that of *Salmonella typhimurium* [34]. This enzyme is readily released from the bacterial periplasm upon osmotic shock [35, 36]. Although the exact functions of FeS assembly ATPase SufC are still unknown, it is found to be an ATPase that is involved in the biosynthesis of Fe-S clusters. Proteins that contain Fe-S clusters fulfill enzymatic or regulatory functions in various cellular processes ranging from respiration to gene expression [37]. Primary sequence analysis of SufC revealed the presence of an ABC transporter signature and the protein does indeed bind and hydrolyze ATP [38]. This class of ATPases is mostly found to be associated with membrane proteins, forming a trans-membrane complex that translocates a wide range of allocrites [39]. Expression of SufC is increased under oxidative stress for the protection of labile [Fe-S] clusters [40, 41]. Thioredoxin reductase is a member of the family of dimeric flavoenzymes that catalyze the transfer of electrons between pyridine nucleotides and disulfide/dithio compounds and promote catalysis via FAD and a redox active disulfide

[42-45]. Thioredoxin reductase plays several key roles in maintaining the redox environment of the cell. The role of these proteins in cryoprotection should still be evaluated.

Four up-regulated proteins were hypothetical, uncharacterized or proteins of unknown functions. Conserved hypothetical protein (68054709) was annotated as phage shock protein A (IM30). The apparent conservation of the PspA protein sequence between closely related bacteria (*E. sibiricum*, *B. cereus*, and *B. anthracis*) as well as distantly related bacterial species such as *Yersinia bercovieri*, *Clostridium perfringens*, and *Trichodesmium erythraeum*, suggests an ancient origin for this protein. So far, the exact function of PspA remains elusive. It is shown in previous research that PspA was induced under various stress conditions including severe heat shock, cold shock, osmotic shock, and exposure to organic solvent (ethanol) [46-48] that all might lead to the dissipation of the proton-motive force. Since, a portion of PspA was observed to be present in the cytoplasmic protein fraction in *E. coli* [49], it is believed that expression of PspA helps to sustain the proton motive force across the plasma membrane [49-51].

3.5 Concluding Remarks

In this study, we have presented the proteomic analysis of the effect of the growth temperature and medium on bacterial cryotolerance in *E. sibiricum* 255-15. It has been found that *E. sibiricum* 255-15 cells have significantly improved cryotolerance after liquid broth growth at 4°C and agar surface growth at both 4°C and 25°C compared with liquid broth growth at 25°C. The use of 2-D liquid separation and MS-based proteomics revealed growth temperature and medium dependent changes in protein synthesis related to the improved cryotolerance. A total of 28 proteins displayed cryoprotection-related

alterations: eight proteins were up-regulated and 20 proteins were down-regulated. This has indicated that the up-regulated eight proteins may induce the improved cryotolerance, and the down-regulation of 20 proteins supply additional freeze-thawing protection [52].

Table 3.1 Cold shock proteins expressed at different growth conditions.

Protein ID (Access. No.)	Exp. *Mr*	Exp. p*I*	Theoretical *Mr*/p*I*	Expression ratio L25/L4 [1]	Expression ratio L25/S25 [1]	Expression ratio L25/S4 [1]
Cold shock protein (68055399)	7154	4.0-4.6	7150/4.5	1.4	5.1	4.6
Cold shock protein (68053997)	7413	4.0-4.6	7409/4.5	1.2	45	36
Cold shock protein (68053999)	7413	4.0-4.6	7409/4.4			

[1] In discussion of expression ratio, L25 indicates 25°C TSB-YE liquid broth growth, L4 indicates 4°C TSB-YE liquid broth growth, S25 indicates 25°C TSA-YE agar surface growth, and S4 indicates 4°C TSA-YE agar surface growth. Protein quantification was based on its peak area from ESI-TOF MS analysis. For example the expression ratio L25/S4 was obtained by dividing the amount of protein expressed at 25°C in liquid broth growth with the amount of the same protein expressed at 4°C on agar surface growth.

Table 3.2 Proteins down-regulated at 4°C TSB-YE, 4°C TSA-YE and 25°C TSA-YE growth that are associated with the improved cryotolerance.

