APPENDIX

SUPPORTING PAPER

1. **Topanurak, S**., Sinchaikul, S., Sookkheo, B., Phutrakul, S., Chen, S. T., Functional proteomics and correlated signaling pathway of the thermophilic bacterium *Bacillus stearothermophilus* TLS33 under cold-shock stress. *Proteomics*, 2005, 5, 4456-4471.
2. **Topanurak, S**., Sinchaikul, S., Phutrakul S., Sookkheo, B., Chen, S. T. Proteomics viewed on stress response of thermophilic bacterium *Bacillus stearothermophilus* TLS 33. *Proteomics*, 2005, 5, 3722-3730.
3. Sinchaikul, S., Sookkheo, B., **Topanurak, S.**, Juan, H.F., Phutrakul, S., Chen, S. T. Bioinformatics, functional genomics and proteomics study of *Bacillus* sp. *J. Chromatogr. B*., 2002, 771, 261-287.
4. Sinchaikul, S., Sookkheo, B., **Topanurak, S**., Pan, F.M., Phutrakul, S., Chen, S. T. Functional and Structural Analysis of *Bacillus* proteome. *Current Proteomics*, 2005, 2, 109-145.
5. **Topanurak, S.,** Sinchaikul, S., Phutrakul, S., Sookkheo, B., Chen, S.T., Differential Gene Expression in Proteome Level of the Thermophillic Bacterium *Bacillus stearothermophilus* TLS33 in Environmental Cold stress. *Thai J. of Biotechnol*., 2005, 6, 6-15

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

REGULAR ARTICLE

Functional proteomics and correlated signaling pathway of the thermophilic bacterium *Bacillus stearothermophilus* TLS33 under cold-shock stress

Supachai Topanurak[1, 2], Supachok Sinchaikul[1], Boonyaras Sookkheo[3], Suree Phutrakul[2] and Shui-Tein Chen[1, 4]

[1] Institute of Biological Chemistry and Genomics Research Center, Academia Sinica, Taipei, Taiwan
[2] Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand
[3] Department of Chemistry, Faculty of Science, Rangsit University, Pathumthani, Thailand
[4] Institute of Biochemical Sciences, College of Life Science, National Taiwan University, Taipei, Taiwan

The thermophilic bacterium *Bacillus stearothermophilus* TLS33 was examined under cold-shock stress by a proteomic approach to gain a better understanding of the protein synthesis and complex regulatory pathways of bacterial adaptation. After downshift in the temperature from 65°C, the optimal growth temperature for this bacterium, to 37°C and 25°C for 2 h, we used the high-throughput techniques of proteomic analysis combining 2-DE and MS to identify 53 individual proteins including differentially expressed proteins. The bioinformatics database was used to search the biological functions of proteins and correlate these with gene homology and metabolic pathways in cell protection and adaptation. Eight cold-shock-induced proteins were shown to have markedly different protein expression: glucosyltransferase, anti-sigma B (σ^B) factor, Mrp protein homolog, dihydroorthase, hypothetical transcriptional regulator in FeuA-SigW intergenic region, RibT protein, phosphoadenosine phosphosulfate reductase and prespore-specific transcriptional activator RsfA. Interestingly, six of these cold-shock-induced proteins are correlated with the signal transduction pathway of bacterial sporulation. This study aims to provide a better understanding of the functional adaptation of this bacterium to environmental cold-shock stress.

Received: March 1, 2004
Revised: January 6, 2005
Accepted: January 16, 2005

Keywords:
Bacillus stearothermophilus / Cold-shock / Signaling pathway / Sporulation / Thermophile

1 Introduction

Prokaryotic microorganisms can be classified according to their temperature ranges of growth. Thermophiles are heat-loving microorganisms with an optimum growth above 50°C, whereas mesophiles grow between 10°C and 50°C [1]. Psychrophiles and psychrotrophs grow at or near 0°C but are differentiated by their optimal and maximal growth temperatures. Despite their differences, common physiological changes occur within all these microorganisms in response to temperature change [1–3]. Study of the cold-shock response is important for understanding cell growth at, or cell tolerance to, low temperature. A downshift in temperature causes a transient induction of a large number of proteins, termed cold-induced proteins (CIPs), but also repression of many proteins that are synthesized under normal condition [4]. Some of these cold-induced proteins are essential for various levels of cellular physiology including metabolism, transcription, translation and protein folding [5, 6]. The most extensive studies on low-temperature adaptation of prokar-

Correspondence: Dr. Shui-Tein Chen, Institute of Biological Chemistry (RM707), Academia Sinica, 128 Yen Chiu Yuan Rd., Sec. II, Nankang, Taipei, 11529, Taiwan
E-mail: bcchen@gate.sinica.edu.tw
Fax: +886-2-27883473

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

WILEY InterScience® DISCOVER SOMETHING GREAT

yotes have been conducted with *E. coli* [7–12] and *B. subtilis* [13–16]. Adaptation and protein expression in response to cold-shock stress has been investigated for several bacteria, such as *B. caldolyticus* [17], *B. caldotenex* [18], *Helicobacter pylori* [19], *Mycoplasma genitalium* [19], *Thermotoga maritime* [20], and *B. stearothermophilus* [21, 22]. Proteomic analysis combining 2-DE, MS and bioinformatics is a powerful approach for studying the alteration of protein expression in organisms under different environmental conditions. These techniques have also been applied to explore structures or functions of specific proteins from microorganisms in sequenced genomes, such as *E. coli* [23, 24], *B. subtilis* [25, 26], and *Saccharomyces cerevisiae* [27].

In this study, we used a proteomic approach combining 2-DE and MS to analyze the differentially expressed proteins of the thermophilic bacterium *B. stearothermophilus* TLS33 under cold-shock stress. After downshift in the temperature from the optimal growth temperature at 65°C to 37°C and 25°C, 2-DE results showed the differential protein expressions at these three temperatures, and image analysis software was used to compare the protein spots on 2-D images and to detect the spots representing differentially expressed protein. We identified a number of protein spots that were differentially expressed when the bacterium encountered in the cold-shock environment. We then used the bioinformatics database to search for protein and gene correlation, biological functions and the relationship of the proteins in signaling pathway of sporulation. Interestingly, we found that eight cold-shock-induced proteins, which differ from those in *B. subtilis* and *E. coli*, play an important role in cell adaptation under cold-shock stress. Moreover, we suggest a model of sigma (σ) signal transduction network correlated with the observed cold-shock-induced proteins in an attempt to further understand the bacterial adaptation under cold-shock stress.

2 Materials and methods

2.1 Bacterial growth and cold-shock experiment

The thermophilic bacterium *B. stearothermophilus* TLS33 was isolated from soil at hot springs in Chiang Mai, Thailand. The bacteria were cultured in 50 mL media containing 0.1% yeast extract, 0.1% tryptone and 0.1% base mixture pH 7.2 (0.8 g/L Titriplex I, 0.36 g/L $CaSO_4.2H_2O$, 1 g/L $MgCl_2.6H_2O$, 0.2 g/L NaOH and 4 mL 0.01 M iron III citrate) [28]. The cultures were incubated in a water bath at 65°C with shaking at 200 rpm (in triplicate for each experiment). After reaching an OD_{620} of 0.5 (the mid-log phase of this bacterium), cold-shock was initiated by transferring the culture flasks to the water baths at 37°C or 25°C. The cell growth profile under cold-shock stress was examined for up to 8 h. For all analytical procedures, the cells cultured 65°C (control), 37°C and 25°C were harvested by centrifugation at 12 000 × *g* at 4°C for 30 min.

2.2 Sample preparation

Cell pellets were resuspended in TE buffer, containing 20 mM Tris-HCl pH 8.0 and 10 mM EDTA, and disrupted on ice by sonication (five cycles of 45 s). Cell debris was removed by centrifugation at 12 000 × *g* at 4°C for 20 min. The sample solutions were precipitated by addition of 10% TCA and 0.1% DTT and stored overnight at −20°C to precipitate the proteins and to remove the salt and nucleic acid. The suspension was centrifuged at 20 000 × *g* for 30 min. Protein pellets were resuspended in ice-cold acetone containing 0.1% DTT and stored at −20°C for at least 30 min. The protein suspension was centrifuged again at 20 000 × *g* for 30 min and the pellets were resuspended in ice-cold acetone without DTT. The protein suspension was stored at −20°C for 30 min and then centrifuged at 20 000 × *g* for 30 min. The obtained pellets were immediately dissolved in lysis buffer (7 M urea, 2 M thiourea, 4% CHAPS, 1% IPG buffer pH 4–7 and 65 mM DTT) [29]. Each sample was sonicated and centrifuged, and the protein concentration was determined using a PlusOne™ 2-D Quant Kit (Amersham Biosciences, Uppsala, Sweden) with bovine albumin (BSA) as a standard.

2.3 2-DE analysis

Each sample was applied onto IPG strips (18 cm, pH 4–7 L; Amersham Biosciences) with a final concentration of 100 μg protein in 350 μL. IPGphor IEF (Amersham Biosciences) was performed under following condition: IPG strips were rehydrated passively and/or actively for 12 h at 30 V followed by ramping to 250 V for 1 h, 500 V for 1 h, 1000 V for 1 h, 3500 V for 3 h, and focusing at 8000 V for up to 7.5 h. After IEF, the IPG strips were equilibrated in equilibration buffer I (50 mM Tris-HCl pH 8.8, 6 M urea, 30% glycerol, 2% SDS, 2% DTT and a trace of bromophenol blue) for 15 min, and then subsequently alkylated in equilibration buffer II (50 mM Tris-HCl pH 8.8, 6 M urea, 30% glycerol, 2% SDS, 2.5% w/v iodoacetamide and a trace of bromophenol blue) for 15 min. Each equilibrated IPG strip was placed on top of the linear gradient of 10–20% polyacrylamide gel (185 × 200 × 1.5 mm) and covered with 0.5% agarose. The second-dimensional separation was performed using the Protean xi Multi-Cells (Bio-Rad) and carried out at 45 mA *per* gel at 15°C until the bromophenol blue dye front reached the bottom of the gel. At the end of each run, the 2-D gels were fixed in 10% methanol/7% acetic acid for 30 min and subsequently stained using the SYPRO Ruby method [30]. The stained gels were scanned using a Typhoon 9200 image scanner (Amersham Biosciences).

2.4 Comparative image analysis

The statistical data (spot detection, spot editing, pattern matching, up- and down-regulations) were acquired and analyzed using the ImageMaster 2D elite software package (Amersham Biosciences) with a high image quality TIF for-

Copyright © by Chiang Mai University
All rights reserved

mat (600 dpi). Statistical analyses were performed in triplicate of three gels from each experimental growth condition to determine spots showing reproducible changes. For gel-to-gel comparison, the 2-D image of cell at 65°C was set as the reference gel image. Before matching the images, background was subtracted (using the lowest-on-boundary method) and normalization was performed to correct for the differences in protein spot intensity. The reference gel image was matched to another gel image, and the matching was manually edited to ensure correct spot matches and to ensure more consistent determination of spot volume. Matched spot data (spot number and volume) was exported to an Excel table (Microsoft) for calculation of spot volume and construction of graphs to compare protein expression. The quantification of each spot was expressed as percent volume, where %V = spot volume/volumes of all spots resolved in the gels.

2.5 Protein digestion

Protein spots were manually excised from the polyacrylamide gels and transferred to 500-μL siliconized Eppendorf tubes. The gel pieces were washed twice with 200 μL 50% ACN/25 mM ammonium bicarbonate buffer pH 8.0 for 15 min each. The gel pieces were then washed once with 200 μL 100% ACN and dried using a SpeedVac concentrator. Dried gel pieces were swollen in 10 μL 25 mM ammonium bicarbonate containing 0.1 μg trypsin (sequencing grade; Promega, Madison, WI, USA). Gel pieces were then crushed with a siliconized blue stick and incubated at 37°C for at least 16 h. Peptides were subsequently extracted twice with 50 μL 50% ACN/5% TFA, and the extracted solutions were combined and dried using SpeedVac concentrator. The peptides or pellets were then resuspended in 20 μL 0.1% TFA and the suspended solutions were purified using ZipTip C18 (Millipore, Bedford, MA, USA). Briefly, 10 μL sample was drawn up and down in the ZipTip ten times, and the ZipTip was washed with 10 μL 0.1% formic acid by drawing up and expelling the washing solution three times. The peptides were eluted with 5 μL 75% ACN/0.1% formic acid.

2.6 MALDI-MS analysis

The samples were mixed in a ratio of 1:1 with matrix solution (5 mg/mL CHCA in 50% ACN, 0.1% TFA and 2% ammonium citrate) and spotted onto the 96-well plate of a PerSeptive Biosystems Voyager DE-RP MALDI-TOF mass spectrometer (PerSeptive Biosystems, Framingham, MA, USA). Adrenocorticotropic hormone (ACTH), 1 pmol/μL, was used as an external calibration. MS analysis was performed using an acceleration voltage of 20 kV [31]. PMF spectra obtained from each digested protein were searched against protein PMF databases *via* the programs Protein Prospector MS-FIT (http://prospector.ucsf.edu/ucsfhtml4.0/msfit.htm) and MASCOT (http://www.matrixscience.com). Search parameters allowed for oxidation of methionine, carbamidomethylation of cysteine, one missed trypsin cleavage, and 50 ppm mass accuracy. Peptides in the mass range of 1000–3500 Da were selectively searched in the database. The remaining autodigestion trypsin and keratin peaks were removed from the mass list before database searching. Protein identification was repeated at least once more using spots from different gels.

2.7 Bioinformatics for functional properties

The SubtiList (http://genolist.pasteur.fr/SubtiList) supplementing with EMBL/GenBank/DDBJ databanks, COGs (http://www.ncbi.nlm.nih.gov/COG), Swiss-Prot (http://us.expasy.org/sprot), TrEMBL (http://www.expasy.ch/sprot) and Micado (http://www-mig.jouy.inra.fr/bdsi/Micado) were searched for sequence homology and biological functions corresponding to identified proteins. In addition, the DIP database (http://dip.doe-mbi.ucla.edu) was used for determining protein-protein interactions.

3 Results and discussion

3.1 Growth profile of *B. stearothermophilus* TLS33 under cold-shock

To investigate transient metabolic adaptation for bacterial survival and its physiological activity, the proteome of thermophile *B. stearothermophilus* TLS33 under cold-shock stress was studied immediately after transferring the bacterial cultures from their optimal growth temperature at 65°C to 37°C or 25°C, which represent the temperature of normal bacterial growth temperature and room temperature, respectively. The growth profile of this bacterium is shown in Fig. 1. We chose the mid-log phase of the bacterial growth at 24 h (OD_{620} ~0.5, T=0 h) for cold-shock experiment because at this time

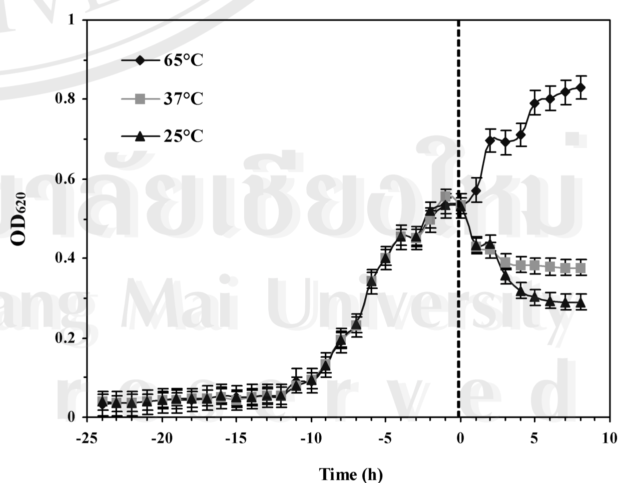


Figure 1. Growth profile of *B. stearothermophilus* TLS33 under cold-shock stress at 37°C and 25°C cultured for 32 h. Cold-shock stress was induced after the 24 h of culture (T = 0 h, starting time of cold-shock stress) for 8 h (T = 8 h).

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

bacterial cells begin to synthesize differential proteins and other components, such as primary and secondary metabolites, leading to more resistance to the non-optimal environmental conditions. Within the first 2 h after cold-shock at 37°C or 25°C, the bacterial growth rate decreased, and reached a constant rate by 4 h. In contrast, bacterial growth at 65°C increased over the 8-h period (T=+8 h). Thus, 2 h after cold shock was considered the appropriate time point for proteomic analysis of this bacterium response to cold-shock stress in this study. We investigated a large number of proteins for which the synthesis and expression levels were changed after cold-shock, including those involved in bacterial adaptation to cold-shock stress and sporulation [28, 32–35].

3.2 2-DE analysis

The cell extracts obtained after a temperature downshift from 65°C to 37°C or 25°C for 2 h or after incubation at 65°C for 2 h (control) (all performed in triplicate) showed reproducible 2-D gel patterns, analysis of which demonstrated the differential protein synthesis of this bacterium under cold-shock stress (Fig. 2). The protein pattern at 65°C showed a number of proteins within a narrow p*I* range of 4–7 and with molecular mass of more than 25 kDa, while the protein patterns at 37°C and 25°C showed fewer proteins. Few proteins with a p*I* in the range 3–4 or 7–11 were observed in any cell extract from any of the experimental temperatures. Thus, we focused our analysis on the proteins in the narrow p*I* range of 4–7, similar to other reports showing that a number of proteins in *B. subtilis* were restricted to a rather narrow p*I* range of 4–7 and a molecular mass range of 5–100 kDa [35–37]. Moreover, the 2-D gel patterns were shown to have a low resolution of protein separation, leading to a low yield of proteins. This may be caused by the low solubility of intracellular proteins from thermophilic cells, which may be due to (1) the high residue hydrophobicity and more charged amino acids, especially Glu, Arg and Lys, necessary to conserve its protein function in high temperature, and (2) the proteins in the cell extracts were precipitated by TCA, which may have resulted in protein aggregation due to the low pH and difficulty with resolubilizing protein precipitate completely. However, we attempted to resuspend the precipitated proteins in lysis buffer homogeneously prior to protein analysis using 2-DE. Although the protein distributions at the three different temperatures appeared similar on 2-D gel patterns, the exact number of protein spot for each temperature were compared on triplicate 2-D gels and quantitated using ImageMaster 2D elite software. The total number of protein spots on 2-D gels detected after culture at 65°C, and after cold shock at 37°C and 25°C were 191, 116 and 115, respectively. The pair-wise comparisons of protein spots from the different temperature groups: (1) 37°C and 65°C, (2) 25°C and 65°C, and (3) 37°C and 25°C, showed the matched proteins of 66, 57 and 75 spots with 34.55%, 29.84% and 64.65% matching, respectively. A twofold increase or decrease in protein synthesis of the differentially expressed protein under cold-shock stress was classified as up-regulation or down-regulation, respectively. The pair-wise comparisons of the proteins in the different temperature groups 1–3,

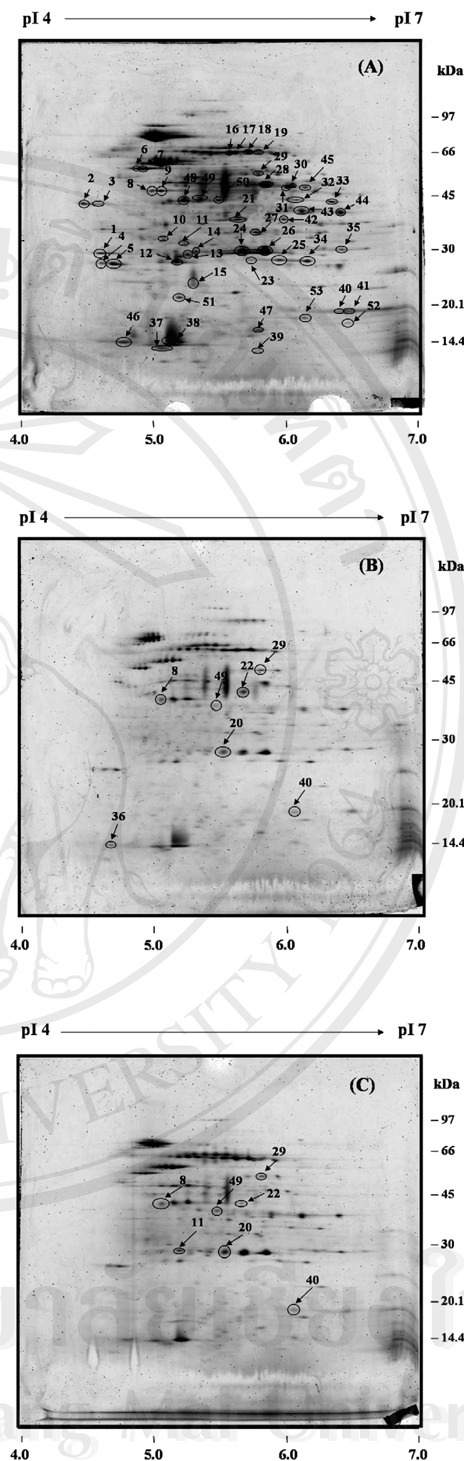


Figure 2. 2-D gel images of cell extracts from *B. stearothermophilus* TLS33 at three different temperatures; (A) 65°C, (B) 37°C, (C) 25°C. The arrows show the spots of proteins, which were identified by MALDI-TOF MS.

Copyright © by Chiang Mai University
All rights reserved

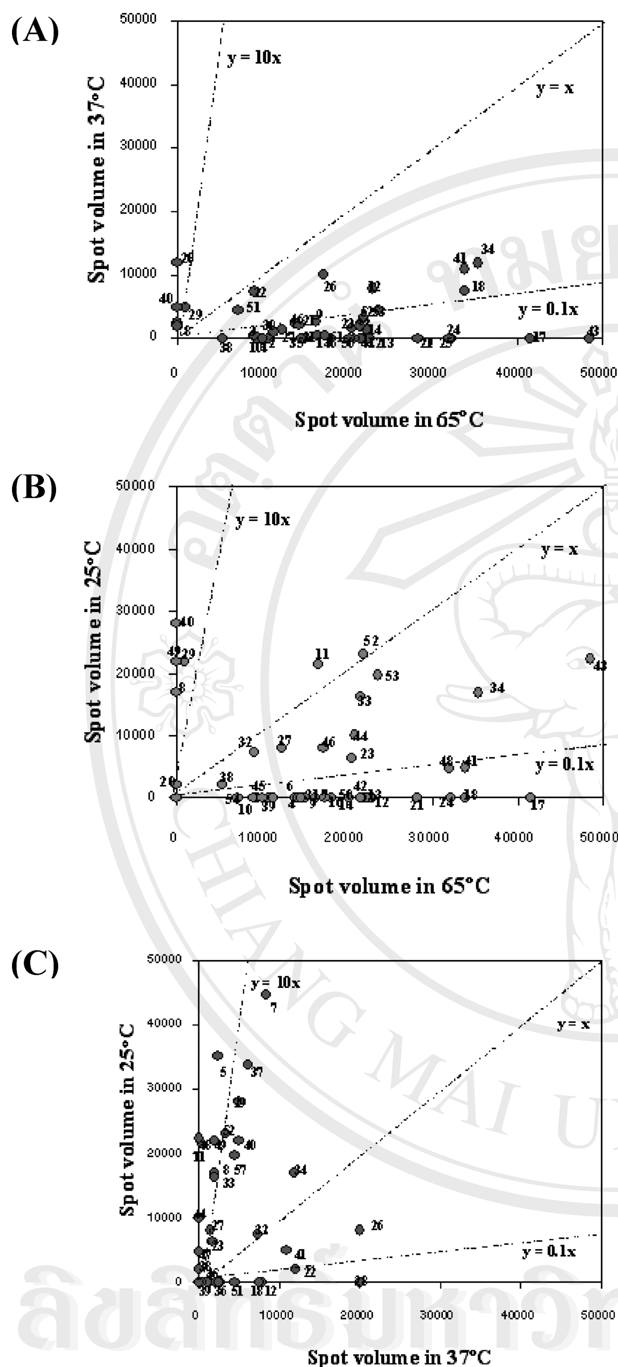


Figure 3. The increase and decrease of protein levels from *B. stearothermophilus* TLS33 after temperature downshift from 65°C to 37°C and to 25°C, demonstrated using ImageMaster 2D Elite software. (A) Comparison between the temperatures at 37°C and 65°C; (B) comparison between the temperatures at 25°C and 65°C; (C) comparison between the temperatures at 25°C and 37°C.

as mentioned above, showed 44, 25 and 10 up-regulated proteins, and 20, 29 and 12 down-regulated proteins, respectively. Furthermore, the differential protein levels were also evaluated using a linear scatter plot, in which y = 10x (up-regulation or increase of protein abundance), y = x (equal) and y = 0.1x (down-regulation or decrease of protein abundance), as shown in Fig. 3A–C. Fifty-three intracellular proteins were detected in the cells of this thermophilic bacterium at these three temperatures and these might represent the housekeeping proteins. Interestingly, only eight major cold-shock-induced proteins were markedly changed by the cold-shock stress.

3.3 Protein identification and differential protein synthesis

Individual protein spots were excised from the 2-D gels, subjected to in-gel digestion and analyzed by MALDI-TOF MS. A total of 53 intracellular proteins were recognized on the basis of the tryptic mass profile comparisons using the SwissProt and NCBI non-redundant database with MS-FIT and MASCOT software (Table 1). In addition, the SubtiList database (http://genolist.pasteur.fr/SubtiList) and EMBL/GenBank/DDBJ databases were used to determine the correlations at the genome level, accession numbers, description and functional categories of the identified proteins (Table 2). Moreover, the intracellular proteins from *B. stearothermophilus* TLS33 were compared with the complete genomes of *B. subtilis* and *B. halodurans* based on the COG (Clusters of Orthologous Groups of protein) database using the COGNITOR program (http://www.ncbi.nlm.nih.gov/COG), which represents a phylogenetic classification in functional annotation (Table 3). The results of the database search showed that most of intracellular proteins from *B. stearothermophilus* TLS33 had biological functions related to cellular process and metabolism, indicating the adaptation or maintenance of this bacterium under cold-shock stress. We found that eight cold-shock-induced proteins had markedly changed at the different temperatures: glucosyltransferase, anti-sigma B (σ^B) factor (RsbT), Mrp protein homolog, dihydroorthase, hypothetical transcriptional regulator in the FeuA-SigW intergenic region, RibT protein, phosphoadenosine phosphosulfate reductase, and prespore-specific transcriptional activator. Only glucosyltransferase, Mrp protein homolog, dihydroorthase, RibT protein, phosphoadenosine phosphosulfate reductase, and prespore-specific transcriptional activator were observed at 37°C and 25°C. Using Image Master 2D Elite software, we also classified the major cold-shock-induced proteins into three groups, depending on the protein synthesis level at different temperatures (Fig. 4). The first group of the protein synthesis at 37°C and 25°C comprised glucosyltransferase, Mrp protein homolog, dihydroorthase, phosphoadenosine phosphosulfate reductase, and prespore-specific transcriptional activator RsfA. The second group of the protein synthesis at only 37°C comprised RibT protein. The last group of the protein synthesis at 65°C and 25°C comprised anti-σ^B factor (RsbT) and the hypothetical transcriptional regulator in FeuA-SigW intergenic region. In addition, the cold-shock-induced proteins at 37°C were Mrp protein homolog, dihydroorthase and RibT pro-

Copyright © by Chiang Mai University
All rights reserved

Table 1. Protein identification of intracellular proteins in *B. stearothermophilus* TLS33, after temperature downshift to 37°C and 25°C, by MS-FIT and MASCOT software, based on Swiss-Prot database (http://www.expasy.ch/sprot)

Spot no.	Protein name	Accession no.	Mol. mass (Da) / p*I*	No. peptide matched	4 sample matched peptides
1	Pts system, fructose-specific IIB component (EIIB-FRU) (Fructose-permease IIB component) (Phosphotransferase enzyme II, B component) (P18) (PTFB)	P26380	35 051 / 4.37	9	MMNIVLAR IDDRFIHGQILTRWIK VHAA-DRIIVVSDDIAQDEMR KTLILSVAPSNVK
2	Stage III sporulation protein AE (S3AE)	P49782	39 191 / 4.55	7	TAASLETDK IGEFWNDIMTEYGGLLPESQK GSLME-FINGDK SFSPQEWLK
3	Inosine-5′-monophosphate dehydrogenase (IMP dehydrogenase) (IMPDH) (IMPD) (Superoxide-inducible protein 12, SOI12) (IMDH)	P21879	39 191 / 4.62	12	DVDLSVELTKTLK GVITNPFFLTPDHQVFDAEHLMGK SGVPIVNNEEDQK LVGIITNRDLRFISDYSMK
4	COMF operon protein 3 (CMF2)	P39147	29 901 / 4.76	9	ALFLLKPDEK VCYSCRSKLK EVWRTRIR SDFSSTFSKVYPDK
5	Unknown protein	–	28 100 / 4.64	–	
6	Flagellar protein (FLIT)	P39740	58 100 / 4.64	14	SMLSHIQNTPESDELLK SIATELQQMKR RVMHTTYLNPYNNITIDGTYYDKR QSIA-TELQQMKRKR
7	DNA topoisomerase I (Omega-protein) (Relaxing enzyme) (Untwisting enzyme, Swivelase) (TOP1)	P39814	58 250 / 4.55	26	TIERYLGKKYK SQMGVDIEQNFEPK NPALPFTTSTLQQEAARK EGTVGLITYMR
8	Probable poly (glycerol-phosphate) alpha-glucosyltransferase (Teichoic acid biosynthesis protein) (TAGE)	P13484	48 700 / 4.44	18	QIPDMDYYFISGGLPSNYGGLTK LFGEECNQNTFFLTFR NNEMSVVYYGDGETIR MYLQEYINDQNQVYLDK
9	Translation initiation factor (IF-2)	P17889	48 100 / 4.64	29	NMDLEVNNHMAMLEEK VVEGE-AGGITQHIGAYQIEENGK ITFLDTPGHAAFTTM-RAR AAEVPIIVAVNK
10	Putative peptidase in gcvT-spoIIIAA intergenic region (YQHT)	P54518	36 067 / 5.10	12	YMTGFTGSAGLAVISGDK SSLPHGVASDK TVAVGQPSDQLK EIYQVVFDAQALGVA-HIKPGMTGK
11	Anti-sigma B factor (RSBT)	P42411	34 803 / 5.14	10	MNDQSCVR IMTEWDIVAAR ELGFGTVDQAR IT-TAISELAR QLGR
12	Extragenic suppressor protein suhB homolog (SUHB)	Q45499	31 533 / 5.20	11	KWIREAGAR ITQSMHESLTIETK SNPNDLVTNIDK NFAISIGIFENGEGK
13	Unknown protein	–	33 103 / 5.30	–	
14	Hypothetical 28.6-kDa protein in recQ-cmk intergenic region precursor (YPBG)	P50733	34 187 / 5.53	9	MYATAKGNHLK THTFPLSKMK LVHFGVPIVFVWGNNDYEVR NAAYLISNGYGTTK
15	Unknown protein	–	24 187 / 5.53	–	
16	GTP-binding protein (LEPA)	P37949	69 284 / 5.49	26	NFSIIAHIDHGK EQLLDSMDLER QEVEDVIGLDA-SEAVLASAK VPAPTGDPEAPLK
17	2,3-Bisphosphoglycerate-independent phosphoglycerate mutase (BPG-independent pgaM) (LEPA)	P39773	69 284 / 5.61	12	NETVGNAVALAK AIQISNTFTNKDFR TYINQLNDQIK SALDVVDDSYANGIYDEFVIPSVITK
18	Unknown protein	–	69 486 / 5.72	–	–
19	Peptide methionine sulfoxide reductase (Protein-methionine-s–oxide reductase) (Peptide met (O) reductase) (MSRA)	P54154	69 284 / 5.81	10	EIATFAGGCFWCMVKPFDEQPGIEK VVSGYTGGH-TENPTYEEVCSETTGHR AEPFYEAEGYHQHFYK NPAHYQRYRTGSGR
20	MRP protein homolog (MRP)	P50863	28 628 / 5.55	11	ALERVGVPYFVSEKPEELK EADA-FILPGVGSFGDAMDNLGYTK GKAVRLKAE-DEKGNKLK EPFLQRPLGELDAVK

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Table 1. Continued

Spot no.	Protein name	Accession no.	Mol. mass (Da) / p*I*	No. peptide matched	4 sample matched peptides
21	Hypothetical 58.2-kDa protein in kdgT-xpt intergenic region (YPWA)	P50848	38 614 / 5.59	12	TGAPK ELIDVLYER ELSLYFLQELGYDFDGGR AIFS-NEVSVEDLPSLWNQK
22	Dihydroorotase (Dhoase) (PYRC)	P25995	39 144 / 5.66	14	NGWILNENGEK VTGETITAIGK LDATDNETVIDAK GGYTTVAAMPNTRPVPDTK
23	Transcription anti-termination protein nusG (NUSG)	Q06795	31 924 / 5.63	9	VIDGPFANFTGSIEEIDYDK VFVNMFGR ETPVE-LEFTQIDK ANLEKR
24	Ferrichrome transport ATP-binding protein (FHUC)	P49938	33 229 / 5.67	10	ITTLIGPNGCGK STILKTMSRIMR WA-LEETGMAEYAERPIEALSGGQR EG-TALEVMTPDILK
25	Protein export protein prsA precursor (PRSA)	P24327	30 331 / 5.85	8	EVIAKTDAGDVTK GELYTNMK EQVKYELLTQKAAK EYSTDSSASK
26	Unknown protein	–	32 802 / 6.16	–	
27	Amidotransferase hisH (HIS5)	O34565	35 465 / 5.84	11	GKAVRLKAEDEKGNKLK LSFHNESPLLTK TEQ-GYAYFVHSYYIDGMEENALLASADYGVR SSTVGMSILTQFTKMAAEQKVK
28	Hypothetical 35.0-kDa protein in rapJ-opuAA intergenic region (YCEB)	O34504	43 485 / 5.75	9	RYWFAEHHSTK GLASTAPEIMIAR IAAQTNTIR QLEALYPNR
29	Hypothetical transcriptional regulator ybbB in feuA-sigW intergenic region (ORF3) (YBBB)	P40408	57 121 / 4.55	18	MQNAVIYQPVQIEYLK ETGQDIFTPCTESELAK LMNVSHIENLAVR LQELAALWNESSQLSQLK
30	Unknown protein	–	58 250 / 6.07	–	
31	Glycine betaine/carnitine/choline transport ATP-binding protein opuCA (OPCA)	O34992	54 178 / 5.85	12	CVIEVQSLIYDLFTASLSDQTDTHSAIEK WTSYYYQ-HYSTDIPVQLSAYR LQELAALWNESSQLSQLK MQNAVIYQPVQIEYLK
32	Penicillin-binding protein 1A/1B (PBP1) [Includes: penicillin-insensitive trans-glycosylase (Peptidoglycan Tgase, PBPA)	P39793	49 081 / 6.01	10	AQPGSTIKPILDYGPVIENK QSVDGGSYSEIQNSSAK WLADYDQQTTAAK TGTGQLAQVPGVEVAGK
33	PBSX phage terminase small subunit (XTMA)	P39785	43 484 / 6.14	14	LSLYFDLFPDQFK DSGLVDGTIVTEAK TIGIWK SFNATQSAIK
34	Probable amino-acid ABC transporter ATP-binding protein in bmrU-ansR intergenic region (YQIZ)	P54537	29 604 / 6.05	15	ENIGMVFQHFHLFPHK CLNLLEKPNGGTITIK DTEITKPK EVLQVMK
35	Hypothetical 19.1-kDa protein in sigD-rpsB intergenic region precursor (ORFC) (YLXL)	P40405	32 236 / 6.30	10	MSTLLWLLSFMLHGVLLYAVIILYTR QILEETEN-TLAAFLLELK ASSASQSDEESQK TEIELFLK
36	RIBT protein (RIBT)	P17622	18 139 / 4.60	12	EDEDIVGAIGVEK DYETDTDR HQGIGK QMMDALK
37	PAL-related lipoprotein precursor (SLP)	P39910	18 200 / 4.77	13	AVFPMLIIVFALSGCTLSTI EGWE-DEIETVPTLMVVDQR EDIIKPLQHVLSK DYPE-QIDK
38	Hypothetical protein YVYF (YVYF)	2897794	17 648 / 5.15	14	
39	Unknown protein	P21468	12 806 / 5.44	10	DGFYQIVNVQSDAAAVQEFDR FNNVLTSNGAEITGTK LAYEINDFR ISDDIIR
40	Phosphoadenosine phosphosulfate reductase (PAPS reductase, Thioredoxin dependent) (PADOPS reductase) (3′-Phosphoadenylsulfate reductase) (PAPS sulfotransferase) (CYH1)	P94498	22 031 / 5.95	9	MLTYDNWEEPTITFPEDDPYK KPDLTLEE-QAEEHGDK DAEIVFLDTGLHFK EALSGH-PAWLSGLR
41	Unknown protein	–	20 546 / 6.05	–	
42	Hypothetical 73.2-kDa protein in sodA-comGA intergenic region (YQGS)	P54496	38 549 / 6.14	11	GAVFFTNAGNQYMAAPEILK ALADSNSLTEIE-NYVTANAK DGSFITDQVVYTDGACYDK NVILVS-LESTQSFVINEK

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Table 1. Continued

Spot no.	Protein name	Accession no.	Mol. mass (Da) / p*I*	No. peptide matched	4 sample matched peptides
43	Aspartokinase 2 (Aspartokinase II) (Aspartate kinase 2) [contains: aspartokinase II alpha subunit; Aapartokinase II beta subunit] (AK2)	P08495	40 764 / 6.32	13	ITDIDTSVLADQLEK GGSDTTAVALAAALK GHQVVVVVSAMGK GIAFEDQITR
44	Hypothetical 45.3-kDa protein in prkA-cspB intergenic region (ORF4) (YHBH)	P45742	43 986 / 5.25	17	HVGQGDGESQVGDVVAR TWNDITKPESK GE-SGGTICSSVYR ELELPNLQQK
45	Unknown protein	O07636	56 284 / 5.35	9	GNVTYPITIDPSVWIFDDR NAEPNS-SATQCVFVTSSGK NEESGVSHQFS LMNGSFAMR
46	Hypothetical 17.9-kDa protein in nprE-pycA intergenic region (YLAL)	–	18 200 / 4.77	–	
47	Stage III sporulation protein AH	P49785	17 587 / 5.64	9	TVVTETADDDLFTTYR TQGYEDALVNAEGDK EEL-NAIVSSDDATAK MTALSEVEGTGK
48	(D49467) Unnamed protein product	P39807	56 667 / 5.20	9	DQMDHLNHEDALK DTYYAYNTK LQTVCQACIK MGELANCPK
49	Prespore specific transcriptional activator (RSFA)	P39650	58 758 / 5.34	11	QDAWSEENDLLLAETVLR FLQNYEGNHEQSSALK LVLFEEDEHASPSFK TIQEDYETLVK
50	Glucose-6-phosphate isomerase (GPI) (Phosphoglucose isomerase) (PGI) (Phosphohexose isomerase) (PHI) (G6PI)	P80860	58 515 / 5.52	21	TGAGSDFLGWVDLPEHYDK DVMDLLEDVDFSIN-VISK GNPQVIFIGNNISSSYMR ALTFFPTE-HELTYLR
51	Unknown protein	–	20 249 / 5.07	–	
52	Hypothetical 21.0-kDa protein in lysS-mecB intergenic region (YACH)	P37569	17 461 / 5.70	10	ELESLIHQEEFENAAHVR MICQECHERPATFHFTK VNMFALLGKPGFEEK STDSEEEQEVNK
53	Molybdopterin-gianine dinucleotide biosynthesis protein B	O31704	18 445 / 5.87	9	LIELYQFLETDCLLIEGFK MAL-VRPFPIVQVVGFQNSGK EDLEALQAVNIIAIIYR AAGADVTAVEGAGVLQLTAR

tein, while the cold-shock-induced proteins at 25°C were glucosyltransferase, anti-σB factor, and hypothetical transcriptional regulator in FeuA-SigW intergenic region, phosphoadenosine phosphosulfate reductase and prespore specific transcriptional activator. This result suggested that the temperatures of cold-shock experiment highly influenced the levels of the protein synthesis.

3.4 Cold-shock effect on sporulation signaling pathway of *B. stearothermophilus* TLS33

To gain a better understanding of the bacterial adaptation under cold-shock stress, advanced bioinformatics based on database searches was used to thoroughly search the specific cold-shock-induced proteins and the relationship between their functions and signal transduction pathways of bacterial adaptation when the bacterium was subjected cold-shock stress. Only six proteins, TagE, RsbT, MrpA, PyrC, YbbB and RsfA, were shown to have the functions correlated to the signaling pathway of sporulation. Regarding sporulation in *B. subtilis*, it has been reported that these proteins were also correlated with the σF and σG factors, which are involved in the 'Forespore' stage of the sporulation process [40]. We propose that the functions of the cold-shock-induced proteins in the sporulation signaling pathway corelates to five routes. In the first route, RsbT is up-regulated when the bacterium is cold-shocked at only 25°C. Generally, RsbT controls the early sporulation in vegetative cell cycle of *Bacillus* sp. by coordination with the σ^B factor [41–45]. When the bacterium is exposed to stress, RsbT from the upstream module is triggered to inactivate the principal negative regulator of RsbS by phosphorylation, and then to activate RsbU. Subsequently, PP2C phosphatase can dephosphorylate the phosphorylated form of RsbV (RsbV-P) to its dephosphorylated form (RsbV) in the downstream module (Fig. 5A). RsbV forms a complex with RsbW and forces the release of σ^B. Thus, RsbT is linked to the upstream and downstream modules, which activates RsbU and stimulates its enzymatic properties towards its substrate [46–53]. In the second route, TagE has two activities in stressed cells, in which it activates the PAS-RsbP to form a complex with RsbW, and alternately activates the σ^F/σ^G in the forespore via RsbW/σ^B route (Fig. 5B). Only this route is activated when the bacterium is cold-shocked at 37°C or 25°C. We suggest that this route may be dormant at 65°C because the protein has not been observed. It has been reported that TagE is involved in the activation of PhoP∼P

Copyright © by Chiang Mai University All rights reserved

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 2. Protein and gene identification, description, and functional category of intracellular proteins in *B. stearothermophilus* TLS33, after downshift temperatures at 37°C and 25°C, by SubtiList and EMBL/GenBank/DDBJ databanks and COG databases

Spot no.	Protein ID/no.	Gene name/no.	Description	Functional category
1	PTFB (P26380)	*lev*E (X56098)	PTS fructose-specific enzyme IIB component	Transport/binding proteins and lipoproteins
2	S3AE (P49782)	*spoIII*AE (U35252)	Mutants block sporulation after engulfment	Sporulation
3	IMDH (P21879)	*gua*B (X55669)	Inosine-monophosphate dehydrogenase	Metabolism and transport of nucleotides and nucleic acids (purine biosynthesis)
4	CMF3 (P39147)	*com*FC (Z18629)	Late competence gene	Transformation /competence
5	Unknown	–	–	–
6	FLIT (P39740)	*fli*T (Z31376)	Flagellar protein	Mobility and chemotaxis
7	TOP1 (P39814)	*top*A (L27797)	DNA topoisomerase I	DNA packing and segregation / DNA replication, recombination
8	TAGE (P13484)	*tag*E (X15200)	UDP-glucose: polyglycerol phosphate glucosyltransferase	Cell wall
9	IF2 (P17889)	*inf*B (M34836)	Initiation factor IF-2 (GTPase)	Initiation, translation factors and enzymes involved in translation
10	YQHT (P54518)	*yqh*T (D84432)	Unknown; similar to Xaa-Pro dipeptidase	Protein modification
11	RSBT (P42411)	*rsb*T (L35574)	Positive regulator of sigma-B activity (switch protein/serine-threonine kinase)	Adaptation to atypical conditions
12	SUHB (Q45499)	*suhB* (AF012285)	Archaeal fructose-1,6-biphosphate and related enzyme of inositol monophosphatase family	Carbohydrate transport and metabolism
13	Unknown	–	–	–
14	YPBG (P50733)	*ypb*G (L47648)	Unknown; similar to unknown proteins	From other organisms
15	Unknown	–	–	–
16	LEPA (P37949)	*lep*A (X91655)	GTP-binding protein	Elongation
17	PMGI (P39773)	*pgm* (L29475)	Phosphoglycerate mutase	Carbohydrate transport and metabolism (gluconeo-genesis)
18	Unknown	–	–	–
19	MSRA (P54154)	*msr*A (L77246)	Peptidyl methione sulfoxide reductase	Detoxification, post-translational modification, protein turnover, chaperone
20	MRP (P50863)	*mrp* (X74737)	Multiple resistance and pH homeostasis	Transport/binding proteins
21	YPWA (P50848)	*ypwA* (L47838)	Unknown; similar to carboxypeptidase	Metabolism of amino acids and related molecule
22	PYRC (P25995)	*pyr*C (BG10714)	Dihydroorthase (Pyrimidine biosynthesis)	Metabolism of nucleotides and nucleic acids
23	NUSG (Q06795)	*nus*G (D13303)	Transcription anti-termination factor	Termination
24	FHUC (P49938)	*fhu*C (X93092)	Ferrichrome ABC transporter (ATP-binding protein)	Transport/binding proteins and lipoproteins
25	PRSA (P24327)	*prs*A (X57271)	Protein secretion (post-translocation molecular chaperone), phosphoribosylpyrophosphate synthetase	Protein secretion, nucleotide and amino acid transport and metabolism
26	Unknown	–	–	–

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Table 2. Continued

Spot no.	Protein ID/no.	Gene name/no.	Description	Functional category
27	HIS5 (O34565)	*his*H (AF017113)	Amidotransferase	Amino acid transport and metabolism (histidine bio-synthesis)
28	YCEB (O34504)	*yce*B (AB000617)	Unknown; similar to unknown proteins	From other organisms
29	YBBB (P40408)	*ybb*B (L19954)	Unknown; similar to transcriptional regulator (AraC/XylS family)	Regulation
30	Unknown	–	–	–
31	OPCA (O34992)	*opu*CA (AF009352)	Glycine betaine/carnitine/ choline ABC trans-porter (ATP-binding protein)	Transport/binding proteins and lipoproteins
32	PBPA (P39793)	*pon*A (U11883)	Penicillin-binding proteins 1A/1B	Cell wall
33	XTMA (P39793)	*xtm*A (Z70177)	Phage PBSX terminase (small subunit)	Phage-related functions, DNA replication, recombination and repair
34	YQIZ (P54537)	*yqi*Z (D84432)	Unknown; similar to amino acid ABC transporter (ATP-binding protein)	Transport/binding proteins and lipoproteins
35	YLXL (P40405)	*ylx*L (Z99112)	Unknown; similar to unknown proteins	From *B. subtilis*
36	RIBT (P17622)	*rib*T (L09228)	Reductase	Metabolism of coenzymes and prosthetic groups, riboflavin biosynthesis
37	SLP (P39910)	*slp* (M57435)	Small peptidoglycan-associated lipoprotein, starvation-inducible outer membrane lipoprotein	Transport/binding proteins and lipoproteins, cell envelope biogenesis, outer membrane
38	YVYF (P39807)	*yvy*F (L14437)	Unknown; similar to flagellar protein	Mobility and chemotaxis
39	Unknown	–	–	–
40	CYH1 (P94498)	*cys*H (U76751)	Phosphoadenosine phosphosulfate	Amino acid transport and metabolism, coenzyme metabolism (FAD biosynthesis)
41	Unknown	–	–	–
42	YQGS (P54496)	*yqg*S (D84432)	Unknown; similar to putative molybdate binding protein	From other organisms
43	AK2 (P08495)	*lys*C (J03294)	Aspartokinase II (alpha and beta subunits)	Amino acid transport and metabolism (threonine and methionine biosynthesis)
44	YHBH (P45742)	*yhb*H (Z99108)	Ribosome-associated protein Y (PSrp-1)	Translation, ribosomal structure and biogenesis
45	Unknown	–	–	–
46	YLAL (O07636)	*yla*L (Z99111)	–	–
47	AH (P49785)	*spoIII*AH (Z99116)	–	–
48	Unnamed protein (BAA24873)	–	Unknown; similar to flagellar proteins	Mobility and chemotaxis
49	RSFA (P39650)	*rsf*A (X73124)	Probable regulator of transcription of sigma-F-dependent genes, leucine zipper motif	Sporulation
50	G6PI (P80860)	*pgi* (Z93936)	Glucose-6-phosphate isomerase	Carbohydrate transport and metabolism (glycolysis, gluconeogenesis)
51	YQBP (P45932)	*yqb*P (D32216)	Unknown; similar to phage-related protein	Phage-related functions
52	YACH (P37569)	*yac*H (D26185)	Unknown	–
53	MOBB (O31704)	*mob*B (AF012285)	Molybdopterin-guanine dinucleotide biosynthesis	Metabolism of coenzymes and prosthetic groups