ID (Access. No.)	Exp. *Mr*	Exp. p*I*	Theoretical *Mr*/p*I*	Expression ratio[1] L25/L4 [2]	L25/S25 [2]	L25/S4 [2]
1, Intermediary metabolism						
1.1 Energy metabolism						
1.1.1 Glycolysis and glycolysis-related pathways						
Enolase (68056418)	46201	4.9-4.6	46329/4.7	2.7	3.1	2.6
Ketose-bisphosphate aldolase, class-II:Fructose-1,6-bisphosphate aldolase, class II (68054326)	30328	4.9-4.0	30458/5.0	3.4	6.2	12
Phosphoglycerate kinase (68056421)	42107	5.2-4.9	42103/5.1	2.2	3.5	22
Pyruvate kinase (68055923)	62770	5.2-4.9	62772/5.2	+++	+++	+++
Glucose-6-phosphate isomerase (68056048)	49322	5.5-5.2	49448/5.2	6.9	+++	+++
Pyruvate dehydrogenase/lipoamide (68056114)	40141	5.2-4.6	40137/5.1	2.7	3.5	2.2
1.1.2 Anaerobic respiration						
Acetate kinase (68055938)	44146	5.2-4.9	44272/5.1	6.7	22	36
1.1.3 ATP-proton motive force interconversion						
NAD+ synthase (68056126)	30070	5.8-5.2	30059/5.3	4.2	2.2	14
ATP synthase F1, beta subunit (68054271)	51563	5.5-5.2	51562/4.9	6.8	5.6	+++
1.1.4 Fermentation						
Iron-containing alcohol dehydrogenase (68053559)	94143	6.1-5.8	94287/6.0	2.1	160	190
1.1.5 Electron transport						
IMP dehydrogenase/GMP reductase:NADH:flavin oxidoreductase/NADH oxidase (68054616)	37115	5.8-5.5	37118/5.5	2.5	2.2	4.9
1.1.6 Metabolism of sugar						
FAD dependent oxidoreductase (68056283)	61679	5.8-5.5	61681/5.7	10	+++	+++
1.2 Metabolism of nucleotides and nucleic acids						
Inosine guanosine and xanthosine phosphorylase:Purine nucleoside phosphorylase I, inosine and guanosine-specific (68055531)	29500	5.2-4.9	29497/4.9	19	4.9	20
Pyrimidine-nucleoside phosphorylase (68055530)	46094	5.2-4.9	46088/5.0	2.9	2.1	6.7
1.3 Metabolism of lipid						
[Acyl-carrier protein] phosphodiesterase (68054042)	23640	5.5-5.2	23639/5.1	2.0	6.9	2.3
MECDP synthase (68056360)	17512	5.2-4.6	17540/5.3	2.5	28	13
1.4 Metabolism of amino acids and related molecules						
Porphobilinogen synthase (68055121)	36246	5.8-5.2	36207/5.3	6.6	7.2	7.3
2. Protein of unknown function						
Hypothetic protein (68055276)	18346	5.8-5.2	18900/5.4	3.4	2.2	2.5
Conserved hypothetical protein (68055483)	21085	5.2-4.6	21042/5.2	2.9	+++	7.6
Conserved hypothetical protein (68056068)	22599	5.5-4.9	22596/4.8	4.2	22	49

[1] The expression ratios of the proteins that were exclusively expressed at 25°C in TSB-YE liquid medium were shown as "+++".In this case, the ratio is apparently infinite because the denominator could not be distinguished from zero.

[2] L25 indicates 25°C TSB-YE liquid broth growth, L4 indicates 4°C TSB-YE liquid broth growth, S25 indicates 25°C TSA-YE agar surface growth, and S4 indicates 4°C TSA-YE agar surface growth. Protein quantification was based on its peak area from ESI-TOF MS analysis.

Table 3.3 Proteins up-regulated at 4°C TSB-YE, 4°C TSA-YE and 25°C TSA-YE growth that are associated with the improved cryotolerance.