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Table 3. Comparison of the functional systems of the *Bacillus* sp. in protein-coding genes using the COG and SubtiList databases. The pathway and functional systems are classified from COG database using COGNITOR program

Functional systems	*B. subtilis*	*B. halodurans*	*B. stearothermophilus* TLS33[a)]
Information storage and processing			
1. Translation, ribosomal structure and biogenesis	152	153	3 (*inf*B, *yhb*H, *ybb*B)
2. Transcription	272	269	2 (*nus*G, *rsf*A)
3. DNA replication, recombination and repair	131	227	2 (*top*A, *xtm*A)
Cellular processes			
1. Cell division and chromosome partitioning	31	32	1 (mrp)
2. Posttranslational modification, protein turnover, chaperones	87	84	2 (*msr*A, *yqh*T)
3. Cell envelope biogenesis, outer membrane	161	115	6 (slp, *spo*IIIAE, *pon*A, *tag*E, *yqi*Z, *rsf*A)
4. Cell motility and secretion	90	90	3 (*lep*A, *fli*T, *yvy*F)
5. Inorganic ion transport and metabolism	148	146	2 (*fhu*C, *opu*CA)
6. Signal transduction mechanisms	122	135	–
Metabolism			
1. Energy production and conversion	164	158	–
2. Carbohydrate transport and metabolism	271	262	3 (*suh*B, pgm, pgi)
3. Amino acid transport and metabolism	293	284	6 (*prs*A, *his*H, *cys*H, *lys*C, *lev*E, *ypw*A)
4. Nucleotide transport and metabolism	82	73	2 (*prs*A, *gua*B)
5. Coenzyme metabolism	109	109	5 (*fhu*C, *cys*H, *mob*B, *opu*CA)
6. Lipid metabolism	84	89	–
7. Secondary metabolites biosynthesis, transport and catabolism	128	110	–
Poorly characterized			
1. General function prediction only	332	329	3 (*com*FC, *ybb*B, *rsb*T)
2. Function unknown	226	246	5 (*ypb*G, *yce*B, *yqg*S, *yhb*H, *yla*L)
Not in COGs	1221	1155	–

a) Some identification

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

under conditions of phosphate starvation since the cold-stress regulon was under the control of the alternative signaling pathway of σ^B [54–56]. In the third route, YbbB activates FeuA to form a complex with RsbW, leading to release the σ^B from RsbW-σ^B complex and subsequently convert to σ^F or σ^G in the forespore (Fig. 5C). In the fourth route, RsfA is also activated when the bacterium is cold-shocked at 37°C or 25°C. RsfA can directly activate σ^F/σ^G in the forespore (Fig. 5D). The activation of σ^F in the forespore could lead to transcription of SpoIIR and SpoIIQ, and disrupt, without preventing, the formation of stress-resistant spores. σ^G could then activate transcription in the engulfed forespore of a large set of genes [40, 57]. In the fifth route, MrpA and PyrC proteins are correlated in the signaling pathway of bacterial sporulation, in which they are down-regulated when the bacterium is cold-shocked at 37°C or 25°C (Fig. 5E). Both of MrpA and PyrC indirectly activate a transcription factor σ^H, leading to σ^F/σ^G production in the forespore stage. The MRP protein homolog is the multi-resistance protein and functions as an Na^+/H^+ antiporter in pH homeostasis, by which it could influence to the post-translational regulation control of σ^H in the early sporulation in cell cycle [58, 59]. On the other hand, PyrC or dihydroorthase could function in the regulation of pyrimidine biosynthesis and the control of gene expression [60]. According to the correlation of the cold-shock-induced proteins in signaling pathway of sporulation,

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

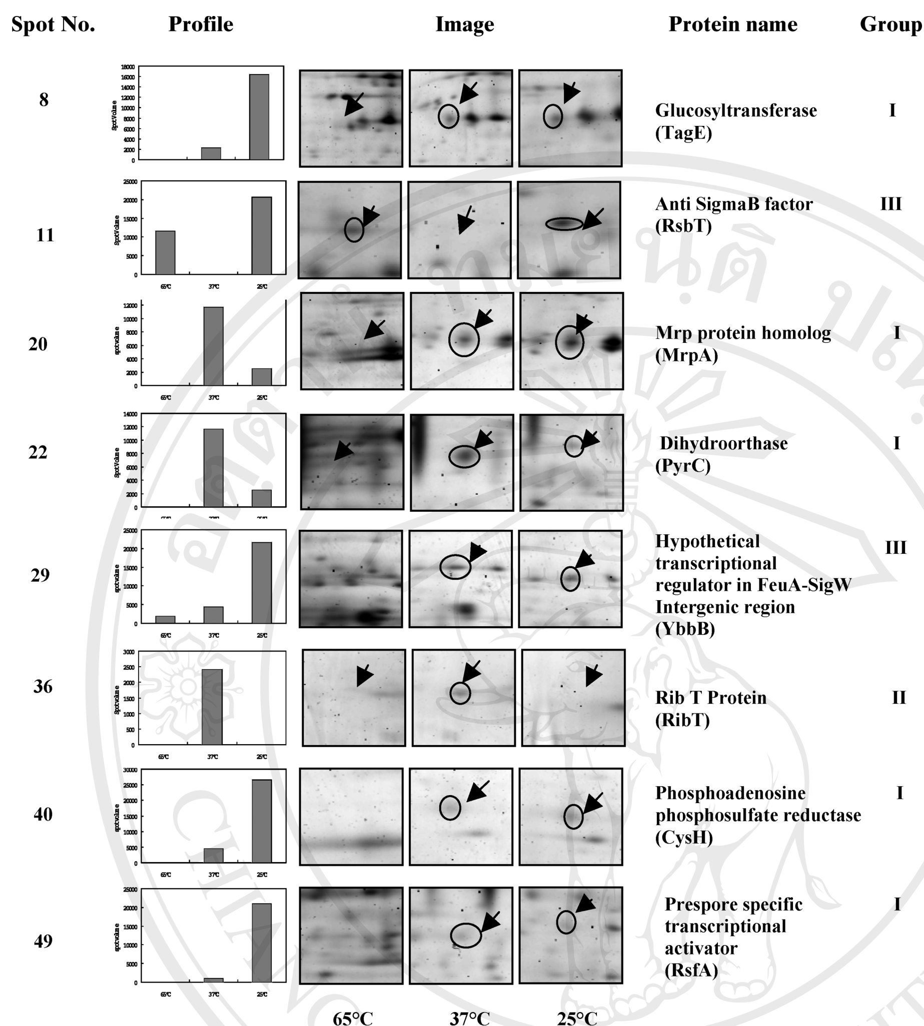


Figure 4. Image analysis of eight major cold-shock-induced proteins of *B. stearothermophilus* TLS33 under cold-shock stress using ImageMaster 2D Elite software. The arrows on 2-D gels show the situation of protein spots. Three groups of cold-shock-induced proteins are classified by ImageMaster 2D Elite software. (I, protein synthesis at 37°C and 25°C; II, protein synthesis at 37°C; III, protein synthesis at 65°C and 25°C).

we presume that up- and down-regulation of these proteins are involved in the σ^F/σ^G production in the forespore (Fig. 6). Thus, the forespore is also considered as an event proceeding bacterial sporulation in the environmental cold-shock stress.

These observations demonstrate the correlation of six major cold-shock-induced proteins when the bacterium encounters cold-shock stress (Fig. 7). TagE, YbbB, RsfA and RsbT, the upstream proteins, are activated in the low temperature at 37°C or 25°C, whereas MrpA and PyrC are inactivated, representing downstream proteins. Surprisingly, RsbT is observed at 65°C and 25°C, but not at 37°C. We suggest that normally, RsbT would be inactive at 65°C, the optimal temperature for this bacterium, and associated with a negative regulator RsbU [49]. To explain the fact that RsbT is not observered by proteomic analysis at 37°C, two hypotheses can be proposed, linking RsbT to upstream and downstream modules of the sporulation signaling pathway. First, the 2-h cold-shock stress at 37°C may initiate the bacterial shock response in which RsbT can form a complex in the upstream module or be subjected to the translational regulation [46]. Second, the temperature at 37°C may not be low enough to induce this cold-shock-induced protein [61]. Although RibT and CysH can be observed at 37°C and 25°C, they are not involved in the signaling pathway of sporulation. RibT protein generally functions as a reductase enzyme in riboflavin or vitamin B_2 biosynthesis and reduction metabolism [62, 63]. However, the function of RibT, which preferentially appeared at only 37°C, has not yet been clarified in *Bacillus* sp. and other bacterial species [64–69]. Likewise, CysH protein, observed at the low temperatures, has also been demonstrated to be involved in the regulation of the sulfur starvation [70]. Thus, RibT and CysH are not part of this signaling pathway. Although other related proteins involved in the TagE, YbbB, RsfA, MrpA and PyrC routes at 65°C can not be identified, this may be due to the limitation of the proteomic analysis in this study. This failure may be caused by post-translational modifications, such as glycosylation or phosphorylation, or by the regulation of gene expression; for example, *tag*E, *ybb*B and *rsf*A genes are turned on at 37°C, while *mrp*A and *pyr*C genes are turned off. If this proposal is valid, it is necessary to determine the unknown genes involved in the upstream module of the

Copyright © by Chiang Mai University
All rights reserved

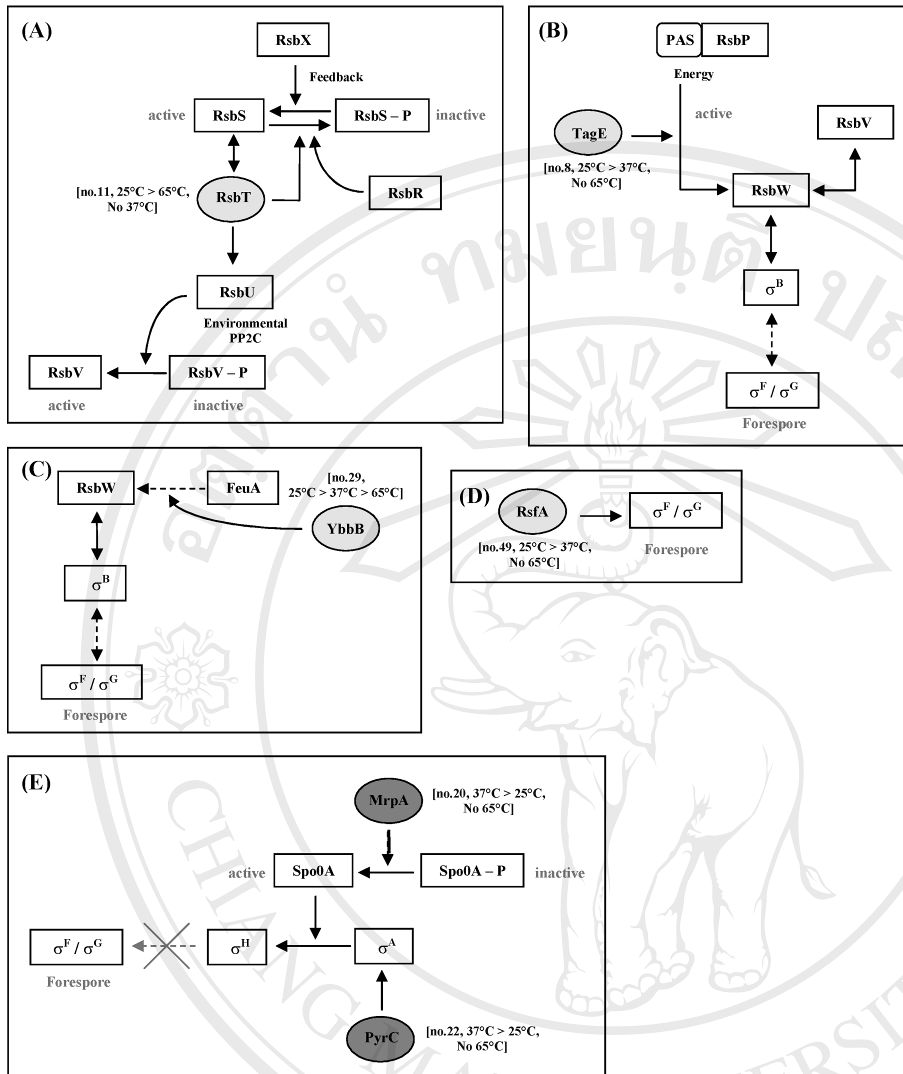


Figure 5. Correlation of the cold-shock-induced proteins in signal transduction pathway of *B. stearothermophilus* TLS33 sporulation under cold-shock stress. (A) RsbT route; (B) TagE route; (C) YbbB route; (D) RsfA route; (E) MrpA and PyrC routes.

signaling pathway. The proteome of *B. stearothermophilus* TLS33 under cold-shock stress is being studied further to obtain a more complete understanding of the biological events.

4 Concluding remarks

In summary, this is the first report of the thermophilic bacterium *B. stearothermophilus* TLS33 proteome, which correlates the signaling pathway of bacterial sporulation, under cold-shock stress. Using a proteomic approach combining 2-DE and MALDI-TOF analysis, individual proteins were identified and shown to have biological functions in the metabolic system of bacterial cell adaptation. Interestingly, six cold-shock-induced proteins were shown to correlate with the sigma B protein, which plays an important role in the signal transduction pathway of sporulation in this bacterium. Thus, this study adds to our understanding of bacterial adaptation under cold-shock stress. However, this thermophile needs to be studied further both with regard to its the biological functions and mechanisms, also at the gene expression level, under different stresses, as well as to the proteins that play an important role in the industrial and pharmaceutical applications.

This report is part of Royal Golden Jubilee Ph.D. project of Mr. Supachai Topanurak supported by the Thailand Research Fund (TRF), Thailand. We gratefully acknowledge the support of the National Research Program for Genomic Medicine of National Science Council, Taiwan (NSC 91–3112–13–001–002) and the support of Academia Sinica, Taiwan for a grant of the Bacillus research. We thank the core facilities for proteomic research, Academia Sinica, Taiwan for proteomics facilities, and the Graduate School, Chiang Mai University, Thailand for

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

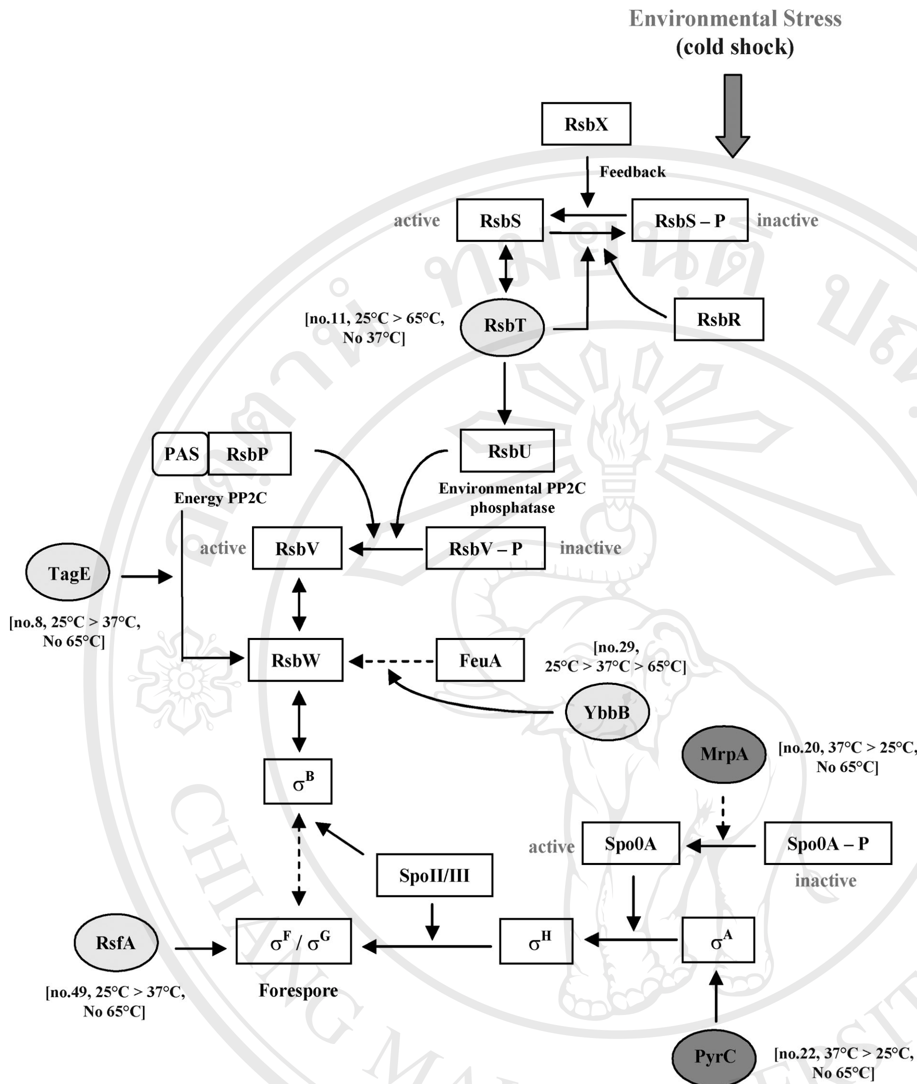


Figure 6. Model summarizing the signal transduction pathway of *B. stearothermophilus* TLS33 sporulation under cold-shock stress.

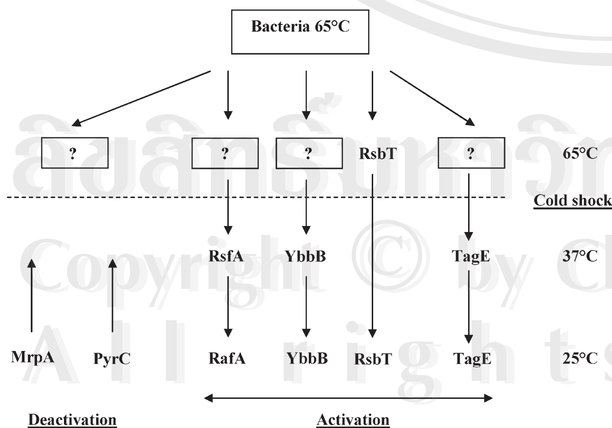


Figure 7. Up- and down-regulations of the six major cold-shock-induced proteins when *B. stearothermophilus* TLS33 encounters cold-shock stress. The direction of arrows represents the change of the protein expression.

accomplishing this project. Finally, we also thank Albert Tseng and John Y. Lin for critical reading and revising of this manuscript.

5 References

[1] Thieringer, H. A., Jones, P. G., Inouye, M., *BioEssays* 1998, *20*, 49–57.

[2] Brock, T. D., Madigan, M. T., *Biology and Microorganisms,* 5th edn., Prentice Hall, Englewood Cliffs 1980, pp. 330–335.

[3] Morita, R. Y., *Bacteriol. Rev.* 1975, *29*, 144–167.

[4] Graumann, P. L., Marahiel, M. A., *Arch. Microbiol.* 1996, *166*, 293–300.

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

[6] La Teana, A., Brandi, A., Falconi, M., Spurio, R., Pon, C. L., Gualerzi, C. O., *Proc. Natl. Acad. Sci. USA* 1991, *88*, 10907–10911.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

[7] Qing, G., Ma, L. C., Khorchid, A., Swapna, G. V., Mal, T. K., Takayama, M. M., Xia, B. *et al., Nat. Biotechnol.* 2004, *22*, 877–882.

[8] Phadtare, S., Inouye, M., *J. Bacteriol.* 2004, *186*, 7007–7014.

[9] Wick, L. M., Engli, T., *Adv. Biochem. Eng. Biotechnol.* 2004, *80*, 1–45.

[10] Mihoub, F., Mistou, M. Y., Guillot, A., Leveau, J. Y., Boubetra, A., Billaux, F., *Int. J. Food Microbiol.* 2003, *89*, 171–184.

[11] Yamanaka, K. J., *Mol. Microbiol. Biotechnol.* 1999, *1*, 193–202.

[12] Phadtare, S., Alsina, J., Inouye, M., *Curr. Opin. Microbiol.* 1999, *2*, 175–180.

[13] Kaan, T., Homuth, G., Mader, U., Bandow, J., Schweder, T., *Microbiol.* 2002, *148*, 3441–3455.

[14] Beckering, C. L., Steil, L., Weber, M. H., Volker, U., Marahiel, M. A., *J. Bacteriol.* 2002, *284*, 6395–6402.

[15] Hecker, M., *Adv. Biochem. Eng. Biotechnol.* 2003, *83*, 57–92.

[16] Petersohn, A., Brigulla, M., Haas, S., Hoheisel. J. D., Volker, V., Hecker, M., *J. Bacteriol.* 2001, *183*, 5617–5631.

[17] Mueller, U., Perl, D., Schmid, F. X., Heinemann, U., *J. Mol. Biol.* 2000, *297*, 975–988.

[18] Hasegawa, Y., Kawada, N., Nosoh, Y., *Arch. Microbiol.* 1980, *126*, 103–108.

[19] Perl, D., Welker, C., Schindler, T., Schroder, K., Marahiel, M. A., Jaenicke, R., Schmid, F. X., *Nat. Struct. Biol.* 1998, *5*, 229–235.

[20] Welker, C., Bohm, G., Schurig, H., Jaenicke, R., *Protein Sci.* 1999, *8*, 394–403.

[21] Sinchaikul, S., Sookkheo, B., Pan, F. M., Phutrakul, S., Chen, S. T., *Proteomics* 2002, *2*, 1316–1324.

[22] Sinchaikul, S., Sookkheo, B., Topanuruk, S., Juan, H. F., Phutrakul, S., Chen, S. T., *J. Chromatogr. B* 2002, *771*, 261–287.

[23] VanBogelen, R. A., Sanker, P., Clark, R. L., Bogan, J. A., Neidhardt, F. C., *Electrophoresis* 1992, *13*, 1014–1054.

[24] Pasquali, C., Frutiger, S., Wilkins, M. R., Hughes, G. J., Appel, R. D., Bairoch, A., Schaller, D. *et al.*, *Electrophoresis* 1996, *17*, 547–555.

[25] Volker, U., Englemann, S., Maul, B., Reithdorf, S., Volker, A., Schmid, R., Mach, H., Hecker, M., *Microbiology* 1994, *140*, 741–752.

[26] Büttner, K., Bernhardt, J., Scharf, C., Schmid, R., Mäder, U., Eymann, C., Antelmann, H. *et al.*, *Electrophoresis* 2001, *22*, 2908–2935.

[27] Shevchenko, A., Jensen, O. N., Podtelejnikov, A. V., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P. *et al.*, *Proc. Natl. Acad. Sci. USA* 1996, *93*, 14440–14445.

[28] Movahedi, S., Waites, W., *J. Bacteriol.* 2002, *184*, 5275–5281.

[29] Rabilloud, T., *Electrophoresis* 1998, *19*, 758–760.

[30] Berggren, K., Chernokalskaya, E., Steinberg, T. H., Kemper, C., Lopez, M. F., Diwu, Z., Haugland, R. P., Patton, W. F., *Electrophoresis* 2000, *21*, 2509–2521.

[31] Dainese, P., Staudenmann, W., Ouadroni, M., Korostensky, C., Gonnet, G., Kertesz, M., James, P., *Electrophoresis* 1997, *18*, 432–442.

[32] Msadek, T., *Trends Microbiol.* 1999, *7*, 201–207.

[33] Kroos, L., Zhang, B., Ichikawa, H., Yu, Y. T. N., *Mol. Microbiol.* 1999, *31*, 1285–1294.

[34] Weber, M. H. W., Marahiel, M. A., *Sci. Prog.* 2003, *86*, 9–75.

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

[36] Graumann, P., Wendrich, T. M., Weber, M. H. W., Svhroder, K., Mirahiel, M. A., *Mol. Microbiol.* 1997, *25*, 741–756.

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

[38] Haney, P. J., Badger, J. H., Buldak, G. L., Reich, C. I., Woese, C. R., Olsen, G. J., *Proc. Natl. Acad. Sci. USA* 1999, *96*, 3578–3583.

[39] Kumar, S., Tsai, C. J., Nussinov, R., *Protein Eng.* 2000, *13*, 179–191.

[40] Price, C. W., Fawcett, P., Ceremonie, H., Su, N., Murphy, C. K., Youngman, P., *Mol. Microbiol.* 2001, *41*, 757–774.

[41] Benson, A. K., Haldenwang, W. G., *J. Bacteriol.* 1992, *174*, 749–757.

[42] Benson, A. K., Haldenwang, W. G., *Proc. Natl. Acad. Sci. USA* 1993, *90*, 2330–2334.

[43] Boylan, S. A., Redfield, A. R., Brody, M. S., Price, C. W., *J. Bacteriol.* 1993, *175*, 7931–7937.

[44] Volker, U., Voelker, A., Haldenwang, W. G., *J. Bacteriol.* 1996, *178*, 5456–5463.

[45] Akbar, S., Gaidenko, T. A., Kang, C. M., O'Reilly, M., Devine, K. M., Price, C. W., *J. Bacteriol.* 2001, *183*, 1329–1338.

[46] Dufour, A., Voelker, U., Voelker, A., Haldenwang, W. G., *J. Bacteriol.* 1996, *178*, 3701–3709.

[47] Voelker, U., Voelker, A., Maul, B., Hecker, M., Dufour, A., Haldenwang, W. G., *J. Bacteriol.* 1995, *177*, 3771–3780.

[48] Kang, C. M., Brody, M. S., Akbar, S., Yang, X., Price, C. W., *J. Bacteriol.* 1996, *178*, 3846–3853.

[49] Yang, X. Y., Kang, C. M., Brody, M. S., Price, C. W., *Genes Dev.* 1996, *10*, 2265–2275.

[50] Kang, C. M., Vijay, K., Price, C. W., *Mol. Microbiol.* 1998, *30*, 189–196.

[51] Hughes, K. T., Mathee, K., *Annu. Rev. Microbiol.* 1998, *52*, 231–286.

[52] Voelker, U., Dulfour, A., Haldenwang, W. G., *J. Bacteriol.* 1995, *177*, 114–122.

[53] Smirnova, N., Scott, J., Voelker, U., Haldenwang, W. G., *J. Bacteriol.* 1998, *180*, 3671–3680.

[54] Qi, Y., Hulett, F. M., *J. Bacteriol.* 1998, *180*, 4007–4010.

[55] Haldenwang, W. G., *Microbiol. Rev.* 1995, *59*, 1–30.

[56] Hecker, M., Volker, U., *Mol. Microbiol.* 1998, *29*, 1129–1136.

[57] Wu, L. J., Errington, J., *J. Bacteriol.* 2000, *182*, 418–424.

[58] Ito, M., Guffanti, A. A., Oudega, B., Krulwich, T. A., *J. Bacteriol.* 1999, *181*, 2394–2402.

[59] Kosono, S., Ohashi, Y., Kawamura, F., Kitada, M., Kudo, T., *J. Bacteriol.* 2000, *182*, 898–904.

[60] Switzer, R. L., Turner, R. J., Lu, Y., *Prog. Nucleic Acid Res. Mol. Biol.* 1999, *62*, 329–367.

[61] Scott, J. M., Mitchell, T., Haldenwang, W. G., *J. Bacteriol.* 2000, *182*, 1452–1456.

[62] Takami, H., Nakasone, K., Takaki, Y., Maeno, G., Sasaki, R., Masui, N., Fuji, F. *et al.*, *Nucleic Acids Res.* 2000, *28*, 4317–4331.

[63] Gusarov, I. I., Kreneva, R. A., Podcharnniaev, D. A., Iomantas, I. V., Akalakina, E. G., Stoinava, N. V., Perumov, D. A., Kozlov, I. I., *Mol. Biol.* 1997, *31*, 446–453.

Copyright© by Chiang Mai University
All rights reserved

[64] Lee, J. M., Zhang, S., Saha, S., Anna, S. S., Jiang, C., Perkins, J., *J. Bacteriol.* 2001, *183*, 7371–7380.

[65] Bresler, S. E., Cherepenko, E. I., Chernik, T. P., Kalinin, V. L., Perumov, D. A., *Genetika* 1970, *6*, 116–124.

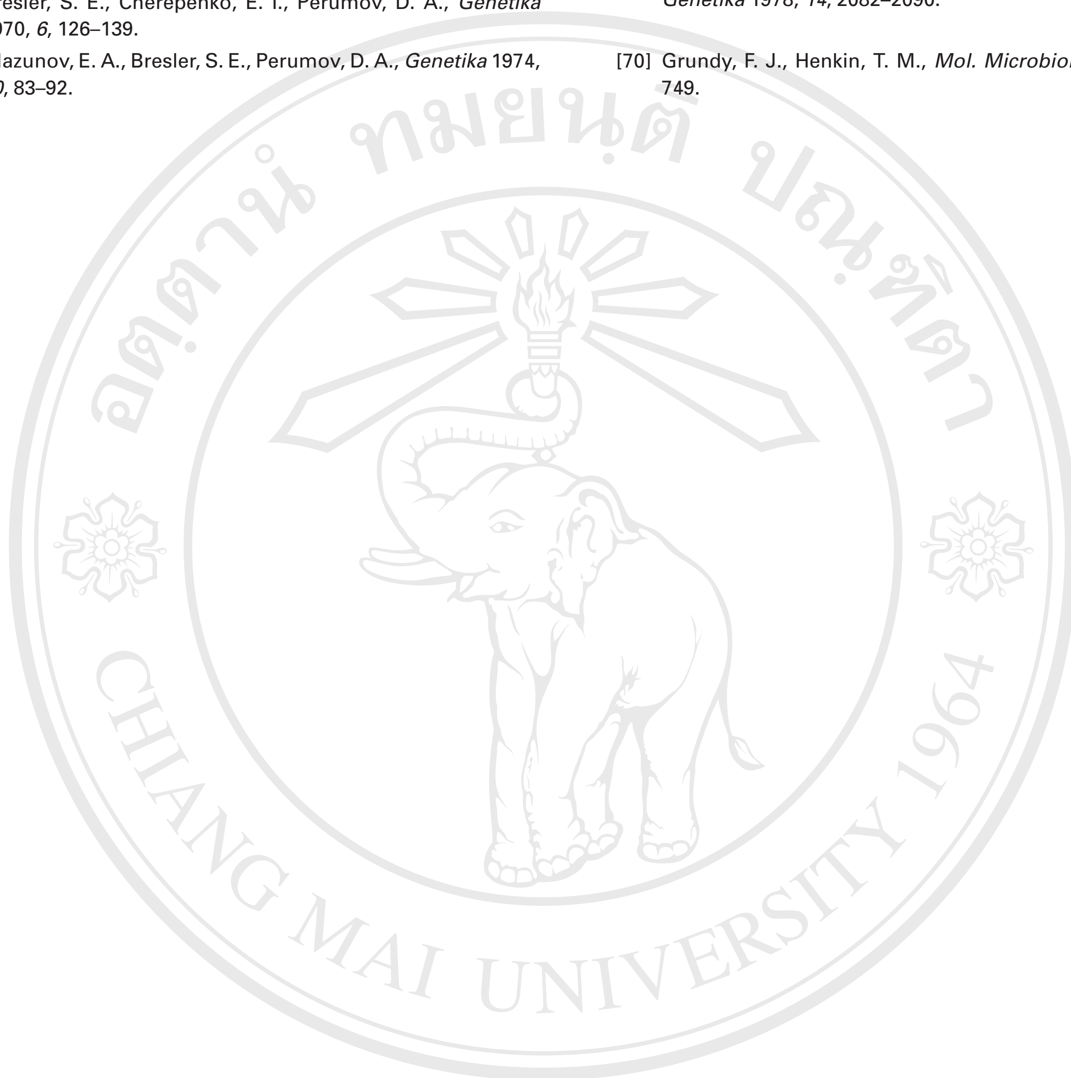
[66] Bresler, S. E., Cherepenko, E. I., Perumov, D. A., *Genetika* 1970, *6*, 126–139.

[67] Glazunov, E. A., Bresler, S. E., Perumov, D. A., *Genetika* 1974, *10*, 83–92.

[68] Bresler, S. E., Glazunov, E. A., Perumov, D. A., Chernik, T. P., *Genetika* 1977, *13*, 2007–2016.

[69] Bresler, S. E., Gorinchuk, G. F., Chernik, T. P., Perumov, D. A., *Genetika* 1978, *14*, 2082–2090.

[70] Grundy, F. J., Henkin, T. M., *Mol. Microbiol.* 1998, *30*, 737–749.



ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

REGULAR ARTICLE

Proteomics viewed on stress response of thermophilic bacterium *Bacillus stearothermophilus* TLS33

Supachai Topanurak[1, 2], *Supachok Sinchaikul*[1], *Suree Phutrakul*[2], *Boonyaras Sookkheo*[3] *and Shui-Tein Chen*[1, 4]

[1] Institute of Biological Chemistry and Genomics, Academia Sinica, Taipei, Taiwan
[2] Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand
[3] Department of Chemistry, Faculty of Science, Rangsit University, Pathunthani, Thailand
[4] Institute of Biochemical Sciences, College of Life Science, National Taiwan University, Taiwan

Thermophilic bacterium *Bacillus stearothermophilus* TLS33, isolated from a hot spring in Chiang Mai, Thailand, usually produces many enzymes that are very useful for industrial applications. However, the functional properties and mechanisms of this bacterium under stress conditions are rarely reported and still need more understanding on how the bacterium can survive in stress environments. In this study, we examined the oxidative stress induced proteins of this bacterium by proteomic approach combining two-dimensional electrophoresis and mass spectrometry. When the bacterium encountered oxidative stress, peroxiredoxin, as an antioxidant enzyme, is one of the interesting stressed proteins which appeared to be systematically increased with different p*I*. There are four isoforms of peroxiredoxin, denoted as Prx I, Prx II, Prx III and Prx IV, which are observed at the same molecular weight of 27 kDa but differ in p*I* values of 5.0, 4.87, 4.81 and 4.79, respectively. The H_2O_2 concentration directly increased Prx II, Prx III and Prx IV intensities, but decreased Prx I intensity. These shifting of peroxiredoxin isoforms may occur by a post-translational modification. Otherwise, the longer time of oxidative stress had not affected the expression level of peroxiredoxin isoforms. Therefore, this finding of peroxiredoxin intends to know the bacterial adaptation under oxidative stress. Otherwise, this protein plays an important role in many physiological processes and able to use in the industrial applications.

Received: June 6, 2004
Revised: December 27, 2004
Accepted: January 17, 2005

Keywords:
Bacillus stearothermophilus / Peroxiredoxin / Stress / Thermophile

1 Introduction

Thermophile organisms grow optimally at temperatures above 60–70°C. Since thermophiles have several unique physiological and metabolic characteristics, they are very usefully exploited in the biotechnological industry for wide applications [1–3]. They also include biotransformations, biocatalysis, bioremediation or bioactive compound production, for example, whereby whole thermophile cell or purified thermostable enzymes. Biocatalytic reactions by microorganisms are also found to be chemoselective, regioselective and stereoselective, therefore allowing production of optically active compounds as well [4]. These abilities are used in the pharmaceutical and agrochemical industries for the biocatalysts of non-natural products. However, thermophiles are often exposed to multiple environmental stresses during bioprocessing; for example, low and high temperatures, low and high pH values, high osmotic pressure, nutrient starvation and oxidation. Among these stress factors, oxidation can be considered one of the most deleterious to the cell, causing cellular damage at both molecular and metabolic levels [5]. In addition, the reactive oxygen species (ROS) derived from the reduction of dioxygen itself is less reactive and from the reduction of dioxygen to superoxide (O_2^-), hydrogen per-

Correspondence: Professor Dr. Shui-Tein Chem, Institute of Biological Chemistry (RM707), Academia Sinica, 128 Yen Chiu Yuan Rd., Sec. II, Nankang, Taipei, 11529, Taiwan
E-mail: bcchen@gate.sinica.edu.tw
Fax: +886-2-27883473

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright by Chiang Mai University
All rights reserved

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

oxide (H_2O_2) or the hydroxyl radical ($OH^{\cdot}$) are highly reactive and able to efficiently damage nucleic acids, membrane lipids and proteins [6]. To eliminate the excess ROS, many organisms use peroxiredoxin system to scavenge the excess ROS which can be harmful to cells [7–15]. Peroxiredoxins (Prx) or alkyl hydroperoxide reductase cysteines (AhpC) are enzymes in a family of thiol-specific peroxides which are found in plants and some microorganisms [16–21]. They are implicated in a wide variety of cellular processes including proliferation [22], differentiation [23], and the immune response [24] as well as in the detoxification of peroxides. They represent a special case which constitutes both the peroxidase and the cosubstrate because the enzyme itself is oxidized upon reaction with the peroxide. Peroxiredoxins have cysteines at their active site which can be oxidized by the peroxide and forms to be cysteine sulphinic acid. Although there have been some reports describing the function of peroxiredoxin [25–27], the changes of their isoforms have never been studied in this bacterium.

At present, proteomic techniques using two-dimensional electrophoresis (2-DE) and mass spectrometry are the credible approaches which are widely used to study the alteration of protein expression in organisms under different environmental conditions and further more understanding of the biology of whole organism cells [28–32]. Thermophilic bacterium *Bacillus stearothermophilus* TLS33, isolated from a natural hot spring in Chiang Mai, Thailand, is an interesting thermophile that usually grow well in a hot environment at 65°C and produce many thermostable enzymes, such as proteases [33], lipase [34] and superoxide dismutase [35], that are very useful for industrial applications. According to its thermostability and applications, it is challenging to investigate the intracellular proteome of thermophilic bacterium *B. stearothermophilus* TLS33 under oxidative stress which is one of the general stresses resulting from extreme conditions from industrial bioprocessing. Using a proteomic approach, we found that peroxiredoxin is the interesting protein which is firstly observed a major difference of the protein expression in this bacterium, responding to oxidative stress. Furthermore, there are four isoforms of peroxiredoxin with the identical molecular weight but different in p*I*. Then, we examined the leveled change of the peroxiredoxin isoforms which may play an important role in a protective mechanism under oxidative stress and able to be used in the bioprocessing and other industrial applications.

Copyright © by Chiang Mai University
All rights reserved

2 Materials and methods

2.1 Bacterial culture and stress experiment

B. stearothermophilus TLS33, isolated from a soil in a hot spring in Chiang Mai, Thailand, was picked a colony from an agar plate and inoculated into 50 mL of nutrient broth (ADSA Micro, Barcelona, Spain) in 250 mL flask as preculture. The preculture was incubated in the water bath at 65°C with shaking at 200 rpm. After 24 h of precultivation, 1 mL of cell suspension was transferred to 100 mL of fresh nutrient broth in 250 mL flask (triplicate for each experiment). The bacterium was cultured at 65°C until mid-log phase (OD_{600}~0.6) and subsequently added the hydrogen peroxide (Riedel de Haën, Seelze, Germany) into the media with the final concentration of 10, 50, 100 and 500 µM in each flask. After half hour interval of H_2O_2 induction, the bacterial cells were observed at 600 nm and harvested by centrifugation at 4°C, 8 500 rpm for 10 min. The survival percentage is defined that the OD_{600} at starting induction of H_2O_2 is set as 100%. The bacterial cells were washed three times with preparation buffer pH 8.0 containing 10 mM Tris-HCl, 1 mM EDTA and 0.1 mM PMSF.

2.2 Sample preparation

The bacterial cells were disrupted by sonication with 4 s pulse for 10 min in preparation buffer pH 8.0 containing 10 mM Tris-HCl, 1 mM EDTA and 0.1 mM PMSF, and the supernatant was obtained by centrifugation at 4°C, 12 000 rpm for 20 min. The sample solutions were precipitated by addition of 10% TCA and 0.1% w/v DTT. The mixture was stored overnight at −20°C and the pellet was obtained by centrifugation at 10 000 × g, 4°C for 15 min. The pellet was washed twice, first with cold acetone containing 20 mM DTE and secondly with cold acetone without DTT, and centrifuged as described above. The supernatant was removed and the pellet was vacuum dried.

2.3 Two-dimensional electrophoresis (2-DE)

Lyophilized protein samples were resolved in lysis buffer containing 7 M urea, 2 M thiourea, 4% CHAPS, 4 mM TCEP, 100 mM DTE and 0.5% IPG buffer pH 4–7L. The protein sample solutions containing 100 µg protein concentration in 350 µL total volume were applied on IPGphor strip (Amersham Biosciences, 18 cm, 4–7L) using IPGphor (Amersham Biosciences, Uppsala, Sweden).

The first-dimensional isoelectric focusing (IEF) on IPGphor was performed under the following condition: 30 V, 12 h (rehydration); 100 V, 3 h; 350 V, 1 h; 500 V, 1 h; 1000 V, 1 h; 5000 V, 1 h; 8000 V, 55 kVh. After IEF, the IPG strips were equilibrated in equilibration buffer I (50 mM Tris-HCl pH 8.8, 6 M urea, 30% v/v glycerol, 2% w/v SDS, 2% w/v DTE and a trace of bromophenol blue) for 15 min, and then subsequently alkylated in buffer II (50 mM Tris-HCl pH 8.8, 6 M, urea, 30% v/v glycerol, 2% w/v SDS, 2.5% w/v iodoacetamide and a trace of bromophenol blue) for 15 min. Each equilibrated IPG strip was placed on top of the 15% homogenous polyacrylamide gel (185 × 200 × 1.5 mm) and covered with 0.5% agarose. The second-dimensional separation using PROTEAN II xi Multi-Cells (Bio-Rad) was carried out at 45 mA *per* gel at 15°C until the bromophenol blue dye front reached the bottom of the gel. At the end of each run, the 2-D gels were stained with SYPRO® Ruby [36] and

scanned using a Typhoon 9200 scanncer (Amersham Biosciences). In addition, the 2-D gel images were exported to the image analysis software program, using ImageMaster™ 2D platinum software (Amersham Biosciences) which provides the 3-D viewing and the protein intensity of each spot.

2.4 In-gel digestion

Protein spots were manually excised from the gels and transferred to 500 μL siliconized Eppendorfs. The gel pieces were washed twice with 200 μL of 50% ACN/25 mM ammonium bicarbonate buffer, pH 8.0, for 15 min each. The gel pieces were then washed once with 200 μL of 100% ACN and dried using a Speed Vac concentrator. Dried gel pieces were swollen in 10 μL of 25 mM ammonium bicarbonate containing 0.1 μg trypsin (Promega, Madison WI, USA). Gel pieces were then crushed with siliconized blue stick and incubated at 37°C for at least 16 h. Peptides were subsequently extracted twice with 50 μL of 50% ACN/5% TFA, then the extracted solutions were combined and dried using a SpeedVac concentrator. The peptides or pellets were then resuspended in 10 μL of 0.1% TFA and the suspended solutions were purified using ZipTip C18 (Millipore, Billerica, MA, USA). Ten microliters of sample were drawn up and down in the ZipTip 10 times and the ZipTip was washed with 10 μL of 0.1% formic acid by drawing up and expelling the washing solution for three times. The peptides were eluted with 5 μL of 75% ACN/0.1% formic acid.

2.5 Protein identification by MALDI-TOF MS/MS

This method was performed as described by Lee *et al.* [37]. Briefly, MALDI-TOF MS/MS were performed on a dedicated Q-Tof Ultima™ MALDI instrument (Micromass) with fully automated data directed acquisition using predefined probe motion pattern and peak intensity threshold for switching over from MS survey scan to MS/MS, and from one MS/MS to another. At a laser firing rate of 10 Hz, individual spectra from 5 s integration period acquired for each of the MS/MS performed were combined, smoothed, deisotoped (fast option) and centroided using the Micromass PGS 2.0 data processing software. All individual MS/MS data thus generated from a particular sample well were then output as a single MASCOT-searchable peak list file and all peak list files generated from each the 1-D μLC fractions as deposited by Probot™ on the corresponding MALDI sample well were manually combined into a single .pkl file before searching against the NCBIr database using MASCOT program.

2.6 Protein modification analysis by LC-ESI nanoMS/MS

The tryptic digested peptides from each Prx isoform were analyzed by 1-D LC-nanoESI MS/MS. The 1-D LC-nanoESI-MS/MS analysis was performed on an integrated nanoLC-MS/MS system (Mircomass) comprising a three-pumping Micromass/Waters CapLC™ system with an autosampler, a stream select module configured for precolumn plus analytical capillary column, and a Micromass Q-Tof Ultima™ API mass spectrometer fitted with nano-LC sprayer, operated under MassLynx™ 4.0 control. Injected samples were first trapped and desalted isocratically on an LC-Packings PepMap™ C18 μ-Precolumn™ Cartidge (5 μm, 300 μm id × 5 mm; Dionex, Sunnyvale, CA, USA) for 2 min with 0.1% formic acid delivered by the auxillary pump at 30 μL/min after which the peptides were eluted off from the precolumn and separated on an analytical C18 capillary column (15 cm × 75 μm id, packed with 5 μm, Zorbax 300 SB C18 paticles; Micro-Tech Scientific, Vista, CA, USA) connected inline to the mass spectrometer, at 300 nL/min using a 40 min fast gradient of 5% to 80% acetonitrile in 0.1% formic acid.

3 Results and discussion

3.1 Bacterial survival of *B. stearothermophilus* TLS33 under oxidative stress

The survival of thermophile *B. stearothermophilus* TLS33 after treatment with different concentrations of H_2O_2 is shown in Fig. 1. In general, hydrogen peroxide (H_2O_2) produced the reactive oxygen species (ROS) which are generated aerobically by auto-oxidation of electron transport chain components and, when present in excess, can damage various components of living cells [7–15]. Thus, the percentage of bacterial survival decreased after addition of H_2O_2 from 50 to 500 μM, except 10 μM H_2O_2. Although all concentrations of H_2O_2 were not shown to be highly affected to the bacterial cells along 120 min of stressed time with the survival higher than 85%, the increase of H_2O_2 concentration effectively decreased the survival of this bacterium.

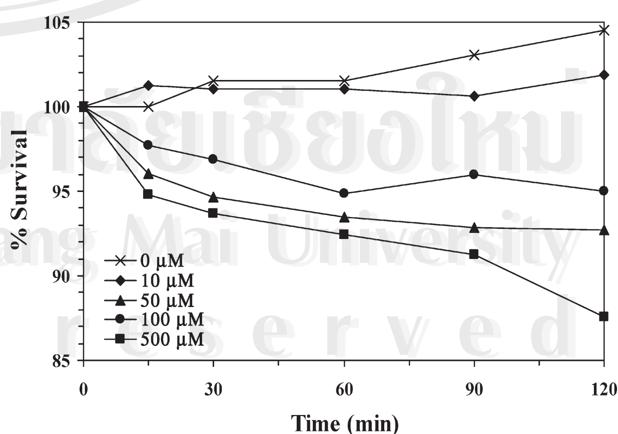


Figure 1. Survival percentage of *B. stearothermophilus* TLS33 after addition of different concentrations of H_2O_2 (10, 50, 100 and 500 μM) for 120 min.

Copyright © by Chiang Mai University
All rights reserved

3.2 2-DE analysis of *B. stearothermophilus* TLS33 cells

Although the regulation in gene level of bacteria under oxidative stress has been reported, the regulation in protein level including protein expression and post-translational modification are rarely studied in this bacterium. In order to better understand the stress response of *B. stearothermophilus* TLS33 under oxidative condition, we used the proteomic approach to detect and identify proteins that are differentially expressed in bacterial cell that is resistant to oxidative stress within a short period time of 15 min. After addition of different concentrations of H_2O_2, 2-DE analysis of the intracellular proteins of *B. stearothermophilus* TLS33 demonstrated a few changes in the protein patterns (Fig. 2). Interestingly, four isoform spots at the low molecular weight approximately 27 kDa were markedly changed. Subsequently, the protein digests of three isoform spots were analyzed by MALDI-TOF MS/MS and identified to be the same protein called peroxiredoxin (Prx) and denoted as Prx I, Prx II, Prx III and Prx IV. Moreover, the probability-based scoring of the mass spectrums of peroxiredoxin isoforms derived from the calculation by Perkins *et al.* [38] informed the high scoring of mass database search which indicated that these four spots were peroxiredoxin reliably (Table 1). In addition, those isoforms Prx I, Prx II, Prx III and Prx IV appeared at the same molecular weight but different in p*I* with 5.0, 4.87, 4.81 and 4.79, respectively. We suggest that a shift to a more acidic spot position may be caused by post-translational modification which might occur in response to oxidative stress when cysteine residues are oxidized to sulphinic acid [39–42].

In general, peroxiredoxins are a ubiquitous family of antioxidant enzymes, which play an important role in many physiopathological processes, including adaptation to oxygen [43, 44]. It is well known that peroxiredoxins control cytokine-induced peroxide levels which mediate signal transduction in mammalian cells. They can be regulated by changes to phosphorylation, redox and possibly oligomerization states [45, 46]. They are also significantly elevated in human thyroid tumors [47] and in the area of brain that are most susceptible to hypoxic and ischemic injury [48]. Recently, the interesting study of peroxiredoxin by genome-based bioinformatics selection showed its function as a vaccine candidate for anthrax disease [49]. In the case of bacteria, the metabolic context of peroxiredoxins and its physiological relevance have not yet been widely studied, especially the alteration of peroxiredoxin isoform depending on the oxidative state. In our results, we first found that thermophile *B. stearothermophilus* TLS33 produced the peroxiredoxins with four isoforms when encountering oxidative stress. In fact, two isoforms of peroxiredoxins I and II (Prx I and II) normally appeared in the cytosol of this bacterium. After addition of H_2O_2 and incubation for 15 min, the peroxiredoxin isoforms were shifted depending on the H_2O_2 concentration. This evidence can be supported by the previous study which reported that the acidic form of peroxiredoxins was an oxidative modified form after organisms have encountered oxidative stress, whereas the basic form was the normal form of peroxiredoxins [41]. We suggest that the reactive oxygen species (ROS) can oxidize the peroxiredoxins and modify it to become four isoforms when the bacterium encountered oxidative stress. Thus, we could observe the increase of Prx II, Prx III and Prx IV in the oxidative stress condition.

3.3 Expression of peroxiredoxin isoforms

Since peroxides are weak oxidizing agents, they can react with cysteinyl-thiols in proteins by formation of disulfide bonds or sulphinic acid derivatives [50]. While cells are growing, significant amounts of O_2^- and H_2O_2 are generated by enzymatic misdirection of electrons to dioxygen. It is assumed that flavin-dependent transfer reactions of respiratory chain are primarily responsible for generation of ROS. If

Table 1. Protein identification of four isoforms of peroxiredoxin from thermophile *Bacillus stearothermophilus* TLS33

Isoform	Mass (kDa)	p*I*	MOWSE score	Peptide matched no	Sequence coverage (%)	Matching sequence
Prx I	27	5.00	347[a)]	4	38	MFDVLDEEQGLAQR
Prx II	27	4.87	252[a)]	4	27	IEYVMIGDPSHQLSR
Prx III	27	4.81	80[a)]	4	7	AQAYHNGEFIEVTEQDFMGK GTFIIDPDGVIQAVEINADGIGR
Prx IV	27	4.78	131[a)]	3	27	MFDVLDEEQGLAQR IEYVMIGDPSHQLSR GTFIIDPDGVIQAVEINADGIGR

a) The score indicates identity or extensive homology. (Protein scores are derived from MASCOT software.)

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

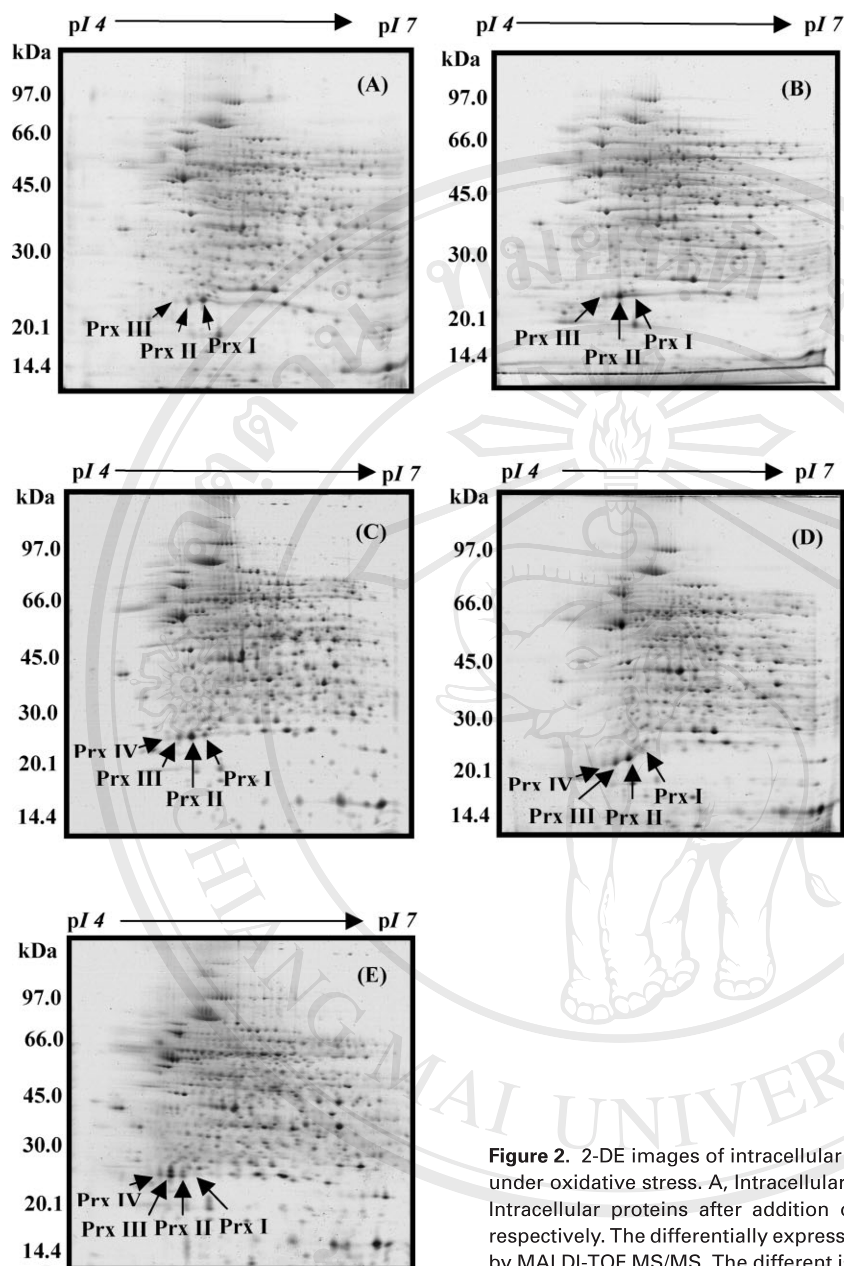
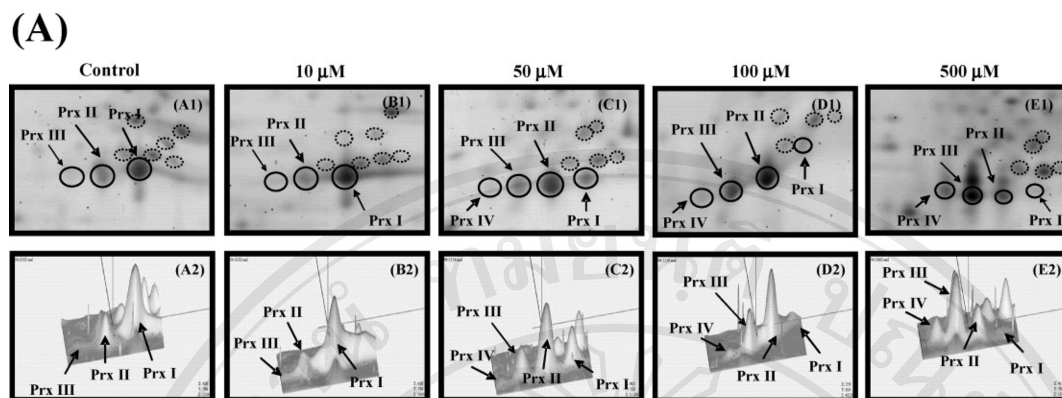


Figure 2. 2-DE images of intracellular proteome of *Bacillus stearothermophilus* TLS33 under oxidative stress. A, Intracellular proteins under control condition; B, C, D and E, Intracellular proteins after addition of 10, 50, 100 and 500 μM of H_2O_2 for 15 min, respectively. The differentially expressed proteins under oxidative stress were analyzed by MALDI-TOF MS/MS. The different isoforms of peroxiredoxin are shown by arrows.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

the amount of ROS increases to toxic levels, cells encounter oxidative stress. In this study, we used H_2O_2 as an oxidative stressor to investigate how bacterial cells protect themselves from harmful conditions of oxidative stress. By the results, peroxiredoxin is one interesting protein which showed a significant difference in protein expression level. In order to conveniently visualize the peroxiredoxin isoform appearance, 3-D viewings of peroxiredoxin area under different concentrations of H_2O_2 were generated by using ImageMaster™ 2D platinum software and used to examine alterations of peroxiredoxin (Fig. 3A). In addition, the volume of peak area in 3-D viewing of peroxiredoxin isoforms can be equivalent to the protein intensity as shown in Fig. 3B. The image viewings showed Prx I normally located on the basic region while Prx II, Prx III and Prx IV located on the acidic region. In the presence of different H_2O_2 concentrations, 2-D and 3-D viewings demonstrated the markedly different expression level of peroxiredoxin isoforms, which are dependent on the concentration of H_2O_2, especially at the high concentration of 500 μM. Interestingly, Prx III slightly appeared in the absence of H_2O_2 and in the low concentration of 10 μM H_2O_2 whereas Prx IV was not observed. The appearance of Prx III and Prx IV occurred in the presence of 50 μM H_2O_2 condition and/or higher concentration of H_2O_2.

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim



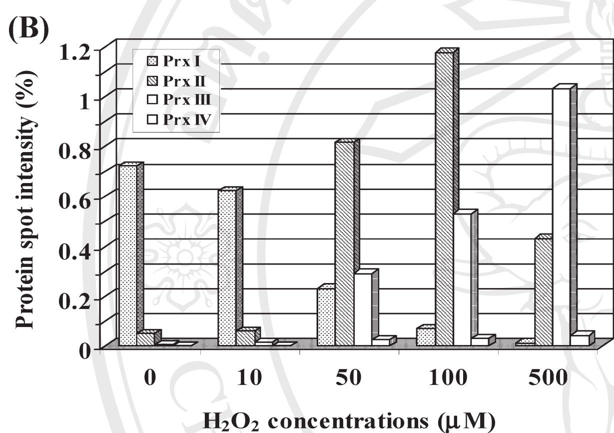


Figure 3. Representative of peroxiredoxin expression in *B. stearothermophilus* TLS33 under oxidative stress by addition of different concentrations of H_2O_2. (A), 2-D and 3-D images covering the peroxiredoxin isoform region under different concentrations of H_2O_2 for 15 min. The volume of peak area represents the intensities of protein spots. A, control (no addition of H_2O_2); B, 10 μM H_2O_2; C, 50 μM H_2O_2; D, 100 μM H_2O_2; E, 500 μM H_2O_2. (B), Percentage of protein spot intensity of peroxiredoxin isoforms on different H_2O_2 concentrations. These percent spot volumes were generated by ImageMaster™ 2D platinum.

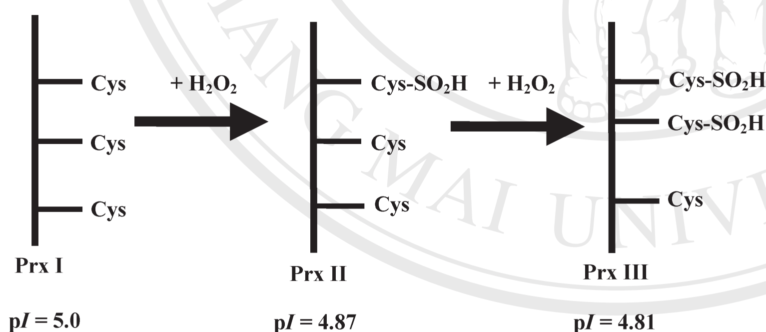


Figure 4. Hypothesis of *B. stearothermophilus* TLS33 peroxiredoxin isoforms that are modified by H_2O_2. The modification of active site cysteines resulted in acidic shift of peroxiredoxin by forming the complex with sulphinic acid at active site residue.

In the absence of H_2O_2, Prx I showed the highest intensity rather than Prx II, Prx III and Prx IV, respectively. When the concentration of H_2O_2 was increased, the Prx I intensity was decreased whereas Prx II, Prx III and Prx IV intensities were increased (Fig. 3B). Thus, we conclude that the intensities of Prx II, Prx III and Prx IV are directly increased to the concentration of H_2O_2 and the appearance of peroxiredoxin isoforms are regulated by H_2O_2. Otherwise, these evidences may be post-translational modification which has also been reported and that the acidic spot was confirmed to be a modified form because the active-site cysteines were oxidized and turned into cysteine sulphinic acid [42, 50, 51]. Thus, we presume that the modified cysteines with sulphinic acids in the protein molecule may formulate the proteins to have positive charge and intend to shift the p*I* of the protein. Therefore, the level of oxidative stress directly affected the modification of peroxiredoxin isoforms and the active-site cysteines may be modified by H_2O_2 to be obtain the acidic form (Fig. 4).

3.4 Post-translational modification by LC-MS/MS analysis

According to our presumption, the Prx isoforms would be modified with ROS or H_2O_2 and formed sulphinic acid in cysteine residues (Cys-SO_2H). To characterize the modifica-

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

tions taking place in the acid spots, both the normal and acidic spots from Prx were analyzed by LC-MS/MS. A modified peptide was found at the LC-MS stage as a peak occurring only in the acidic spots of Prx II and Prx III and not in the basic one of Prx I (Fig. 5). There were 598 *m/z* peaks in Prx II, and 598 *m/z* and 809 *m/z* peaks in Prx III, in which these peaks are different from its native form (Prx I). To confirm the hypothesis, these tryptic digestion peptides were analyzed by collision-induced dissociation. From MS/MS spectrum in Fig. 5, it indicates the H_2O_2 oxidized SH group in cysteine residue of both Prx II and Prx III. A mass difference of 135 absolute mass units was detected between the b3 and b4 ion in 598 *m/z* peak of both of Prx II and III and between b6 and b7 in 809 *m/z* peak of Prx III, indicative of the presence of a cysteinyl residue modified by two oxygen atoms. According to our purpose the H_2O_2 would oxidize disulfide group in another cysteine residue in Prx II and cause the protein to shift to a more acidic region which formed Prx III, herein this was proved by the presence of 598 *m/z* and 809 *m/z* peak in Prx III but it was found that only 598 *m/z* is present in Prx II. However, due to the high sequence conservation between Prx II and III, we speculate that the acidic Prx III spot also corresponds to an oxidized form at the active site. This has also been reported in the similar modification [42]. For Prx IV, we could not find the mass difference of 135 mass units. However, we will investigate the fourth isoform in a further study.

3.5 Effect of stressed time on modification of peroxiredoxin isoforms under oxidative stress

According to the results of different concentration of H_2O_2, Prx II showed the highest intensity at 100 μM of H_2O_2 and subsequently decreased at 500 μM, while Prx III and Prx IV were directly variable to the increase of H_2O_2 concentrations. Hence, we examined the effect of stressed time on the irreversible oxidation of peroxiredoxin isoforms by adding the high concentration of 500 μM H_2O_2. At 120 min of stress period, Prx I showed the decreased intensity whereas Prx III showed the increased intensity (Fig. 6). Surprisingly, peroxiredoxin isoform called Prx IV eventually appeared in the rather acidic region (p*I* 4.79) when the stressed time was longer. The longer stressed time period had no affect on the reversible oxidation of peroxiredoxin isoforms to return to its original form. This finding indicated that when the bacterium encountered higher oxidative stress, the oxidized forms of Prx III and Prx IV were not altered themselves to any reduced forms. According to the previous reports, the oxidation of the active site cysteine to the sulphinic state has also been considered to be irreversible [52–54]. Otherwise, the reversibility of Prx III to its original form (Prx I) is probably difficult because of the stability of sulphinic acid (RSO_2H) which was obtained from modification of cysteine residue. Nevertheless, it is necessary to further study that whether H_2O_2 is not presented in all oxidized surrounding conditions of bacterial cells, these proteins would be reduced and re-

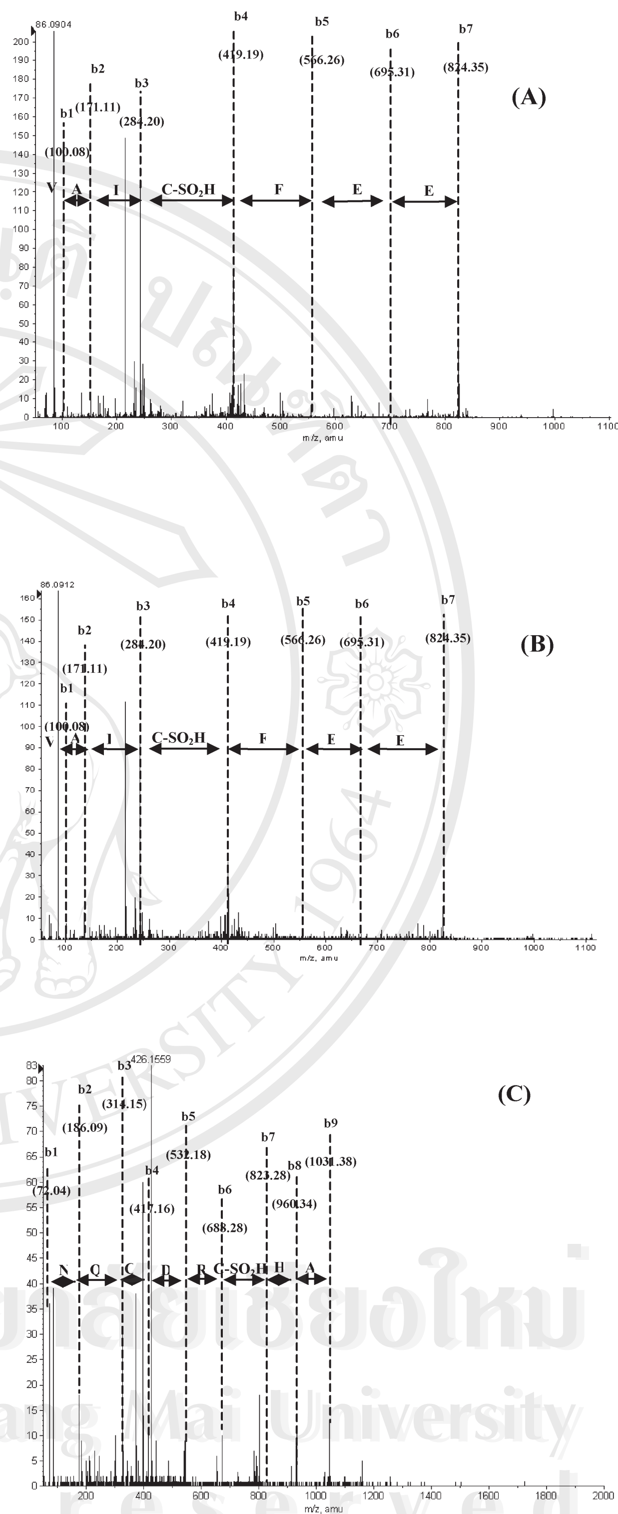


Figure 5. Collision-induced dissociation spectrum of the Prx II – 598 *m/z* peak (A) and Prx III – 598 *m/z* peak (B) and 809 *m/z* peak (C). The *m/z* ion peaks are shown on the top of peak. The amino acid sequences were calculated from the difference of mass units of b ion series. The difference of mass units of 135 was derived from the calculation of cysteine molecular weight (103) plus two oxygen atoms (32).

Copyright © by Chiang Mai University
All rights reserved

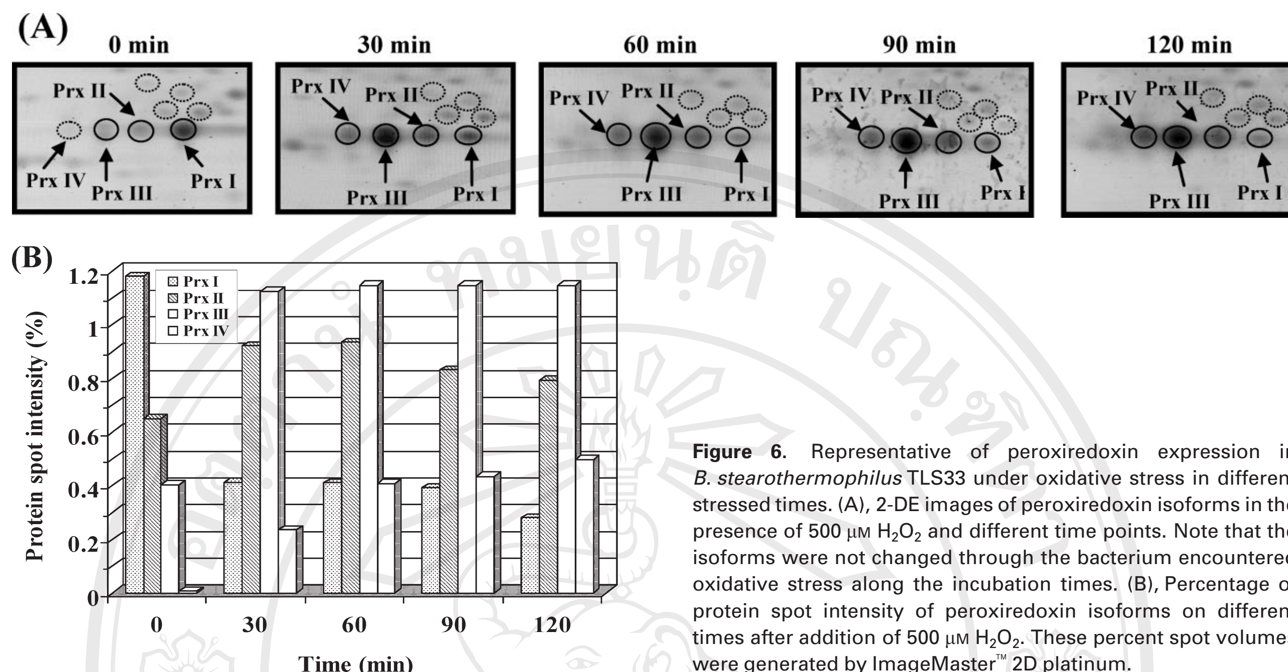


Figure 6. Representative of peroxiredoxin expression in *B. stearothermophilus* TLS33 under oxidative stress in different stressed times. (A), 2-DE images of peroxiredoxin isoforms in the presence of 500 μM H_2O_2 and different time points. Note that the isoforms were not changed through the bacterium encountered oxidative stress along the incubation times. (B), Percentage of protein spot intensity of peroxiredoxin isoforms on different times after addition of 500 μM H_2O_2. These percent spot volumes were generated by ImageMaster™ 2D platinum.

turned to its original form. For this reason, the regulatory pathway of these proteins and other proteins under oxidative stress will be further studied.

4 Concluding remarks

In summary, this study is the first finding of peroxiredoxin in the thermophilic bacterium *B. stearothermophilus* TLS33 that produced many isoforms to reduce ROS when it encountered oxidative stress. By proteomic approach combining 2-DE and mass spectrometry, we detected an alteration of the peroxiredoxin isoforms upon oxidative stress, showing the increased intensity of acidic form of peroxiredoxin. Analysis of tryptic peptides generated from the peroxiredoxin spots in the basic and acidic areas by mass spectrometry provided the different isoforms of peroxiredoxin that had the same molecular weight but different p*I*. This modification by p*I* shifting may be caused by adding a negative charge to form a complex with the protein and shifting the p*I* to be more acid. This evidence is possible to use this protein as diagnostic indicators or protein markers in physiological condition or use in the industrial applications.

This report is part of Royal Golden Jubilee Ph.D. project of Mr. Supachai Topanurak supported by the Thailand Research Fund (TRF), Thailand. We gratefully acknowledge the support of the National Research Program for Genomic Medicine of National Science Council, Taiwan (NSC 91-3112-13-001-002) and the support of Academia Sinica, Taiwan for a grant of the Bacillus research. We also thank the core facilities for proteomic research, Academia Sinica, Taiwan for proteomics facilities, and the Graduated School, Chiang Mai University, Chiang Mai, Thailand, for supporting this work. Finally, we appreciatedly thank Dr. Hsueh-Fen Juan, Department of Life Science, Institute of Molecular and Cellular Biology, National Taiwan University, Taipei, Taiwan, for supporting of MALDI-TOF-MS/MS and the suggestions to us.

5 References

[1] Schallmey, M. Singh, A. Ward, O. P. *Can. J. Microbiol.* 2004, *50*, 1–17.

[2] Sinchaikul, S., Sookkheo, B., Topanuruk, S., Juan, H. F., Phutrakul, S., Chen, S. T., *J. Chromatogr. B* 2002, *771*, 261–287.

[3] Demirjian, D. C., Morís-Varas, F., Cassidy, C. S., *Curr. Opin. Chem. Biol.* 2001, *5*, 144–151.

[4] Bustard, M. T., Burgess, J. G., Meeyoo V., Wright, P. C., *J. Chem. Tech. Biotech.* 2000, *75*, 1095–1109.

[5] Miyoshi, A., Rochat, T., Gratadoux, J. J., Loir, L. L., Costa, O., Langella, P., Azevedo, V., *Genet. Mol. Res.* 2003, *2*, 348–349.

[6] Mostertz, J., Scharf, C., Hecker, M., Homuth, G., *Microbiology* 2004, *150*, 497–512.

[7] Fratelli, M., Demol, H., Puype, M., Casagrande, S., Villa, P., Eberini, I., Vadekerckhove, J., Gianazza, E., Ghezzi, P., *Proteomics* 2003, *3*, 1154–1161.

[8] Kim, D. M., Swartz, J. R., *Biotech. Bioeng.* 2004, *2*, 123–129.

[9] Klichko, V., Radyuk, S. N., Orr, W. C., *Arch. Insect Biochem. Physio.* 2004, *56*, 34–50.

[10] Gromer, S., Urig, S., Becker, K., *Med. Res. Rev.* 2004, *24*, 40–89.

Copyright © by Chiang Mai University
All rights reserved

[11] Sarafian, T. A., Verity, M. A., Vinters, H. V., Shih, C. Y., Shi, L., Xiang, D. J., Dong, L., Shau, H., *J. Neurosci. Res.* 1999, *56*, 206–212.

[12] Evrard, C., Capron, A., Marchand, C., Clippe, A., Waittiez, R., Soumillion, P., Knoops, B., Declercq, J. P., *J. Mol. Biol.* 2004, *337*, 1079–1090.

[13] Isermann, K., Liebau, E., Roeder, T., Bruchhaus, I., *J. Mol. Biol.* 2004, *338*, 745–755.

[14] Budanov, A. V., Sabina, A. A., Feinstein, E., Koonin, E., Chumakov, P. M., *Science* 2004, *304*, 597–560.

[15] Li, H., Singh, A. K., Mcityre, L. M., Sherman, L. A., *J. Bacteriol.* 2004, *186*, 3331–3345.

[16] Dietz, K. J., *Annu. Rev. Plant Biol.* 2003, *54*, 93–107.

[17] Horling, F., Lamkemeyer, P., Konig, J., Finkemeier, I., Kandlbinder, A., Baier, M., Diestz, K. J., *Plant Phys.* 2003, *131*, 317–325.

[18] Nishimaya, Y., Massey, V., Takeda, K., Kawasaki, S., Sato, J., Watanabe, T., Nuumura, Y., *J. Bacteriol.* 2001, *183*, 2431–2438.

[19] Logan, C., Mayhew, S., *J. Biol. Chem.* 2000, *275*, 30019–30028.

[20] Wilkinson, S. R., Temperton, N. J., Mondragon, A., Kelly, J., *J. Biol. Chem.* 2000, *275*, 8220–8225.

[21] Antelmann, H., Engelmann, S., Schmid, R., Hecker, M., *J. Bacteriol.* 1996, *178*, 6571–6578.

[22] Prosperi, M. T., Ferbus, D., Karczinshi, I., Goubin, G., *J. Biol. Chem.* 1993, *268*, 11050–11056.

[23] Nemato, Y., Yamamoto, T., Takada, S., Matsui, Y., Obinata, M., *Gene* 1990, *91*, 261–265.

[24] Shau, H., Butterfield, L. H., Chiu, R., Kim, A., *Immunogenetics* 1994, *40*, 129–134.

[25] Konig, J., Baier, M., Horling, F., Kahmann, U., Harris, G., Schurmann, P., Dietz, K. J., *Proc. Natl. Acad. Sci. USA* 2002, *99*, 5738–5743.

[26] Dietz, K. J., Horing, F., Konig, J., Baier, M., *J. Exp. Bot.* 2002, *53*, 1321–1329.

[27] Prouzet-Mauleou, V., Monribot-Espagne, C., Boucherie, H., Lagniel, G., Lopez, S., Labrre J., Garin, J., Lauquin, G. J. M., *J. Biol. Chem.* 2002, *277*, 4823–4830.

[28] Eymann, C., Dreisbach, A., Albrecht, D., Bernhardt, J., Becher, D., Gentner, S., Tam, L. T., Büttner, K., Buurman, G., Scharf, C., Venz, S., Völker, U., Hecker, M., *Proteomics* 2004, *4*, 2849–2876.

[29] Voigt, B., Schweder, T., Becher, D., Ehrenreich, A., Gottschalk, G., Feesche, J., Maurer, K. H., Hecker, M., *Proteomics* 2004, *4*, 1465–1490.

[30] Bernhardt, J., Volker, U., Volker, A., Antelmann, H., Schmid, R., Mach, H., Hecker, M., *Microbiology* 1997, *143*, 999–1017.

[31] Gohar, M., økstad, O. A., Gilois, N., Sanchis, V., Kolst, A. B., Lereclus, D., *Proteomics* 2002, *2*, 784–791.

[32] Büttner, K., Bernhardt, J., Scharf, C., Schmid, R., Mäder, U., Eymann, C., Antelmann, H., Volker, A., Volker, U., Hecker, M., *Electrophoresis* 2001, *22*, 2908–2935.

[33] Sinchaikul, S., Sookkheo, B., Phutrakul, S., Pan, F. M., Chen, S. T., *Protein Expres. Purif.* 2001, *22*, 388–398.

[34] Sookkheo, B., Sinchaikul, S., Phutrakul, S., Chen, S. T., *Protein Expres. Purif.* 2000, *20*, 142–151.

[35] Sookkheo, B., Sinchaikul, S., Thannan, H., Thongprasong, O., Phutrakul, S., Chen, S. T., *Proteomics* 2002, *2*, 1311–1315.

[36] Berggren, K., Chernokalskaya, E., Steinberg, T. H., Kemper, C., Lopez, M. F., Diwu, Z., Haugland, R. P., Patton, W. F., *Electrophoresis* 2000, *21*, 2509–2521.

[37] Lee, C. L., Hsiao, H. H., Lin, C. W., Wu, S. P., Huang, S. Y., Wu, C. Y., Wang, A. H. J., Khoo, K. H., *Proteomics,* 2003, *3*, 2472–2486.

[38] Perkins, D. N., Pappin, D. J. C., Crasy, D. M., Cottrell, J. S., *Electrophoresis* 1999, *20*, 3551–3567.

[39] Weber, H., Engelmann, S., Becher, D., Hecker, M., *Mol. Microbiol.* 2004, *52*, 133–140.

[40] Yang, K. S., Kang, S. W., Woo, H. A., Hwang, S. C., Chae, H. Z., Kim, K., Rhee, S. G., *J. Biol. Chem.* 2002, *277*, 38029–28036.

[41] Wagner, E., Luche, S., Penna, L., Chevallet, M., Dorsselaser, A. V., Leize-Wagner, L., Rabilloud, T., *Biochem. J.* 2002, *366*, 777–785.

[42] Rabilloud, T., Heller, M., Gasneir, F., Luche, S., Rey, C., Aebersold, R., Benahmed, M., Louisot, P., Lunardi, J., *J. Biol. Chem.* 2002, *22*, 19396–19401.

[43] Sayed, A. A., William, D. L., *J. Biol. Chem.* 2004, *in press*.

[44] Chae, H. Z., Robison, K., Poole, L. B., Church, G., Storz, G., Rhee, S. G., *Proc. Natl. Acad. Sci. USA* 1994, *91*, 7017–7021.

[45] Hofmann, B., Hecht, H. J., Flohe, L., *Biol. Chem.* 2002, *383*, 347–364.

[46] Wood, Z. A., Schroder, E., Harris, J. R., Poole, L. B., *Trends Biochem. Sci.* 2003, *28*, 32–40.

[47] Yanagawa, T., Ishikawa, T., Ishii, T., Tabuchi, K., Iwasa, S., Bannai, S., Omura, K., Suzuki, H., Yoshida, H., *Cancer Lett.* 1999, *145*, 127–132.

[48] Ichimiya, S., Davis, J. G., O'Rourke, D. M., Katsumata, M., Greene, M. I., *DNA Cell Biol.* 1997, *16*, 311–321.

[49] Ariel, N., Zvi, A., Makarova, K. S., Chitlaru, T., Elhanany, E., Velen, B., Cohen, S., Friedlander, A. M., Shaffermann, A., *Infect. Immun.* 2003, *71*, 4563–4579.

[50] Cabiscol, E., Tamarit, J., Ros, J., *Int. Microbiol.* 2000, *3*, 3–8.

[51] Wood, Z. A., Poole, L. B., Karplus, A., *Science* 2003, *300*, 650–653.

[52] Woo, H.A., Chae, H.Z., Hwang, S. C., Yang, K. S., Kang, S. W., Kim, K., Rhee, S. G., *Science* 2003, *300*, 656–653.

[53] Claiborne, A., Yeh, J. I., Mallett, T. C., Luba, J., Crane, E. J., Charrier, V., Parsonge, D., *Biochemistry* 1999, *38*, 15407–15416.

[54] Schroder, E., Littlechild, J. A., Lebedev, A. A., Errington, N., Vagin, A. A., Isonpv, M. N., *Struct. Fold Des.* 2000, *8*, 605–615.

ELSEVIER

Journal of Chromatography B, 771 (2002) 261–287

JOURNAL OF
CHROMATOGRAPHY B

www.elsevier.com/locate/chromb

Review

Bioinformatics, functional genomics, and proteomics study of *Bacillus* sp.

Supachok Sinchaikul[a], Boonyaras Sookkheo[b], Supachai Topanuruk[a,c], Hsueh-Fen Juan[a], Suree Phutrakul[c], Shui-Tein Chen[a,*]

[a]*Institute of Biological Chemistry, Academia Sinica, 128 Yen Chiu Yuan Rd., Sec II, Nankang, Taipei, Taiwan 11529*
[b]*Chemistry Unit, School of Science, Rangsit University, Patumthani, Thailand 12000*
[c]*Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand 50200*

Abstract

The ability of bioinformatics to characterize genomic and proteomic sequences from bacteria *Bacillus* sp. for prediction of genes and proteins has been evaluated. Genomics coupling with proteomics, which is relied on integration of the significant advances recently achieved in two-dimensional (2-D) electrophoretic separation of proteins and mass spectrometry (MS), are now important and high throughput techniques for qualifying and analyzing gene and protein expression, discovering new gene or protein products, and understanding of gene and protein functions including post-genomic study. In addition, the bioinformatics of *Bacillus* sp. is embraced into many databases that will facilitate to rapidly search the information of *Bacillus* sp. in both genomics and proteomics. It is also possible to highlight sites for post-translational modifications based on the specific protein sequence motifs that play important roles in the structure, activity and compartmentalization of proteins. Moreover, the secreted proteins from *Bacillus* sp. are interesting and widely used in many applications especially biomedical applications that are the highly advantages for their potential therapeutic values. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Reviews; Bioformatics; Proteomics; Functional genomics; *Bacillus*

Contents

1. Introduction 262
2. Genome and differentially expressed genes of *Bacillus* sp. under different conditions 262
   2.1. Functional classification of gene products 263
   2.2. Differentially expressed genes under different conditions 263
3. Proteomics analysis of *Bacillus* sp. proteins 265
   3.1. Secrete to extracellular 268
   3.2. On the cell membrane 271
   3.3. In the cytosol 271

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

*Corresponding author. Tel.: +886-2-278-55981, ext. 7071; fax: +886-2-278-83473.
E-mail address: bcchen@gate.sinica.edu.tw (S.-T. Chen).

1570-0232/02/$ – see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S1570-0232(02)00054-5

4. Post-translational modification of gene product in *Bacillus* sp. 273
5. Bioinformatics to link differential display genes and proteomic identified proteins 277
6. Biomedical applications of *Bacillus* sp. 278
7. Conclusions 285
Acknowledgements 285
References 286

1. Introduction

Over the past few years, up to 56 microorganisms have had their genomes completely sequenced, with another 170 or in progress (http://www.tigr.org). The genome of *Bacillus* sp. has been found only two strains that were *B. subtilis* and *B. halodurans* [1,2]. It does have the other strains that have not been yet completely determined their genomes such as *B. stearothermophilus* strain 10, *B. cereus* and *B. licheniformis*. In the generally characteristics, the members of the bacterial genus *Bacillus* sp. more than 368 strains are Gram-positive, endospore forming, rod-shaped bacteria that grow aerobically [3]. They exhibit a wide range of physiological abilities that allow them to live in every natural environment. In addition, they are important organisms for researches and industrial applications such as they are used in many medical, pharmaceutical, agricultural, and industrial processes that take advantage of their wide range of physiological characteristics and their ability to produce a host of enzymes, antibodies and other metabolites. Due to more availability and applications of *Bacillus* sp., the study of genome is increasingly important and need to know the biological system of these bacteria, which intends to understand the gene function, metabolic pathway and proteomics including protein structure. For this review, it is discovered the bioinformatics, functional genomics and proteomics study of *Bacillus* sp. It intends to be a literature search for *Bacillus* strains that provides a guide to choosing the most efficient way to analyze a new sequence or to collect information on a gene or protein of interest by applying current publicly available databases and Web servers, and also help to analyze genomic and proteomic data clearly. In addition, the known functional genomics will underlie both a better understanding of cellular function and a more efficient application of biotechnology. Moreover, the thermophilic bacteria *B. stearothermophilus* P1 and TLS33 isolating from a hot spring in Chiang Mai, Thailand, are another strains that used to study the proteomic analysis in both extracellular and cytosol, and present as the first report. It is very interesting to their identified proteins and functions that will be useful for the other strains of *Bacillus* sp. to get more information.

2. Genome and differentially expressed genes of *Bacillus* sp. under different conditions

The outcome of genomics is the discipline of functional genomics whereby one uses DNA-based technologies to make inferences about organism structure and behavior. The complete genomic nucleotide sequence of the spore-forming bacterium *Bacillus* sp. has a few strains and the well-known strain is *B. subtilis*. The competition of the *B. subtilis* genomic nucleotide sequence has revealed 345 small polypeptide open-reading frames (of 85 codons or less) and comprises a total of approximately 4100 protein-coding genes with 4.2 Mbp in size and 42% unknown function [1]. Of these protein-coding genes, 53% are represented once, while a quarter of the genome corresponds to several gene families that have been greatly expanded by gene duplication, the largest family containing 77 putative ATP-binding transport proteins. Otherwise, the complete genome sequence of the alkaliphilic bacterium *B. halodurans* C-125 has been known since 2000 year [2]. Its 4 202 353-bp genome contains 4066 predicted protein-coding sequences (CDSs), 2141 (52.7%) of which have functional assignments, 1182 (29%) of which are conserved CDSs with unknown function and 743 (18.3%) of which have no match to any protein databases. Among the total CDSs, 8.8% match sequences of proteins found only in *B. subtilis* and 66.7% are widely conserved in comparison with the proteins of various organisms, including *B. subtilis*.

Copyright © by Chiang Mai University
All rights reserved

Although the bacterial genus *Bacillus* sp. has many strains but the genetic and physical maps may be different; for example, the assigned positions of the genes on the physical map of *B. halodurans* C-125 were compared with those on the genetic and physical maps of *B. subtilis* and showed that the positions of some putative genes of *B. halodurans* C-125 differed completely from *B. subtilis* orthologues, suggesting that genome organization is not conserved between these two strains, although the size and GC content of the C-125 chromosome are almost identical to those of *B. subtilis* (4.2 Mb and 43.7 mol%, respectively) [4]. In addition, the comparison of completely genomes between *B. subtilis and B. halodurans* also shows in Table 1.

2.1. Functional classification of gene products

Genes were classified, which based on the representation of cells as Turing machines in which one distinguishes between the machine and the program using BLAST2P software running against a composite protein databank compound of SWISS PROT, TREMBL and *B. subtilis* proteins [1]. It is well known that 58% of the *B. subtilis* proteins having known a function but up to 42% of the gene products have never been predicted by similarity to proteins of known function: 16% of the proteins are similar to unknown proteins from *B. subtilis* and some other organism; and 26% of the proteins are not significantly similar to any other proteins in databanks. The example of the functional classification devised for *B. subtilis* protein coding genes is shown in Table 2, which based on the one hand on the distinction between the machinery of the chemical reactions taking place in the cell [5].

2.2. Differentially expressed genes under different conditions

Most of *Bacillus* sp. initiates a series of transitional responses that are designed to maintain or restore growth. These include the induction of macromolecular hydrolases such as proteases, lipases and polysaccharidases, chemotaxis and mobility, and competence (i.e., the ability to take up DNA from the environment). If these response fail to reestablish growth, sporulation is induced. The ability of *Bacillus* sp. to form highly resistant endospores imparts an enormous competitive advantage in environments

Table 1
Comparison of the general features of the *B. subtilis* and *B. halodurans* genomes

Genome features	*B. subtilis*	*B. halodurans*
Size (bp)	4 214 814	4 202 353
G+C content (mol%)		
Total genome	43.5	43.7
Coding region	44.3	44.4
Non-coding region	39.3	39.8
Open reading frames		
Percent of genome (coding)	87	85
Protein coding genes	4104	4066
Conserved with function assigned	2379	2144
Conserved with unknown function	668	1182
Non-conserved	1053	743
Percent AUG initiation codons	78	78
Percent GUG initiation codons	9	12
Percent UUG initiation codons	13	10
Insertion elements		
PBSX prophage-related protein	27	2
Transposase and related protein	10	112
RNA elements		
Stable RNA (percent of genome)	1.27	1.02
16S, 23S and 5S rRNA	10	8
tRNA	86	78

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Table 2
Functional classification of *B. subtilis* protein coding genes

Category	Number
1. Cell envelope and cellular processes	867 (21%)
1.1 Cell wall	93
1.2 Transport/binding proteins and lipoproteins	381
1.3 Sensors (signal transduction)	38
1.4 Membrane bioenergetics (electron transport chain and ATP synthase)	78
1.5 Mobility and chemotaxis	55
1.6 Protein secretion	18
1.7 Cell division	21
1.8 Sporulation	139
1.9 Germination	23
1.10 Transformation/competence	21
2. Intermediary metabolism	742 (18%)
2.1 Metabolism of carbohydrates and related molecules	261
2.1.1 Specific pathways	214
2.1.2 Main glycolytic pathways	28
2.1.3 TCA cycle	19
2.2 Metabolism of amino acids and related molecules	205
2.3 Metabolism of nucleotides and nucleic acids	83
2.4 Metabolism of lipid	77
2.5 Metabolism of coenzymes and prosthetic groups	99
2.6 Metabolism of phosphate	9
2.7 Metabolism of sulfur	8
3. Information pathways	482 (12%)
3.1 DNA replication	22
3.2 DNA restriction/modification and repair	39
3.3 DNA recombination	17
3.4 DNA packaging and segregation	10
3.5 RNA synthesis	244
3.5.1 Initiation	19
3.5.2 Regulation	213
3.5.3 Elongation	8
3.5.4 Termination	4
3.6 RNA modification	19
3.7 Protein synthesis	96
3.7.1 Ribosomal proteins	56
3.7.2 Aminoacyl-tRNA synthetases	25
3.7.3 Initiation	6
3.7.4 Elongation	6
3.7.5 Termination	3
3.8 Protein modification	27
3.9 Protein folding	8
4. Other functions	289 (7%)
4.1 Adaptation to atypical conditions	72
4.2 Detoxification	68
4.3 Antibiotic production	30
4.4 Phage-related functions	83
4.5 Transposon and IS	10
4.6 Miscellaneous	26
5. Similar to unknown proteins	667 (16%)
5.1 From *B. subtilis*	177
5.2 From other organisms	490
6. No similarity	1053 (26%)

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

such as soil, where long periods of drought and nutrient deprivation are common. At least 4% of the *B. subtilis* genome is dedicated to the processes of sporulation, germination and outgrowth.

The *Bacillus* sp. genome encodes more than 40 temperature-shock and general stress proteins, which presumably contribute to this organism's ability to survive shorter periods of adversity. It encodes numerous ''heat shock inducible'' genes, including chaperones and proteases [6]. Proteins induced in response to lower temperatures include cold acclimatization proteins, for those synthesized continuously at low temperature, and cold shock proteins (Csp), for those synthesized in response to a sudden temperature downshift [7]. The adaptation of bacteria to highly alkaline environments was interesting and also studied in *Bacillus* sp. because many alkaliphilic *Bacillus* strains produce unique alkaline enzymes such as xylanases, proteases and amylases [8]. One of the most commonly encountered is likely to be osmotic stress resulting from frequent wetting and drying of its habitat [9]. For example, the specific and general stress proteins in *B. subtilis* in which response to heat shock, salt and ethanol stress, and glucose and phosphate starvation was analyzed using 2-D PAGE analysis [10]. The induction of at least 42 general stress proteins absolutely required the alternative sigma factor σ^B except in oxidative stress condition. However, at least seven stress proteins, among them ClpC, ClpP, Sod, AhpC and AhpF, remained stress inducible in a *sigB* mutant. Such a detailed analysis also permitted the description of subgroups of general stress proteins, which are subject to additional regulatory circuits, indicating a very thorough fine-tuning of this complex response. The relative synthesis rate of the general stress proteins constituted up to 40% of the total protein synthesis of stressed cell and thereby emphasizes the importance of the stress regulation. Moreover, the oxidative stress stimulation in *B. subtilis* was also found more than 50 expressed proteins by proteomic approach using 2-D PAGE [11]. Not only the cytoplasmic stress response systems direct the upregulation genes encoding heat shock proteins, molecular chaperones and proteases but the extracytoplasmic secretion stress system also regulates the genes for the survival of severe secretion stress. In recently, a novel two-component regulatory system in *B. subtilis*, named CssR–CssS, was required for the cell to survive to severe secretion stress caused by a combination of high-level production of the α-amylase AmyQ and reduced levels of the extracytoplasmic folding factor PrsA [12].