Protein ID (Access. No.)	Exp. *Mr*	Exp. p*I*	Theoretical *Mr*/p*I*	Expression ratio L25/L4 [1]	L25/S25 [1]	L25/S4 [1]
Protein of unkown function (68054621)	14582	4.3-4.0	14578/4.3	0.49	0.31	0.24
2',5' RNA ligase (68056138)	19752	5.8-5.2	19743/5.4	0.22	0.045	0.049
Hypothetical protein (68055956)	20237	4.6-4.0	20367/4.7	0.057	0.48	0.028
Hypoxanthine phosphoribosyl transferase (68055907)	20294	5.5-4.9	20292/4.9	0.17	0.057	0.043
Conserved hypothetical protein (68054709)	24467	4.9-4.3	24460/4.9	0.013	0.14	0.16
FeS assembly ATPase SufC (68056341)	28809	5.2-4.9	28807/4.9	0.46	0.29	0.031
Uncharacterized protein conserved in bacteria (680550340)	35710	4.6-4.0	35845/4.5	0.1	0.0084	0.0086
Thioredoxin reductase (68055606)	36634	5.8-5.2	36622/5.3	0.034	0.0306	0.011

[1] L25 indicates 25°C TSB-YE liquid broth growth, L4 indicates 4°C TSB-YE liquid broth growth, S25 indicates 25°C TSA-YE agar surface growth, and S4 indicates 4°C TSA-YE agar surface growth. Protein quantification was based on its peak area from ESI-TOF MS analysis.

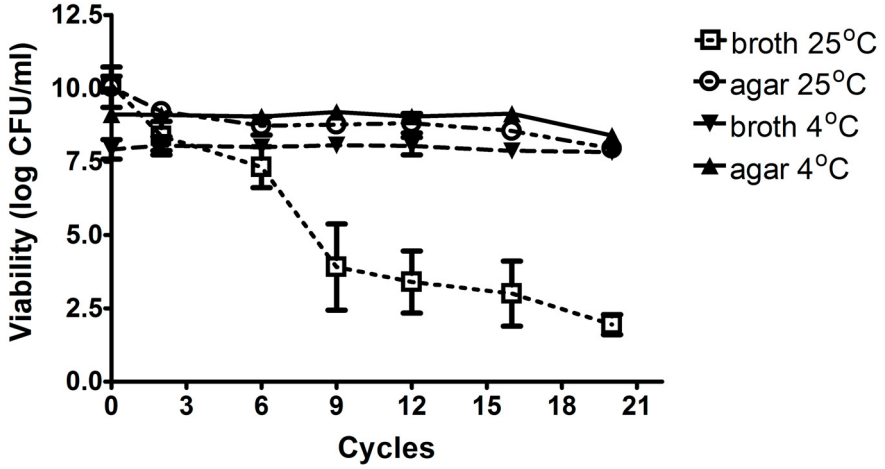

Figure 3.1 Viability of *E. sibiricum* 255-15 cells grown in liquid broth or on solid agar medium at both 25°C and 4°C after repetitive cycles of freeze-thaw treatments.