In addition, the differentially expressed genes of *Bacillus* sp. can be found in substrate utilization such as carbohydrate, ammonium, opines and inorganic phosphate. For example, the ability of *Bacillus* sp. to secrete a wide variety of extracellular macromolecular depolymerases is a major factor contributing to their colonization of soil [13]. Genes encoding secreted amylases, arabinases, chitonases, mannanases, cellulases and xylanasess are evident in the genome sequence. Proteases and lipases are also frequently encountered, both intracellular and extracellular, the latter allowing proteins to provide sources of both carbon and nitrogen.

3. Proteomics analysis of *Bacillus* sp. proteins

Proteomics analysis is concerned with the global changes in protein expression as most commonly accomplished by a combination of two-dimensional gel electrophoresis (2-DE) to separate and visualize proteins and mass spectrometry (MS) for protein identification. Emerging technologies for proteome analysis also include microarray and microfluidic devices. There are called proteomics, which is the large-scale analysis of proteins, is contributed greatly to understanding gene function in post-genomic era even through the protein chemistry. Important technical advances related on 2-DE and protein MS have increased sensitivity, reproducibility and throughput of proteome analysis while creating an integrated technology.

In addition, artificial neural network were trained on the prediction of the subcellular location of bacterial proteins. A cross-validated average prediction accuracy of 93% was reached for distinction between cytoplasmic and non-cytoplasmic proteins, based on the analysis of protein amino acid composition [14]. The greatest differences in residue frequencies between cytoplasmic and non-cytoplasmic sequences are found for Glu, Leu, Ile and Arg (predominance in cytoplasmic sequences), and for Ser, Asn, Gly, Thr and Ala (predominance in

Copyright © by Chiang Mai University
All rights reserved

Table 3
Extracellular proteins of *B. subtilis* 168; the proteins were identified by mass spectrometry

Proteins	Function/similarity
Metabolism of carbohydrates	
AbnA	Arabinan-endo 1,5-α-L-arabinase
AmyE[M]	α-Amylase
BglC	Endo-1,4-β-glucanase, cellulase
BglS	Endo-β-1,3-1,4 glucanase
CitH	Malate dehydrogenase
Csn[H]	Chitosanase
Eno	Enolase
PdhB	Pyruvate dehydrogenase (E1 β subunit)
PdhD	Pyruvate dehydrogenase (E3 subunit)
Pel[H]	Pectate lyase
PelB	Pectate lyase
XynA	Endo-1,4-β-xylanase
XynD[lipo 2 TM H]	Endo-1,4-β-xylanase
YnfF	Endo-xylanase
YvgN	Similar to plant-metabolite dehydrogenase
YvpA	Pectate lyase
YwjH	Similar to transaldolase (pentose phosphate)
YxlA	Arabinan-endo 1,5-α-L-arabinase
Metabolism of amino acids	
RocA	Pyrroline-5 carboxylate dehydrogenase
RocF	Arginase
Metabolism of proteins	
AprE	Serine alkaline protease (subtilin E)
Bpr	Bacillopeptidase F
Epr	Minor extracellular serine protease
Ggt[TM]	γ-Glutamyltranspeptidase
Mpr	Extracellular metalloprotease
NprE	Extracellular neutral metalloprotease
Vpr[H]	Extracellular serine protease
WprA[RR W 3H]	Cell wall-associated protein precursor
YwaD	Aminopeptidase
Metabolism of nucleotides and nucleic acids	
YfkN[RR TM pst]	2′,3′-Cyclic-nucleotide 2′-phosphodiesterase
YhcR[RR TM]	5′-Nucleotidase
Yurl	Ribonuclease
Metabolism of lipids	
LlpA[RR]	Lipase
Metabolism of phosphate	
GlpQ[pst]	Glycerophosphoryl diester phosphodiesterase
PhoA[pst]	Alkaline phosphatase A
PhoB[pst]	Alkaline phosphatase III
PhoD[RR pst]	Phosphodiesterase/alkaline phosphatase D
Metabolism of the cell wall	
CwlC[W ex]	*N*-Acetylmuramoyl-L-alanine amidase
LytD[W]	*N*-Acetylglucosaminidase (major autolysin)
PbpA[TM 3]	Penicillin-binding protein 2A
PbpC[lipo lgt]	Penicillin-binding protein 3
PbpX[RR]	Penicillin-binding protein
WapA[RRW≠ 4 H]	Cell wall-associated protein precursor
Todj[lipo lgt]	D-Alanyl-D-alanine carboxypeptidase
YvcE[W]	Cell wall-binding protein
YwtD[W H]	Similar to murein hydrolase
Mobility and chemotaxis	
Hag[dual H]	Flagellin protein
FlgK[ex]	Flagellar hook-associated protein 1 (HAP1)
FliD[ex]	Flagellar hook-associated protein 2 (HAP2)

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Table 3. Continued

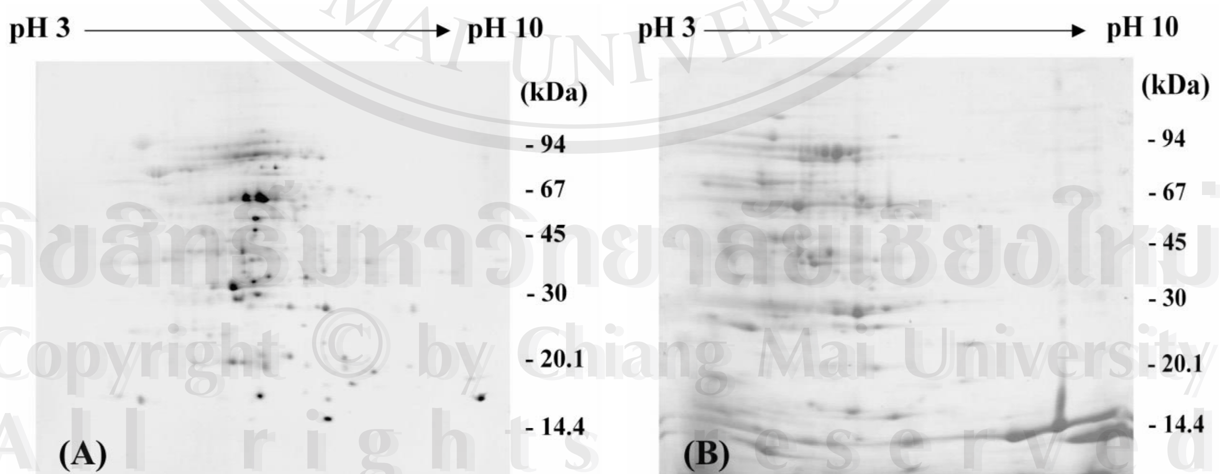
Proteins	Function/similarity
Protein synthesis (Elongation)	
Fus, Ef-G	Elongation factor G
Protein folding	
GroEL	Class I heat-shock protein (chaperonin)
Transport/binding proteins and lipoproteins	
FeuA$^{\text{lipo lgt}}$	Iron-uptake system (binding protein)
FhuD$^{\text{lipo lgt}}$	Ferrichrome ABC transporter (binding protein)
MntA$^{\text{lipo}}$	Manganese ABC transporter (binding protein)
MsmE$^{\text{lipo lgt}}$	Multiple sugar ABC transporter (binding protein)
OppA$^{\text{lipo}}$	Oligopeptide ABC transporter (binding protein)
OpuAC$^{\text{lipo pst lgt}}$	Glycine betaine ABC transporter(binding protein)
PstS$^{\text{lipo pst}}$	Phosphate ABC transporter (binding protein)
RbsB$^{\text{lipo lgt}}$	Ribose ABC transporter (binding protein)
YcdH$^{\text{lipo pst}}$	Zinc ABC transporter (binding protein)
YclQ$^{\text{lipo H}}$	Ferrichrome ABC transporter (binding protein)
YdhF$^{\text{RR lipo pst}}$	Similar to unknown proteins from *B. subtilis*
YflY$^{\text{lipo lgt}}$	Iron (III) didtrate ABC transporter (binding protein)
YfmC$^{\text{lipo}}$	Ferrichrome ABC transporter (binding protein)
YqlX$^{\text{lipo pst}}$	Amino acid ABC transporter (binding protein)
YrpE$^{\text{lipo pst lgt}}$	Similar to unknown proteins from *B. subtilis*
YusA$^{\text{lipo lgt}}$	Probably part of the S box regulon
YxeB$^{\text{lipo lgt}}$	ABC transporter (binding protein)
Sporulation	
TasA$^{\text{H}}$	Translocation-dependent antimicrobial spore component
RNA synthesis/regulation	
Ywt$^{\text{F}}$	Transcriptional regulator
Detoxification	
KatA$^{\text{H}}$	Vegetative catalase 1
PenP$^{\text{H}}$	β-Lactamase precursor
SodA$^{\text{H}}$	Superoxide dismutase
Ybxl	Similar to β-lactamase
YceD	Similar to tellurium resistance protein
Phage-related functions	
XepA$^{\text{ex}}$	PBSX prophage lytic exoenzyme
XkdG$^{\text{ex H}}$	PBSX prophage gene
XkdK$^{\text{ex}}$	PBSX prophage gene
KkdM$^{\text{ex}}$	PBSX prophage gene
XlyA$^{\text{W ex}}$	*N*-Acetylmuramoyl-L-alanine amidase
Similar to unknown proteins	
YbdN	Unknown
Yfnl$^{\text{TM}\neq\text{H}}$	Probable transmembrane glycoprotein
YlqBH	Unknown
YncM$^{\neq\text{H}}$	Similar to unknown proteins from *B. subtilis*
YolA	Unknown
YoaW	Unknown
YrpD$^{\text{lipo 2}\neq}$	Similar to unknown proteins from *B. subtilis*
YuaB	Unknown
YvgO	Unknown
YweA$^{3\neq}$	Similar to unknown proteins from *B. subtilis*
YwoF	Unknown
YxaK$^{\text{H}}$	Similar to unknown proteins from *B. subtilis*
YxkC$^{3\text{ H}}$	Unknown

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

periplasmic and extracellular sequences). On average, these differences are more significant for extracellular proteins than for periplasmic proteins. Periplasmic proteins seem to contain a higher amount of Pro, Lys, Cys and Met compared to extracellular sequences.

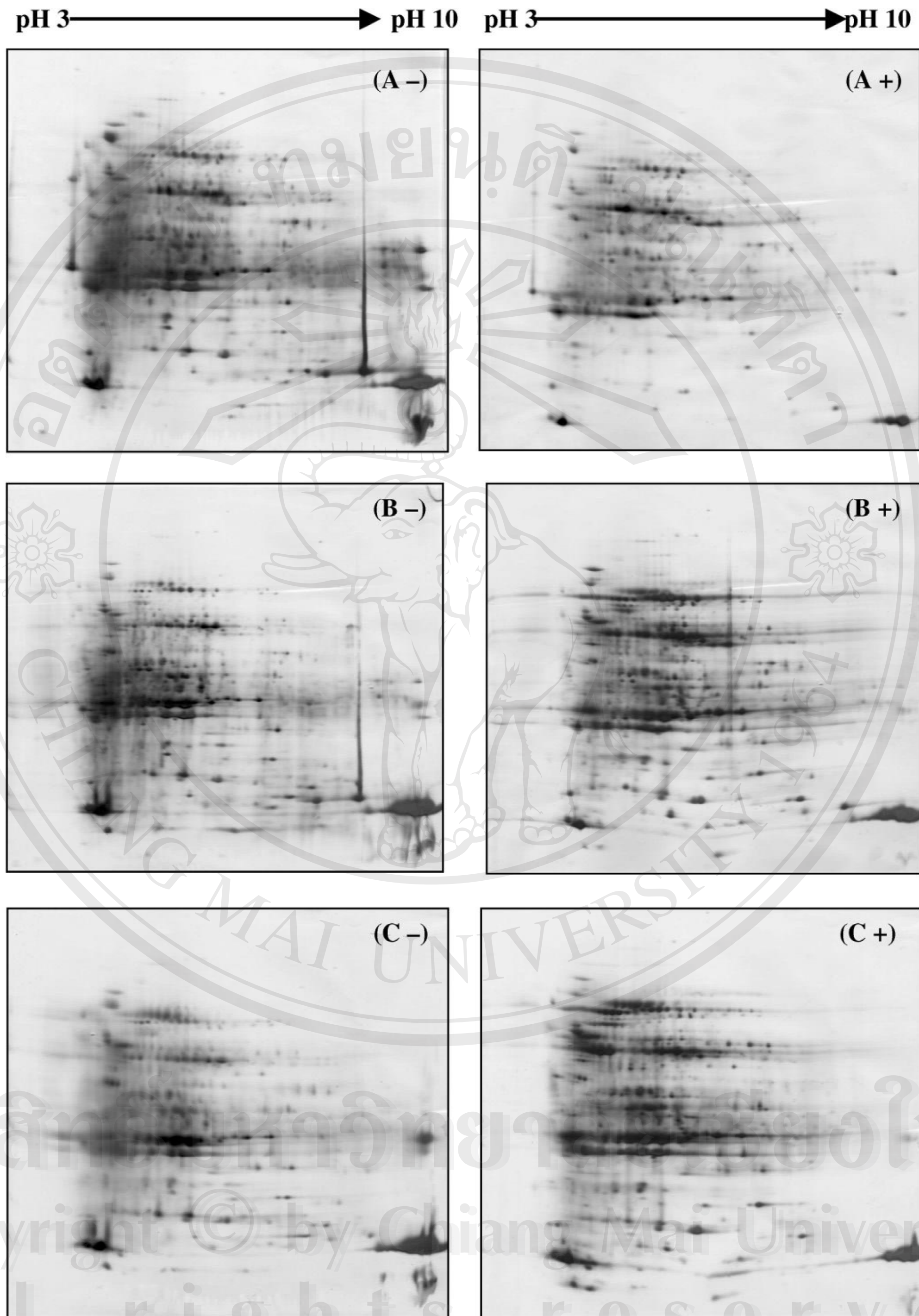
3.1. Secrete to extracellular

It is well known that most species of *Bacillus* strains have a high capacity to secrete proteins into the culture medium such as *B. subtilis*, *B. licheniformis*, *B. amyloliquefaciens*, *B. firmus*, *B. stearothermophilus*, *B. catenulatus* and *B. thermoleovolans*. However, the properties of their proteins vary from strain to strain. In *B. subtilis* genome, several genes encoding proteins of the major secretion pathway have been identified: *secA*, *secD*, *secE*, *secF*, *secY*, *ffh* and *ftsY*. In addition, the proteomic approaches were used to identify the extracellular complement of the *B. subtilis* secretome and showed that using different growth conditions and a hypersecreting mutant, ~200 extracellular proteins were visualized by 2-D PAGE of which 82 identified proteins were identified by MALDI-TOF mass spectrometry [15]. The 82 identified extracellular proteins include 50 proteins to which a function had been assigned previously (Table 3). In addition, the secreted proteins from xylan-induced *B. firmus* K-1 were analyzed using 2-D PAGE and found that five stimulatory secreted proteins containing xylose isomerase, xylanases, GltC regulatory protein and 3-dehydroquinate dehydratase were expressed under substrate induction using xylan [16]. Moreover, the study of extracellular proteins from *B. stearothermophilus* using 2-D PAGE that have never been reported also showed in this review. In our previously reports, the bacterium *B. stearothermophilus* strain P1 and TLS33 have been known that they could secrete the extracellular lipase and proteases, respectively [17,18]. It is therefore studied the extracellular proteins using proteomic approach and the preliminary result of 2-D gel comparing between extracellular and intracellular proteins showed the different protein patterns (Figs. 1 and 2). Otherwise, it also found the extracellular superoxide dismutase (SOD) with the molecular mass approximately 28 kDa and p*I* 5.8 [19]. Furthermore, the substrate-induced lipase from *B. stearothermophilus* P1 using olive oil induction not only was studied the proteomic analysis using 2-D PAGE but it was also analyzed using Agilent 2100 Bioanalyzer and Protein 200 LabChip kit (AgilentTechnologies, USA), which is the new technique for protein sizing and chip-based separation. The analyzed proteins under substrate induction were shown the high-resolution separation on



ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Fig. 1. Comparison of the cytosolic proteins (A) and extracellular proteins (B) from a thermophilic bacterium *B. stearothermophilus* TLS33 by 2-D PAGE analysis using a wide pH range of IPG strip 3–10, 18 cm.



ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Fig. 2. 2-D PAGE gels of the extracellular proteins after culturing the thermophilic bacterium *B. stearothermophilus* P1 in the presence (+) and absence (—) of 0.5% (v/v) olive oil at 65 °C for 12 (A), 24 (B), 36 (C), 48 (D), 72 (E) and 96 h (F). Dying with silver staining.

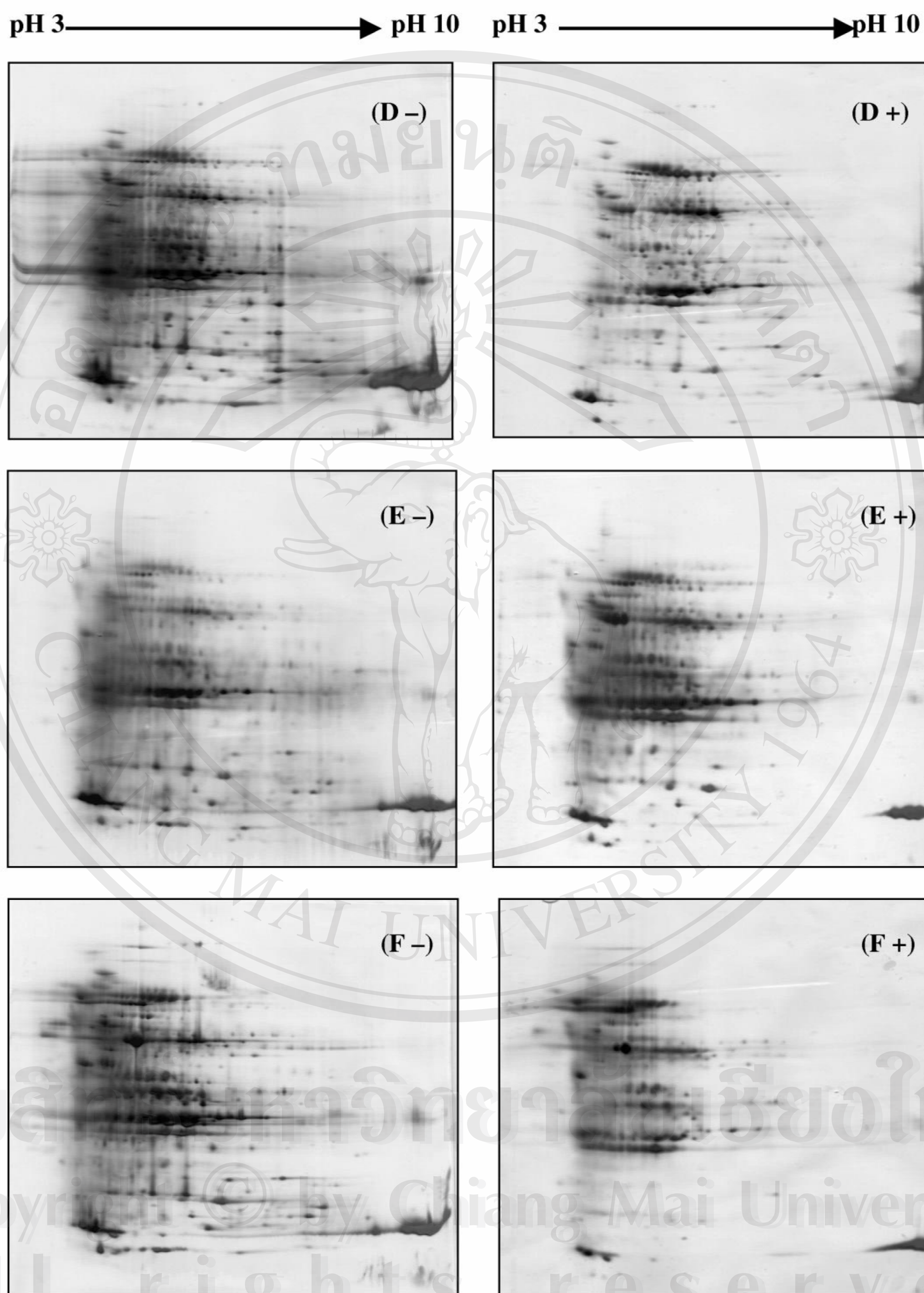


Fig. 2. (*continued*)

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

SDS–PAGE with the different protein patterns (Fig. 3) and reviewed in real-time as separation on different electropherograms (Fig. 4).

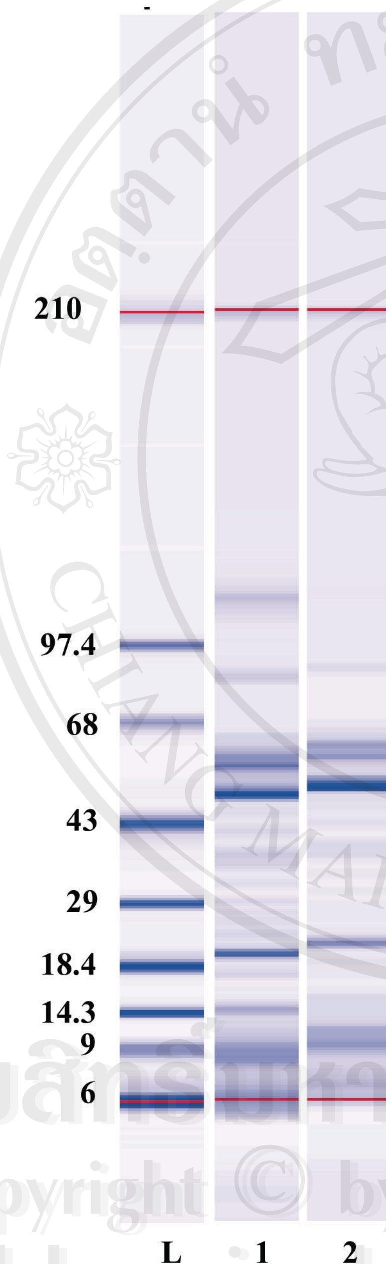


Fig. 3. Analysis of olive oil-induced lipase from *B. stearothermophilus* P1 with a linear gel, gradient gel, and the Protein 200 assay using Agilent 2100 Bioanalyzer. Labels: (L) protein marker (kDa); (1) 0.5% (v/v) olive oil induction; (2) non-induction.

3.2. On the cell membrane

The membrane proteins are found in a few strains of *Bacillus* sp., for example, the membrane proteinase from *B. cereus* [20], δ-endotoxin and insecticidal crystal protein Cry1C from *B. thuringiensis* [21,22], and penicillinase from *B. licheniformis* [23]. Most Gram-positive *Bacillus* sp. are found to be extracellular proteins and have a few proteins in membrane portions. The membrane proteins in *Bacillus* sp. have rarely been identified using 2-D PAGE analysis. Otherwise, the membrane-associated or -bound proteins are difficult to underrepresent on 2-D gels because they are generally low abundant, alkaline p*I* and poorly soluble in aqueous media used for isoelectric focusing. However, it has also found in the genome of *B. subtilis* with 29.2% membrane proteins and the proteomic approach using 2-D PAGE found some membrane proteins such as YfnI, LytC, MurG and Ffh [24,25]. Furthermore, the hydropathy profile alignment can be introduced as a tool to search for structural homologs of membrane proteins and the *B. subtilis* genome was found 109 sequence homologs of membrane proteins [26].

3.3. In the cytosol

Bacillus species have also been used to produce a number of intracellular proteins on an industrial scale, including glucose isomerase [27], thermostable enzymes such as glucokinases [28], and a variety of restriction endonucleases such as *Bam*HI, *Bci*I, *Bgl*II and *Bst*I [29]. The proteomic analysis of cytosolic proteins in *Bacillus* sp. had a few study; for example, the biofilm proteome of whole cell proteins from *B. cereus* strain 5 was investigated using 2-D PAGE and 10 proteins were synthesized as result of surface attachment of which four were unique to biofilm profile [30], the cytosolic proteins of *B. subtilis* 168 was also identified up to 50 proteins using 2-D PAGE [31,32], and the cold shock protein in a thermophilic bacterium *B. stearothermophilus* P1 was also found with the molecular mass approximately 7.3 and p*I* 4.5 using 2-D PAGE (Fig. 5) and the comparison of downshift temperatures from 65 to 37 °C and 25 °C by ImageMaster 2D Elite software showed the different protein patterns (Fig. 6) [33]. In addition, the new protein analysis technique of

Copyright © by Chiang Mai University
All rights reserved

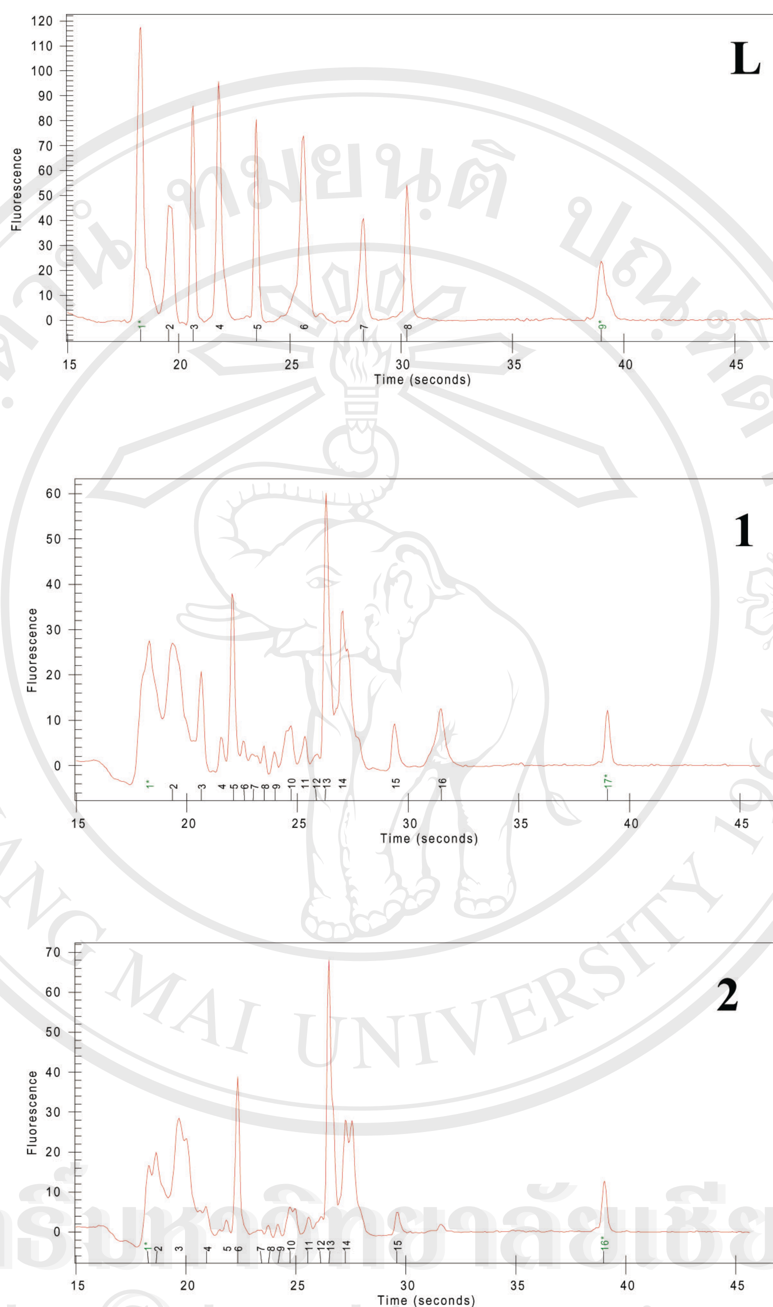


Fig. 4. Electropherogram of olive oil-induced lipase from *B. stearothermophilus* P1 demonstrating the resolution of the Protein 200 assay using Agilent 2100 Bioanalyzer. Labels: (L) protein marker (kDa); (1) 0.5% (v/v) olive oil induction; (2) non-induction.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

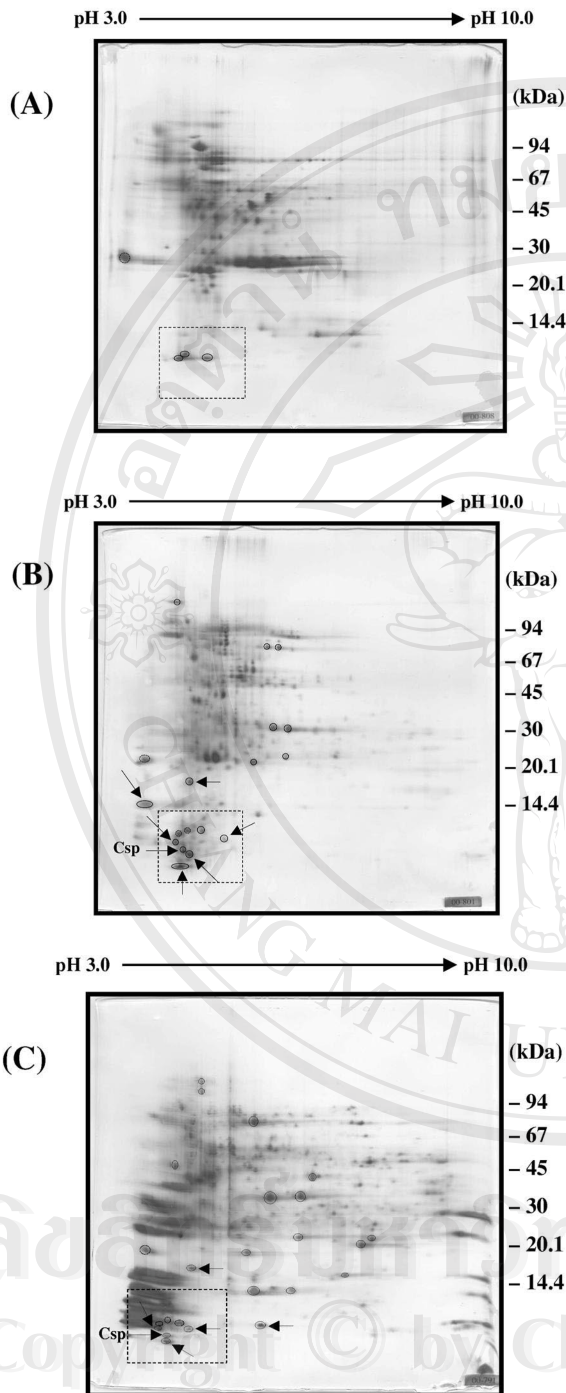


Fig. 5. 2-D PAGE mapping of cold shock protein induction varies with time. The low-molecular mass region of 2-D gels of protein extracts from *B. stearothermophilus* P1 after different times of cold shock induction at 37 and 25 °C. (A) Protein pattern at 65 °C as reference; (B) protein pattern after shift from 65 to 37 °C; (C) protein pattern after shift from 65 to 25 °C.

Agilent 2100 Bioanalyzer using Protein 200 assay kit also analyzed the cold shock stress proteins. The protein resolution by Protein 200 assay showed the high-resolution separation of different protein patterns on SDS–PAGE (Fig. 7) and different electropherograms (Fig. 8). Moreover, the cytosolic proteins of *B. stearothermophilus* strain TLS33 after the downshift temperatures at 37 and 25 °C was also identified using 2-D PAGE that have never been reported before and the cytosolic proteins on 2-D gels among different temperatures were different protein patterns (Fig. 9). The visualized analysis of different proteins after downshift temperatures was analyzed using ImageMaster 2D Elite software (Fig. 10) and could be also plotted to be the analytical graphs between the downshift temperatures at 37 or 25 °C and control at 65 °C (Fig. 11). It indicated that if the proteins at 37 °C near the function $y = 10x$, these proteins were significantly differential expressed from the control at 65 °C. In addition, the 53 detectable proteins in cell extract were identified with MALDI-TOF spectrometry and the data shows in Table 4.

4. Post-translational modification of gene product in *Bacillus* sp.

Post-translational modification of proteins is an important key for regulation in many cellular processes including recognition, signaling, targeting and metabolism. The post-translational modifications are generally divided into eight types, which lead to a charge-dependent change to a protein, such as acylation, alkylation, carboxylmethylation, phosphorylation, sulfation, carboxylation, sialylation and proteolytic processing [34]. The differences of peptide mass fingerprinting data including isoelectric point (p*I*) and molecular masses between experimental and theoretical peptides can be also provided information about the exact mass and p*I* of the post-translational modification in addition to theirs sequence and the research community could annotate such information as such in the databases for further use.

The proteome can be compartmentalized into different organelles or into essentially three categories of protein that are intracellular, membrane and extracellular. Extracellular and transmembrane

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

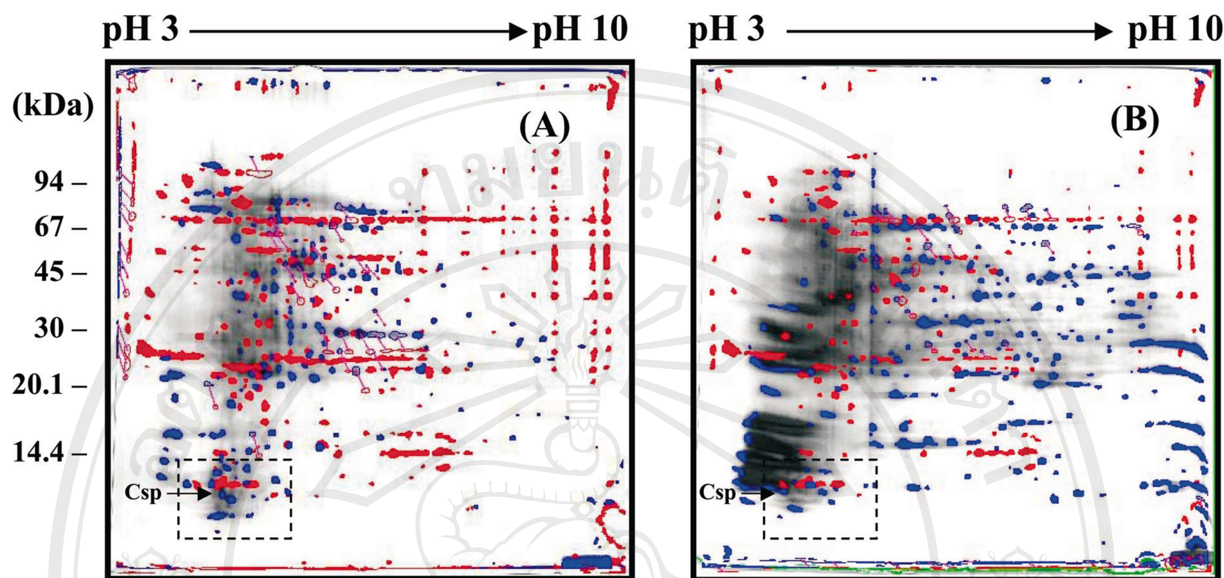


Fig. 6. 2-D Gel image comparison of the protein contents from *B. stearothermophilus* P1 after temperature downshifts from 65 to 37 °C (A) and 25 °C (B) by ImageMaster 2D Elite software. The same analysis parameters show the sensitivity of 8475, operator size of 31, noise factor of 5, background factor of 1 and split level of 7. Labels: blue color spot, protein contents at 37 and 25 °C; red color spot, protein contents at 65 °C as a reference; no color spot with the line, similar protein spot.

proteins are frequently glycosylated, whereas intracellular signaling molecules can be basically or inducible phosphorylated. It is thought that at any one time approximately one-third of all proteins in eukaryotic cells are phosphorylated. Although it is more difficult to detect and analyze post-translational modification containing peptides, post-translational modifications are frequently discovered during large-scale proteomic experiments. In the study of post-translational modifications of *Bacillus* sp., it still has a few reports; for example, Lentz et al. [35] studied the modification of fatty acid specificity of cytochrome P450 BM-3 (CYP102) from *B. megaterium* by directed evolution using a validated assay. The CYP102 enzyme catalyzing the subterminal hydroxylation of fatty acids with a chain length of 12–22 carbons was mutated variants towards caprylic, capric and lauric acid by site-specific random mutagenesis at position 87 of the mutant F87A(LARV). The best mutants, F87V(LAR) and F87V(LARV), showed a higher catalytic activity toward ω-(*p*-nitrophenoxy)decanoic acid (10-*p*-NCA) than F87A(LARV). In addition, they proved capable of hydroxylating ω-(*p*-nitrophenoxy)octanoic acid (8-*p*-NCA) which the wild-type enzyme is unable to do. Both variants catalyzed hydroxylation of capric acid, which is not a substrate for the wild-type, with a conversion rate of up to 57%. The chain length specificity of the mutants in fatty acid hydroxylation processes shows a good correlation with their activity towards *p*-NCA pseudosubstrates. The *p*-NCA assay therefore, allows high-throughput screening of large mutant libraries for the identification of enzyme variants with the desired catalytic activity towards fatty acids as the neutral substrates. In another study by Lessard et al. [36], the gene encoding the dihydrolipoyl acetyltransferase (E2) and dihydrolipoyl dehydrogenase (E3) components of the pyruvate dehydrogenase (PDH) multienzyme complex from *B. stearothermophilus* were overexpressed in *E. coli* and used the post-translational modification for improvement of stoichiometry of the subunit interaction in assembly in vitro. The lipoyl domain of the recombinant E2 protein was found to be unlipoylated, but it could be correctly post-translationally modified in vitro using a recombinant lipoate protein ligase from *E. coli*. The lipoylated E2 component was able to bind recombinant E1 and E3

Copyright © by Chiang Mai University
All rights reserved

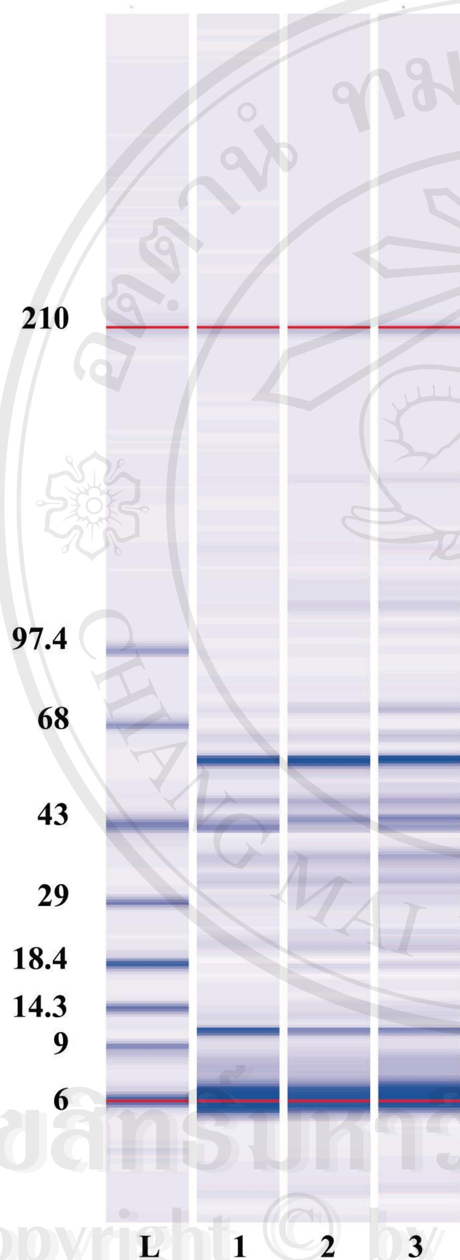


Fig. 7. Analysis of cold shock-stress proteins from *B. stearothermophilus* P1 after downshift temperatures from 65 to 37 °C and 25 °C with a linear gel, gradient gel, and the Protein 200 assay using Agilent 2100 Bioanalyzer. Labels: (L) protein marker (kDa); (1) control at 65 °C; (2) downshift temperature at 37 °C; (3) downshift temperature at 25 °C.

components in vitro to generate a PDH complex with a catalytic activity comparable with that of the wild-type enzyme. Therefore, the use of post-translational modification could be assembly of a fully active PDH complex from recombinant proteins in vitro. On the other hand, Schreier et al. [37] also studied the *B. subtilis glnR* mutants defective in regulation by post-translational modification. The *B. subtilis glnR* gene encoded a 135-amino acid repressor, GlnR, regulates *glnRA* transcription in response to nitrogen levels in the growth medium. Two glnR mutants unable to repress under nitrogen excess conditions were mutated at Gly72 that lie within α-helix-turn-α-helix (HTH) motif by site-directed mutagenesis and affected regulation, suggesting that amino acids within the putative HTH region were critical for GlnR function and might be involved in DNA binding. The mutations within the C-terminal region of GlnR at amino acid 110, 116, 123 and 129 were found to affect regulation because the production of truncated proteins were constitutively repressed and the substitution of Asp129 with Asn also led to loss of repression. In addition, the study of Sakamoto et al. [38] reported that the *meso*-diaminopimelate dehydrogenase gene from *B. sphaericus* was expressed in *E. coli* and the N-terminal amino acid of both the enzyme from native and transformant were serine, indicating that the N-terminal methionine is removed by post-translational modification. Even though the post-translational modifications cannot currently be studied at high throughput, certain categories such as phosphorylation, acylation or alkylation also influence protein structure and function, and the known protein sequence specificity.

Furthermore, the post-translational modification databases and resources can be found on the Internet with websites:

(1) DSDBASE, disulfide database derived from 3D data (http://www.ncbs.res.in/~faculty/mini/dsdbase/dsdbase.html),
(2) GlycoSuiteDB, database of glycan structures (http://www.glycosuite.com),
(3) *O*-GlycBase, *O*-glycosylated protein database (http://www.cbs.dtu.dk/databases/OGLYCBASE),

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

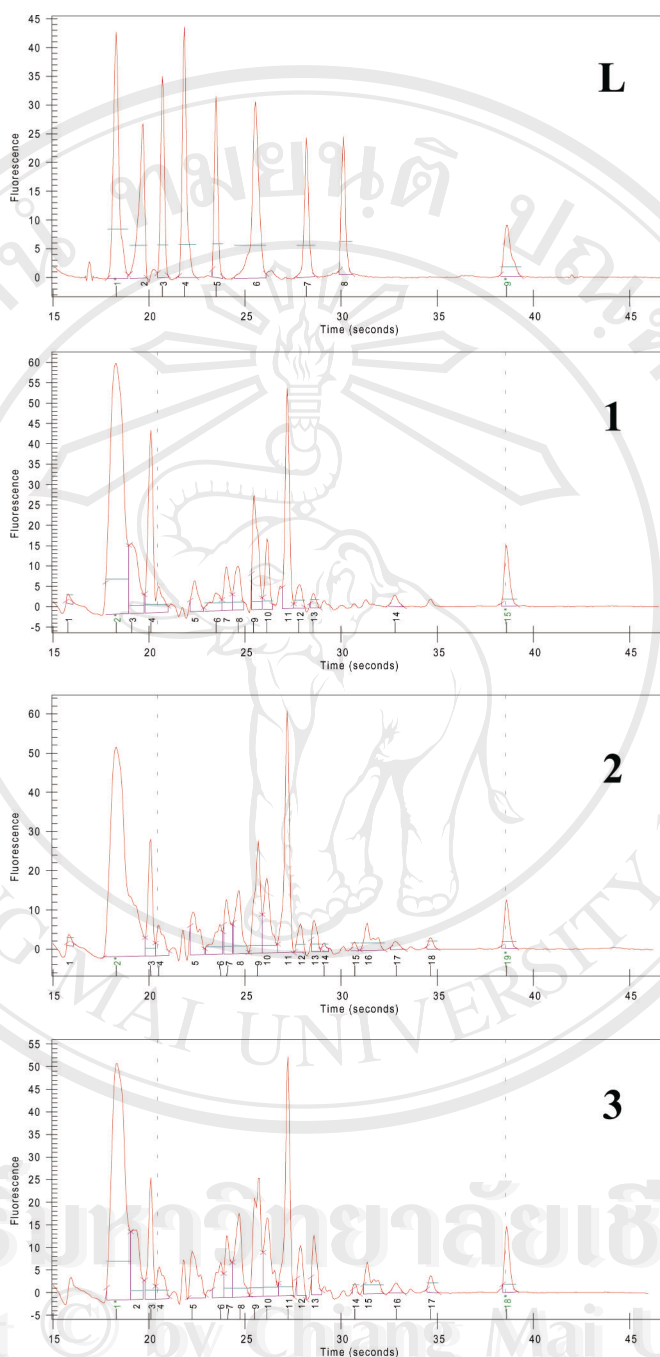


Fig. 8. Electropherogram of cold shock-stress proteins from *B. stearothermophilus* P1 after downshift temperatures from 65 to 37 °C and 25 °C demonstrating the resolution of the Protein 200 assay using Agilent 2100 Bioanalyzer. Labels: (L) protein marker (kDa); (1) control at 65 °C; (2) downshift temperature at 37 °C; (3) downshift temperature at 25 °C.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

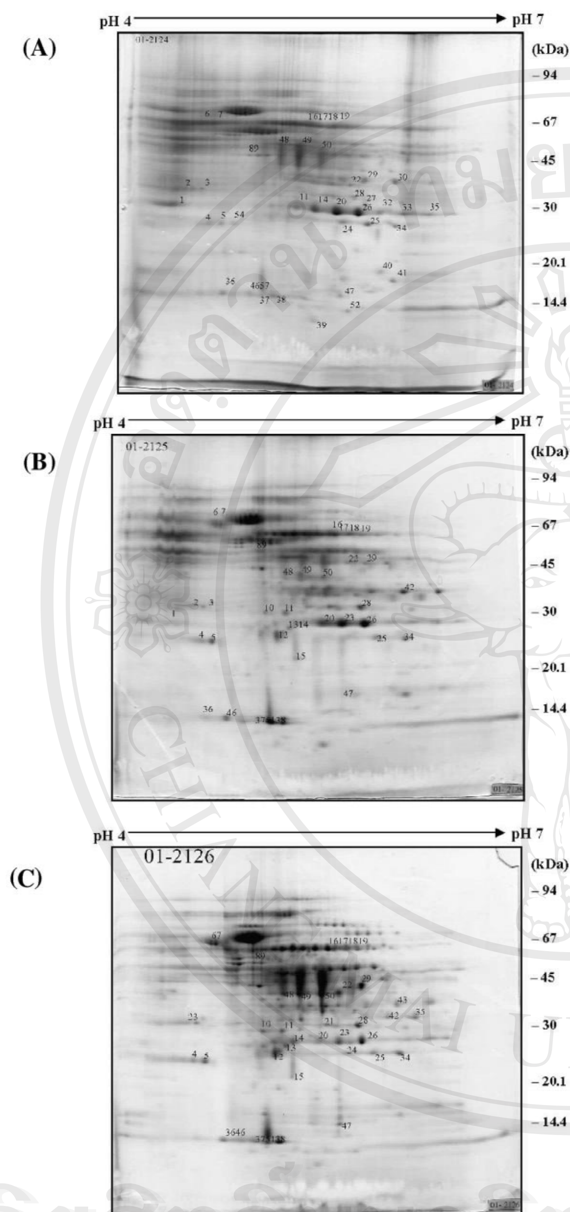


Fig. 9. The cytosolic proteins in *B. stearothermophilus* TLS33 after downshift temperatures to 37 °C (B) and 25 °C (C) analyzed by 2-D PAGE using a narrow pH range of IPG strip 4–7, 18 cm. The control is used the cell extract at 65 °C (A).

(4) PhosphoBase, phosphorylation site database (http://www.cbs.dtu.dk/database/PhosphoBase), and
(5) RESID, database of amino acid modifications (http://www-nbrf.georgetown.edu/pirwww/dbinfo/resid.html)

5. Bioinformatics to link differential display genes and proteomic identified proteins

Bioinformatics are now important and very useful for searching the unknown sequence including gene function. An up-to-date list of the databases of *Bacillus* sp. has several websites and shows in Table 5. Some of these are available on the Internet with links to many provided from the ExPASy proteomics server including protein identification and characterization programs (http://www.expasy.ch/tools.html). These allow not only identification of proteins but further characterization ranging from the calculation of basic physiochemical properties to the prediction of potential post-translational modifications and three-dimensional structures. Annotated protein and two-dimensional electrophoresis databases is the bioinformatic core of proteome research. Swiss-Prot is a typical example of such as annotated database. Many proteome projects are now underway, resulting in the generation of two-dimensional electrophoresis database that are accessible on the Internet and can be browsed with interactive software and integrated with in-house results. SubtiList database is a one server that provides a complete dataset of DNA and protein sequences derived from the paradigm strain *B. subtilis* 168, linked to the relevant annotations and functional assignments, and available at http://genolist.pasteur.fr/SubtiList. It allows one to easily browse through these data and retrieve information, using various criteria (such as gene names, location, keyword, functional category, etc.). This server supplementing with EMBL/GenBank/DDBJ databanks is also used for identification of the cytosolic proteins in *B. stearothermophilus* TLS33 showing the identified protein and gene including accession numbers, description and functional category (Table 6). Another database of Clusters of Orthologous Groups of protein (COG) is a new database search and represents an attempt on a phylogenetic classification of proteins from complete genome (http://www.ncbi.nlm.nih. gov/COG) [39,40]. It is to serve as a platform for functional annotation of newly sequence genomes and for

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

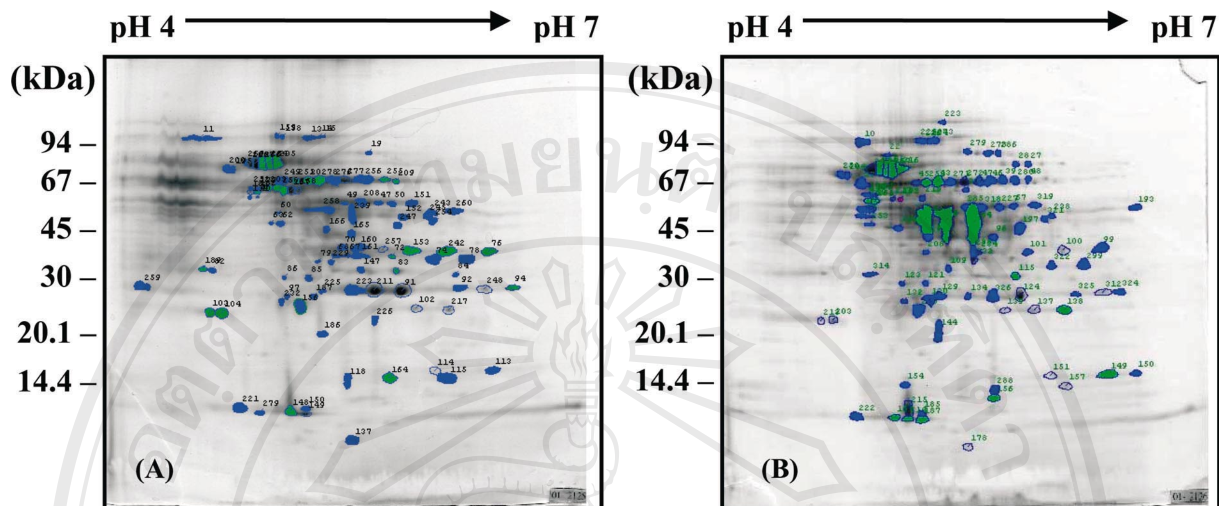


Fig. 10. 2-D Gel image comparison of the cytosolic proteins from *B. stearothermophilus* TLS33 after downshift temperatures from 65 to 37 °C (A) and 25 °C (B) by ImageMaster 2D Elite software. Labels: blue color spots, decreased protein spots; red color spot, increased protein spots; no color spot, no change of proteins.

studies on genome evolution. The protein sequences encoded in two complete genomes of *B. subtilis* and *B. halodurans* were compared by COG database using COGNITOR program (Table 7). It can be used to search the BLAST protein homologs and identified the protein functions. Not only the COG database, the proteome analysis of *B. subtilis* and *B. halodurans* is also identified and shown in the Proteome Analysis@EBI database from website http://www.ebi.ac.uk/proteome that the SWISS-PROT/TrEMBL complete non-redundant proteome sets are constructed by selecting entries from SPTR, which is a comprehensive protein sequence database consisting of SWISS-PROT, TrEMBL and TrEMBLnew (Table 8) [41]. Otherwise, the InterPro comparative analysis of the proteome of *B. subtilis* with another available proteome of *B. halodurans* or *E. coli* can be investigated and shown the functional classification of organism using Gene Ontology (GO) with the general statistics. Therefore, it is very useful and can be used as a bioinformatic tool. Furthermore, we have also developed Bulk Gene Search System (BGSS) that is able to correlated gene identifying from cDNA microarray with their corresponding protein functions. The developed database search of BGSS, which is composed of the UniGene, LocusLink and Proteome databases, can be easily used to find the associated protein functions and related information in the worldwide databases by applying only accession numbers [42]. The BGSS database of *Bacillus* derived from Swiss-Prot and SubtiList database is now available at http://bio-.clkao.org/perl/genequery2.pl that is very useful for analyzing the whole *Bacillus* genomes in the nearly future.

6. Biomedical applications of *Bacillus* sp.

In the present, the effect of genome and proteome based approaches on biomedical research has not yet been achieved. However, the exciting progress is being made, and brief overviews of several biomedical areas are given below to illustrate the potential of this approach. The role of *Bacillus* sp. in biomedical applications is higher important coupled with knowing genome of some species as shown in Table 9. For example, Shih et al. [43] reported the chemistry and biosynthesis of poly-γ-glutamic acid (γ-PGA) produced by various strains of *Bacillus* sp. such as *B. anthracis*, *B. natto*, *B. subtilis*, *B. anthracis* and *B. licheniformis*. γ-PGA is an unusual anionic, naturally occurring homo-polyamide that is made of D- and L-glutamic acid units connected by

Copyright © by Chiang Mai University
All rights reserved

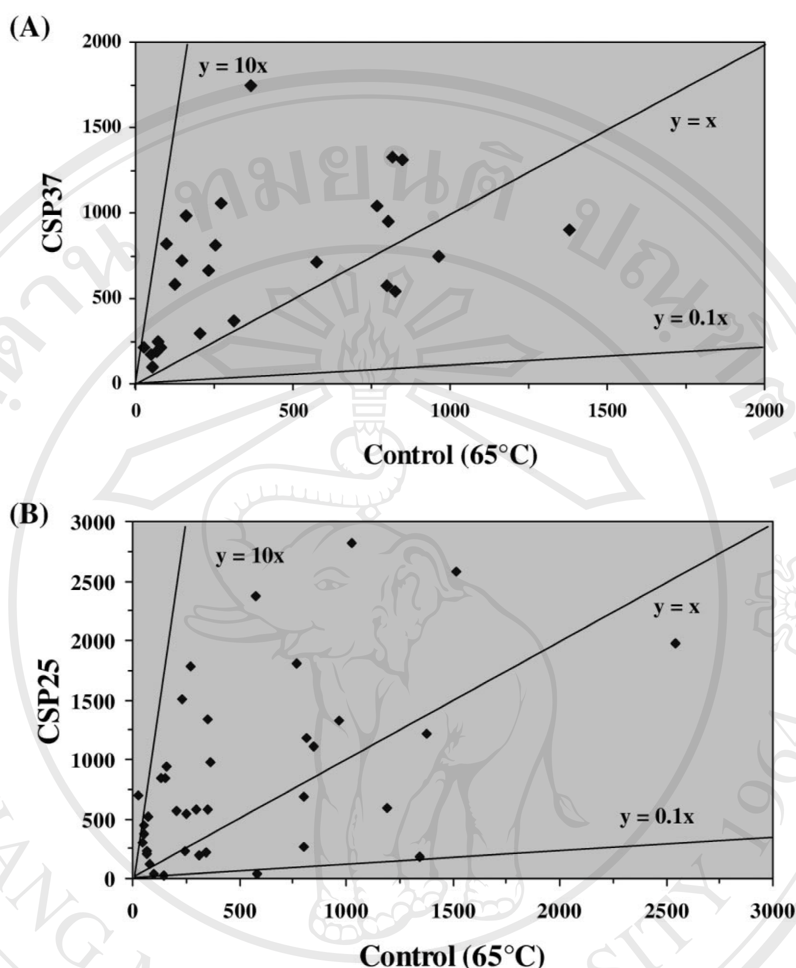


Fig. 11. The relationships of cytosolic proteins from *B. stearothermophilus* TLS33 between control at 65 °C and downshift temperature at 37 °C (A) or 25 °C (B).

amide linkages between α-amino and γ-carboxylic acid groups. Due to suitable properties of γ-PGA such as water-soluble, biodegradable, edible and non-toxic toward humans and the environment, their potential applications have been of interest in the past few years in a broad range of industrial fields such as fields of food, cosmetics, medicine and water treatments. γ-PGA and its derivatives also offer a wide range of unique applications including being used as thickener, cryoprotectant, bitterness relieving agent, sustained release materials, humectant, drug carrier, biodegradable fibers, biological adhesive, highly water-absorbable hydrogels, biopolymer flocculants, heavy metal absorber and animal feed additives. Most applications of γ-PGA and its derivatives in medicine are used as drug carrier and biological adhesives. For example, polyglutamic pacitaxel (PG-TXL) or Taxol is an anticancer agent that exhibited markedly antitumor activity against murine tumors and human tumor xenografts [44,45], prostaglandin E_1 (PGE_1) is clinically used to treat peripheral vascular disturbance, in addition, it is also effective on fulminant or subfulminant viral hepatitis by means of its cytoprotective activity [46], the cross-linking of gelatin and poly-(L-glutamic acid) has been shown to be promising as a surgical adhesive and hemostatic agent [47–50], the cured gel or gelatin–PGA aqueous solutions is lowly biodeg-

Copyright © by Chiang Mai University
All rights reserved

Table 4
Protein identification of cytosolic proteins in *B. stearothermophilus* TLS33 after downshift temperatures to 37 and 25 °C by MS-FIT software (http://prospector.ucsf.edu/ucsfbin3.4/msfit.cgi) and SwissProt and TrEMBL (http://www.expasy.ch/sprot)

Spot no.	App. M_w (Da)/pI	MOWSE score	Matched	Theoretical M_w (Da)/pI	Accession no.	Protein name
1	35051/4.37	207	9/29 (31%)	18225.3/8.96	P26380	PTS system, fructose-specific IIB component (EIIB-FRU) (fructose-permease IIB component) (phosphotransferase enzyme II, B component) (P18)
2	39191/4.55	4.32	4/14 (28%)	43843.1/6.55	P49782	Stage III sporulation protein AE
3	39191/4.62	13.2	5/17 (29%)	55725.3/6.31	P21879	Inosine-5′-monophosphate dehydrogenase (IMP dehydrogenase) (IMPDH) (IMPD) (superoxide-inducible protein 12) (SOI12)
4	29901/4.76	15.9	4/23 (17%)	26413.8/9.11	P39147	COMF operon protein 3
5	78100/4.64	–	–	–	–	Unknown
6	78100/4.64	118	4/27 (14%)	13536.7/5.71	P39740	Flagellar protein flit
7	78250/4.55	8.66	26/47 (55%)	79078.9/8.97	P39814	DNA topoisomerase I (ω-protein) (relaxing enzyme) (untwisting enzyme) (swivelase)
8	78700/4.44	2.72	5/18 (27%)	78304.4/6.07	P13484	Probable poly (glycerol-phosphate) α-glucosyltransferase (teichoic acid biosynthesis protein E)
9	78100/4.64	3.74	29/40 (72%)	78622.2/5.40	P17889	Translation initiation factor IF-2
10	36067/5.1	43.1	6/30 (20%)	38120.6/5.01	P54518	Putative peptidase in GCVT-SPOIIIAA intergenic region
11	34803/5.14	31.8	4/30 (13%)	14349.6/6.30	P42411	Anti-σB factor RSBT
12	31533/5.2	71.5	11/32 (34%)	29759.9/5.24	Q45499	Extragenic suppressor protein SUHB homolog
13	33103/5.3	–	–	–	–	Unknown
14	49187/5.53	51.6	5/22 (22%)	28560.1/6.56	P50733	Hypothetical 28.6-kDa protein in RECQ-CMK intergenic region precursor
15	49187/5.53	51.6	11/37 (29%)	56309.1/5.34	P39773	2,3-bisphosphoglycerate-independent phosphoglycerate mutase (phosphoglyceromutase) (BPG-independent PGAM) (vegetative protein 107) (VEG107)
16	69284/5.49	5.57	26/46 (56%)	68587.5/5.17	P37949	GTP-binding protein LEPA
17	69284/5.61	157	12/38 (31%)	56309.1/5.34	P39773	2,3-bisphosphoglycerate-independent phosphoglycerate mutase (phosphoglyceromutase) (BPG-independent PGAM) (vegetative protein 107) (VEG107)
18	69486/5.72	–	–	–	–	Unknown
19	69284/5.81	7.31	4/11 (36%)	20182.6/5.75	P54154	Peptide methionine sulfoxide reductase (protein-methionine-s-oxide reductase) (peptide MET(O) reductase)
20	686284/5.55	29.6	11/28 (39%)	38639.1/5.39	P50863	MRP protein homolog
21	68614/5.59	137	12/31 (38%)	58175.0/4.95	P50848	Hypothetical 58.2-kDa protein in KDGT-XPT intergenic region
22	39144/5.66	64.2	14/30 (46%)	46609.5/5.52	P25995	Dihydroorotase (dhoase)
23	41924/5.63	186	8/27 (29%)	20126.3/5.43	Q06795	Transcription antitermination protein NUSG
24	33229/5.67	16	4/29 (13%)	29955.0/6.46	P49938	Ferrichrome transport ATP-binding protein FHUC
25	30331/5.85	4.12	5/16 (31%)	32510.1/8.77	P24327	Protein export protein PRSA precursor
26	29802/6.16	84.1	6/24 (25%)	23303.0/5.78	O34565	Amidotransferase HISH
27	35465/5.84	84.1	6/24 (25%)	23303.0/5.78	O34565	Amidotransferase HISH
28	3485/5.75	8.46	6/36 (16%)	35012.1/6.47	O34504	Hypothetical 35.0-kDa protein in RAPJ-OPUAA intergenic region
29	37121/4.55	15.6	18/36 (50%)	60763.5/6.42	P40408	Hypothetical transcriptional regulator in FEUA-SIGW intergenic region (ORF3)
30	78250/6.07	–	–	–	–	Unknown
31	41178/5.85	23.6	8/30 (26%)	43247.7/6.51	O34992	Glycine betaine/carnitine/choline transport ATP-binding protein OPUCA
32	32081/6.01	2.09	6/21 (28%)	99563.2/4.94	P39793	Penicillin-binding protein 1A/1B (PBP1) [includes: penicillin-insensitive transglycosylase (peptidoglycan TGASE); penicillin-sensitive transpeptidase (DD-transpeptidase)]
33	32484/6.14	5.76	4/37 (10%)	30244.1/9.15	P39785	PBSX phage terminase small subunit
34	29604/6.05	1.86	4/18 (22%)	26949.8/8.79	P54537	Probable amino acid ABC transporter ATP-binding protein in BMRU-ANSR intergenic region
35	32236/6.3	303	4/26 (15%)	19119.6/4.62	P40405	Hypothetical 19.1-kDa protein in SIGD-RPSB intergenic region precursor (ORFC)
36	18139/4.6	48.4	4/33 (12%)	14573.7/5.79	P17622	RIBT protein
37	18200/4.77	21.1	4/26 (15%)	14538.0/5.54	P39910	PAL-related lipoprotein precursor
38	17648/5.15	91.7	4/33 (12%)	6189.2/4.93	2897794	(D49467) Unnamed protein product

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Table 4. Continued

Spot no.	App. M_w (Da)/pI	MOWSE score	Matched	Theoretical M_w (Da)/pI	Accession no.	Protein name
39	12806/5.44	48.8	4/24 (16%)	11124.6/5.23	P21468	30S Ribosomal protein S6 (BS9)
40	22031/5.95	125	5/28 (17%)	26974.8/5.58	P94498	Phosphoadenosine phosphosulfate reductase (PAPS Reductase, Thioredoxin-dependent) (padops Reductase) (3′-phosphoadenylylsulfate reductase) (PAPS sulfotransferase)
41	20546/6.05	–	–	–	–	Unknown
42	28549/6.14	47.6	11/27 (40%)	73207.7/5.01	P54496	Hypothetical 73.2-kDa protein in SODA-COMGA intergenic region
43	40764/6.32	34.1	13/23 (56%)	43836.3/4.84	P08495	Aspartokinase 2 (aspartokinase II) (aspartate kinase 2) [contains: aspartokinase II α subunit; aspartokinase II β subunit]
44	43986/5.25	202	17/36 (47%)	45333.3/6.12	P45742	Hypothetical 45.3-kDa protein in PRKA-CSPB intergenic region (ORF4)
45	16284/5.35	17	6/29 (20%)	17895.3/5.37	O07636	Hypothetical 17.9-kDa protein in NPRE-PYCA intergenic region
46	18200/4.77	–	–	–	–	Unknown
47	17587/5.64	5.69	4/10 (40%)	23805.8/4.78	P49785	Stage III sporulation protein AH
48	61667/5.2	103	9/28 (32%)	16231.7/8.38	P39807	Hypothetical 16.2-kDa protein in COMF-FLGM intergenic region
49	58758/5.34	181	9/32 (28%)	29728.7/5.31	P39650	Prespore-specific transcriptional activator RSFA
50	58515/5.52	240	21/35 (60%)	50527.5/5.04	P80860	Glucose-6-phosphate isomerase (GPI) (phosphoglucose isomerase) (PGI) (phosphohexose isomerase) (PHI) (vegetative protein 54) (VEG54)
51	20249/5.07	–	–	–	–	Unknown
52	14461/5.7	129	4/33 (12%)	21023.0/6.59	P37569	Hypothetical 21.0-kDa protein in LYSS-MECB intergenic region
53	18445/5.87	22.1	5/23 (21%)	19145.1/5.79	O31704	Molybdopterin-guanine dinucleotide biosynthesis protein B

Table 5
Database searches of *Bacillus* sp.

Server	Web sites
General	
1. LBSN	http://www.bacterio.cict.fr/b/bacillus.html
2. DSMZ	http://www.dsmz.de/bactnom/nam0379.htm
Genome project	
1. Entrez	http://www.ncbi.nlm.nih.gov/Entrez
2. *B. stearothermophilus* Genome Sequencing–Strain 10	http://www.genome.ou.edu/bstearo.html
3. *B. subtilis* Genome Sequencing Project	http://genolist.pasteur.fr/SubtiList/help/project.html
4. *B. halodurans* Genome Project–JAMSTEC	http://www.sun01.hydra.mki.co.jp:8093/jamstec/micrHome.html
Genome analysis (DNA sequence data)	
1. DDBJ	http://www.ddbj.nig.ac.jp
2. EMBL	http://www.ebi.ac.uk/embl/index.html
3. Genbank	http://www.ncbi.nlm.nih.gov/Genbank/Genbank/GenbankSearch.html
4. SubtiList	http://genolist.pasteur.fr/SubtiList
5. BSORF (JAFAN)	http://bacillus.genome.ad.jp/BSORF-DB.html
6. NRSub (Non-Redundant *B. subtilis* Database)	http://pbil.univ-lyon1.fr/nrsub/nrsub.html
7. Micado	http://locus.jouy.inra.fr/cgi-bin/genmic/madbase/progs/madbase.operl
8. PEDANT	http://pedant.gsf.de/index.html
9. EBI GeneQuiz	http://jura.ebi.ac.uk:8765/ext-genequiz/genomes/bs0005
10. GIB (DDBJ)	http://gib.genes.nig.ac.jp/Bs
11. TIGR	http://www.tigr.org/tdb/mdb/mdb/mdbcomplete.html
Proteome analysis (Protein sequence data)	
1. Proteome Analysis@EBI	http://www.ebi.ac.uk/proteome/BACSU (*B. subtilis*), or BACHD (*B. halodurans*)
2. COGs	http://www.ncbi.nlm.nih.gov/COG
3. Protein Data Bank	http://www.rcsb.org/pdb
4. SwissProt	http://kr.expasy.org/sprot
5. Sub-2D	http://microbio2.biologie.uni-greifswald.de:8880

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Table 6
Protein and gene identification, description, and functional category of cytosolic proteins in *B. stearothermophilus* TLS33 after downshift temperatures at 37 and 25 °C by SubtiList databases, supplemented with EMBL/GenBank/DDBJ databanks

Spot no.	Protein ID/no.	Gene name/no.	Description	Functional category
1	PTFB_BACSU (P26380)	*levE* (X56098)	PTS fructose-specific enzyme IIB component	Transport/binding proteins and lipoproteins
2	S3AE_BACSU (P49782)	*spoIIIAE* (U35252)	Mutants block sporulation after engulfment	Sporulation
3	IMDH_BACSU (P21879)	*guaB* (X55669)	Inosine-monophosphate dehydrogenase	Metabolism of nucleotides and nucleic acids
4	CMF3_BACSU (P39147)	*comFC* (Z18629)	Late competence gene	Transformation/competence
5	Unknown	–	–	–
6	FLIT_BACSU (P39740)	*fliT* (Z31376)	Flagellar protein	Mobility and chemotaxis
7	TOP1_BACSU (P39814)	*topA* (L27797)	DNA topoisomerase I	DNA packing and segregation
8	TAGE_BACSU (P13484)	*tagE* (X15200)	UDP-glucose: polyglycerol phosphate glucosyltransferase	Cell wall
9	IF2_BACSU (P17889)	*infB* (M34836)	Initiation factor IF-2	Initiation
10	YQHT_BACSU (P54518)	*yqhT* (D84432)	Unknown; similar to Xaa-Pro dipeptidase	Protein modification
11	RSBT_BACSU (P42411)	*rsbT* (L35574)	Positive regulator of σB activity (switch protein/serine-threonine kinase)	Adaptation to atypical conditions
12	SUHB_BACSU (Q45499)	*suhB* (AF012285)	–	–
13	Unknown	–	–	–
14	YPBG_BACSU (P50733)	*ypbG* (L47648)	Unknown; similar to unknown proteins	From other organisms
15	PMGI_BACSU (P39773)	*pgm* (L29475)	Phosphoglycerate mutase	Main glycolytic pathways
16	LEPA_BACSU (P37949)	*lepA* (X91655)	GTP-binding protein	Elongation
17	PMGI_BACSU (P39773)	*pgm* (L29475)	Unknown	–
18	Unknown	–	–	–
19	MSRA_BACSU (P54154)	*msrA* (L77246)	Peptidyl methione sulfoxide reductase	Detoxification
20	MRP_BACSU (P50863)	*mrp* (X74737)	–	–
21	YPWA_BACSU (P50848)	*ypwA* (L47838)	Unknown; similar to carboxypeptidase	Metabolism of amino acids and related molecule
22	Unknown	–	–	–
23	NUSG_BACSU (Q06795)	*nusG* (D13303)	Transcription antitermination factor	Termination
24	FHUC_BACSU (P49938)	*fhuC* (X93092)	Ferrichrome ABC transporter (ATP-binding protein)	Transport/binding proteins and lipoproteins
25	PRSA_BACSU (P24327)	*prsA* (X57271)	Protein secretion (post-translocation molecular chaperone)	Protein secretion
26	Unknown	–	–	–
27	HIS5_BACSU (O34565)	*hisH* (AF017113)	Amidotransferase	Metabolism of amino acid and related molecules

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Table 6. Continued

28	YCEB_BACSU (O34504)	*yceB* (AB000617)	Unknown; similar to unknown proteins	From other organisms
29	YBBB_BACSU (P40408)	*ybbB* (L19954)	Unknown; similar to transcriptional regulator (AraC/XylS family)	Regulation
30	Unknown	–	–	–
31	OPCA_BACSU (O34992)	*opuCA* (AF009352)	Glycine betaine/carnitine/choline ABC transporter (ATP-binding protein)	Transport/binding proteins and lipoproteins
32	PBPA_BACSU (P39793)	*ponA* (U11883)	Penicillin-binding proteins 1A/1B	Cell wall
33	XTMA_BACSU (P39793)	*xtmA* (Z70177)	PBSX terminase (small subunit)	Phage-related functions
34	YQIZ_BACSU (P54537)	*yqiZ* (D84432)	Unknown; similar to amino acid ABC transporter (ATP-binding protein)	Transport/binding proteins and lipoproteins
35	YLXL_BACSU (P40405)	*ylxL* (Z99112)	Unknown	–
36	RIBT_BACSU (P17622)	*ribT* (L09228)	Reductase	Metabolism of coenzymes and prosthetic groups
37	SLP_BACSU (P39910)	*slp* (M57435)	Small peptidoglycan-associated lipoprotein	Transport/binding proteins and lipoproteins
38	YVYF_BACSU (P39807)	*yvyF* (L14437)	Unknown; similar to flagellar protein	Mobility and chemotaxis
39	Unknown	–	–	–
40	CYH1_BACSU (P94498)	*cysH* (U76751)	Phosphoadenosine phosphosulfate	Metabolism of amino acids and related molecules
41	Unknown	–	–	–
42	YQGS_BACSU (P54496)	*yqgS* (D84432)	Unknown; similar to putative molybdate binding protein	From other organisms
43	AK2_BACSU (P08495)	*lysC* (J03294)	Aspartokinase II (α and β subunits)	Metabolism of amino acids and related molecules
44	YHBH_BACSU (P45742)	*yhbH* (Z99108)	Unknown; similar to unknown proteins	From other organisms
45	Unknown	–	–	–
46	Unknown	–	–	–
47	Unknown	–	–	–
48	YLAL_BACSU (O07636)	*ylaL* (Z97025)	Unknown; similar to unknown proteins	From *B. subtilis*
49	RSFA_BACSU (P39650)	*rsfA* (X73124)	Probable regulator of transcription of σF-dependent genes	Sporulation
50	G6PI_BACSU (P80860)	*pgi* (Z93936)	Glucose-6-phosphate	Main glycolytic pathways
51	YQBP_BACSU (P45932)	*yqbP* (D32216)	Unknown; similar to phage-related protein	Phage-related functions
52	YACH_BACSU (P37569)	*yacH* (D26185)	Unknown	–
53	MOBB_BACSU (O31704)	*mobB* (AF012285)	Molybdopterin-guanine dinucleotide biosynthesis	Metabolism of coenzymes and prosthetic groups

raded in the body without inducing any problematic inflammatory response, and a potential biological adhesive made from porcine collagen and poly-(L-glutamic acid) has been developed, which is superior to fibrin in sealing air leakage from the lungs [51]. Moreover, most drugs are metabolized in the liver via the cytochrome P-450 family of enzymes, encompassing a wide array of overlapping specificities for different substrates [52,53]. They also play a major role in metabolism that can lead to the production and removal of toxic species, and in some instances it is possible to correlate the ability of failure to remove such a toxin with specific P-450 or subgroup. Not only cytochrome P-450 was found in

Copyright © by Chiang Mai University
All rights reserved

Table 7
Comparison of the function systems of the *Bacillus* sp. in protein coding genes by COG database (http://www.ncbi.nlm.nih.gov/COG)

Genome functions	*B. subtilis*	*B. halo-durans*	*B. stearo-thermophilus* TLS33[a]
Information storage and processing			
1. Translation, ribosomal structure and biogenesis	152	153	–
2. Transcription	272	269	–
3. DNA replication, recombination and repair	131	227	–
Cellular processes			
1. Cell division and chromosome partitioning	31	32	–
2. Posttranslational modification, protein turnover, chaperones	87	84	–
3. Cell envelope biogenesis, outer membrane	161	84	–
4. Cell motility and secretion	90	90	–
5. Inorganic ion transport and metabolism	148	146	–
6. Signal transduction mechanisms	122	135	–
Metabolism			
1. Energy production and conversion	164	158	2
2. Carbohydrate transport and metabolism	271	262	5
3. Amino acid transport and metabolism	293	284	15
4. Nucleotide transport and metabolism	82	73	5
5. Coenzyme metabolism	109	109	–
6. Lipid metabolism	84	89	3
7. Secondary metabolites biosynthesis, transport and catabolism	128	110	4
Poorly characterized			
1. General function prediction only	332	329	–
2. Function unknown	226	246	–
Not in COGs	1221	1155	–

[a] Some identifications.

human but it was also found in microorganism; for example, Hopkins et al. [54] and English et al. [55] reported the induction of cytochrome P-450$_{BM-3}$ (CYP102) by barbiturates, peroxisome proliferators and non-steroidal anti-inflammatory drugs in *B. megaterium*, in which cytochrome *P*-450s are a superfamily of mono-oxygenases involved in the oxidation of a wide range of compounds, including steroids, fatty acids and xenobiotics such as drugs and environmental chemicals. The inducers interacted with the repressor protein Bm3R1 that also interacted with fatty acids, causing it to dissociate with the operator of the CYP102 gene and allowing transcription to occur. Otherwise, they also play a key part in various steps of primary and secondary metabolism, as well as in drug detoxification [56]. Schiavon et al. [57] used the uricase from *B. fastidiosus* (UC) conjugate to linear or branched poly(ethylene glycol) and poly(*N*-acryloymorpholine) for improved therapy with uricase and prevented gout disease or hyperuricemia and the uricase could catalyse the oxidation of uric acid, a final product of purine catabolism, to allantoin which is more easily excreted than the starting compound.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Table 8
Comparison of proteome analysis of *B. subtilis* and *B. halodurans* by Proteome Analysis@EBI database based on high-level terms of the Gene Ontology (GO) data that has been assigned to InterPro entries and shown the general statistics for the number of proteins in the proteome

GO Classification	*B. subtilis*	*B. halodurans*
Term		
Nucleic acid binding	475 (11.5%)	435 (10.8%)
DNA binding	370 (8.9%)	347 (8.6%)
DNA repair protein	22 (0.5%)	19 (0.4%)
DNA replication factor	6 (0.1%)	9 (0.2%)
Transcription factor	212 (5.1%)	217 (5.4%)
RNA binding	100 (2.4%)	87 (2.1%)
Structural protein of ribosome	60 (1.4%)	59 (1.4%)
Translation factor	12 (0.2%)	12 (0.2%)
Motor	6 (0.1%)	5 (0.1%)
Enzyme	1310 (31.7%)	1192 (29.7%)
Peptidase	92 (2.2%)	77 (1.9%)
Protein kinase	56 (1.3%)	63 (1.5%)
Protein phosphatase	20 (0.4%)	12 (0.2%)
Signal transducer	95 (2.3%)	115 (2.8%)
Receptor	5 (0.1%)	0 (0%)
Transmembrane receptor	5 (0.1%)	0 (0%)
Cell adhesion molecule	2 (0%)	4 (0%)
Structural protein	70 (1.6%)	70 (1.7%)
Transporter	372 (9.0%)	348 (8.6%)
Ion channel	8 (0.1%)	10 (0.2%)
Neurotransmitter transporter	2 (0%)	3 (0%)
Ligand binding or carrier	446 (10.8%)	458 (11.4%)
Electron transporter	97 (2.3%)	93 (2.3%)
Molecular function unknown	39(0.9%)	35 (0.8%)
Unclassified	**1409 (34.1%)**	**1472 (36.7%)**
Total	**4121 (100%)**	**4008 (100%)**

Tomita et al. [58] studied the stabilized enzymatic reagents for measuring glucose, creatine kinase and γ-glutamyltransferase (γ-GT) with thermostable enzymes from a thermophilic bacterium *B. stearothermophilus*. These enzymes are glucokinase (GlcK, EC 2.7.1.2) and alanine dehydrogenase (AlaDH, EC 1.4.1.1), which the long-term stability in using GlcK for measurement of glucose and creatine kinase (CK) activity, and in using AlaDH for γ-glutamyltransferase (γ-GT) activity. In addition, Yamashita et al. [59] found the Cry and Cyt proteins of *B. thuringiensis* exhibiting cytocidal activity against human leukemic T cells and uterus cervix cancer (HeLa) cells but not against normal T cells. Moreover, the *Bacillus* strains could be also used as a host in expression of proteins; for example, Heikkila et al. [60] studied the immunization with meningoccal class 1 outer membrane protein (BacP1) produced in *B. subilis* and reconstituted in the presence of Zwittergent or Triton X-100. The use of adjuvant improved the immunogenicity of the BacP1-detergent preparations and such liposomes were reproducibly immunogenic in mice and guinea pigs at a low dose indicating its immunizing properties. Otherwise, the superoxide dismutase (SOD) in *Bacillus* sp. has also useful in biomedical application such as it can alleviate a wide range of effects produced by exposure to O_2^--generating systems and it has an anti-inflammatory property which was discovered well before the protein was identified as an enzyme. SOD administered intravenously along with catalase, was found to protect against oxygen toxicity and reduced inflammation caused by irradiation. Orgotein is a sample of the generic name adopted for drug versions of highly purified SOD and was reported to have anti-inflammatory and anti-viral activities [61]. On the other hand, the lipase secreting from the same species *B. stearothermophilus* P1 could catalyze the esterification of ibuprofen, which is an important group of non-steroidal anti-inflammatory drugs and used to reduced swelling and inflammation, representing an effective way to prepare the prodrug [62].

7. Conclusions

Genome and proteome analysis of *Bacillus* sp. provide a powerful set of tools for study of functional genomics thoroughly to the biochemistry of proteins and the combination with bioinformatics is also increasingly useful for interpretation of comparative-genomic, functional genomic and proteomic results from many species. Therefore, the proteomics has unique and significant advantages as an important complement to a genomics approach, and applications for this technology are readily apparent.

Acknowledgements

We gratefully thank the National Taiwan Science Council and Academia Sinica, Taiwan for supporting a grant of the *Bacillus* research. This review is also part of the Royal Golden Jubilee Ph.D. project of Mr.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Table 9
Biomedical applications of *Bacillus* sp.

Proteins	Sources	Functions	Applications
1. Poly-γ-glutamic acid (γ-PGA) [43]	*B. anthracis* *B. natto* *B. subtilis* *B. thuringiensis*	Drug carrier or sustained release materials, curable biological adhesive and hemostatic or medical bonding kit suture thread	Use for gene therapy; cancer drug, substitutes for fibrin
2. Cytochrome P450 (CYP102) [54,55]	*B. megaterium*	Regulation of transcription and/or fatty acid hydroxylation, play a key part in various steps of primary and secondary metabolism, as well as in drug detoxification	Use in oxidation of a wide range of compounds including steroids, fatty acids and xerobiotics such as drugs and environmental Chemical
3. Uricase [56]	*B. fastidiosus*	Catalyse the oxidation of uric acid, a final product of purine catabolism, to allantoin which is more easily excreted than the starting compound	Use as therapeutic proteins
4. GlcK and AlaDH [58]	*B. stearothermophilus*	GlcK is used for measurement of glucose and creating kinase (CK) activity and AlaDH is used for measurement of γ-GT activity	Stabilize enzymatic reagents for measuring some components in biological fluids
5. Cry and Cyt [59]	*B. thuringiensis*	Exhibit cytocidal activity against human leukaemic T cells and uterus cervix cancer (HeLa) cells	Anti-cancer cell drug
6. BacP1 [60]	*B. subtilis*	Induce the synthesis of bactericidal antibodies reacting with P1 epitopes exposed on meningococcal surface	Use as meningococcal disease vaccine
7. Superoxide dismutase (SOD) [61]	*B. stearothermophilus*	Catalyze the dismutation od superoxide radicals to oxygen and hydrogen peroxide.	Anti-inflammatory drug and antiviral activities
8. Lipase [62]	*B. stearothermophilus*	Catalyze the esterification process	Synthesis of ester prodrug or or non-steroidal anti-inflammatory drugs such as ibuprofen and naproxen

Supachai Topanurak supported by the Thailand Research Fund (TRF), Thailand.

References

[1] F. Kunst, N. Ogasawara, I. Moszer, A.M. Albertini, G. Alloni, V. Azevedo, M.G. Bessieres, A. Bolotin, S. Borchert, R. Borriss, L. Boursier, A. Brans, M. Braun, S.C. Brignell, S. Born, S. Brouillet, C.V. Bruschi, B. Caldwell, V. Capuano, N.M. Carter, S.K. Choi, J.J. Codani, I.F. Connerton, A. Danchin et al., Nature 390 (1997) 249.

[2] H. Takami, K. Nakasone, Y. Takaki, G. Maeno, R. Sasaki, N. Masui, F. Fuji, C. Hirama, Y. Nakamura, N. Ogasawara, S. Kuhara, K. Horikoshi, Nucleic Acids Res. 28 (2000) 4317.

[3] F.G. Priest, M. Goodfellow, C. Todd, J. Gen. Microbiol. 134 (1988) 1847.

[4] H. Takami, K. Horikoshi, Extremophiles 4 (2000) 99.

[5] I. Moszer, FEBS Lett. 430 (1998) 28.

[6] M. Hecker, W. Schumann, U. Volker, Mol. Microbiol. 19 (1996) 417.

[7] P. Graumann, T.M. Wendrich, M.H.W. Weber, K. Schroder, M.A. Marahiel, Mol. Microbiol. 25 (1997) 741.

[8] H. Takami, K. Nakasone, C. Hirama, Y. Takaki, N. Masui, F. Fuji, Y. Nakamura, A. Inoue, Extremophiles 3 (1999) 21.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

[9] J. Boch, B. Kempf, E. Bremer, J. Bacteriol. 178 (1996) 5121.
[10] J. Bernhardt, U. Volker, A. Volker, H. Antelmann, R. Schmid, H. Mach, M. Hecker, Microbiology 143 (1997) 999.
[11] J. Bernhardt, K. Buttner, C. Scharf, M. Hecker, Electrophoresis 20 (1999) 2225.
[12] H.L. Hyyrylainen, A. Bolhuis, E. Darmon, L. Muukkonen, P. Koski, M. Vitikainen, M. Sarvas, Z. Pragai, S. Bron, J.M. van Dijl, V.P. Kontinen, Mol. Microbiol. 41 (2001) 1159.
[13] A. Wipat, C.R. Harwood, FEMS Microbiol. Ecol. 28 (1999) 1.
[14] G. Schneider, Gene 237 (1999) 113.
[15] H. Antelmann, H. Tjalsma, B. Voigt, S. Ohlmeier, S. Bron, J.M. van Diji, M. Hecker, Genome Res. 11 (2001) 1484.
[16] P.W. Chu, M.N. Yap, C.Y. Wu, C.M. Huang, F.M. Pan, M.J. Tseng, S.T. Chen, Electrophoresis 21 (2000) 1740.
[17] B. Boonsinthai, S. Phutrakul, J. Sci. Fac. CMU 23 (1991) 9.
[18] B. Sookkheo, S. Sinchaikul, S. Phutrakul, S.T. Chen, Protein Expr. Purif. 20 (2000) 142.
[19] B. Sookkheo, S. Sinchaikul, H. Thannan, O. Thongprasong, S. Phutrakul, S.T. Chen, Proteomic (2002) in press.
[20] B. Fricke, T. Buchmann, S. Friebe, J. Chromatogr. A 715 (1995) 247.
[21] H. Ihara, T. Uemure, M. Masuhara, S. Ikawa, K. Sugimoto, A. Wadano, M. Himeno, Comp. Biochem. Phys. B 120 (1998) 197.
[22] M. Villalon, V. Vachon, R. Brousseau, J.L. Schwartz, R. Laprade, Biochim. Biophys. Acta 1368 (1998) 27.
[23] A. Helenius, M. Sarvas, K. Simons, Eur. J. Biochem. 116 (1981) 27.
[24] T.J. Stevens, I.T. Arkin, Proteins 39 (2000) 417.
[25] S. Ohlmeier, C. Scharf, M. Hecker, Electrophoresis 21 (2000) 3701.
[26] J.S. Lolkema, D.J. Slotboom, FEMS Microbiol. Rev. 22 (1998) 305.
[27] I. Diers, in: A.C.R. Dean, D.C. Ellwood, C.G.T. Evans, J. Melling (Eds.), Continuous Culture, Ellis Horwood, Chichester, 1976, p. 208.
[28] C.R. Goward, R. Hartwell, T. Atkinson, M.D. Scawen, Biochem. J. 237 (1986) 415.
[29] R.J. Sharp, M.D. Scawen, T. Atkinson, in: C.R. Harwood (Ed.), Biotechnology Handbooks, Vol. 2: Bacillus, Plenum, New York, 1989, p. 255.
[30] M.C. Oosthuizen, B. Steyn, D. Lindsay, V.S. Brozel, A. von Holy, FEMS Microbiol. Lett. 194 (2001) 47.
[31] K. Buttner, J. Bernhardt, C. Scharf, R. Schmid, U. Mader, C. Eymann, H. Antelmann, A. Volker, U. Volker, M. Hecker, Electrophoresis 22 (2001) 2908.
[32] R.D. Appel, A. Bairoch, J.C. Sanchez, J.R. Vargas, O. Golaz, C. Pasquali, D.F. Hochstrasser, Electrophoresis 17 (1996) 540.
[33] S. Sinchaikul, B. Sookkheo, S. Phutrakul, S.T. Chen, Proteomic (2001) submitted for publication.
[34] A.A. Gooley, N.H. Packer, The importance of protein co- and post-translational modifications in proteome projects, in: M.R. Wilkins, K.L. Williams, R.D. Appel, D.F. Hochstrasser (Eds.), Proteome Research: New Frontiers in Functional Genomics, Springer, Berlin, 1997, p. 65, Chapter 4.
[35] O. Lentz, Q.S. Li, U. Schwaneberg, S.L. Wahl, P. Fischer, R.D. Schmid, J. Mol. Catal. B 15 (2001) 123.
[36] I.A.D. Lessard, G.J. Domingo, A. Borges, R.N. Perham, Eur. J. Biochem. 258 (1998) 491.
[37] H.J. Schreier, C.A. Rostkowski, Gene 161 (1995) 51.
[38] S. Sakamoto, M. Seki, S. Nagata, H. Misono, J. Mol. Catal. B 12 (2001) 58.
[39] R.L. Tatusov, E.V. Koonin, D.J. Lipman, Science 278 (1997) 631.
[40] R.L. Tatusov, D.A. Natale, I.V. Garkavtsev, T.A. Tatusova, U.T. Shankavaram, B.S. Rao, B. Kiryutin, M.Y. Galperin, N.D. Fedorova, E.V. Koonin, Nucleic Acids Res. 29 (2001) 22.
[41] A. Bairoch, R. Apweiler, Nucleic Acids Res. 28 (2000) 45.
[42] H.F. Juan, C.L. Kao, C.Y. Kao, S.T. Chen, Bulk Gene Sercch System: a facile bioinformatic system to search and classify gene and protein functions from microarray, Bioinformatics (2002) submitted for publication.
[43] I.L. Shih, Y.T. Van, Biores. Technol. 79 (2001) 207.
[44] C. Li, D.F. Yu, A. Newman, F. Cabral, C. Stephens, N. Hunter, L. Milas, S. Wallace, Cancer Res. 58 (1998) 2404.
[45] C. Li, J.E. Price, L. Milas, N.R. Hunter, S. Ke, W. Tansey, C. Charnsagavej, S. Wallace, Clin. Cancer Res. 5 (1999) 891.
[46] M. Hashida, K. Akamutsu, M. Nishikawa, F. Yamashita, Y. Takakura, J. Control. Release 62 (1999) 253.
[47] Y. Otani, Y. Tabata, Y. Ikada, J. Biomed. Mater. Res. 31 (1996a) 157.
[48] Y. Otani, Y. Tabata, Y. Ikada, Biomaterials 19 (1996b) 1387.
[49] Y. Otani, Y. Tabata, Y. Ikada, Biomaterials 19 (1998a) 2167.
[50] Y. Otani, Y. Tabata, Y. Ikada, Biomaterials 19 (1998b) 2091.
[51] T. Sekine, T. Nakamura, Y. Shimizu, H. Ueda, K. Matsumoto, Y. Takimoto, T. Kiyotani, J. Biomed. Mater. Res. 35 (2000) 305.
[52] F.P. Guengerich, A. Parikh, Curr. Opin. Biotechnol. 8 (1997) 623.
[53] S. Rendic, F.J. Di Carlo, Drug. Metab. Rev. 29 (1997) 413.
[54] N.E. Hopkins, N. English, V. Hughes, C.W. Rowley, C.R. Wolf, W.L. Alworth, Biochem. Biophys. Res. Commun. 244 (1998) 868.
[55] N. English, V. Hughes, C.R. Wolf, Biochem. J. 316 (1996) 279.
[56] U. Schwaneberg, D. Appel, J. Schmitt, R.D. Schmid, J. Biotechnol. 84 (2000) 249.
[57] O. Schiavon, P. Caliceti, P. Ferruti, F.M. Veronese, Il Farmaco 55 (2000) 264.
[58] K. Tomita, K. Nomura, H. Kondo, K. Nagata, H. Tsubota, J. Pharm. Biomed. Anal. 13 (1995) 477.
[59] S. Yamashita, T. Akao, E. Mizuki, H. Saitoh, K. Higuchi, Y.S. Park, H.S. Kim, M. Ohba, Can. J. Microbiol. 46 (2000) 913.
[60] I.I. Heikkila, E. Wahlstrom, S. Muttilainen, M. Nurminen, H. Kayhty, M. Sarvas, P.H. Makela, Vaccine 14 (1996) 886.
[61] W.P. Michalski, J. Chromatogr. B 684 (1996) 59.
[62] S. Sinchaikul, B. Sookkheo, S. Phutrakul, F.M. Pan, S.T. Chen, Protein Expr. Purif. 22 (2001) 388.

Copyright © by Chiang Mai University
All rights reserved

Functional and Structural Analysis of *Bacillus* Proteome

Supachok Sinchaikul[1], Boonyaras Sookkheo[2], Supachai Topanurak[1,3], Fu-Ming Pan[1], Suree Phutrakul[3] and Shui-Tein Chen[1,4,*]

[1]Institute of Biological Chemistry and Genomics Reseach Center, Academia Sinica, Taipei, Taiwan 11529, [2]Department of Chemistry, Faculty of Science, Rangsit University, Patumthani, Thailand 12000, [3]Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand 50200 and [4]Institute of Biochemical Sciences, College of Life Science, National Taiwan University, Taiwan 106

Abstract: This review summarizes the current status of application of functional and structural proteomic strategies for bacteria belonging to the species *Bacillus* and their applications in life science research. Structural proteomics plays a crucial role for studying the relationship between structures and functions of proteins. In addition, bioinformatics is very useful for rapid search of unknown proteins including their functions, and interpretation of data in both genomics and proteomics. Expression proteome analysis based on advanced and high throughput technologyis available now is important for quantifying and analyzing protein or gene expression and understanding their biological functions. Such analysis is currently possible by the application of various multidimensional techniques, such as two-dimensional (2-D) gel electrophoresis, mass spectrometry (MS), capillary electrophoresis (CE) and other related techniques. Proteome analysis has been extensively used to study the physiology as well as molecular mechanisms of responses of *Bacillus* under various stress condition such as heat, cold, acid, alkali, osmotic pressure and starvation to obtain important information for comprehensive study of protein and gene functions. The applications of proteomic analysis for biomarkers and drug discovery for *Bacillus* sp. and the future applications of proteomics for *Bacillus* have been discussed.