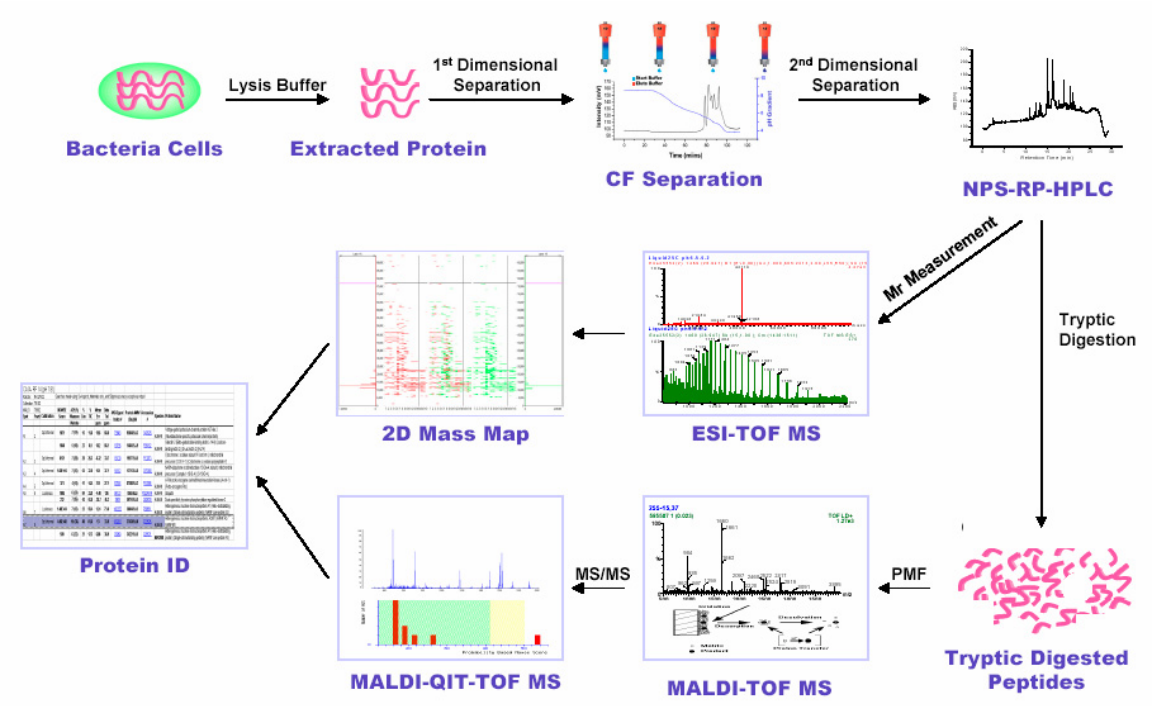

Figure 3.2 Experimental overview of 2-D liquid phase separation combined with MS for proteomic analysis of cryotolerance in *E. sibiricum* 255-15.

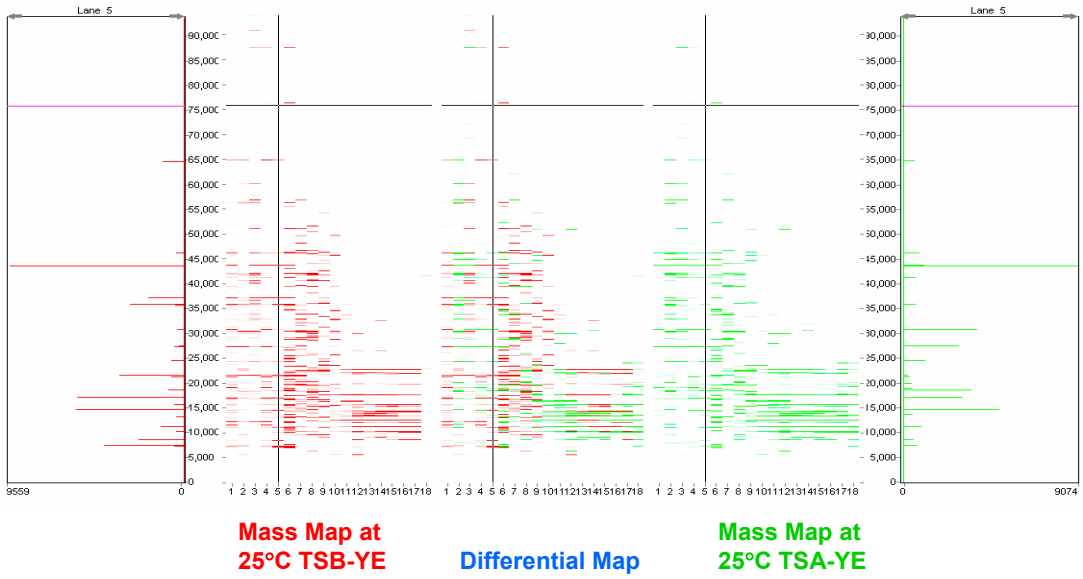

Figure 3.3 2-D mass map of *E. sibiricum* 255-15 after 25°C TSB-YE liquid broth growth (left) and 25°C TSA-YE agar surface growth (right) of all CF fractions in *M*r range of 5-95 kDa with the mass map of lane 5 from each sample highlighted on the left and right plot respectively. Differential mass map shown in the middle was generated through point-by-point subtraction.

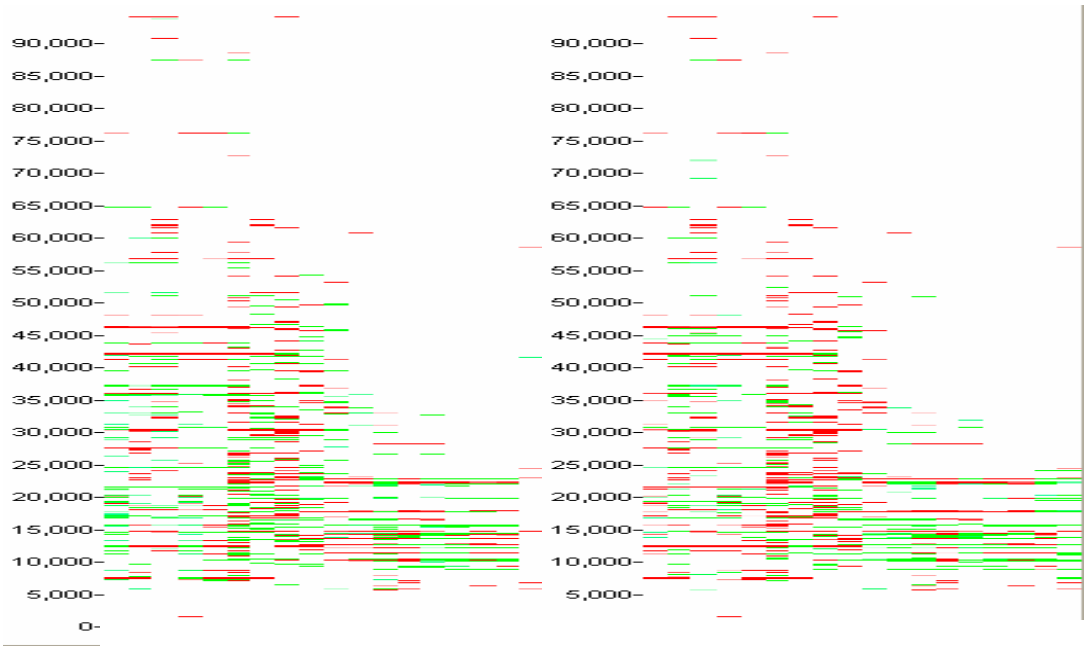

Figure 3.4 Reproducibility demonstration of protein profiling by 2-D mass mapping is generated by two independent CF/NPS-RP-HPLC/ESI-TOF MS experiments with two groups of samples cultured separately. Examples shown in this figure are the comparisons of 2-D mass maps between *E. sibiricum* 255-15 after 25°C (green) and 4°C (red) TSB-YE liquid broth growth.