Key Words: *Bacillus*, proteome, functional analysis, structural analysis.

1. INTRODUCTION

The complete genome sequences of several *Bacillus* sp. have been determined, such as *B. subtilis* (Kunst *et al*., 1997), *B. halodurans* (Takami *et al*., 2000), *B. cereus* (Ivanova *et al*., 2003) and *B. anthracis* (Read *et al*., 2003). The Proteome Analysis database (http://www.ebi.ac.uk/proteome) provides comprehensive statistical and comparative analyses of the predicted proteomes of these fully sequenced *Bacillus* sp. Those show 4105, 4007, 5239 and 5310 proteins which correspond to their genomes and molecular functions of 2515, 2454, 3051 and 2281 of those proteins, respectively, are known. In addition, the comparison of complete genomes of *B. subtilis*, *B. halodurans*, *B. cereus* and *B. anthracis* is shows in Table **1**. However, the complete genomes of some species have not been determined yet such as *B. stearothermophilus*, *B. licheniformis*, *B. formis* and *B. thuringiensis*. Although the genomics of *Bacillus* sp. has been studied and the genome sequences have been determined, the related studies at the level of the proteome have not yet been carried out in comparable details. Proteomics methods are currently being used to generate a more detailed knowledge base, including post-translational modifications and interaction components that modulate the activity of the target. Proteomics not only includes the identification and quantification of proteins, it also involves a study of their structure, localization, modification, interactions, activities and functions. A wide range of technologies are used to perform these tasks and the data that is generated is of an extremely large-scale as well as multi-dimensional. It is convenient to classify proteomic approaches into four groups:

1. protein characterization,
2. differential display or expression proteomics,
3. protein-protein interaction studies (also known as functional proteomics), and
4. structural proteomics.

In this review, the main focus is on the expression and functional proteomics describing the changes in protein expression during differentiation, proliferation and signaling of cells. In addition, brief descriptions of structural proteomics describing the elucidation of the primary structure of the components in the proteomic pattern have been provided. However, both functional and structural proteomics are the emerging paradigms that are gaining importance in the post-genomic era as a valuable discipline to process the protein target information being deciphered.

2. TECHNIQUES FOR PROTEOME ANALYSIS

It is well known that protein analysis has different methods to determine differentially expressed proteins and the most popular methods that are widely used for proteome analysis are two-dimensional polyacrylamide gel electrophoresis (2-D PAGE) and mass spectrometry (MS). Combina-

*Address correspondence to this author at the Institute of Biological Chemistry (RM707), Academia Sinica, 128 Yen Chiu Yuan Rd., Sec II, Nankang, Taipei, Taiwan 11529; Tel: 886-2-27855981 Ext.7071; Fax: 886-2-27883473; E-mail: bcchen@gate.sinica.edu.tw

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

1570-1646/05 $50.00+.00 ©2005 Bentham Science Publishers Ltd.

Table 1. Comparison of the General Features of the Complete Genomes of *B. Subtilis*, *B. Halodurans*, *B. Cereus* and *B. Anthracis*

Genome features	*B. subtilis*	*B. halodurans*	*B. cereus* ATCC14579	*B. anthracis* Ames
Size (bp)	4,214,814	4,202,353	5,426,909	5,227,293
G+C content (mol%)				
Total genome	43.5	43.7	35.3	35.4
Coding region	44.3	44.4	84.0	37
Non-coding region	39.3	39.8	38.0	39.0
Open reading frames				
Percent of genome (coding)	87.0	85.0	84.0	84.3
Protein coding genes	4,104	4,066	5,366	2,762
Conserved with function assigned	2,379	2,144	3,839	1,212
Conserved with unknown function	668	1,182	1,481	657
Non-conserved	1,053	743	142	877
Percent AUG initiation codons	78	78	-	-
Percent GUG initiation codons	9	12	-	-
Percent UUG initiation codons	13	10	-	-
Insertion elements				
PBSX prophage-related protein	27	2	-	62
Transposase and related protein	10	112	78	18
RNA elements				
Stable RNA (percent of genome)	1.27	1.02	1.00	3.00
16S, 23S and 5S rRNA	10	8	13	11
tRNA	86	78	108	95

tions of various orthogonal techniques that can be applied in a high-throughput format tend to provide powerful approaches for protein identification at the proteome level.

2.1. Two-Dimensional Polyacrylamide Gel Electrophoresis (2-D PAGE)

The widely established technique for protein separation, 2-D PAGE permits the separation and detection of proteins from a wide variety of sources without the need for any prior knowledge of function. The first dimension of separation is based on differences in the charge of polypeptides and the second dimension is a molecular separation according to the mass of polypeptide molecules. Based on the location of the protein molecules within the cell, the proteome of *Bacillus* species are classified into three groups: extracellular, cell membrane and cytosolic proteins. For the proteomic study of extracellular proteins, the extracellular proteins samples were prepared from culture media and examined by 2-D PAGE. The 2-D PAGE result of *B. subtilis* 168 extracellular proteins demonstrated approximately 100-110 spots in a 2-D PAGE gel (Hirose *et al*., 2000). Among 23 proteins whose N-terminal amino acid sequences were determined, 17 spots were extracellular proteins having signal peptides in their precursor form. The effect of carbon source on the protein composition of the extracellular preparations was also examined by 2-D PAGE analysis and 19 additional extracellular proteins were detected in cultures maintained in cellobiose, maltose and soluble starch. The results of 2-D PAGE could indicate that most extracellular proteins target to and translocate across the cytoplasmic membrane by co-operation between the signal-recognition particle and Sec protein-secretion pathways. Due to many virulence factors secreted by *Bacillus* sp., most of them are regulated by the transcriptional activator, PlcR, which is maximally expressed at the beginning of the stationary phase. The proteomics approach by 2-D PAGE has also been used to study the impact of the PlcR regulon on the secreted proteins of *Bacillus* sp. A typical example of this approach is the comparison of the extracellular proteomes of *B. cereus* strains ATCC14579 and ATCC14579 *plc*R, in which *plc*R has been disrupted (Gohar *et al*., 2002). The comparative gel analysis showed that most of the large extracellular protein spots in the wild-type strain were affected by the mutation in *plc*R: some of them completely disappeared, whereas others were strongly decreased in amount in the mutant strain. In addition, the study of extracellular proteins from the thermophilic bacterium *B. stearothermophilus* has a few reports of proteomic analysis using 2-D PAGE. In our previous reports, the thermophilic bacterium *B. stearothermophilus* P1 and TLS33 have been known to secrete the extracellular protease, lipase and superoxide dismutase (Sookkheo *et al*., 2000; Sinchaikul *et al*., 2002a; Sookkheo *et al*., 2002). Therefore the extracellular and intracellular proteins were studied using proteomic approach and the preliminary result of 2-D PAGE comparing the extracellular and cytosolic proteins in *B. stearothermophilus* P1 showed different protein patterns (Fig. **1**). We have also studied the extracellular and cytosolic proteins in this bacterium under substrate-inductions, such as olive oil-induction (for lipase study) and skim milk-induction (for protease study) (Fig. **2**). The preliminary study of the substrate-induced lipase from this bacterium by addition of olive

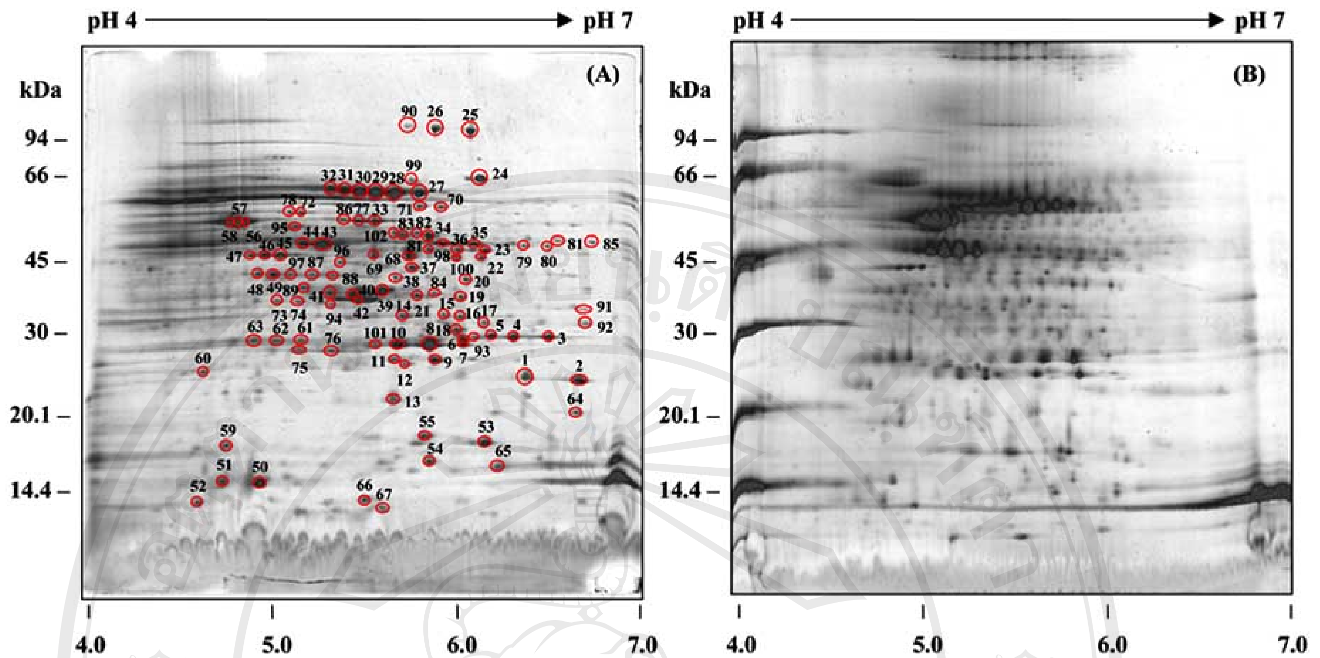


Fig. (1). Comparison of extracellular proteins (A) and cytosolic proteins (B) from a thermophilic bacterium *B. stearothermophilus* P1 by 2-D PAGE analysis using IPG strip with a narrow pH range (4-7), 18 cm. long. One hundred and two extracellular proteins (circled and labeled spots) were identified by MALDI-TOF mass spectrometry analysis.

oil in the culture medium showed the different protein patterns, analyzed by 2-D PAGE in both wide p*I* range of 3-10 and narrow p*I* range of 4-7 (Fig. **3**). Most of the high abundant extracellular proteins were observed in a narrow p*I* range of 4-7 and under olive oil induction, indicating that many proteins could be induced by olive oil. The 2-D gel patterns between control (no-induction) and olive oil induction could be compared by using ImageMaster 2-D gel image analysis software, which can quantitate the protein spots in 2-D gels and allow analysis of up or down regulation of expressed proteins. Furthermore, using 2D-PAGE, 168 protiens were identified from the cell wall of *B subtilis*. It was found that most abundant cell wall binding proteins (CWBPs) resulted from the WapA and WprA precursor processing and some of cell wall proteins were also components of the extracellular proteome (Antelmann *et al*., 2002). Membrane-associated or membrane-bound proteins are often difficult to analyze by 2-D PAGE due to their low abundance, alkaline p*I* and poor solubility. However, it has found that the genome of *B. subtilis* contained 29.2% membrane proteins and the proteomic approach using 2-D PAGE found some membrane proteins such as Ffh, LytC, MurG and YfnI (Stevens and Arkin, 2000; Ohlmeier *et al*., 2000). The membrane-bound LytC protein identified in the alkaline p*I* range has one theoretical hydrophobic transmembrane domain (Blackman *et al*., 1998). Ffh and MurG were also identified in the alkaline pI range. It has been found that Ffh, as a part of the signal recognition particle, is a membrane-associated protein, and MurG might also be membrane-associated (Nakamura *et al*., 1994). In addition, the membrane fractions of *B. pseudofirmus* OF4 grown to the mid-logarithmic phase at pH 7.5 and 10.5 were examined by 2-D PAGE and showed 148 spots, which were resolved in this analysis with similar protein patterns. Of these 148, the densities of 19 spots were increased more than two-fold when grown at pH 10.5, whereas the densities of 18 were decreased more than two-fold (Gilmour *et al*., 2000). Recently, the membrane protein fraction of *B. anthracis* (Ames strain) was analyzed by 2-D PAGE and MALDI-TOF MS, and found to represent close to 100 spots on 2-D gels but only 32 membrane proteins were identified, which indicated the limitation of proteomic analysis for membrane proteins (Ariel *et al*., 2003). Proteomic study of cytosol or intracellular proteins has also been carried out. For example, the biofilm proteome of whole cell proteins from *B. cereus* strain DL5 was investigated using 2-D PAGE and 10 proteins were found to be synthesized as a result of surface attachment of which four were unique to biofilm profile (Oosthuizen *et al*., 2001). The results also revealed distinct and reproducible phenotypic differences between 2-h and 18-h-old biofilm and planktonic cells. The 2-h-old biofilm proteome indicated expression of 15 unique proteins whereas the 18-h-old biofilm proteome contained 7 uniquely expressed proteins. Differences between the micro-colony (2-h) proteome and the more developed biofilm (18-h) proteome were largely due to up- and down-regulation of the expression of a multitude of proteins (Oosthuizen *et al*., 2002). The cytosolic proteins of *B. subtilis* 168 have also been identified. Up to 1,800 proteins were detected using 2-D PAGE, in which more than 300 new entries were found and putative post-translational modifications in an extensive scale were observed, with 47 proteins forming more than one spot (Buttner *et al*., 2001). The cytosolic proteins of thermo-

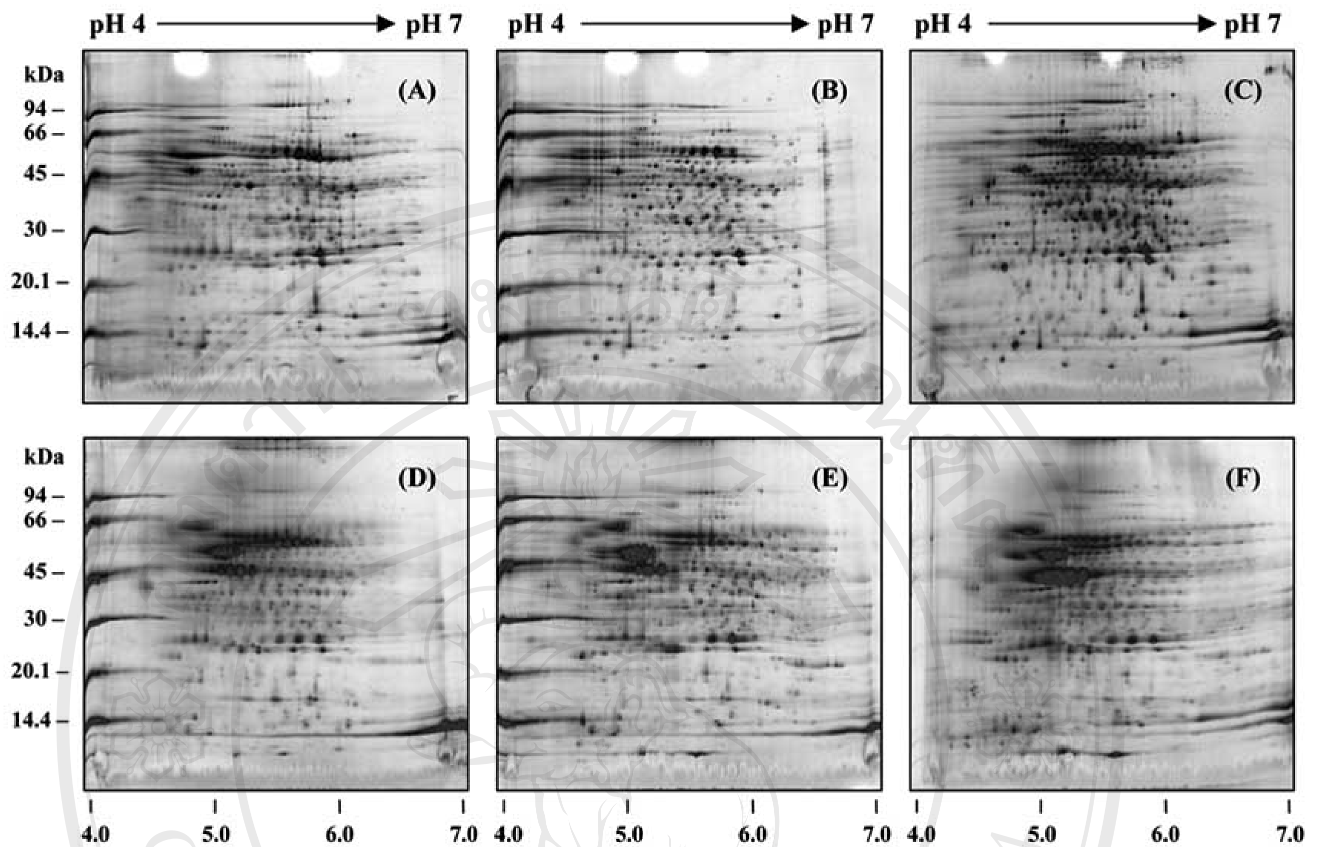


Fig. (2). Two dimensional polyacrylamide gel electrophoresis (2-D PAGE) of the extracellular and cytosolic proteins in *B. stearothermophilus* P1 under substrate inductions (0.5% v/v olive oil or skim milk), analyzed by using a IPG strip with narrow pH range (4-7), 18 cm long. A, extracellular proteins-no induction (control); B, extracellular proteins-olive oil induction; C, extracellular proteins-skim milk induction; D, cytosolic proteins-no induction (control); E, cytosolic proteins-olive oil induction; F, cytosolic proteins-skim milk induction.

philic bacterium *B. stearothermophilus* strain P1 and TLS33 were also identified under cold shock stress using 2-D PAGE. This led to the detection of the cold shock protein (CSP) with the molecular mass approximately 7.3 kDa and p*I* 4.5, and the 8 significant cold shock-induced proteins related to the sporulation pathway (Sinchaikul *et al*., 2002). The 2-D PAGE patterns of cytosolic proteins in *B. stearothermophilus* TLS33 under various stresses such as cold shock at 25°C, heat shock at 80°C, salt (10% NaCl), 10% ethanol, 1% H_2O_2 and UV exposure for 1h (Fig. **4**). For better understanding of the bacterial adaptation under stresses, we have initially focused on the proteome profile of *B. stearothermophilus* TLS33 under cold shock stress at 25°C and 37°C followed by 2-D PAGE and mass spectrometric analysis. Interestingly, it was found that eight cold shock-induced proteins were significantly different in protein expression levels, such as glucosyltransferase, anti-sigma B factor, Mrp protein homolog, dihydroorthase, hypothetical transcriptional regulator in Few-SigW intergenic region, RibT protein, phosphoadenosine phosphosulfate reductase and prespore specific transcriptional activator (Fig. **5**). Moreover, the summary of the most abundant proteins in *B. subtilis* 168 and *B. stearothermophilus* TLS33 cytosolic proteins are shown in Table **2**. Furthermore, the outermost proteinaceous layer of bacterial spores, called the coat, is currently under study because it is critical for survival, germination as well as for pathogenicity of spores and hence for disease. The spore coats of *B. subtilis* and *B. anthracis* were identified using 2-D PAGE and found to consist of: 38 *B. subtilis* spore proteins containing 12 known coat proteins and 11 of *B. anthracis* spore proteins containing 6 candidate coat or exosporium proteins (Lai *et al*., 2003).

2.2. Mass Spectrometry

Mass spectrometry has become the method of choice for protein identification and characterization following separation by 2-D PAGE. This is achieved by either peptide mass fingerprint analysis using MALDI-TOF MS (Gevaert and Vanderkerckhove, 2000; Jungblut and Thiede, 1997; Mortz *et al*., 1994) or *de novo* sequencing using liquid chromatography/electrospray ionization mass spectrometry (LC/ESI-MS) (Chalmers and Gaskell, 2000; Choudhary *et al*., 2001; Link *et al*., 1999). MS is not only used for identification of protein spots from 2-D PAGE, but it also used for investigation of protein modifications. Using high mass accuracy electrospray tandem mass MS (LC-MS/MS) and nanoelectrospray tandem MS, seven phosphorylated threonine and

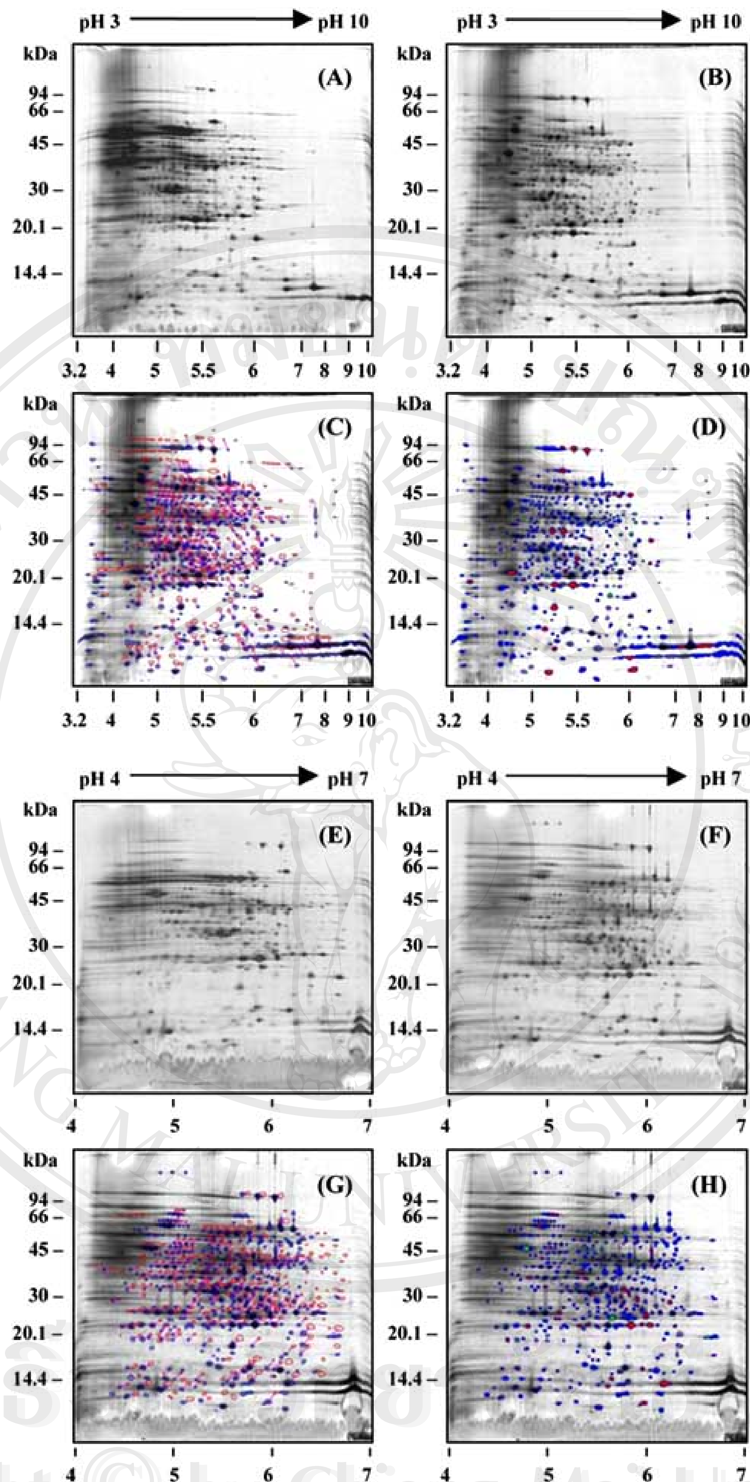


Fig. (3). Two dimensional polyacrylamide gel electrophoresis (2-D PAGE) of the extracellular proteins after culturing the thermophilic bacteria *B. stearothermophilus* P1 in the absence and presence of 0.5% (v/v) olive oil. Two dimensional PAGE analysis was done by using the wide and narrow pH ranges of 3-10 (A,B) and 4-7 (E,F), respectively. Comparison of the 2-D gel images of extracellular proteins were carried out using the ImageMaster software package with the same analysis parameters (sensitivity of 8652, operator size of 31, noise factor of 5, background factor of 1 and split level of 7). While using IPG strip of wide pH range (3-10), 370 and 441 protein spots were detected in control (no-induction) (A) and olive oil induction (B) with 212 matching proteins (48.07%) (C). While using IPG strip of narrow pH range (4-7), 331 and 393 protien spots were detected in control (no-induction) (E) and olive oil induction (F) with 189 matching proteins (48.09%) (G). Labels: red color spot, control (no induction) as a reference; blue color spot, olive oil induction; no color spot with the line, similar protein spot. Differential expression of proteins is shown in D and H: green; up-regulated proteins; red, down-regulated proteins; transparent blue, unchanged protein spot; opaque blue, unmatched protein spot.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

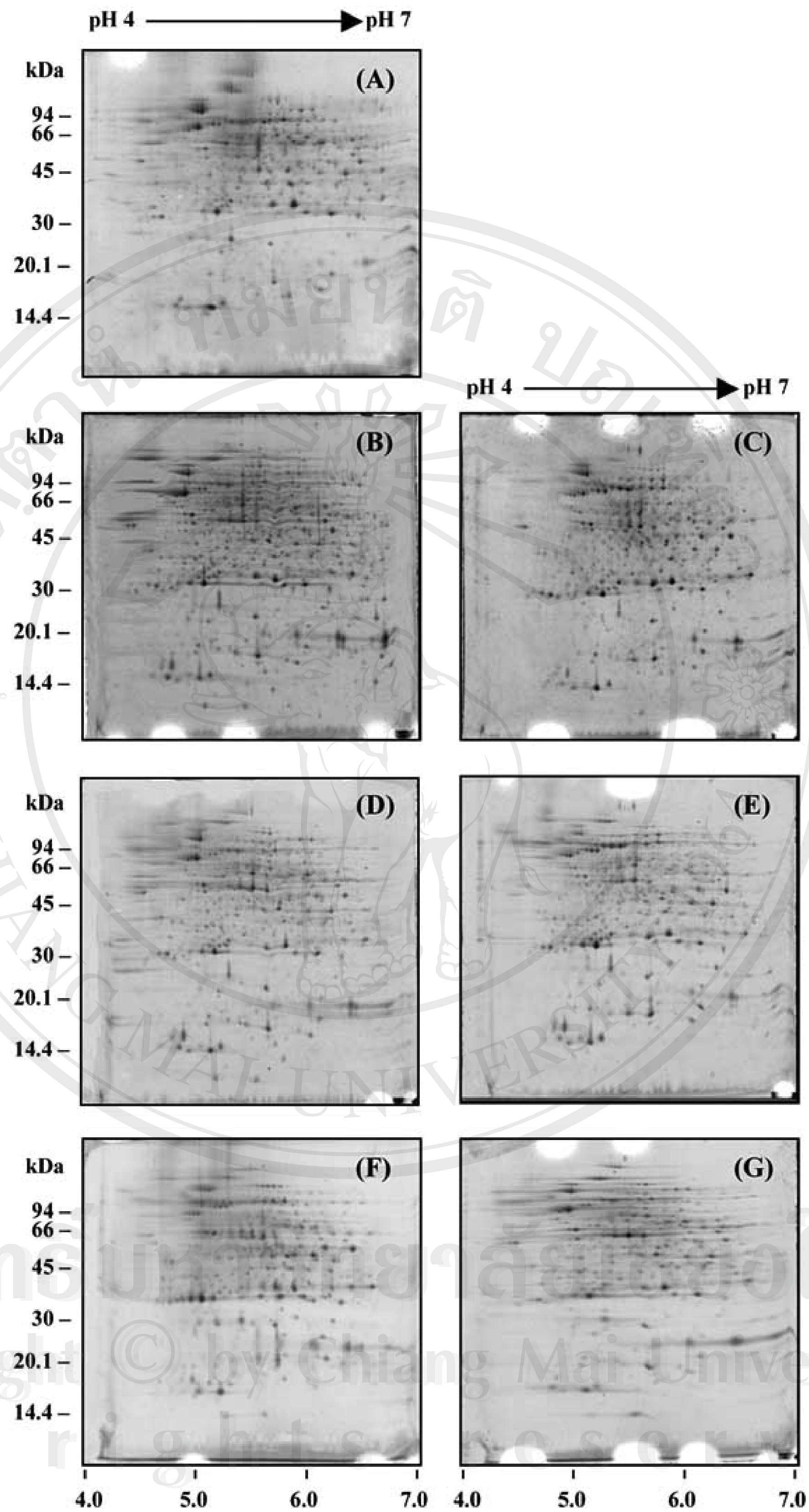


Fig. (4). Two dimensional polyacrylamide gel electrophoresis (2-D PAGE) of differentially expressed proteins in the thermophilic bacteria *B. stearothermophilus* TLS33 under various stresses. A, control; B, cold shock at 25°C; C, heat shock at 80°C; D, 10% salt (NaCl); E, 10% ethanol; F, 1% H_2O_2; G, 1 hr-UV.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

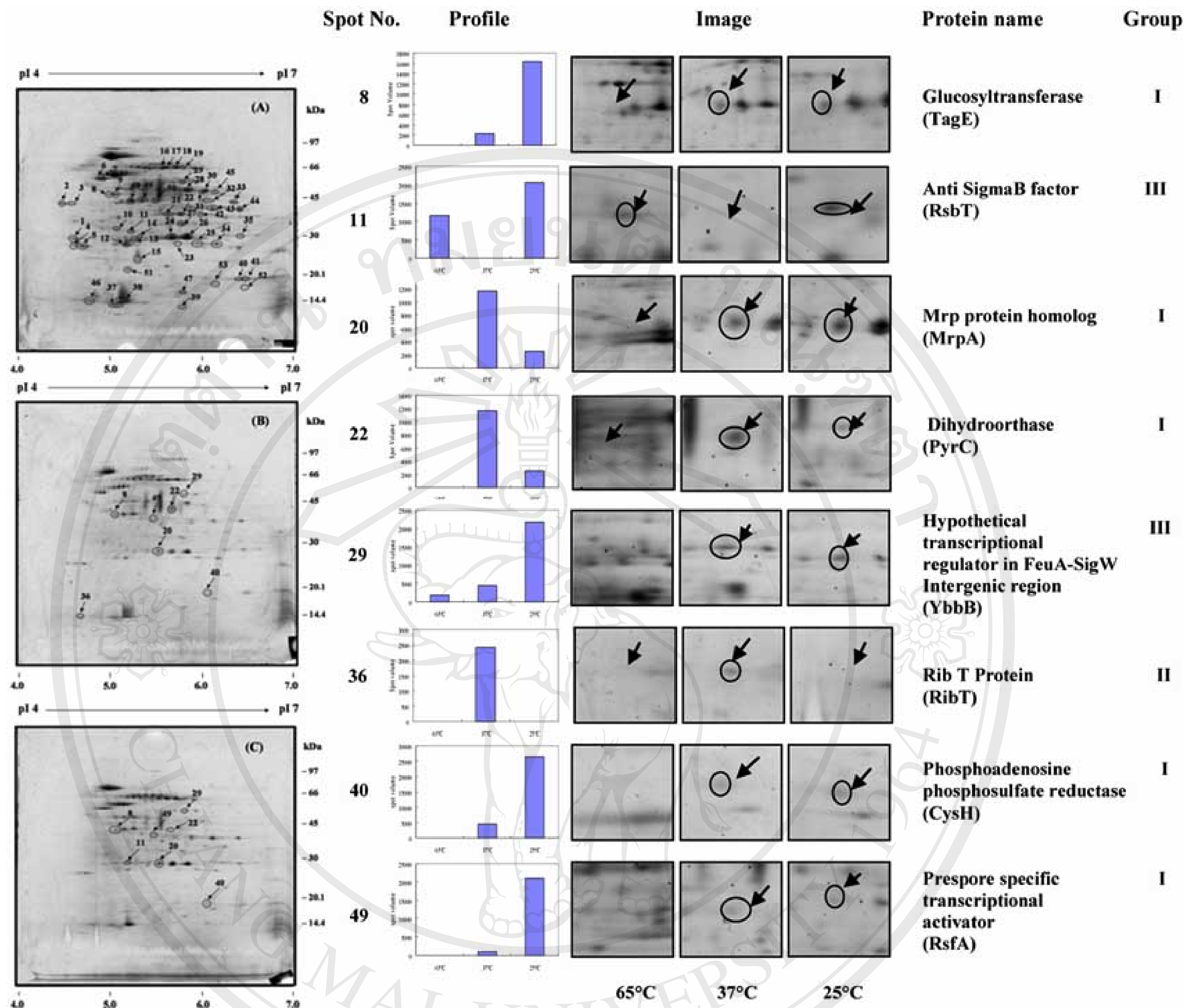


Fig. (5). Two dimensional polyacrylamide gel electrophoresis (2-D PAGE) of cytosolic proteins in *B. stearothermophilus* TLS33 at different temperatures (A, 65°C; B, 37°C; C, 25°C). The arrows in the 2-D gels on the left hand side of the figure show the proteins, which have been identified by MALDI-TOF mass spectrometry analysis. Right hand side of the Figure 5 shows the zoomed 2-D gel images of eight cold shock-induced proteins and the image data analysis by ImageMaster 2D Elite software. The arrows on the zoomed 2-D images show the location of the proteins. Three groups of cold shock-induced proteins are classified by ImageMaster 2D Elite software, based on the temperature of protein expression (I, protein expression at 37°C and 25°C; II, protein expression at 37°C; III, protein expression at 65°C and 25°C).

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University

one phosphorylated serine residues were identified in PrkCc, a Ser/Thr kinase from *B. subtilis* (Madec *et al*., 2003). MS was also used for studying the enzyme-induced covalent modification of methionyl-tRNA synthetase from *B. stearothermophilus* by methionyl-adenylate, which was identified as the labeled amino acid residues (Hountondji *et al*., 2000). In recent years, there have been many reports using MS for the identification and characterization of *Bacillus* species. For example, MALDI-TOF MS was used to rapidly characterize the spores of 14 microorganisms of the *Bacillus* group including *B. cereus*, *B. subtilis*, *B. anthracis*, *B. mycoides* and *B. thuringiensis*. Based on the mass spectra, unique patterns of biomarkers were used to distinguish *Bacillus* species members from one another, such as the differentiation at the strain level of *B. cereus* and *B. thuringiensis*, and these biomarkers differentiated the spores of *B. cereus* group from those of *B. subtilis* and *B. globigii* (Hathout *et al*., 1999; Ryzhov *et al*., 2000). Likewise, analysis by MALDI-TOF MS was also applied for the characterization of *B. anthracis* spore biomarkers, which corresponded to molecular weights of 2528.3, 2792.4, 3077.4 and 3590.7 Da, and have not been observed in extracts of the three closely related *Bacillus* species – *B. cereus*, *B. thuringiensis* and *B. mycoides* (Elhananny *et al*., 2001). Recently, it has been reported that identification of *Bacillus* spores is possible through the combination of a miniaturized MALDI-TOF

Table. 2. Summary of the Most Abundant Proteins in *B. Subtilis* 168 and *B. Stearothermophilus* TLS33 Cytosolic Extracts

***B. subtilis* 168**		***B. stearothermophilus* TLS33**	
Proteins	**Description**	**Proteins**	**Description**
Metabolism of carbohydrate			
Eno	Enolase in Glycolysis	G6pI	Glucose-6-phosphate
PdhD	Pyruvate dehydrogenase / 2-oxoglutarate dehydrogenase	PmgI	Phosphoglycerate mutase
PdhA	Pyruvate dehydrogenase (E1 alpha sub-unit)	SuhB	Archaeal fructose-1,6-bisphosphatase and related enzymes of inositol monophos-phatase family
PdhB	Pyruvate dehydrogenase (E1 beta subunit)		
PdhC	Pyruvate dehydrogenase (dihydrolipo-amide acetyltransferase E2 subunit)		
Pgk	Phosphoglycerate kinase		
SucC	Succinyl-CoA synthetase (beta subunit)		
Tkt	Transketolase		
YdjL	Similar to L-iditol 2-dehydrogenase		
YwjH	Similar to transaldolase (pentose phosphate)		
Metabolism of proteins/amino acids			
ArgF	Ornithine carbamoyltransferase	Ak2	Aspartokinase II (and subunits)
AroA	3-Deoxy-D-arabino-heptulosonate 7-phosphate synthase / chorismate mutase isozyme 3	Cyh1	Phosphoadenosine phophosulfate
Asd	Aspartate-semialdehyde dehydrogenase	His5	Amidotransferase
CysK	Cysteine synthetase A	YpwA	Similar to carboxypeptidase
DnaK	Class I heat-shock protein (molecular chaperone)	YqhT	Similar to Xaa-Pro dipeptidase
GlnA	Glutamine synthetase		
GlyA	Serine hydroxymethyltransferase		
GroEL	Class I heat-shock protein (chaperonin)		
IlvC	Ketol-acid reductoisomerase (Isoleucine, leucine, valine biosynthesis)		
IlvD	Dihydroxy-acid dehydratase		
LeuB	3-Isopropylmalate dehydrogenase		
MetC	Amino acid metabolism		
RocD	Ornithine aminotransferase		
SerA	Phosphoglycerate dehydrogenase		
TrxA	Thioredoxin		
YwaA	Similar to branched-chain amino acid aminotransferase		
Metabolism of nucleotides and nucleic acids			
EF-G	Translational apparatus	GuaB	Inosine-monophosphate dehydrogenase
EF-Ts	Translational apparatus	ImdH	Inosine-monophosphate dehydrogenase
EF-Tu	Translational apparatus	MobB	Molybdopterin-guanine dinucleotide biosynthesis
GuaB	Inosine-monophosphate dehydrogenase	Top1	DNA topoisomerase I
PurB	Adenylosuccinate lyase		
RplJ	Ribosomal protein L10 (BL5) in Translational apparatus		
RplL	Ribosomal proteins-large subunit		
RpsF	Ribosomal proteins-small subunit		
RpsB	Ribosomal protein S2 in Translational apparatus		
YpfD	Similar to ribosomal protein S1 homolog		
Metabolism of cell wall			
		PbpA	Penicillin-binding proteins 1A/1B
		TagE	UDP-glucose: polyglycerol phosphate glucosyltransferase

(Table 2) contd….

Copyright © by Chiang Mai University
All rights reserved

(Table 2) contd….

***B. subtilis* 168**		***B. stearothermophilus* TLS33**	
Proteins	**Description**	**Proteins**	**Description**
Transport/binding proteins and lipoproteins			
		CitH	Secondary transporter of divalent metal ions/citrate complexes
		FhuC	Ferrichrome ABC transporter (ATP-binding protein)
		OpcA	Glycine betaine/carnitine/cholic ABC transporter (ATP-binding protein)
		PrsA	Protein secretion (post-translocation molecular chaper-one)
		PtfB	PTS fructose-specific enzyme IIB component
		Slp	Small peptidoglycan-associated lipoprotein
		YqgS	Similar to putative molybdate binding protein
Mobility and chemotaxis			
Hag	Flagellin protein	FliT	Flagellar protein for mobility and chemotaxis
		LepA	GTP-binding protein
		YvyF	Similar to flagellar protein
Sporulation			
		RsfA	Probable regulator of transcription of ^{F}-dependent genes
		S3AE	Mutants block sporulation after engulfment
RNA synthesis/regulation			
		IF2	Initiation factor IF-2
		NusG	Transcription antitermination factor
		RsbT	Positive regulator of ^{B} activity
		YbbB	Similar to transcriptional regulator (AraC/XylS family)
Detoxification			
AhpC	Alkyl hydroperoxide reductase (small subunit)		
SodA	Superoxide dismutase	MsrA	Peptidyl methione sulfoxide reductase
Others			
AtpA	ATP synthase (subunit alpha)		
YkrA	Similar to unknown proteins	Cmf3	Late competence/transformation
YjlD	Similar to NADH dehydrogenase	XtmA	PBSX terminase (small subunit)
Yqjl	Similar to unknown proteins from B. subtilis	YqbP	Similar to phage-related protein
YxjH	Similar to unknown proteins from B. subtilis		

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่

mass spectrometer, *in situ* proteolysis, peptide mass mapping, use of a database limited to peptides derived from small acid soluble protein (SASP) families, and probability testing calculations (English *et al*., 2003). Comparison of the spectra of peptides to that of intact proteins offer the advantages of more sensitivity, better reproducibility and more accurate mass analysis, and thus used for *Bacillus* spore identification (English *et al*., 2003). In addition, it has been reported that MS could be applied to detect the protein-protein interactions within the intact cells of *B. subtilis*. These analyses showed that the identified patterns of associated proteins in the ribosome of *B. subtilis* strain ATCC633 were homologous with the ribosomal proteins of *Haloacula marismortui* and *Thermus thermophilus*, and other proteins involved in protein biosynthesis were shown to be associated with ribosomal proteins (Winters and Day, 2003). Furthermore, we have also identified 102 protein spots of extracellular proteins in *B. stearothermophilus* P1 in the stationary phase (48-h) by using MALDI-TOF MS followed by the database search engine MASCOT (www.matrixscience.com) (Table **3**). Work on identification of extracellular proteins in this thermophile in the stationary phase is in progress. We are also going to analyze the proteins in other phases of growth, under various conditions of stress and including the cytosolic proteins for understanding the global proteome profile of this bacterium.

Table 3. Identification of Extracellular Proteins in *B. Stearothermophilus* P1 in the Stationary Phase (48 h) Using MASCOT and the Mowse Score (http://www.matrixscience.com) from SwissProt and TrEMBL Databases (http://www. expasy.ch/sprot)

Spot no.	Approx. MW (Da)/pI	Mowse score	Matched (%)	Theoretical MW (Da)/pI	Accession no.	Protein name
1.	21,500 / 6.40	25	5 (29%)	19,707 / 5.79	P39667	Hypothetical protein yrxA (ORF1) (YrxA)
2.	21,000 / 6.68	22	6 (19%)	29,104 / 6.26	P31112	Heptaprenyl diphosphate synthase component I (EC 2.5.1.30) (HEPPP synthase subunit 1) (Spore germination protein C1) (Hep1)
3.	27,000 / 6.52	48	2 (10%)	28,264 / 6.46	P37543	Hypothetical protein yabB (YabB)
4.	27,000 / 6.32	16	5 (16%)	28,278 / 6.60	P40746	Protein-glutamine gamma-glutamyltransferase (EC 2.3.2.13) (Transglutaminase) (TGase) (Tgl)
5.	27,200 / 6.22	19	1 (19%)	31,038 / 6.27	O34500	Manganese transport system membrane protein mntD (MntD)
6.	26,700 / 6.05	19	2 (19%)	25,144 / 6.19	P53558	Dethiobiotin synthetase (EC 6.3.3.3) (DTB synthetase) (DtbS)
7.	26,000 / 6.05	20	6 (18%)	27,707 / 6.22	P36840	Acetylglutamate kinase (EC 2.7.2.8) (NAG kinase) (AGK) (N-acetyl-L-glutamate 5-phosphotransferase) (ArgB)
8.	26,000 / 5.87	14	2 (11%)	23,466 / 5.88	P16440	Riboflavin synthase alpha chain (EC 2.5.1.9) (RisA)
9.	24,000 / 5.90	22	4 (36%)	19,133 / 5.79	O31704	Molybdopterin-guanine dinucleotide biosynthesis protein B (MobB)
10.	26,000 / 5.70	26	3 (21%)	27,699 / 5.36	P13792	Alkaline phosphatase synthesis transcriptional regulatory protein phoP (PhoP)
11.	24,000 / 5.68	23	5 (24%)	21,679 / 5.46	P81100	Stress response protein SCP2 (Scp2)
12.	23,300 / 5.73	19	2 (15%)	26,822 / 5.35	P39639	Hypothetical protein ywfC (YwfC)
13.	19,300 / 5.68	21	4 (32%)	18,124 / 5.87	P46910	Hypothetical protein ywiD (YwiD)
14.	30,000 / 5.72	20	2 (18%)	29,612 / 5.61	P53559	6-carboxyhexanoate-CoA ligase (EC 6.2.1.14) (Pimeloyl-CoA synthase) (BioW)
15.	30,500 / 5.95	15	5 (14%)	31,373 / 5.84	P45856	Proable 3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157) (Beta-hydroxybutyryl-CoA dehydrogenase) (Hbd)
16.	30,000 / 6.03	14	2 (12%)	30,690 / 6.76	Q45581	Putative HTH-type transcriptional regulator ybbH (YbbH)
17.	29,000 / 6.17	17	2 (15%)	30,130 / 5.95	P37495	Hypothetical protein yybI (YybI)
18.	28,200 / 6.02	36	8 (34%)	28,264 / 6.46	P37543	Hypothetical protein yabB (YabB)
19.	33,000 / 6.03	22	7 (27%)	31,639 / 6.22	P39759	Hypothetical protein ykqA (YkqA)
20.	36,500 / 6.07	14	2 (11%)	30,169 / 6.04	Q45480	Hypothetical protein ylyB (YlyB)
21.	33,500 / 5.81	19	7 (11%)	36,617 / 5.92	P54556	Pantothenate kinase (EC 2.7.1.33) (Pantothenic acid kinase) (CoaA)
22.	40,300 / 6.15	23	3 (11%)	44,930 / 6.05	P13799	Sensor protein (DegS)
23.	41,300 / 6.18	15	6 (21%)	40,973 / 6.03	P54453	Hypothetical protein yqeH (YqeH)
24.	62,500 / 6.15	17	3 (10%)	50,828 / 6.05	P05648	Chromosomal replication initiator protein dnaA (DnaA)
25.	> 94,000 / 6.10	18	5 (6%)	134,546 / 5.51	P23477	ATP-dependent nuclease subunit B ATP-dependent nuclease subunit B (AddB)
26.	> 94,000 / 5.90	11	5 (8%)	119,394 / 5.79	O08394	Probable bifunctional P-450:NADPH-P450 reductase 1 (CypD)
27.	57,500 / 5.82	20	7 (20%)	61,478 / 5.93	Q45493	Hypothetical protein ykqC (YkqC)
28.	57,500 / 5.68	22	4 (10%)	59,841 / 5.76	P40871	2,3-dihydroxybenzoate-AMP ligase (EC 6.3.2.-) (Dihydroxybenzoic acid-activating enzyme) (DhbE)
29.	57,500 / 5.58	17	3 (10%)	52,051 / 5.86	O32193	Sensor protein (CssS)
30.	58,000 / 5.50	21	7 (22%)	55,436 / 5.27	P42176	Nitrate reductase beta chain (EC 1.7.99.4) (NarH)
31.	58,000 / 5.42	15	6 (12%)	65,940 / 5.39	P45740	Thiamine biosynthesis protein (ThiC)

(Table 3) contd….

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

(Table 3) contd….