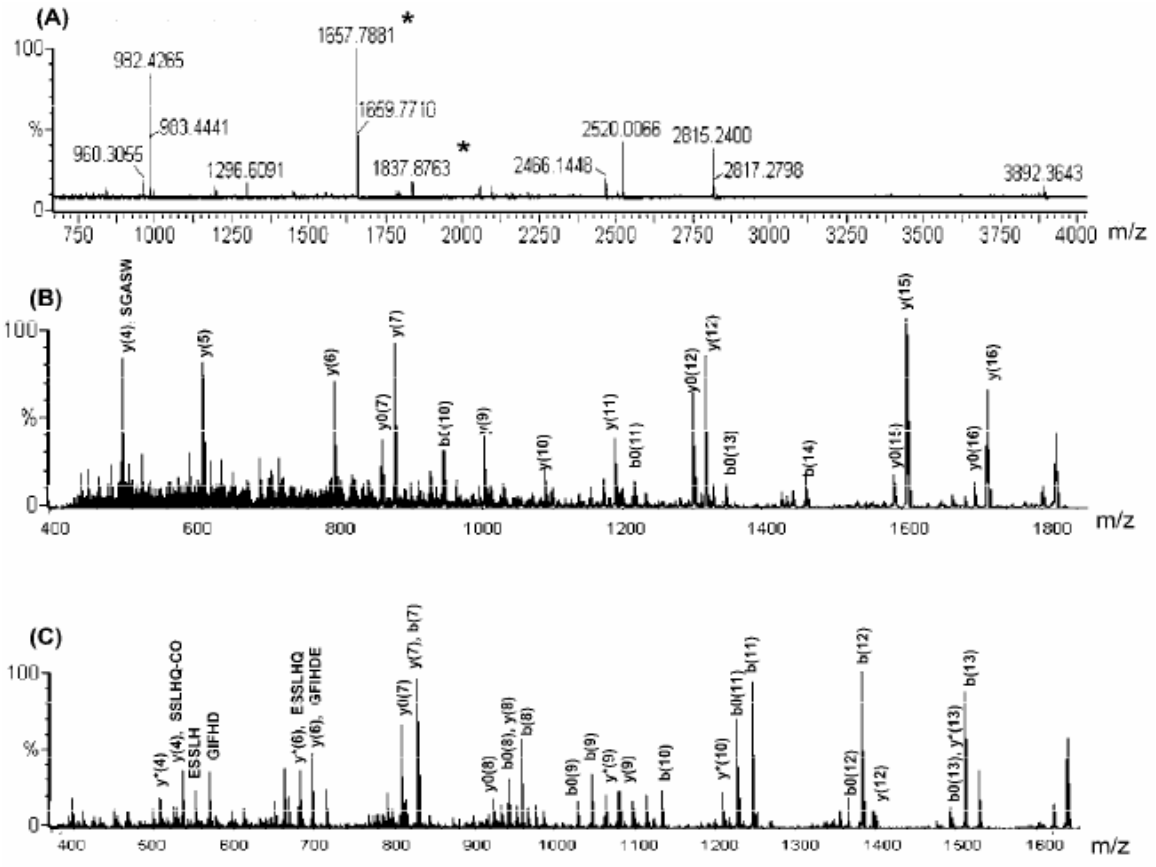

Figure 3.5 (A) MALDI-TOF MS spectrum of protein (18.900kDa, p*I* 5.44). MALDI-QIT-TOF MS/MS fragmentation of peptide 1837.8 (DDATDETSGASWIDQVK) was shown in (B) and peptide 1657.9 (FIGIFHDESSLHQK) in (C). Both MS/MS spectra confirm the protein identity. In this case, $y^*(i)=y(i)–NH_3$, $y0(i)=y(i)–H_2O$, $b0(i)=b(i)–H_2O$.

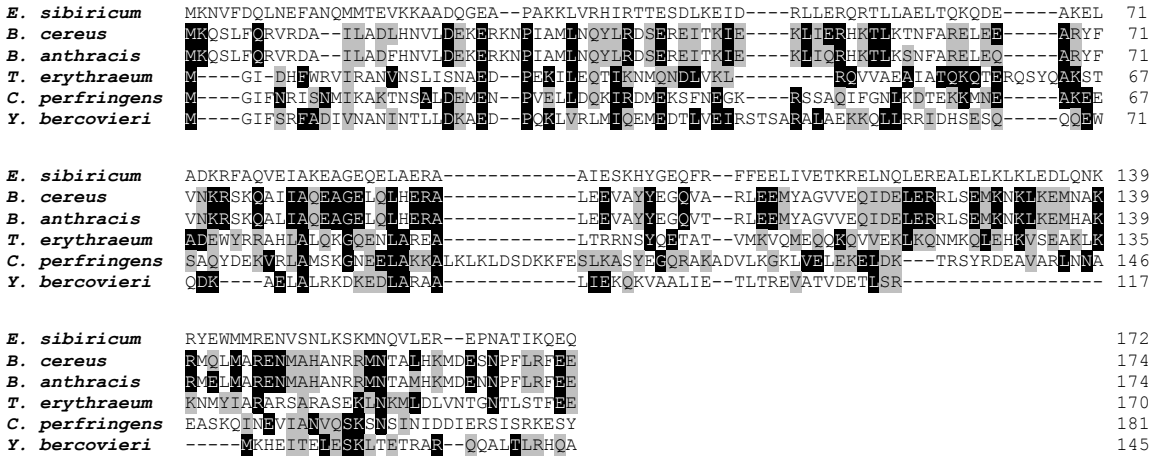

Figure 3.6 Alignment of hypothetic protein of 24460 Da in *E. sibiricum* 255-15 with putative homologous phage shock protein A from *Bacillus cereus*, *Bacillus anthracis*, *Trichodesmium erythraeum*, *Clostridium perfringens* and *Yersinia bercovieri*. Identical residues are boxed in black and similar residues are boxed in gray.

3.6 References:

[1] Storad, B. C., *ASU Research* 1990, 5, 22-25.

[2] Harris, S. A., *The Permafrost Environment*, Barnet's &Noble Books: Towowa New York 1986.

[3] Williams, P. J. and Smith, M. W., *The Frozen Earth: Fundamentals of Geocryology*, Cambridge Univ. Press, New York 1989, p. 306.

[4] Vishnivetskaya, T., Kathariou, S., McGrath, J., Gilichinsky, D. and Tiedje, J. M., *Extremophiles* 2000, 4, 165-173.

[5] Vorobyova, E., Soina, V., Gorlenko, M., Minkovskaya, N.*, et al.*, *FEMS Mcirobiol. Rev.* 1997, 20, 277-290.

[6] Gilichinsky, D. A., Vorobyova, E. A., Erokhina, L. G., Fyordorov-Dayvdov, D. and Chaikovskaya, N., *Adv. Space Res.* 1992, 12, 255-263.

[7] Zvyagintsev, D. G., Gilichinskii, D. A., Blagodatskii, S. A., Vorob'eva, E. A.*, et al.*, *Microbiology* 1985, 54, 131-136.

[8] Khlebnikova, G. M., Gilichinskii, D. A., Fedorov-Davydov, G.;, D. and Vorob'eva, E. A., *Microbiology* 1990, 59, 106-112.

[9] Vishnivetskaya, T. A., Siletzkya, R., Jefferiesa, N., Tiedjeb, J. M. and Kathariou, S., *Cryobiology* 2007, 54, 234-240.

[10] Vishnivetskaya, T. A., Petrova, M. A., Urbance, J., Ponder, M.*, et al.*, *Astrobiology* 2006, 6, 400-414.

[11] Ponder, M. A., Gilmour, S. J., Bergholz, P. W., Mindock, C. A.*, et al.*, *FEMS Micriobiol. Ecol.* 2005, 53, 103-115.

[12] Qiu, Y., Kathariou, S. and Lubman, D. M., *Proteomics 33* 2006, 5221-5233.

[13] Jensen, O. N., Podtelejnikov, A. V. and Mann, M., *Anal. Chem.* 1997, 69, 4741-4750.

[14] Chintalapati, S., Kiran, M. D. and Shivaji, S., *Cell Mol Biol* 2004, 50, 631-642.

[15] Thammavongs, B., Corroler, D., Panoff, J. M., Auffray, Y. and Boutibonnes, P., *Lett. Appl. Microbiol.* 1996, 23, 398-402.

[16] Willimsky, G., Bang, H., Fischer, G. and Marahiel, M., *J. Bacteriol.* 1992, 174, 6326-6335.

[17] Goldstein, J., Pollitt, N. S. and Inouye, M., *Proc. Natl. Acad. Sci USA* 1990, 87, 283-287.

[18] Kim, W. S. and Dunn, N. W., *Curr. Microbiol.* 1997, 35, 59-63.

[19] Panoff, J. M., Thammavongs, B., Laplace, J. M., Hartke, A.*, et al.*, *Cryobiology* 1995, 32, 516-520.

[20] Kim, W. S., Khunajakr, N. and Dunn, N. W., *Cryobiology* 1998, 37, 86-91.

[21] Wouters, J. A., Jeynov, B., Rombouts, F. M., de Vos, W. M., *et al.*, *Microbiology* 1999, 145, 3185-3194.

[22] Wouters, J. A., Frenkiel, H., De Vos, W. M., Kuipers, O. P. and Abee, T., *Appl. Environ. Microbiol.* 2001, 67, 5171-5178.

[23] Sardesai, N. and Babu, C. R., *Folia Microbiol (Praha).* 2000, 45, 177-182.

[24] Wouters, J. A., Mialhes, M., Rombouts, F. M., De Vos, W. M., *et al.*, *Appl. Environ. Microbiol.* 2000, 66, 3756-3763.