Spot no.	Approx. MW (Da)/pI	Mowse score	Matched (%)	Theoretical MW (Da)/pI	Accession no.	Protein name
32.	58,000 / 5.33	9	2 (5%)	59,681 / 5.31	P13242	CTP synthase (EC 6.3.4.2) (UTP--ammonia ligase) (CTP synthetase) (PyrG)
33.	47,500 / 5.58	14	4 (16%)	53,007 / 5.61	O06491	Glutamyl-tRNA(Gln) amidotransferase subunit A (EC 6.3.5.-) (Glu-ADT subunit A) (GatA)
34.	44,000 / 5.87	22	7 (25%)	49,453 / 5.88	P12047	Adenylosuccinate lyase (EC 4.3.2.2) (Adenylosuccinase) (ASL) (Glutamyl-tRNA synthetase regulatory factor) (Pur8)
35.	42,500 / 6.12	17	4 (22%)	39,432 / 6.04	P39820	Glutamate 5-kinase 1 (EC 2.7.2.11) (Gamma-glutamyl kinase 1) (GK 1) (ProB)
36.	41,500 / 6.03	22	3 (15%)	40,973 / 6.03	P54453	Hypothetical protein yqeH (YqeH)
37.	38,500 / 5.78	19	6 (21%)	40,195 / 5.64	P80862	Phosphoserine aminotransferase (EC 2.6.1.52) (PSAT) (Vegetative protein 234) (VEG234) (SerC)
38.	36,800 / 5.68	14	2 (12%)	40,709 / 5.67	P37573	Hypothetical protein yacK (YacK)
39.	34,500 / 5.62	18	5 (20%)	32,668 / 5.65	P37887	Cysteine synthase (EC 4.2.99.8) (O-acetylserine sulfhydrylase) (O-acetylserine (Thiol)-lyase) (CSase) (Superoxide-inducible protein 11) (SOI11) (CysK)
40.	33,500 / 5.45	17	6 (21%)	32,680 / 5.50	Q45582	Hypothetical protein ybbI (YbbI)
41.	33,500 / 5.45	23	7 (29%)	34,665 / 5.52	O06974	Hypothetical UPF0052 protein (YvcK)
42.	32,500 / 5.48	15	6 (17%)	31,279 / 5.29	O34423	Hypothetical protein yjqC (YjqC)
43.	42,500 / 5.30	16	5 (23%)	45,606 / 5.20	P94415	Response regulator aspartate phosphatase C (EC 3.1.-.-) (RapC)
44.	42,500 / 5.18	16	9 (16%)	41,098 / 5.24	P37570	Hypothetical ATP:guanido phosphotransferase yacI (EC 2.7.3.-) (YacI)
45.	40,500 / 5.07	35	4 (21%)	40,330 / 5.01	P17865	Cell division protein ftsZ (FtsZ)
46.	40,500 / 4.98	15	5 (19%)	40,330 / 5.01	P17865	Cell division protein ftsZ (FtsZ)
47.	40,500 / 4.92	18	5 (24%)	39,312 / 5.04	O34640	Hypothetical protein yerI (YerI)
48.	37,300 / 4.93	16	5 (21%)	35,289 / 4.81	O32436	Transcriptional activator protein med precursor (Med)
49.	37,000 / 5.02	13	5 (19%)	38,597 / 5.02	P40399	Sigma factor sigB regulation protein rsbU (RsbU)
50.	< 14,400 / 4.95	17	4 (22%)	16,156 / 4.91	O32163	NifU-like protein (NifU)
51.	< 14,400 / 4.75	17	2 (16%)	13,104 / 4.99	P80241	General stress protein 17M (GSP17M) (G17m)
52.	< 14,400 / 4.62	12	1 (35%)	6,263 / 4.89	P45940	Hypothetical protein yqcE (YqcE)
53.	16,500 / 6.17	25	5 (36%)	16,762 / 5.91	Q45478	Hypothetical protein ylyA (ORF-T) (YlyA)
54.	15,500 / 5.87	22	4 (35%)	15,857 / 5.01	P96720	Hypothetical protein ywqH (YwqH)
55.	17,000 / 5.83	26	3 (24%)	17,684 / 5.57	O32068	Hypothetical protein ytzF (YtzF)
56.	47,500 / 4.88	16	6 (15%)	43,809 / 4.84	P08495	Aspartokinase 2 (EC 2.7.2.4) (Aspartokinase II) (Aspartate kinase 2) [Contains: Aspartokinase II alpha subunit; Aspartokinase II beta subunit] (Ak2)
57.	47,500 / 4.83	28	8 (25%)	43,809 / 4.84	P08495	Aspartokinase 2 (EC 2.7.2.4) (Aspartokinase II) (Aspartate kinase 2) [Contains: Aspartokinase II alpha subunit; Aspartokinase II beta subunit] (Ak2)
58.	47,500 / 4.78	30	9 (25%)	43,809 / 4.84	P08495	Aspartokinase 2 (EC 2.7.2.4) (Aspartokinase II) (Aspartate kinase 2) [Contains: Aspartokinase II alpha subunit; Aspartokinase II beta subunit] (Ak2)
59.	17,200 / 4.77	27	4 (28%)	18,244 / 5.39	P52035	Glutathione peroxidase homolog bsaA (BsaA)
60.	22,000 / 4.63	23	3 (20%)	23,773 / 4.63	P23449	Probable flagellar assembly protein fliH (FliH)
61.	26,500 / 5.17	17	2 (16%)	26,196 / 5.24	O32192	Transcriptional regulatory protein cssR (CssR)
62.	26,500 / 5.03	18	2 (14%)	29,431 / 5.01	P07601	Tryptophan synthase alpha chain (EC 4.2.1.20) (TrpA)

(Table 3) contd….

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

(Table 3) contd….

Spot no.	Approx. MW (Da)/pI	Mowse score	Matched (%)	Theoretical MW (Da)/pI	Accession no.	Protein name
63.	26,500 / 4.92	16	2 (10%)	28,812 / 5.20	O34409	YflN protein (YflN)
64.	18,300 / 6.67	17	4 (32%)	20,481 / 6.59	P49854	Hypothetical protein ykkA (YkkA)
65.	15,200 / 6.23	18	4 (24%)	16,735 / 5.69	P54512	Transcriptional regulator mntR (Manganese transport regulator) (MntR)
66.	< 14,400 / 5.52	23	5 (37%)	7,749 / 5.65	O35011	Probable DNA-directed RNA polymerase omega chain (EC 2.7.7.6) (Transcriptase omega chain) (RNA polymerase omega subunit) (RpoZ)
67.	< 14,400 / 5.62	23	2 (46%)	8,615 / 5.29	P54494	Hypothetical protein yqgQ (YqgO)
68.	40,500 / 5.75	39	4 (24%)	43,048 / 5.76	Q45460	Choline transport ATP-binding protein opuBA (OpbA)
69.	40,500 / 5.58	19	6 (18%)	42,608 / 5.55	P39131	Putative UDP-N-acetylglucosamine 2-epimerase (EC 5.1.3.14) (UDP-GlcNAc-2-epimerase) (MnaA)
70.	52,500 / 5.92	26	4 (14%)	48,078 / 6.45	O31440	Cytochrome P450 152A1 (EC 1.14.-.-) (P450BsBeta) (Fatty acid beta-hydroxylase) (CypC)
71.	52,500 / 5.82	18	3 (9%)	57,080 / 5.69	P35136	D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) (PgdH)
72.	51,000 / 5.18	18	3 (8%)	53,874 / 5.23	Q06752	Cysteinyl-tRNA synthetase (EC 6.1.1.16) (Cysteine--tRNA ligase) (CysRS) (Syc)
73.	32,500 / 5.05	28	3 (24%)	28,114 / 5.08	P35146	3-Dehydroquinate dehydratase (EC 4.2.1.10) (3-dehydroquinase) (Type I DHQase) (AroD)
74.	32,500 / 5.15	22	3 (14%)	35,833 / 4.98	P37941	2-Oxoisovalerate dehydrogenase beta subunit (EC 1.2.4.4) (OdbB)
75.	25,300 / 5.17	18	2 (20%)	25,974 / 5.07	P31749	Uridylate kinase (EC 2.7.4.-) (UK) (Uridine monophosphate kinase) (UMP kinase) (KraC)
76.	25,000 / 5.32	15	2 (22%)	24,646 / 5.34	O06738	Probable 2-phosphosulfolactate phosphatase (EC 3.1.3.-) (ComB)
77.	47,500 / 5.50	19	5 (14%)	45,808 / 5.47	P37942	Lipoamide acyltransferase component of branched-chain alpha-keto acid dehydrogenase complex (EC 2.3.1.-) (E2) (Dihydrolipoamide branched chain transacylase) ST (Odb2)
78.	51,000 / 5.12	15	5 (13%)	49,692 / 5.31	P54444	Hypothetical protein yrkQ (YrkQ)
79.	42,000 / 6.38	24	3 (12%)	41,358 / 5.85	P45938	Hypothetical protein yqcC (YqcC)
80.	42,000 / 6.52	14	2 (10%)	44,837 / 6.21	P53554	Biotin biosynthesis; cytochrome P450-like enzyme (BioI)
81.	43,000 / 6.57	22	9 (15%)	44,534 / 6.84	Q8RKI8	Putative CDP-glycerol:glycerophosphate glycerophosphotransferase (TarB)
82.	44,500 / 5.80	17	2 (11%)	47,881 / 5.67	P29726	Adenylosuccinate synthetase (EC 6.3.4.4) (IMP--aspartate ligase) (AdSS) (purA)
83.	44,000 / 5.72	19	2 (10%)	47,840 / 5.75	P23445	Flagellum-specific ATP synthase (FliI)
84.	34,000 / 5.90	22	5 (20%)	37,197 / 6.21	P50740	Isopentenyl-diphosphate delta-isomerase (Idi2)
85.	43,000 / 6.75	20	3 (10%)	40,179 / 6.54	P31104	Chorismate synthase (EC 4.2.3.5) (5-enolpyruvylshikimate-3-phosphate phospholyase) (AroC)
86.	48,000 / 5.42	21	6 (22%)	45,960 / 5.05	P16263	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex (Odo2)
87.	37,200 / 5.23	33	4 (16%)	40,258 / 5.35	P22322	Germination protease precursor (EC3.4.24.78) (Spore protease) (GPR endopeptidase) (Gpr)
88.	37,000 / 5.35	14	2 (12%)	32,152 / 5.30	P39842	Multidrug-efflux transporter 2 regulator Multidrug-efflux transporter 2 regulator (BltR)
89.	35,000 / 5.18	15	4 (15%)	32,419 / 5.17	O05510	Putative fructokinase (EC 2.7.1.4) (ScrK)

(Table 3) contd….

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

(Table 3) contd….

Spot no.	Approx. MW (Da)/pI	Mowse score	Matched (%)	Theoretical MW (Da)/pI	Accession no.	Protein name
90.	> 94,000 / 5.75	14	9 (12%)	105,966 / 5.87	P54397	Probable ATP-dependent helicase dinG homolog Probable ATP-dependent helicase dinG homolog (DinG)
91.	31,000 / 6.72	23	3 (17%)	31,115 / 5.25	P50843	4-Deoxy-L-threo-5-hexosulose-uronate ketol-isomerase (KduI)
92.	29,000 / 6.72	16	2 (15%)	26,938 / 6.40	P54527	Hypothetical protein yqiK Hypothetical protein yqiK (YqiK)
93.	27,000 / 6.12	24	3 (15%)	30,623 / 6.08	P54374	Shikimate 5-dehydrogenase (EC 1.1.1.25) (AroE)
94.	32,000 / 5.33	24	3 (14%)	30,982 / 5.46	P28822	Dihydropteroate synthase (EC 2.5.1.15) (DhpS)
95.	46,000 / 5.15	15	2 (9%)	50,364 / 5.04	P80860	Glucose-6-phosphate isomerase (GPI) (Phosphoglucose isomerase) (Pgi)
96.	39,300 / 5.38	14	2 (9%)	43,539 / 5.25	O34714	Oxalate decarboxylase (OxdC)
97.	37,200 / 5.12	14	4 (21%)	37,560 / 5.28	P54550	Probable NADH-dependent flavin oxidoreductase (YqjM)
98.	41,500 / 6.02	19	3 (10%)	40,868 / 5.95	P36839	Acetylornithine aminotransferase (EC 2.6.1.11) (ACOAT) (ArgD)
99.	62,000 / 5.78	16	7 (17%)	62,684 / 5.27	P46906	Arginyl-tRNA synthetase (EC 6.1.1.19) (Arginine--tRNA ligase) (ArgRS) (Syr)
100.	40,000 / 6.02	26	7 (26%)	39,003 / 6.09	P25499	Heat-inducible transcription repressor (HrcA)
101.	26,000 / 5.58	17	2 (18%)	25,958 / 5.69	Q9F4F7	4'-Phosphopantetheinyl transferase (Ffp)
102.	44,500 / 5.68	13	2 (9%)	45,713 / 5.54	O32148	Purine catabolism protein (PucG)

2.3. Capillary Electrophoresis

Capillary electrophoresis (CE) is one of the techniques, which provides a simple and rapid analysis along with high resolution of separation. The high accuracy and rapid separation possible with CE may at times be more effective than high performance liquid chromatography (HPLC). CE is thus quite often used for analysis and quantitation of proteins. For example, CE was used for analyzing the amount of parasporal crystal protein (-endotoxin) from the fermentation broth of *B. thuringiensis* after it was dissolved in a reducing agent, such as -mercaptoethanol. The results indicated that the CE could be more accurate and rapid than HPLC (Liu and Tzeng, 2001). A combination of capillary electrophoresis and mass spectrometry or 'CE-MS' has emerged as a powerful tool for the quantitative metabolome analysis, which allows quantitative analysis of metabolites. These data are then compared with the complementary data obtained from gene expression and proteome analysis to elucidate the change in biological functions. More than 500 positively charged metabolites including amino acids and amines have been separated by CE and detected by MS. Simultaneous analysis of hundreds of negatively charged species such as carboxylic acids, phosphorylated carboxylic acids, phosphorylated saccharides, nucleotides and CoA compounds in glycolysis, pentose phosphate and the TCA cycle pathways is thus possible. Quantitative analysis of metabolites of *B. subtilis* cell by CE and has enabled the determination of 1692 metabolites from cell extracts. This has revealed significant changes in metabolites during sporulation of the bacteria (Soga *et al*., 2003). A dynamic pH junction-sweeping CE with laser-induced fluorescence (LIF) detection has also been applied as a robust single method to analyze sub-micromolar amounts of three flavin derivatives, riboflavin, flavin mononucleotide (FMN), and flavin adenine dinucleotide (FAD), from several types of samples, including cell extracts of *B. subtilis* (Britz-McKibbin *et al*., 2003). The application of CE with sweeping by borate complexation has been used to analyze a group of seven pyridine and adenine nucleotide metabolites derived from *B. subtilis* cell extracts. This showed that concentration of pyridine and adenine nucleotides in cell extracts varied widely from 78.6 μM for nicotinamide-adenine dinucleotide in malate to 0.66 μM for nicotinamide-adenine dinucleotide phosphate in glucose culture medium (Markuszewski *et al*., 2003).

2.4. Other Proteomic Techniques: DIGE, ICAT, Aptamers, Protein Chip

Over the past couple of years, proteomics techniques have been developed to obtain more efficient methods for the separation and analysis of mixtures of proteins to study the differential expression of the same protein in two different protein extracts. The difference gel electrophoresis (DIGE) technique allows the multiplex analysis of two or three samples containing a mixture of proteins on the same 2-D PAGE, which is an advantage of this technique (Unlu *et al*., 1997). The protein extracts under comparison are covalently labeled with different fluorescent CyDye™, which are N-hydroxy succimidyl ester derivatives of Cy2, Cy3 and Cy5 and have their own characteristic excitation and emission spectra. Different images of the same 2-D PAGE gel can be obtained using different filters according to the excitation and emission spectra of different fluorescent dyes used to label the extracts. This technique was used successfully to examine the effect of Cry toxins of *B. thuringiensis* (Bt) on the gut proteins from the larvae of an Indian-meal moth (IMM, *Plodia interpunctella*) colony. Two unique protein

Copyright © by Chiang Mai University
All rights reserved

spots were identified that represent a highly distinguishable profile for the Bt-susceptible (S) and Bt-resistant (R) IMM larvae (Candas *et al*., 2003).

The Isotope Coded Affinity Tag (ICAT) method is also used to study the differential expression of proteins. The protein extracts from control and experimental samples are differentially labeled by the addition of suitable isotopically "light" and "heavy" ICAT reagents that reacts, for example, with cysteine residues of the proteins. The ICAT labeled protein extracts are then mixed and subjected to trypsin digestion followed by affinity isolation of the ICAT labeled peptides and multidimensional HPLC with on-line nano-ESI tandem mass spectrometric (MS-MS) analysis. The "light" and "heavy" isotopically labeled forms of each peptide usually coelute from the HPLC and are then distinguished in the MS analysis due to the mass difference inherent in the isotope tags. The MS analysis data for each of the peptide pairs are used for identification of the original protein and their quantification (Gygi *et al*., 1999).

Dual channel image analysis also offers the opportunity to visualize the content and synthesis rate of a whole set of bacterial proteins on a single electropherogram. This novel technique is useful for rapid identification of proteins that belong to different stimulons or regulons. This approach has been tested for the identification of proteins of *B. subtilis* under oxidative stresses (Bernhardt *et al*., 1999; Hecker, 2003). The global regulation of differential stresses of *B. subtilis* was investigated by pulse-labeling with L-[^{35}S] methionine. In this technique, the protein synthesis pattern (red color) and the protein level pattern (green color) can be directly compared. Thus, the advantage of this technique is the matching of protein patterns on one single gel, which are one of the bottlenecks in data analysis.

Protein-protein interactions by NMR chemical shift mapping and bacterial two-hybrid system have also been used to study the functional specificity of protein interactions. The chemical shift mapping approach was able to reproduce the known functional specificities among pairs of closely related proteins of the phosphoenolpyruvate: sugar phosphotransferase systems of *B. subtilis* and the tyrosyl-tRNA synthetase of *B. stearothermophilus* (Karimova *et al*., 2001; Rajagopal *et al*., 1997).

Another new approach, aptamer technology has the potential for use as inexpensive *in vitro*-generated receptors for biosensors in biological warfare detection and other areas as well. Based on systematic evolution of ligands by exponential enrichment (SELEX) method used for selection and PCR amplification, the DNA aptamers were able to bind to and detect the nonpathogenic Sterne strain *B. anthracis* spores (Bruno and Kiel, 1999). The high affinity spore surface bound aptamers were detectable with a dynamic range equivalent to <10 - $>6 \times 10^6$ anthrax spores.

Protein microarray technology based on protein chips and biosensors is also a new, powerful tool for applications in proteomics research and diagnostics. Antibodies, which offer specificity and high affinity, are the reagents of choice for protein microarrays. Antibody microarrays have been used successfully to detect and quantify specific target proteins in complex mixtures (Haab *et al*., 2001), to screen and identify differentially expressed proteins produced by recombinant antibody technology (de Wildt *et al*., 2000), and to detect clinically relevant cytokines (Huang *et al*., 2001; Moody *et al*., 2001). Microarray system was employed to detect protein and analytes of *B. globigii* and levels as low as 6.2×10^4 cfu/ML with 15 min-completed assays were detected (Delehanty and Ligler, 2002). Based on the same principle, the protein chip technology of Agilent 2100 Bioanalyzer and Protein 200 LapChip kit (AgilentTechnologies, USA) is also a new choice for protein sizing and chip-based separation. It has been used for analyzing proteins from *B. stearothermophilus* under substrate induction and cold shock stress, demonstrating the high-resolution separation on image-like gel with the different protein patterns and electropherograms (Sinchaikul *et al*., 2002b). MicroSERS is a new biochip technology that uses surface-enhanced Raman scattering (SERS) microscopy for label-free transduction. The biochip itself comprises of pixels of capture biomolecules immobilized on a SER-active metal surface. Once the biochip has been exposed to the sample and the capture biomolecules have selectively bound their ligands, a Raman microscope is used to collect SERS fingerprints from the pixels on the chip. The initial study has shown that *Bacillus* spores of six species of *B. brevis*, *B. cereus*, *B. coagulans*, *B. stearothermophilus*, *B. subtilis* and *B. thuringiensis* could often be identified at the subspecies/strain levels on the basis of SERS fingerprints collected from single organism. This technique could also detect the difference on spore surface between viable and heat-killed spores of *B. thuringiensis* (Grow *et al*., 2003).

3. COMPREHENSIVE STUDY OF PROTEIN FUNCTIONS OF *BACILLUS* SP.

A key goal of proteomics is to assign function to proteins and many such functions are related to the levels of the proteins.

3.1. Secretion Network

In general, *Bacillus* species secrete numerous proteins into the medium. Many of the secretory proteins as well as their export signals and their processing steps during secretion have been characterized; in contrast, the molecular mechanisms of protein secretion have been relatively poorly characterized. However, the basic principles of protein secretion in *Bacillus* sp. are similar to those in other organisms such as *E. coli* that involve the late stages in the secretion process following protein translocation across the cytoplasmic membrane. In general, proteins to be translocated across the membrane are synthesized with an N-terminal extension of 15-30 amino acids, the signal peptide. This signal peptide is required for maintaining the translocation-competent state of the precursor, its targeting to the membrane and the initiation of translocation across the membrane, which is mediated by complex machinery consisting of the so called 'Sec'-proteins (Tjalsma *et al*., 2000; Bron *et al*., 1998; von Heijne, 1990; Pugsley, 1993; Driessen, 1994). The model of protein secretion pathway in *Bacillus* sp. is shown in Fig. **6**. Briefly, the precursor proteins in the cytoplasm become associated with factors, which assist in maintaining the export-competent state and the targeting to translocase complex in the membrane. These cytoplasmic factors (C/T) are chaper-

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

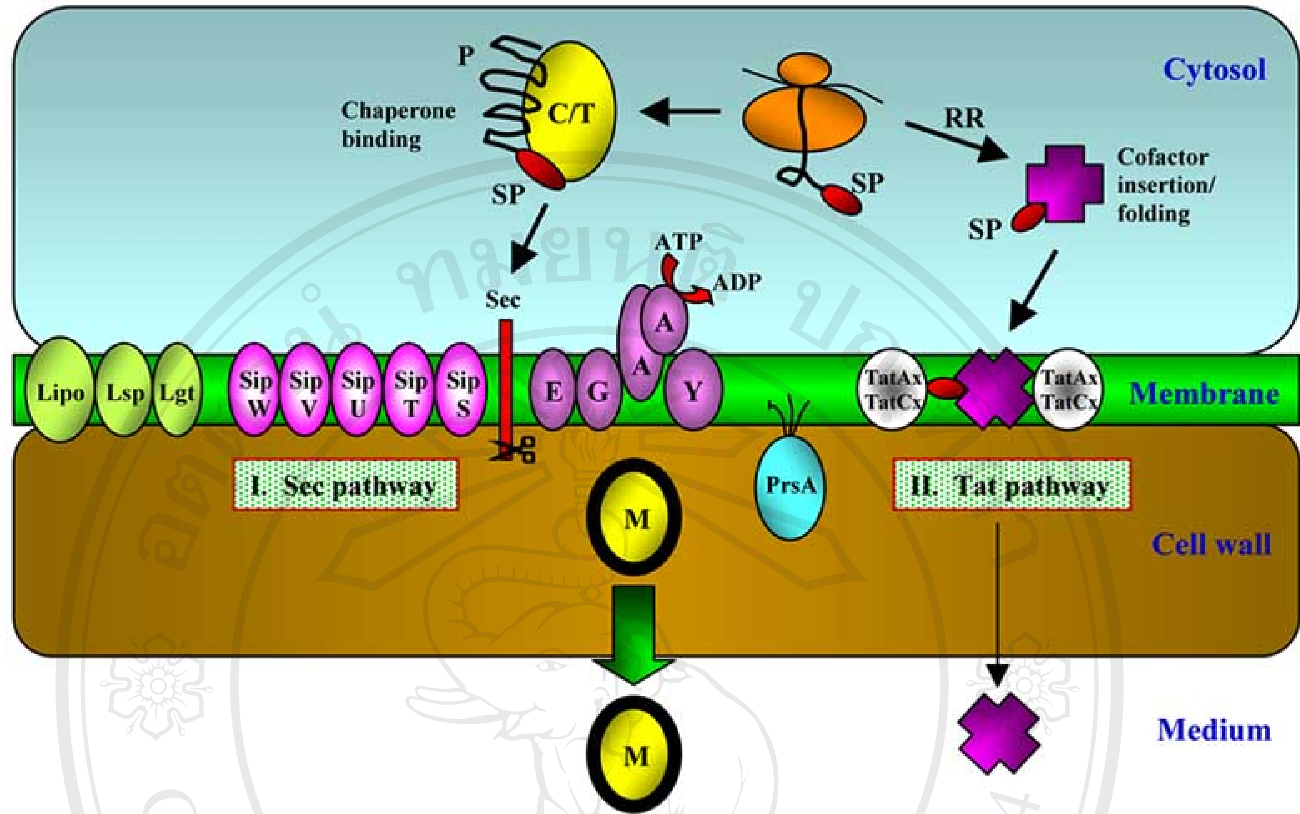


Fig. (6). Model of the protein secretion in *Bacillus* sp. by two pathways. I. Sec pathway: P, precursor protein; C/T, chaperone and targeting factors; SP, signal peptide; A,Y, E and G, the SecA, SecY, SecE and SecG components of the translocase complex; Sip, signal peptidases type I; Lsp, signal peptidase type II (lipo-type); Lgt, lipoprotein diaccylglyceryl transferase; Lipo, lipoproteins; PrsA, extracytoplasmic folding factor; M, mature protein. II. Tat pathway: RR, twin-arginine signal peptide; Tat, Tat complexes (TatAx: TatAd & TatAy, TatCd and TatCy).

ones and components of the signal recognition particle (SRP)-like pathway. SecA is a peripheral membrane-associated ATPase with affinity for both the precursor/chaperone complex and the translocase, of which SecY and SecE are major components. SecD, SecG and SecF also play a role in the translocation process. During or shortly after translocation, the signal peptide is removed by a signal peptidase (Sip), which has been identified in *B. subtilis* (van Dijl *et al*., 1992; Bolhuis *et al*., 1996; Tjalsma *et al*., 1997). Cleavage of the signal peptide is required for release of the protein from the cell. The subsequent folding of the mature protein into a stable and active conformation usually requires additional factors, such as the PrsA protein (Kontinen and Sarvas, 1998) or metal ions (Petit-Glatron *et al*., 1993). For late stages in the secretion process, the paralogous sip gene family of *B. subtilis* consists of five chromosomal genes, denoted sipS, sipT, sipU, sipV and sipW. In addition, certain *B. subtilis* strains contain plasmid-borne sipP genes. All identified sip genes are type I signal peptidases (SPases) that remove the signal peptides from secretory proteins. Interestingly, cells lacking both SipS and SipT were not viable, which may be due to jamming of the secretion machinery with secretory pre-proteins, showing that these two SPases are of major importance for protein secretion. The newly identified SPase, SipW is highly similar to SPases from archaea and the endoplasmic reticular (ER) membrane of eukaryotes, suggesting that these enzymes form a sub-family of the type I Spases, which are conserved in all forms of life. Lipoproteins are also directed into the general (Sec) pathway for protein secretion by their signal peptides, which show similar structural characteristics as the signal peptides of secretory proteins: a positively charged amino terminus, a hydrophobic core region, and a carboxyl-terminal region containing the cleavage site for signal peptidase. The major difference between signal peptides of lipoproteins and secretory proteins is the presence of a well conserved "lipobox", containing a conserved cysteine residue in the signal peptide which is the target for lipid modification by the lipoprotein diacylglyceryl transferase (Lgt) and subsequent processing by the lipoprotein-specific type signal peptidase (SPase) II (Lsp). After processing, the diacylglyceryl residue is anchored into the cytoplasmic membrane (Tjalsma *et al*., 1999; Lesela *et al*., 1999). Inspection of the genome sequence revealed 114 putative pre-lipoproteins containing the characteristic lipobox. Eight potential lipoproteins were found in the extracellular space. All of them lacked the cysteine residue at the N-terminus, indicating a proteolytic shaving after their processing by SPase II or an alternative protease (Antelman *et*

al., 2001). However, the detailed analysis of heterologous protein secretion in *Bacillus* sp. has allowed identification of several bottlenecks in the secretion pathway that, alone or in combination, can drastically reduce the amount of secreted proteins in the supernatant. The major bottlenecks were found to be: (1) inefficient translocation of the protein across the plasma membrane; (2) inefficient folding of the protein on the trans-side of the plasma membrane; (3) degradation of the translocated protein by cell wall-associated proteases; and (4) degradation of the protein by secreted proteases in the supernatant. Although limitations still exist, the high technology approaches have extensively characterized these bottlenecks at the molecular level and secretory networks will be better understood in the near future.

Recently, an alternative Sec-independent protein export pathway was found to exist in *Bacillus* bacteria for a class of redox cofactor containing proteins, which allows the translocation of these proteins in a folded state. The corresponding signal peptides contain a specific twin arginine motif and therefore, the new pathway has been designated as "Tat" (Twin-arginine-translocation) pathway (van Dijl *et al*., 2002). Most exported proteins appear to use the Sec pathway, 69 of these proteins appear to use the Tat pathway, as their signal peptides contain twin-arginine (RR) or lysine-arginine (KR) motifs. Precursors with a twin-arginine signal peptide have the potential to fold in the cytoplasm before their translocation by the Tat machinery in the membrane. Known components of the *B. subtilis* Tat pathway are the TatAc/Ad/Av and TatCd/Cy proteins, which are specified by two paralogous gene families, respectively. TatCd and TatCy are determinants for pathway specificity but function of TatAc is presently unknown. Furthermore, proteomic techniques were applied to verify how many extracellular *B. subtilis* proteins follow the Tat pathway and the results showed that the extracellular accumulation of 13 proteins with potential RR/KR-signal peptides was Tat-independent and only phosphodiesterase PhoD was shown to be secreted in a strictly Tat-dependent manner (Jangbloed *et al*., 2002). On the other hand, pre-PhoD is processed at a very low rate and substantial amounts of this protein can be detected at the membrane-cell wall interface (Muller and Wagner, 1999). As no Tat-dependent exported proteins are detectable on the cell wall proteome of *B. subtilis*, the goal is to investigate the Tat pathway involved in the biogenesis of membrane proteins.

3.2. Regulatory Network of *Bacillus* sp. Under Stresses

The bacteria can respond by embarking upon one of two distinct developmental pathways: (1) natural competence for genetic transformation, in which specific proteins for DNA uptake, repair and recombination are synthesized, or (2) sporulation, a complex genetic program during which the cell differentiates into a dormant, heat- and stress-resistant form (the spore), giving its genome the ultimate shelter from a hostile environment. The induction of stress proteins is an important component of the adaptational network of a non-growing cell of *Bacillus* sp. A diverse range of stresses such as heat and cold shocks, salt stress, ethanol and starvation for oxygen or nutrients induce the same set of proteins, called general stress proteins. Although the adaptive functions of these proteins are largely unknown, they are proposed to provide general and rather non-specific protection of the cell under these adverse conditions. Moreover, the crucial steps in exploring adaptational networks are to define the genes induced by a single stress or starvation stimulus, to identify and analyze the corresponding proteins, and to understand their adaptive function in response to stress or starvation. Proteomics is an excellent tool for elucidating this network, dissecting it into its individual components and studying the various adaptive functions of these components. The basic steps for exploring the modules of the entire network are: (1) The definition of stimulons which is the entire set of proteins/genes induced or repressed by one stimulus – all proteins induced by a specific stimulus contribute to stress adaptation, and therefore defining the size and structure of a stimulon represents the first step in elucidating adaptation to the stimulus. (2) The dissection of stimulons into single regulons, the basic modules of global gene expression (a regulon consists of a set of genes distributed on the genome, but controlled by a global regulator); (3) the analysis of regions that overlap between different regulons (modulons), an essential step towards exploring complex adaptational networks. However, only dissection of the entire genome into single regulon is not yet sufficient for understanding global gene regulation because single regulons do not exist independently from one another but are tightly connected, forming a complex adaptational network (Hecker and Volker, 2001). In the present, by using the high throughput techniques of proteome analysis coupled with comprehensive computer-aided data analysis, it is possible to proceed from 2-D PAGE analysis of proteins from growing, starved, or stressed *B. subtilis* cells index to a more global analysis and description of the gene regulation map of a cell (Antelmann *et al*., 1997a, 2001). The 2-D PAGE analysis was also used to investigate the changes in the protein synthesis profile in *B. subtilis* wild-type strains and *sig*B mutants in response to heat shock, salt and ethanol stress, and glucose or phosphate starvation (Bernhardt *et al*., 1997). The results showed that at least 42 increased stress proteins absolutely required the alternative sigma factor ^{B}, while at least seven stress proteins, among them ClpC, ClpP, Sod, AhpC and AhpF, remained stress-inducible in a *sig*B mutant. It indicated that the subgroups of general stress proteins act through fine-tuning of this complex response. The relative synthesis rate of the general stress proteins constituted up to 40% of the total protein synthesis of stressed cells and thereby emphasizes the importance of the stress regulon. Furthermore, the regulations of the stress response in *Bacillus* sp. are mostly regulated by sigma factor B (^{B}), which is one of the first alternate sigma factors (Haldenwang and Losick, 1979) and is the master regulator of a large number of general stress genes induced in response to heat, ethanol, acid or salt stress as well as by starvation for glucose, phosphate and oxygen. The inducers of ^{B}-dependent genes have been divided into two main pathways (Alper *et al*., 1994; Voelker *et al*., 1996; Kang *et al*., 1996; Price *et al*., 2001; Price, 2002). The first pathway contains extracellular signals that trigger a drop in the cellular ATP level. Starvation for oxygen, glucose or phosphate induces the sigB gene *via* ATP depletion. The second pathway of stress factors includes heat, cold, salt and ethanol. The regulatory network leading to ^{B} activation and repression functions by a so-called partner switching mechanism, in which the sigB gene is encoded by the sigma-B operon consisting of eight genes in the order rsbR-rsbS-rsbT-rsbU-

rsbV-rsbW-sigB-rsbX, where rsb denotes regulation of sigma B (^{B}) (Akbar *et al*., 1997; Schumann, 2000). By proteomics, transcriptional analysis, transposon mutagenesis and consensus promoter-based screening, 75 genes of *B. subtilis* had previously been described as ^{B}-dependent general stress genes. Gene array-based analysis confirmed 62 of these already known general stress genes and detected 63 additional genes subject to control by the stress sigma factor ^{B} (Petersohn *et al*., 2001). In addition, it has been reported that *B. licheniformis* cells, contain a protein that closely resembles *B. subtilis* ^{B} in size and antigenic properties and the level of this protein rapidly increases following environmental or energy stress, and its genome encodes a homologue of the sigB general stress operon, including the ^{B} structural gene and seven rsb regulatory genes (Brody and Price, 1998). Finally, we have summarized the model regulation of sigma factor signal transduction network of *Bacillus* sp. *via* ^{B}-general stress and Pho regulons as shown in Fig. **7** (A, B).

3.2.1 Temperature Stresses

Temperature is an important environmental factor, which requires adaptive responses from bacterial cells. While a sudden increase in the growth temperature induces a heat shock response, a decrease in temperature results in a cold shock response. Both responses involve a transient increase in a set of genes called heat and cold shock genes, respectively, and the transient enhanced synthesis of the corresponding proteins allows the stressed cells to adapt to the new situation.

For heat shock response, an immediate increase in the growth temperature results in the unfolding of proteins and hydrophobic amino acid residues normally buried within the interior of the proteins and become exposed on their surface. These hydrophobic residues often constitute hydrophobic surface proteins and can interact and form aggregates, which may become life threatening. Regulation of the heat shock response has been studied most thoroughly in *B. subtilis* and

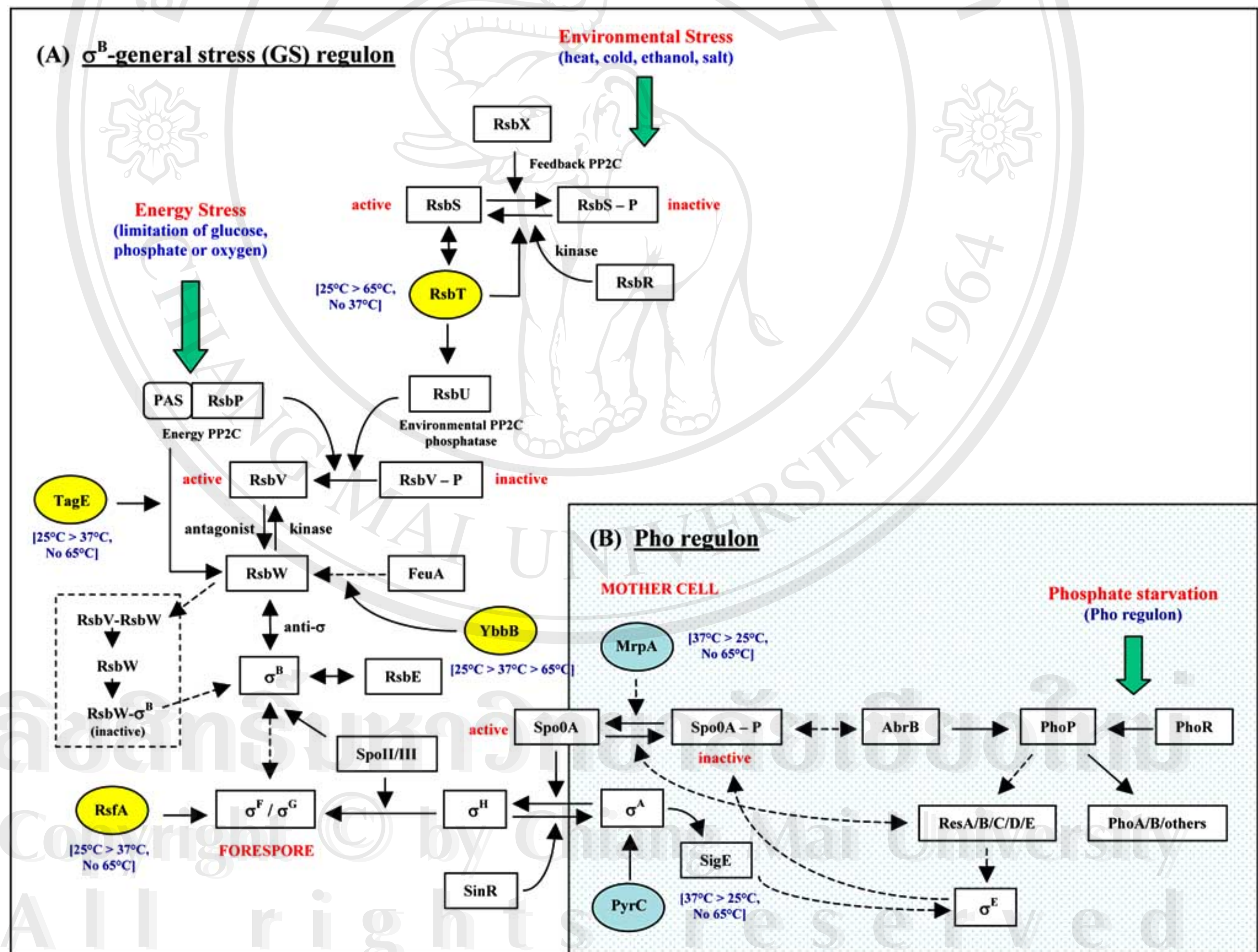


Fig. (7). Model of signal transduction pathway of *Bacillus* sp. in response to various stresses *via* ^{B}-general stress regulon (A) and Pho regulon (B). Six cold shock-induced protein encoding genes of thermophilic bacterium *B. stearothermophilus* TLS33 under cold shock stresses at 37°C and 25°C is correlated to the pathway and shown in six circles.

Copyright © by Chiang Mai University
All rights reserved

turns out that (1) all known heat shock genes are induced at the level of transcription; (2) they are either under the positive control of an alternate sigma factor or under negative control of a transcriptional repressor, (3) the heat shock response is transient; and (4) the heat shock genes present within one bacterial species form several regulons regulated by different mechanisms (Schumann, 2000). In general, the regulation of the heat shock response in *Bacillus* sp. is mainly regulated by sigma factor B (^{B}) with the transcriptional repressors or operators such as HrcA/CIRCE controlling inverted repeat of chaperone expression, and CtsR/CtsR box, class three stress gene repressor. In *B. subtilis*, it has been found that at least three different classes of heat-inducible genes can be defined by their common regulatory characteristics. Class I genes, as exemplified by the *dna*K and *gro*E operons, are most efficiently induced by heat stress. This expression involves an ^{A}-dependent promoter, an inverted repeat (called the CIRCE element) highly conserved among eubacteria, and probably a repressor interacting with the CIRCE element. Class II genes are the majority of general stress genes (more than 40 genes) and induced at ^{B}-dependent promoters by different growth-inhibiting conditions. The activation of ^{B} by stress or starvation is the crucial event in the induction of this large stress regulon. Class III genes, as encoded by a few genes including *Ion*, *clp*C, *clp*P and *fts*H, can respond to different stress factors independently of ^{B} or CIRCE. Stress induction of these genes occurs at promoters presumably recognized by ^{A} and probably involves additional regulatory elements, which remain to be defined (Hecker *et al*., 1996). Moreover, there have been studies conducted with other *Bacillus* species. For example, the proteins of *B. cereus* ATCC 14579 under heat stress was analyzed by 2-D PAGE and 31 heat-induced proteins were observed, which revealed induction of stress proteins (CspB, CspE and SodA), proteins involved in sporulation (SpoVG and AldA), metabolic enzymes (FolD and Dra), heat-induced proteins in related organisms (DnaK, GroEL, ClpP, RsbV, HSP16.4, YflT, PpiB and TrxA), and other proteins (MreB, YloH and YbbT) (Periago *et al*., 2002). These observations indicated that heat adaptation of *B. cereus* involved proteins that function in a variety of cellular processes. Otherwise, a gene cluster encoding the alternative sigma factor ^{B}, three predicted regulators of ^{B} (RsbV, RsbW and RsbY), and one protein whose function is not known (Orf4) were identified in the genome sequence of the food pathogen *B. cereus* ATCC 14579 (van Schaik *et al*., 2004). Western blotting with polyclonal antibodies raised against ^{B} revealed that there was 20.1-fold activation of ^{B} after a heat shock from 30 to 42°C. Analysis of transcription of the sigB operon by Northern blotting and primer extension revealed the presence of a ^{B}-dependent promoter upstream of the first operon reading frame (rsbV) of the sigB operon, indicating that transcription of sigB is autoregulated.

For cold shock response, it evokes two major threats to the cells, namely a drastic reduction in membrane fluidity of cytoplasmic membrane that influences membrane-associated cellular functions such as active transport and protein secretion, and a transient complete stop of translation that affect the efficiencies of DNA replication and recombination. Immediately after cold shock, bacteria either completely or partly stop growth for a period of time and then cells adapt to the lower temperature, and this lag-phase has been designated as the acclimation phase (Schumann, 2000; Weber and Marahiel, 2003). The cold shock response has been studied extensively in *B. subtilis* (Graumann *et al*., 1996,1997; Graumann and Marahiel, 1999) and to some extents in *B. stearothermophilus* and *B. cereus* (Sinchaikul *et al*., 2002a, 2002b; Francis *et al*., 1998). Following a rapid decrease in temperature, the physiology of *B. subtilis* cells changed profoundly. Cold shock adaptation has been monitored at the level of membrane composition, adjustment in DNA topology, and change in cytosolic protein synthesis/composition. Some major players in these processes (cold-stress induced proteins and cold acclimatization proteins, CIPs and CAPs) have been identified and mechanisms in cold shock acclimatization have begun to emerge. However, important questions regarding their cellular function still need to be answered. Most published studies deal with the regulatory role of CspA after a temperature down-shock. There have been a number of studies on the cold shock response in *B. subtilis* by proteomic and genomic analysis. For example, after a temperature shift from 37° to 15°C, the synthesis of a majority of proteins in *B. subtilis* was repressed; in contrast, 37 proteins analyzed by 2-D PAGE were synthesized at rates higher than preshift rates (Graumann *et al*., 1996). The 2-D PAGE analysis of a cspB null mutant revealed that CspB affects the level of induction of several cold induced proteins (CIPs). Other identified CIPs function at various levels of cellular physiology, such as chemotaxis (CheY), sugar uptake (Hpr), translation (ribosomal proteins S6 and L7/L12), protein folding (PPiB), and general metabolism (CysK, IlvC, Gap and triosephosphate isomerase). Also, the transcriptome of *B. subtilis* was analyzed at different time points (30, 60 and 90 min) after a temperature downshift from 37° to 18°C using DNA microarray (Kaan *et al*., 2002). Around 50 genes exhibited an increased or decreased mRNA level under cold-shock conditions and many of the repressed genes encoded enzymes involved in the biosynthesis of amino acids, nucleotides and coenzymes, indicating metabolic adaptation of the cells to the decreased growth rate at the low temperature. The cold shock-induced increase of mRNA levels of the classical cold shock genes cspB, cspC and cspD were verified and it was found that the amount of mRNA specified by the operon ptb-bcd-buk-lpd-bkdA1-bkdA2-bkdB, which encodes enzymes involved in degradation of branched-chain amino acids, also increases after temperature downshift. As cells utilize the isolucine and valine degradation intermediates for synthesis of branched-chain fatty acids, this finding reflects the adaptation of membrane lipid composition, ensuring the maintenance of appropriate membrane fluidity at low temperatures. Likewise, the *B. subtilis* two-component systems of cold-induced proteins (CIPs) and cold acclimatization proteins (CAPs) encoded by the desKR operon has been demonstrated to be essential for the cold-induced expression of the lipid-modifying desaturase Des, which is required for efficient cold adaptation of the membrane in the absence of isoleucine (Beckering *et al*., 2002). In addition, we also studied the functional proteomics of thermophilic bacterium *B. stearothermophilus* TLS33 under cold shock stress. It was observed that after the temperature downshift from 65°C to 37 ° or 25°C for 2 h, six cold shock-induced proteins were markedly expressed under cold stress, such as glucosyltransferase (TagE), anti-sigma B factor (RsbT), Mrp

protein homolog (MrpA), dihydroorthase (PyrC), hydroadenosine phosphosulfate reductase (YbbB) and prespore specific transcriptional activator RsfA (RsfA), whose encoded genes were correlated with the sporulation signaling pathway as shown in Fig. **7A** (Topanurak *et al*., 2005). It was also found that the regulation of this bacterial cell under cold shock could be divided to 2 routes of activation (RsbT, TagE, YbbB and RsbT) and deactivation (MrpA and PyrC). However, the biological functions and mechanisms in the gene expression level of this thermophile under different stress conditions will be clarified in the future.