[25] Galkin, M. A., Ishmukhametov, R. R. and Vik, S. B., *Biochim Biophys Acta.* 2006, 1757, 206-214.

[26] Bakermans, C., Ayala-del-Río, H. L., Ponder, M. A., Vishnivetskaya, T., *et al.*, *Int. J. Syst. Evol. Microbiol.* 2006, 56, 1285-1291.

[27] Pooley, L. and Brown, T. A., *FEMS Mcirobiol. Lett.* 1991, 77, 229-232.

[28] Korona, R., Nakatsu, C. H., Forney, L. J. and Lenski, R. E., *Proc. Natl. Acad. Sci USA* 1994, 91, 9037-9041.

[29] Wang, Y., Delettre, M., Guillot, A., Corrieu, G. and Beal, C., *Cryobiology* 2005, 50, 294-290.

[30] Arn, E. A. and Abelson, J. N., *J. Biol. Chem.* 1996, 271, 31145-31153.

[31] Sato, M. and Takahash, H., *J. Gen. Appl. Microbiol.* 1971, 16, 279-290.

[32] Senft, A. W., Miech, R. P., Brown, P. R. and Senft, D. G., *Int. J. Parasitol* 1972, 2, 249-260.

[33] Dovey, H. F., McKerrow, J. H. and Wang, C. C., *Mol. Biochem. Parasitol.* 1984, 11, 157-167.

[34] Lee, C. C., Craig, S. P. I. and Eakin, A. E., *Biochemistry* 1998, 37, 3491-3498.

[35] Hochstadt-Ozer, J. and Stadtman, E. R., *J. Biol.Chem* 1971, 246, 5294-5303.

[36] Cai, Q., Ferraris, J. D. and Burg, M. B., *Am. J. Physiol. Renal. Physiol.* 2004, 286, F58-67.

[37] Beinert, H., Holm, R. H. and Munck, E., *Science* 1997, 277, 653-659.

[38] Rangachari, K., Davis, C. T., Eccleston, J. F., Hirst, E. M., *et al.*, *FEBS Lett.* 2002, 514, 225-228.

[39] Young, J. and Holland, I. B., *Biochim Biophys Acta.* 1999, 1461, 177-200.

[40] Zheng, M., Wang, X., Templeton, L. J., Smulski, D. R., *et al.*, *J. Bacteriol.* 2001, 183, 4562-4570.

[41] Nachin, L., Loiseau, L., Expert, D. and Barras, F., *EMBO J.* 2003, 22, 427-437.

[42] Dai, S., Saarinen, M., Ramaswamy, S., Meyer, Y., *et al.*, *J.Mol.Biol.* 1996, 264, 1044-1057.

[43] Williams, C. H., in: Muller, F. (Ed.), *Chemistry and Biochemistry of Flavoenzymes*, CRC Press, Boca Raton 1992, pp. 121-211.

[44] Brown, D. M., Upcroft, J. A. and Upcroft, P., *Mol. Biochem. Parasitol.* 1996, 83, 211-220.

[45] Ben-Menachem, G., Himmelreich, R., Herrmann, R., Aharonowitz, Y. and Rottem, S., *Microbiology* 1997, 143, 1933-1940.

[46] Brissette, J. L., Russel, M., Weiner, L. and Model, P., *Proc. Natl. Acad. Sci USA* 1990, 87, 862-866.

[47] Model, P., Jovanovic, G. and Dworkin, J., *Mol. Microbiol.* 1997, 24, 255-261.

[48] Kleerebezem, M. and Tommassen, J., *Mol. Microbiol.* 1993, 7, 947-956.

[49] Brissette, J. L., Weiner, L., Ripmaster, T. L. and Model, P., *J.Mol.Biol.* 1991, 220, 35-48.

[50] Bergler, H., Abraham, D., Aschauer, H. and Turnowsky, F., *Microbiology* 1994, 140, 1937-1944.

[51] Kleerebezem, M., Crielaard, W. and Tommassen, J., *EMBO J.* 1996, 15, 162-171.

[52] Wang, N., Yamanaka, K. and Inouye, M., *J. Bacteriol.* 1999, 181, 1603-1609.