3.2.2. Glucose, Phosphate and Nitrogen Starvations

In general, the first essential glucose starvation-specific response in switching to utilization of alternative carbon sources is mediated mainly by the catabolite control protein CcpA (Stülke and Hillen, 2000). Cells growing on excess glucose synthesize ATP mainly *via* substrate phosphorylation, and enzymes of the TCA cycle are expressed at a relatively low level. Excess glucose intermediates do not enter the repressed TCA cycle but are excreted as acetoin, lactate, acetate, or other metabolites, resulting in an acidification of the extracellular medium. After exhaustion of glucose, the glycolytic pathway is repressed because of the need for a high glucose concentration for expression of the gapA operon (Tobisch *et al*., 1999; Luwig and Stülke, 2001). During entry into the stationary phase as triggered by glucose starvation, the visualization of the overall regulation of protein synthesis in growing *B. subtilis* cells was done by 2-D PAGE and found that 150 proteins were synthesized *de novo* and cessation of the synthesis of almost 400 proteins occurred (Bernhardt *et al*., 2003). Proteins induced following glucose starvation belong to two main regulaton groups: general stress/starvation responses induced by different stresses or starvation stimuli (^{B}-dependent general stress regulon, stringent response, sporulation), and glucose-starvation-specific responses (drop in glycolysis, utilization of alternative carbon sources, gluconeogenesis). The combination of transcriptome and proteome analysis using 2-D PAGE and microarray has been used for studying the regulatory genes under glucose repression in *B. subtilis* and found that 2-D PAGE allowed identification of 11 proteins and the synthesis of four (IolA, I, S and PckA) of them was under CcpA-independent control. However, the microarray analysis could detect 66 glucose-repressive genes, 22 of which (glmS, acoA, C, yisS, sped, gapB, pckA, yvdR, yxeF, iolA, B J, R, S and yxbF) were at least partially under CcpA-independent control (Yoshida *et al*., 2001). CcpA and IolR, a repressor of the iol divergon, were involved in the glucose repression of the synthesis of inositol dehydrogenase encoded by iolG included in the above list. Then, the CcpA-independent glucose repression of the iol genes appeared to be explained by inducer exclusion. Furthermore, *B. subtilis* also responds to phosphate starvation by inducing genes of the phosphate stimulon, comprising: (1) the phosphate starvation-specific Phoregulon, (2) the ^{B}-dependent general stress (^{B}-GS) regulon, and (3) PhoP-PhoR/ ^{B}-independent phosphate starvation-inducible genes (Hulett, 1996; Antelmann *et al*., 1997a, 1997b; Gaidenko and Price, 1998; Hecker and Volker, 1998; Pragai *et al*., 2004). The ^{B}-GS regulon has ~200 members, while the Pho regulon presently has 31 members (Petersohn *et al*., 2001; Price *et al*., 2001). During phosphate starvation, genes of the Pho regulon are regulated by the PhoP-PhoR two-component signal transduction system. The PhoP response regulator is activated by its cognate sensor kinase, PhoR. Phosphorylated PhoP (PhoP~P) induces the expression of the phoPR operon about three-fold from a low constitutive level of expression and it is required for the induction or repression of other members of the Pho regulon. In the absence of the regulators of one of these regulons, the expression of the other regulon is activated to a higher level. For maximal induction of the Pho regulon, the respiration signal transduction system, ResD-ResE, is required. If, despite these responses, phosphate starvation persists, a third response regulator, Spo0A, initiates sporulation and terminates the phosphate response by repressing *phoPR* transcription *via* AbrB and ResD-ResE (Fig. **7B**). The phosphate starvation response in *B. subtilis* was also analyzed by 2-D PAGE of cell extracts and supernatants from phosphate-starved cells (Antelmann *et al*., 2000). Most of the phosphate starvation-induced proteins were under the control of ^{B}, the activity of which is increased by energy depletion. By MALDI-TOF MS analysis, two alkaline phosphatases (APases, PhoA and PhoB), and the lipoprotein YdhF were identified as very strongly induced PhoPR-dependent proteins secreted into the extracellular medium. In the cytoplasmic fraction, PstB1, PstB2 and TuaD were identified as already known PhoPR-dependent proteins. For nitrogen starvation, DNA microarrays were used to analyze the changes in gene expression in *B. subtilis* strain 168 under nitrogen limiting (glutamate) and nitrogen excess (ammonium plus glutamate) conditions. It was observed that among more than 100 genes that were significantly induced the comG, comF, come, nin-nucA and ComK transcription units together with recA were detected (Jarmer *et al*., 2002)

3.2.3. Acid and Alkali Stresses

It has been found that the acid stress can damage or disrupt the biological processes thorough the acid-induced death (Bearson *et al*., 1998; Foster and Hall, 1991). Severe acidic pH creates a situation whereby protons leak across the membrane faster than housekeeping pH homeostasis (O'Driscoll *et al*., 1997). Weak acids in their unprotonated form can diffuse into the cell and dissociate, thereby lowering the intracellular pH and resulting in the inhibition of various essential metabolic and anabolic processes. In response to encounters with acids, microorganisms have evolved complex inducible acid survival strategies (Abee and Wouters, 1999). The effect of acid shock in sporulating *B. subtilis* cells on the production of acid shock proteins has been recently studied by proteomic technique using 2-D PAGE and mass spectrometry. It has been shown that acid shock at pH 5 increased the heat resistance of spores subsequently formed in *B. subtilis* wild type (Lee *et al*., 2003). The catalase, KatA, was induced by acid shock early in sporulation. Since it was degraded later in sporulation, it appears to act to increase heat resistance by altering spore structure. For the alkali stress, the physiological prerequisites for adaptation and growth of different bacilli at alkaline pH are under investigation, mainly with the alkaliphilic strains *B. halodurans* C-125 (Hamamoto *et al*., 1994) and *B. pseudofirmus* OF4 (Krulwich, 1998). There is strong evidence that monovalent cation/proton antiporters are essential for maintaining a neutral cytoplasmic pH and, therefore, for growth under alkaline

conditions (Krulwich, 1995). Recently, there have been studies on the response of *B. subtilis* cells towards an alkali shock at the transcriptional level by using DNA microarray technique and the ydjF gene was used as a model gene for an alkali-inducible gene. The results showed that more than 80 genes were induced by a sudden increase in the external pH values from 6.3 to 8.9 (Wiegert *et al*., 2001). It was also discovered that a large subset of these genes belong to the ^{W} regulon, which was confirmed by the analysis of a sigW knockout. A comparison of *B. subtilis* wild type with the congenic sigW knockout also led to the discovery of new member of the ^{W} regulon. These results are also similar to the study of the pst operon of *B. subtilis* induced by alkali stress (Atalla and Schumann, 2003). During a sudden increase in the external pH values to 8.9, *B. subtilis* cells induced about 80 genes, which can be divided into two classes. Most of these genes are members of the ^{W} regulon, while some are under the control of so-far-unknown transcriptional regulators. The genes of the pst operon belong to the second class. Most interestingly, increasing the phosphate concentration within the medium prevented alkali induction of the pst operon, and phoA, another member of the PhoRP regulon, did not respond to alkali stress.

3.2.4. Oxidative Stress

The gram-positive bacterium *B. subtilis* responds to oxidative stress by the activation of different cellular defense mechanisms. These are composed of scavenging enzymes such as catalase (KatA) as well as protection and repair systems organized in highly sophisticated networks. Recently, the peroxide and the superoxide stress stimulons of *B. subtilis* were characterized by means of transcriptomics and proteomics and the results demonstrated that oxidative stress-responsive genes could be classified into two groups (Mostertz *et al*., 2004). One group encompasses genes, which show similar expression patterns in the presence of both reactive oxygen species. Examples are members of the PerR and the Fur regulon which were induced by peroxide and superoxide stress. The second group is composed of genes primarily responding to one stimulus, like the members of the SOS regulon which were particularly upregulated in the presence of peroxide, and many genes involved in sulfate assimilation and methionine biosynthesis, which were only induced by superoxide. Also, studies have been carried out on the global transcriptional profiling to monitor the magnitude and kinetics of change in the mRNA population in *B. subtilis* after exposure to either hydrogen peroxide (H_2O_2) or *tert*-bytyl peroxide (t-buOOH) (Helmann *et al*., 2003). The peroxide stimulons could be largely accounted for by three regulons controlled by the PerR, ^{B}, and OhrR transcription factors. Three members of the PerR regulon (katA, mrgA and zosA) were strongly induced by H_2O_2 and weakly induced by t-buOOH. The remaining members of the PerR regulon genes were only modestly upregulated by peroxide treatment. The ^{B} regulon was activated by H_2O_2 and t-buOOH. Apart from the ^{B} regulon, there was a single gene of ohrA that was strongly and rapidly induced by t-buOOH. Moreover, it has been found that SigM, an extracytoplasmic function sigma factor of *B. subtilis* involved in maintaining membrane and cell wall integrity, was activated not only in response to superoxide stress but also in response to antibiotics, ethanol, heat and acid (Thackray and Moir, 2003).

3.2.5 Other Stresses

Limited information is available about the genes and enzymes involved in sulfur assimilation in *B. subtilis*, or about the regulation of their expression or activity. However, 2-D PAGE was used to compare the proteome of a wild-type strain grown with either sulfate or glutathione as sole sulfur source and 15 proteins were identified whose synthesis was modified under those two conditions (Coppée *et al*., 2001). In the presence of sulfate, an increased amount of proteins involved in the metabolism of C1 units (SerA, GlyA, FolD) and in the biosynthesis of purines (PurQ, Xpt) and pyrimidines (Upp, PyrAA, PyrF) was observed. In the presence of glutathione, the synthesis of two uptake systems (DppE, SsuA), and oxygenase (SsuD), cysteine synthase (CysK) and two proteins of unknown function (Ytml, YurL) were increased. DNA microarray and proteomic techniques revealed 350 genes that were induced during growth arrest by disulfide stress by greater than three-fold after addition of 1 mM diamide as the specific thiol oxidant diamide to an exponentially growing culture of *B. subtilis* (Leichert *et al*., 2003). Strongly induced genes included known oxidative stress genes that are under the control of the global repressor PerR and heat shock genes controlled by the global repressor CtsR. Other genes that were strongly induced encode putative regulators of gene expression and proteins protecting against toxic elements and heavy metals. Many genes were substantially repressed by disulfide stress; among them were most of the genes belonging to the negative stringent response. Therefore, the global response of *B. subtilis* to disulfide stress showed a close relationship to hydrogen peroxide-induced oxidative stress and an even closer relationship to paraquat-induced oxidative stress; there was also a clear overlap with the heat shock stress, the stringent response, and the heavy metal response. For the osmotic stress, the biophysical behavior of the *B. subtilis* membrane under hypertonic conditions containing 1.5 M NaCl was studied. It was observed that the hydrophobicity of the culture medium increases, the phospholipids and fatty acid (FA) compositions show important differences: a higher cardiolipin (CL) content (at the expense of phosphatidyl glycerol) and a higher unsaturated and straight chain fatty acid content (Lopez *et al*., 2000, 2002). The fluidity of the membranes, determined with fluorescent probes, indicates that membrane proteins increase the lipid packing and keep the membrane more sensitive to temperature changes. The proteome analysis of *B. subtilis* cell grown at low and high salinity revealed the induction of 16 protein spots and the repression of 2 protein spots, respectively (Hoffmann *et al*., 2002). Four of the 16 high-salinity-induced proteins corresponded to DhbA, DhbB, DhbC and DhbE, enzymes that are involved in the synthesis of 2,3-dihydroxybenzoate (DHB) and its modification and esterification to the iron siderophore bacillibactin. These proteins are encoded by the *adh*ACEBF operon, which is negatively controlled by the central iron regulatory protein Fur and is derepressed upon iron limitation. The expression of several genes and operons encoding putative iron uptake systems was increased upon salt stress. Moreover, a genome-wide transcriptional profiling analysis of adaptation of *B.*

subtilis to high salinity was carried out to investigate both the initial reaction to a sudden increase in salinity elicited by the addition of 0.4 M NaCl and the cellular adaptation reactions to prolonged growth at high salinity (1.2 M NaCl) (Steil *et al*., 2003). Following salt shock, a sigB mutant displayed immediate and transient induction and repression of 75 and 51 genes, respectively. Continuous propagation of this strain in the presence of 1.2 M NaCl triggered the induction of 123 genes and led to the repression of 101 genes. Moreover, these studies revealed (1) immediate and transient induction of the SigW regulon following salt shock, (2) a role of the DegS/DgU two-component system in sensing high salinity, (3) a high-salinity-mediated iron limitation, and (4) a repression of chemotaxis and motility genes by high salinity, causing servere impairment of the swarming capability of *B. subtilis* cells.

3.3. Protein-Protein Interactions of *Bacillus* sp

Protein-protein interactions are intrinsic to virtually every cellular process. In any important biological process such as DNA replication, transcription, translation, splicing, secretion, cell cycle control, signal transduction and intermediary metabolism, protein complexes have been implicated as essential components (Phizicky and Fields, 1995). Many different classes of proteins are composed of more than one subunit. Some of the best characterized proteins are multisubunit proteins containing two or more different components. Since these proteins purify as multisubunit complexes, their protein-protein interactions were self-evident. Protein-protein interactions can have a number of different effects: (1) they can alter the kinetic properties of proteins, (2) they can have one common mechanism to allow for substrate channeling, (3) they can result in the formation of a new binding site, (4) they can inactivate a protein, and (5) they can change the specificity of an enzyme for its substrate. Methods for detection and analysis of protein-protein interactions have been widely used and can be classified into four different methods: (1) physical methods to select and detect proteins that bind another protein, such as protein affinity chromatography, affinity blotting, immunoprecipitation, cross-linking; (2) library-based methods, such as protein probing, phage display, yeast two-hybrid system; (3) genetic methods, such as extragenic suppressors, synthetic lethal effects, overproduction phenotypes, unlinked noncomplementation; (4) popular methods to estimate and determine binding constants, such as binding to immobilized proteins, sedimentation through gradients, gel filtration chromatography, sedimentation equilibrium.

There are some recent reports describing protein-protein interactions in *Bacillus* sp. For example, the protein-protein interactions were used to pre-detect the intact cell of *B. subtilus* (ATCC 6633) by cyanogen treatment. Cyanogen readily permeates cell walls and membranes and six cross-linked proteins were isolated and identified by MALDI-MS analysis (Winters and Day, 2003). The *B. subtilis* Rap family of proteins were also characterized by protein-protein interaction modules containing the so-called tetratricopeptide repeats (TPRs) and revealed that RapC interaction with ComA inhibits the response regulator's ability to bind its target DNA promoter but does not affect its phosphorylation state (Core and Perego, 2003). Likewise, protein-protein interactions were used to study the *B. subtilis* glutamine synthetase controlling gene expression *via* transcription factor TnrA. It was observed that the feedback-inhibited form of glutamine synthetase directly interacts with TnrA and blocks the DNA binding activity of TnrA (Wray *et al.*, 2001). The protein-protein interactions that regulate the energy stress activation of ^{B} in *B. subtilis* were also determined by surface plasmon resonance. It was observed that the dissociation constants of the RsbW-RsbV and RsbW-^{B} interactions were similar (63 and 92 nM, respectively) (Delumeau *et al*., 2002). Nonetheless, an analysis of the complexes by nondenaturing PAGE in competition assays suggested that the affinity of $RsbW_2$ for RsbV is much higher than that for ^{B}. The result of protein-protein interactions would suggest higher affinity of RsbW for RsbV than for ^{B}, rather than a difference in the concentrations of RsbV and ^{B}, is the driving force that is responsible for the switch of RsbW to unphosphorylated RsbV. Also the yeast two-hybrid system was used to detect the interactions between *B. subtilis* ^{B} regulators relating to seven genes of rsbR, S, T, U, V, W, and X (Voelker *et al*., 1996). The yeast reporter system was activated when RsbS was paired with either RsbR or RsbT, RsbR was paired with RsbT, and RsbV was paired with either RsbU or RsbW. Likewise, using the yeast two-hybrid system and purified FtsZ, and full-length and truncated SpoIIE proteins in *B. subtilis*, it was found that the two proteins interact directly and that domain II and possibly domain I of SpoIIE are required for interaction (Lucet *et al*., 2000). It was also found that SpoIIE interacts with each other and possibly the self-interaction plays a role in assembly of SpoIIE into the division machinery. The protein-protein interactions are not only studied in *B. subtilis*, but also in other species of *Bacillus* for example, the interaction of the *B. stearothermophilus* ribosomal protein S15 with 16S rRNA was studied and it was found that S15 is a primary ribosomal protein that interacts specifically with a three-way junction in the central domain of 16S rRNA and its binding induces a conformational change in the RNA (Batey and Williamson, 1996). The bacterial two-hybrid system (BACTH) was also used to study protein-protein interactions such as the interactions between various subdomians of the dimeric tyrosyl-tRNA synthetase (TyrRS) of *B. stearothermophilu*. The BACTH system could confirm the known interactions of the / domains and those between the / domain and the domain that could be anticipated from the three dimensional structure of TyrRS (Karimova *et al*., 2001). Also, the interactions of the peripheral subunit-binding domain (PSBD) of the dihydrolipoyl acetyltransferase component in the assembly of the pyruvate dehydrogenase multienzyme complex of *B. stearothermophilus* were examined by using directed mutagenesis, surface plasmon resonance detection and isothermal titration microcalorimetry. The enzymes pyruvate decarboxylase (E1) and dihydrolipoyl dehydrogenase (E3) bind tightly but in a mutually exclusive manner to PSBD of dihydrolipoyl acetyltransferase in the pyruvate dehydrogenase multienzyme complex of *B. stearothermophilus*. It was found that several positively charged residues of the PSBD, most notably Arg135, play an important part in the interaction with both E1 and E3, whereas Met131 makes a significant contribution to the binding of E1 only (Jung *et al*., 2003). This indicated that the binding sites for E1 and E3 on the PSBD were overlapping but probably significantly different, and that addi-

tional hydrophobic interactions might be involved in binding E1 compared to E3. Also, protein-protein interactions of *B. amyloloquefaciens* were studied by genetic selection for compensating mutations at the barnase-barstar interface, in which barnase and barstar are trivial names of the extracellular RNase and functions as the intracellular inhibitor (Jucovic and Hartley, 1996).

3.4. Functional Modifications of *Bacillus* sp. Including Post-Translational Modifications

The functional modifications of *Bacillus* sp. have been widely studied, especially by chemical modifications. For example, Matsumoto *et al*., (2002) studied the chemically modified "polar patch" mutants of serine protease subtilisin Bacillus lentus (SBL) for application in peptide synthesis. Using the strategy of combined site-directed mutagenesis and chemical modification to create chemically modified mutant (CMM) enzymes by introduction of polar and/or homochiral auxiliary substituents, the CMMs are capable of catalyzing the coupling reactions of not only L-amino acid esters but also of D-amino acid esters as acyl donors with glycinamide to give the corresponding dipeptides in good yield. Khajeh *et al*., (2001) have studied the chemical modification of lysine residues in two bacterial -amylases from mesophile *B. amyloliquefaciens* (BAA) and thermophile *B. licheniformis* (BLA) using citraconic anhydride and found that the modification brought about dramatic enhancement of thermal stability of BAA and catalytic activity of BLA. Using site-directed mutagenesis combined with chemical modification, Lloyd *et al*., (2000) have studied the site-selective glycosylation of serine protease from *B. subtilisin* and found that the total 48 glycosylated forms of the serine protease subtilisin *Bacillus* lentus (SBL) provided 22 enzymes having increased esterase activities; and all glycosylations at position 217, in the S1' pocket, increased kcat/KM up to 8.4-fold compared to the wild type. In the study of surface-accessible residues in the monomeric and assembled forms of the S-layer protein SbsB of *B. stearothermophilus* PV72/p2 by Howorka *et al*., (2000), 75 selected amino acids (mainly serine, threonine, and alanine), located throughout the primary structure, were replaced by cysteine and 72 out of 75 mutants formed regularly-structured self-assembly products identical to wild type. This suggests that the replacement of most of the selected amino acids by cysteine does not dramatically alter the structure of the protein. However, the three defective mutants, which showed a greatly reduced ability to self-assemble, were successfully incorporated into S layers of wild type proteins. In another study for therapeutic application (Schiavon *et al*., 2000), the therapeutic protein of uricase from *B. fastidiosus* (UC) was chemical modified by covalently linking to linear (poly) ethylene glycol (PEG-1), branched PEG (PEG-2) and poly (N-acryloylmorpholine) (PAcM). Pharmacokinetic investigations in mice demonstrated increased residence time in blood for all the conjugates as compared with native uricase. Among the UC derivatives, the uricase conjugated PAcM was the longest lasting in blood. This indicated the use of the less known amphiphilic polymer PAcM as an alternative to PEGs in modification of enzymes devised for therapeutic applications.

The post-translational modifications (PTMs) are becoming an important research tool, not only for studying modifications of a protein, but also for large-scale systems biology studies, e.g. large-scale protein function prediction and metabolism. The comparative distribution of open reading frame (ORF) function amongst selected free-living organisms by COG database (http://www.brgene.lncc.br/COG_ClassificationComparative.htm) and Neurogadget Bioinformatics (http://www.neurogadgets.com/bioinformatics.php) shows that *B. subtilis* has the functional category of PTMs, protein turnover and chaperones about 98 and 113 numbers of ORFs, respectively. Thus, few PTMs are observed and it is difficult to detect and analyze the modifications, especially for the small peptides. There are few reports on the studies of PTMs of *Bacillus* sp., for example, the modification of fatty acid specificity of cytochrome P450 BM-3 (CYP102) from *B. megaterium* (Lentz *et al*., 2001), the modification of gene encoding dihydrolipoyl acetyltransferase (E2) and dihydrolipoyl dehydrogenase (E3) components of the pyruvate dehydrogenase (PDH) multienzyme complex from *B. stearothermophilus* for improvement of stoichiometry of the subunit interaction in assembly *in vitro* (Lessard *et al*., 1998; Wallis and Perham, 1994), the PTM of *gln*R for studying the defective regulation of *gln*RA transcription in response to nitrogen levels in the growth medium (Schreier and Rostkowski, 1995). The PTM of thermophilic cytochrome c-551 in *B. subtilis* for improvement of expression, signal cleavage and N-terminus function (Kai *et al*., 1997).

4. STRUCTURAL PROTEOMICS OF *BACILLUS* SP.

Structural proteomics plays a crucial role in assigning function to sequenced proteins, defining pathways in which the targets are involved, and understanding structure-function relationships of the protein targets. The aim of this field is to identify the molecular structure, i.e., the amino acid sequence of the protein entities involved in a given process and to relate this information to the database of identified genes. The protein production for structural proteomics also includes selection of target protein, cloning and expression of recombinant proteins, high-density fermentation for the production of the protein, labeling with selenium and/or stable isotopes required for NMR studies, purification, analytical characterization, and the process of organizing these results and reagents into databases and reagent libraries (Lefkovits, 2003; Jung and Lee, 2004). With the advent of proteomics, attempts have been initiated to predict 3-D structure on the basis of primary structure and sequence homology. The most powerful method that has revolutionized proteomics is mass spectrometric analysis, which is generally used for large scale identification of proteins (Godovac-Zimmermann and Brown, 2001; Mann *et al*., 2001). Another powerful method for studying structural proteomics is X-ray crystallization. The availability of three-dimensional structures of the proteins allows for a detailed understanding of their functions as well as their mechanisms of action and helps to understand the general processes in which these proteins are involved. Such detailed structural knowledge about the exact position of specific atoms of these proteins provides the ability to influence or modify the function and mechanism of these proteins and therefore is helpful for developing novel vaccines or antibacterial drugs. In addition to X-ray crystallography, other methods for structural characterization of proteins include nuclear mag-

netic resonance (NMR) spectrometry, electron crystallography and tomography, site-directed mutagenesis, affinity purification with mass spectrometry and fluorescene resonance energy transfer (FRET) (Sali *et al*., 2003; Galvao-Botton *et al*., 2003; Deutzmann, 2004; Ippel *et al*., 2004). Several applications of such structural proteomics techniques for *Bacillus* sp. proteins will be discussed now. The three-dimensional fold of the *B. subtilis* peptide antibiotic subtilosin A was determined by two-dimensional ^{1}H-NMR with spectra recorded at 600 MHz. Based on the backbone conformation, an obtained structure for subtilosin A has been presented which is characterized by three inter-residue bridges where two cysteines are linked with two phenylalanine residues, respectively, and a third cysteine is bound to a threonine residue (Marx *et al*., 2001). Another example of application of structural proteomic techniques relates to the extracellular lipase P1 from thermophilic bacterium *B. stearothermophilus* P1, for which the steps starting from the growth of bacteria, proteomic analysis, molecular cloning and expression, purification, characterization and x-ray crystallographic studies have been summarized in Fig. **8** (Sinchaikul *et al*., 2001a, 2001b, 2002b, 2002c; Tyndal *et al*., 2002). Based on the recent developments of high throughput proteomics analysis using mass spectrometry and two major complementary techniques, NMR spectrometry and X-ray crystallography, are playing a central role in structural proteomic research. However, an integration of these methods with comparative modeling and electron microscopy would speed up the goal for completing a full database for protein folding in the near future.

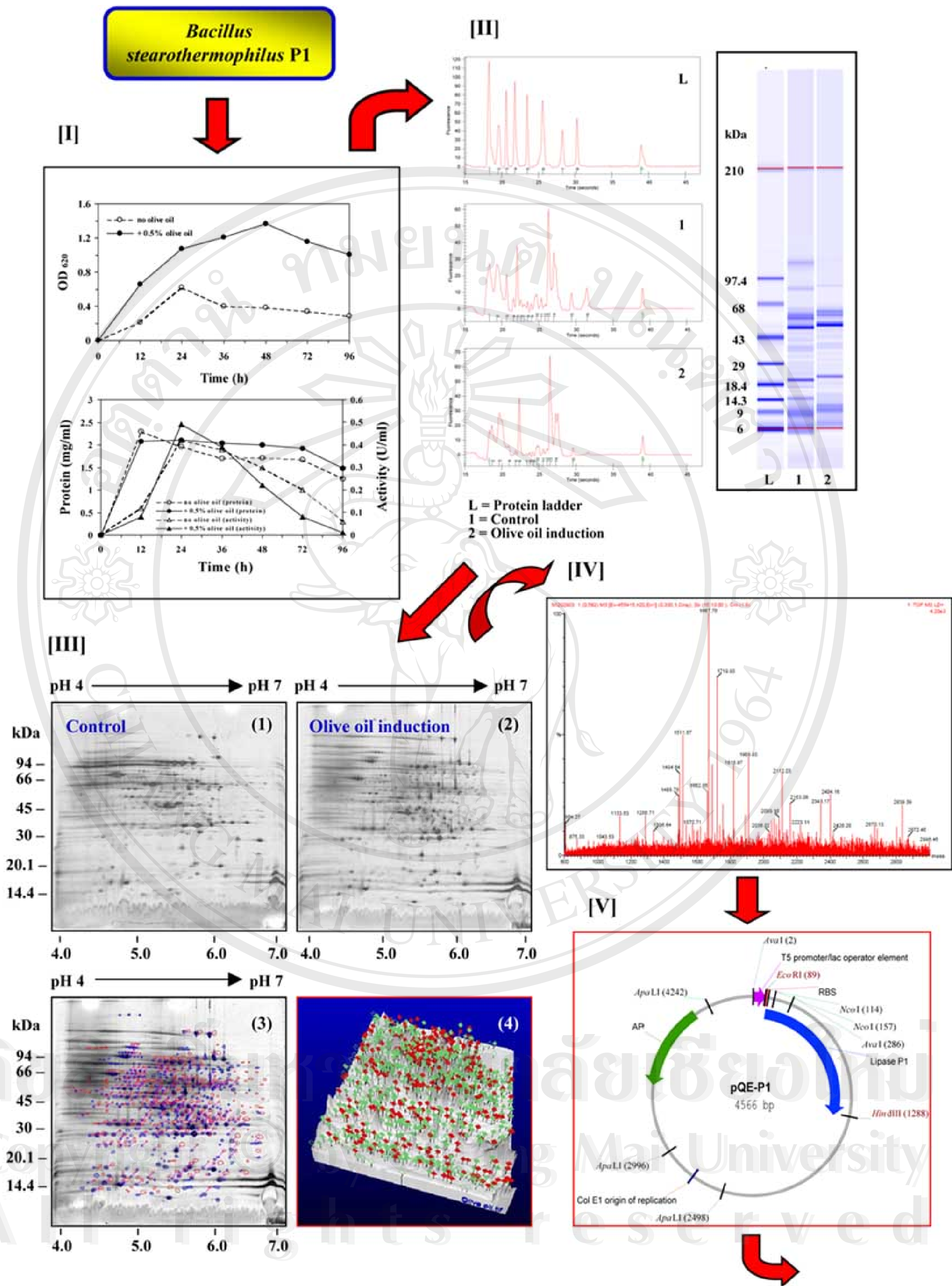
5. BIOINFORMATICS FOR *BACILLUS* PROTEOME

Bioinformatics of *Bacillus* sp. for the characterization of protein encoding genes (open reading frames), expression and function, modifications, interactions and structure, can be done by searching many databases available on the World Wide Web (WWW). The summarized list of the currently available databases for *Bacillus* proteome is shown in Table **4**. In this review, we will describe some useful databases for

Table 4. Bioinformatics of *Bacillus* Proteome Based on Database Searches

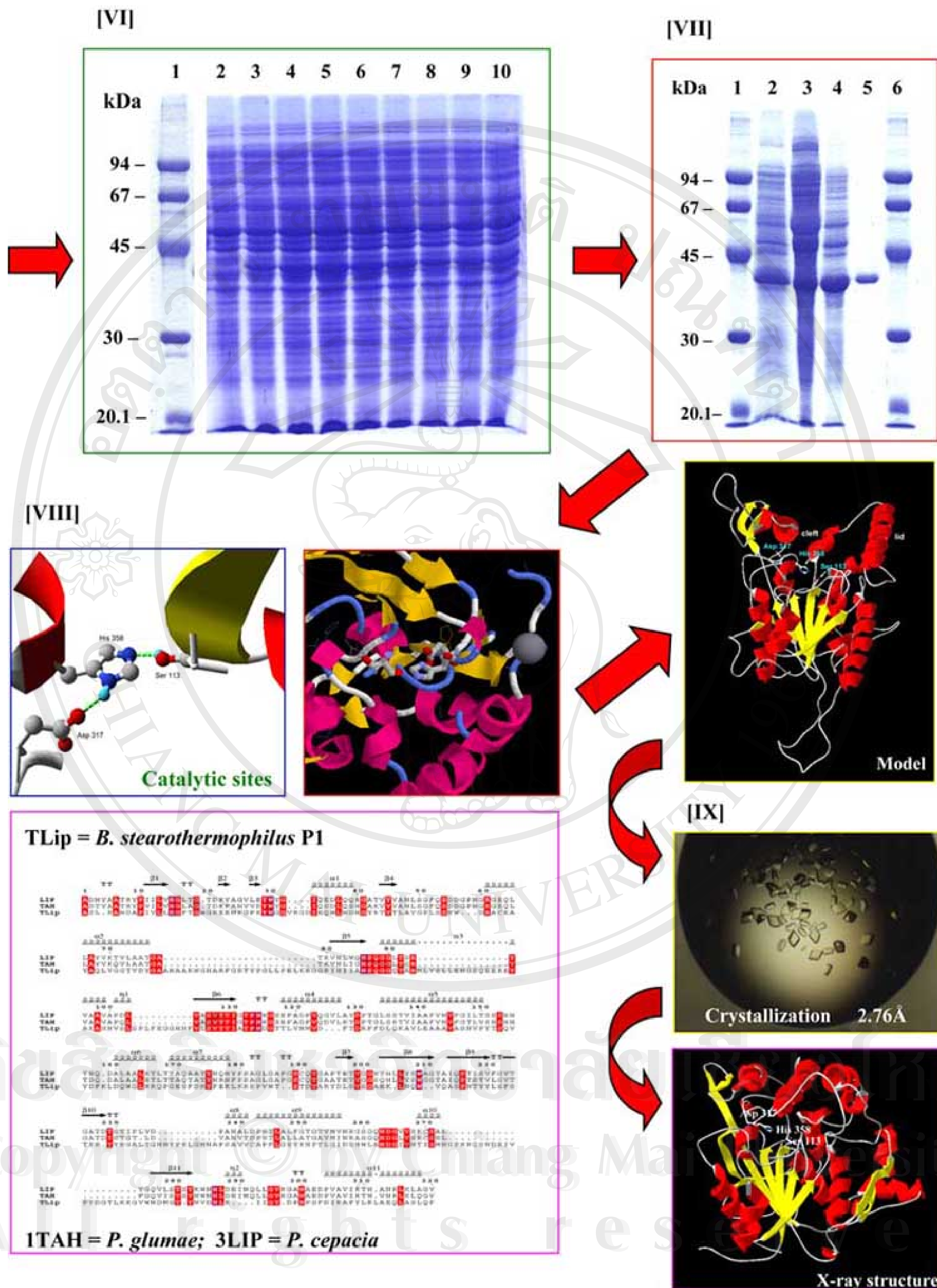
Server	Websites
Protein annotations	
1. IntroPro	http://www.ebi.ac.uk/interpro/
2. iPROCLASS	http://pir.georgetown.edu/iproclass/
3. PIR (Protein Information Resource)	http://pir.georgetown.edu/
4. Swiss-Prot and TrEMBL	http://us.expasy.org/sprot/
5. SubtiList	http://genolist.pasteur.fr/SubtiList/
Proteomics resources	
1. GelBank – 2D gels	http://gelbank.anl.gov
2. Proteome analysis@EBI	http://www.ebi.ac.uk/proteome/
3. Sub2D	http://microbio2.biologie.uni-greifswald.de:8880/
Functional classifications	
1. COG	http://www.ncbi.nlm.nih.gov/COG
2. DBTBS	http://dbtbs.hgc.jp/
3. GeneQuiz	http://jura.ebi.ac.uk:8765/ext-genequiz/genomes/bs0005/
4. Micado	http://genome.jouy.inra.fr/micado/
5. TIGRFAMs	http://www.tigr.org/TIGRFAMs/
Post-translational modifications	
1. DSDBASE	http://www.ncbs.res.in/~faculty/mini/dsdbase/dsdbase.html
2. O-GlycBase	http://www.cbs.dtu.dk/databases/OGLYCBASE
3. GlycoSuiteDB	http://www.glycosuite.com
4. PhosphoBase	http://www.cbs.dtu.dk/databases/PhosphoBase/
Protein-protein interactions, regulation and pathways	
1. BIND	http://bind.ca
2. DIP	http://dip.doe-mbi.ucla.edu/
3. BioCye	http://biocyc.org/
4. DBTBS	http://dbtbs.hgc.jp/
5. KEGG	http://www.genome.ad.jp/kegg/kegg2.html
6. WIT	http://wit.mcs.anl.gov/WIT2/
Structural classifications	
1. 3D-GENOMICS	http://www.sbg.bio.ic.ac.uk/3dgenomics
2. PDB (The Protein Data Bank)	http://www.rcsb.org/pdb/
3. PDBsum	http://www.biochem.ucl.ac.uk/bsm/pdbsum/index.html
4. SCOP	http://scop.berkeley.edu/

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved



Copyright © by Chiang Mai University
All rights reserved

(Fig. 8) contd….



Copyright © by Chiang Mai University
All rights reserved

Fig. (8). Overview of structural proteomics of thermostable lipase in *B. stearothermophilis* strain P1 by various analysis techniques. I, bacterial growth; II, protein chip analyzer, III, 2-D PAGE; IV, MS analysis; V, molecular cloning; VI, expression; VII, Purification; VIII, structure prediction and modeling; IX, Crystallography.

searching information on the *Bacillus* proteome. In general, a collection of popular and very useful databases used for studying the proteome or the genome is available through the ExPASy Molecular Biology Server available at http:// expasy.org/. These allow not only identification of proteins but further characterization ranging from the calculation of basic physiochemical properties to the prediction of potential PTMs and three-dimensional structures. This server can also link to the Swiss-Prot and TrEMBL databases that are very useful for searching the protein annotations such as protein sequence, description of the function of a protein, its domains structure, PTMs, variants, etc (Boechmann *et al*., 2003). Swiss-Prot and TrEMBL databases provide a number of *Bacillus* proteins with 7,973 and 63,647 entries, respectively. SubtiList database is also a popular server that provides a complete dataset of DNA and protein sequences derived from the paradigm of *B. subtilis* 168. It is linked to the relevant annotations and functional assignments, and available at http://genolist.pasteur.fr/SubtiList/ (Moszer *et al*., 1995). Currently this database contains 4,214,630 bp of DNA sequence representing the whole *B. subtilis* chromosome and 4,106 protein genes, which is supplemented with information from the *B. subtilis* entries present in the EMBL/GenBank/DDBJ databanks. Also, this server is used for the identification of the extracellular proteins in *B. stearothermophilus* P1 showing the identified proteins and gene including accession numbers, description and functional category (Table **5**). For the database search as part of proteomic analysis, GelBank database from Argonne National Lab Protein Mapping Group provides an interactive interface for the comparison of 2-DE patterns in the context of proteome sequence queries demonstrating the computerized image of 2-D PAGE with *in silico* 2-DE gels and data analysis to characterize normal protein expression and to detect altered protein expression in biological systems responding to environmental or pathological stresses. It is available on http://gelbank.anl.gov. The *in silico* 2-DE gels of *Bacillus* groups such as *B. subtilis*, *B. halodurans*, *B. cereus* (ATCC14579) and *B. anthracis* (Ames) with 4106, 4066, 5255 and 5311 annotated proteins, respectively show a wide range of pI (1-14) and molecular mass (100-100,000 Da). COG (Clusters of Orthologous Groups of proteins) is a new database search to study the functional classification of proteins, and represents an attempt for the phylogenetic classification of proteins encoded in complete genomes. It currently consists of 3079 *B. subtilis* proteins and 3105 *B. halodurans* proteins. Among those proteins, about 716 and 729 proteins from *B. subtilis* and *B halodurans* respectively are of known functions (http://www.ncbi.nlm.nih.gov/COG) (Tatusov *et al*., 1997, 2001). It also contains proteins of metabolic pathways and major functional systems, each linked to the subset of COGs that contribute to it. The protein sequences encoded in two complete genomes of *B. subtilis* and *B. halodurans*, and partially in genes of *B. stearothermophilus* TLS33 were compared by COG database using COGNITOR program showing the general functional categories (Table **6**). The proteome analysis of *B. subtilis*, *B. halodurans*, *B. cereus* ATCC 14579 and *B. anthracis* (strain Ames) is also shown in the Proteome Analysis@EBI database from website http://www.ebi.ac.uk/proteome. The SWISS-PROT/TrEMBL complete non-redundant proteome sets are constructed by selecting entries from SPTR, which is a comprehensive protein sequence database consisting of SWISS-PROT, TrEMBL and TrEMBLnew (Table **7**) (Bairoch and Apweiler, 2000). To understand protein-protein interactions, BIND (Biomolecular Interaction Network Database) is useful for searching the protein-protein interactions with three classifications for molecular associations: molecules that associate with each other to form interactions, molecular complexes that are formed from one or more interaction(s) and pathways that are defined by a specific sequence of two or more interactions (Bader *et al*., 2003). It is available at http://www.bind.ca. This database provides the protein interaction (2157 interactions) of *Bacillus* sp. and also presents the description, molecular function, cellular component and biological process. To study the structural classification of proteins, the SCOP (Structural Classification of Proteins) database is a comprehensive ordering of all proteins of known structure, according to their evolutionary and structural relationships, and can be accessed at http:// scop.mrc-lmb.cam.ac.uk/scop (Murzin *et al*., 1995; Lo Conte *et al*., 2002). Protein domains in SCOP are grouped into species and hierarchically classified into families, superfamilies, folds, classes and also provides for each entry links to coordinates, images of the structure, interactive viewers, sequence data and literature references. This database includes approximately 401 *Bacillus* protein domains of known structures. Another interesting database to compare structural and functional annotations of proteins between sequenced genomes is 3D-GENOMICS (http://www.sbg.bio.ic.ac.uk/ 3dgenomics), which is achieved through the inclusion of the PDB (Berman *et al*., 2002), SCOP (*via* the ASTRAL database) (Lo Conte *et al*., 2002; Chandonia *et al*., 2002) and Pfam (Bateman *et al*., 2002) databases from which domain information and common ancestry can be inferred (Fleming *et al*., 2004). This database provides 3,331 protein sequences of *B. subtilis* and 2,463 protein sequences of *B. halodurans*, describing a summary of annotations, annotated closest SwissProt/PIR homologue, PFAM short description and SCOP fold/superfamily/ family.

6. PROTEOMIC APPLICATIONS IN BIOMARKERS AND DRUG DISCOVERY FROM *BACILLUS* SP.

The applications of proteomics analysis are coming into widespread use. One of the applications is the investigation and detection of biomarkers in *Bacillus* sp. The members of the genus *Bacillus* are widely distributed in soil, water and air. The resistant property of their spores to environmental stress is of considerable importance in the food processing industry, in the preparation of sterile products, and especially for the prevention of air pollution. The fully automated and analytical methods of curie-point pyrolysis mass spectrometry (PyMS) and diffuse reflectance-absorbance Fourier-transform infrared spectroscopy (FT-IR) were used for rapid detection of the dipicolonic acid (DPA) as biomarker in *Bacillus* spores from thirty six strains of aerobic endospore-forming bacteria to find out whether a *Bacillus* culture is sporulated or not (Goodacre *et al*., 2000). Moreover, MALDI-TOF MS with a nanospray source was used to characterize the primary structures of three hydrolyzed forms of lipopeptide biomarkers in *B. globigii* (*B. subtilis niger*) spores, which often have antimicrobial and antifungal properties (William *et al*., 2002). Likewise, the small acid soluble

Copyright © by Chiang Mai University
All rights reserved

Table 5. Protein and Gene Identification, Description and Functional Category of Extracellular Proteins in *B. Stearothermophilus* P1 in the Stationary Phase (24 h) Using SubtiList Databases, Supplemented with EMBL/GenBank/DDBJ Databanks

Spot no.	Protein ID / no.	Gene name / no.	Description	Functional category
1	YrxA (P39667)	yrxA (BG10865)	Unknown; similar to transcriptional regulator	Transcription regulation
2	Hep1 (P31112)	hepS (BG10279)	Heptaprenyl diphosphate synthase component I	Metabolism of coenzymes and prosthetic groups
3	YabB (P37543)	yabB (BG10098)	Unknown; similar to unknown proteins	From other organisms
4	Tgl (P40746)	tgl (BG10946)	Transglutaminase - cross-links in spore coat proteins	Sporulation
5	MntD (O34500)	mntD (BG13854)	Manganese ABC transporter –manganese uptake	Transport/binding proteins and lipoproteins
6	DtbS (P36840)	bioD (BG11526)	Dethiobiotin synthetase	Metabolism of coenzymes and prosthetic groups
7	ArgB (P36840)	argB (BG10193)	N-acetylglutamate 5-phosphotrans-ferase	Metabolism of amino acids and related molecules
8	RisA (P16440)	ribE (BG10519)	Riboflavin synthase (alpha subunit)	Metabolism of coenzymes and prosthetic groups
9	MobB (O31704)	mob (BG12622)	Molybdopterin-guanine dinucleotide biosynthesis	Metabolism of coenzymes and prosthetic groups
10	PhoP (P13792)	phoP (BG10363)	Two-component response regulator involved in phos-phate regulation	Transcription regulation
11	SCP2 (P81100)	yceC (BG12767)	Unknown; similar to tellurium resistance protein	Detoxification
12	YwfC (P39639)	ywfC (BG10627)	Unknown	No similarity
13	YwiD (P46910)	ywiD (-)	-	-
14	BioW (P53559)	bioW (BG11529)	6-Carboxyhexanoate-CoA ligase	Metabolism of coenzymes and prosthetic groups
15	Hbd (P45856)	mmgB (BG11320)	3-Hydroxybutyryl-CoA dehydrogenase	Metabolism of lipids
16	YbbH (Q45581)	ybbH (BG11569)	Unknown; similar to unknown proteins	From other organisms
17	YybI (P37495)	yybI (BG10022)	Unknown	No similarity
18	YabB (P37543)	yabB (BG10098)	Unknown; similar to unknown proteins	From other organisms
19	YkqA (P39759)	ykqA (BG10990)	Unknown; similar to unknown proteins	From other organisms
20	YlyB (Q45480)	ylyB (BG11796)	Unknown; similar to pseudouridylate synthase	RNA modification
21	CoaA (P54556)	coaA (BG11748)	Probable pantothenate kinase	Metabolism of coenzymes and prosthetic groups
22	DegS (P13799)	degS (BG10392)	Two-component sensor histidine kinase involved in degradative enzyme and competence regulation	Sensors (signal transduction)
23	YqeH (P54453)	yqeH (BG11636)	Unknown; similar to unknown proteins	From other organisms
24	DnaA (P05648)	dnaA (BG10065)	Initiation of chromosome replication	DNA replication
25	AddB (P23477)	addB (BG10465)	ATP-dependent deoxyribonuclease (subunit B): involved in initiation stage of recombination	DNA recombination
26	CypD (O08394)	cypD (BSU07250)	Probable bifunctional P-450:NADPH-P450 reductase 1	Metabolism of lipids
27	YkqC (Q45493)	ykqC (BG11813)	Unknown; similar to unknown proteins	From other organisms
28	DhbE (P40871)	dhbE (BG11020)	2,3-Dihydroxybenzoate-AMP ligase (enterobactin synthetase component E)	Metabolism of coenzymes and prosthetic groups
29	CssS (O32193)	cssS (BG14132)	Two-component sensor histidine kinase involved in the control of cellular responses to protein secretion stress	Sensors (signal transduction)
30	NarH (P42176)	narH (BG11082)	Nitrate reductase (beta subunit)	Membrane bioenergetics (electron transport chain and ATP synthase)
31	ThiC (P45740)	thiC (BG11246)	Biosynthesis of the pyrimidine moiety of thiamin	Metabolism of coenzymes and prosthetic groups

(Table 5) contd….

Copyright © by Chiang Mai University
All rights reserved

(Table 5) contd….

Spot no.	Protein ID / no.	Gene name / no.	Description	Functional category
32	PyrG (P13242)	pyrG (BG10410)	CTP synthetase (requirement for cytidine in the absence of ammonium ion)	Metabolism of nucleotides and nucleic acids
33	GatA (O06491)	gatA (BG12839)	Glutamyl-tRNA(Gln) amidotransferase (subunit A) (formation of correctly charged Gln-tRNA(Gln) through transamidation of misacylated Glu-tRNA(Gln))	Aminoacyl-tRNA synthetases
34	Pur8 (P12047)	purB (BG10702)	Adenylosuccinate lyase	Metabolism of nucleotides and nucleic acids
35	ProB (P39820)	proB (BG10963)	Glutamate 5-kinase (praline biosynthesis)	Metabolism of amino acids and related molecules
36	YqeH (P54453)	yqeH (BG11636)	Unknown; similar to unknown proteins	From other organisms
37	SerC (P80862)	serC (BG12673)	Phosphoserine aminotransferase	Metabolism of amino acids and related molecules
38	YacK (P37573)	yacK (BG10150)	Unknown; similar to unknown proteins	From *B. subtilis*
39	CysK (P37887)	cysK (BG10136)	Cysteine synthetase A	Metabolism of amino acids and related molecules
40	YbbI (Q45582)	ybbI (BG11570)	Unknown; similar to unknown proteins	From other organisms
41	YvcK (O06974)	yvcK (BG12401)	Unknown; similar to unknown proteins	From other organisms
42	YjqC (O34423)	yjqC (BG13220)	Unknown; similar to manganese-containing catalase	Detoxification
43	RapC (P94415)	rapC (BG11966)	Response regulator aspartate phosphatase	Sporulation
44	YacI (P37570)	yacI (BSU00850)	Hypothetical ATP: guanido phosphotransferase yacI	Metabolism of phosphate
45	FtsZ (P17865)	ftsZ (BG10232)	Cell-division initiation protein (septum formation) - required for septum formation during sporulation	Cell division
46	FtsZ (P17865)	ftsZ (BG10232)	Cell-division initiation protein (septum formation) - required for septum formation during sporulation	Cell division
47	YerI (O34640)	yerI (BG12837)	Unknown; similar to unknown proteins	From *B. subtilis*
48	Med (O32436)	med (BG13126)	Positive regulator of comK	Transformation/competence
49	RsbU (P40399)	rsbU (BG11057)	Indirect positive regulator of sigma-B activity (serine phosphatase)	Adaptation to atypical conditions
50	NifU (O32163)	nifU (BSU32680)	NifU-like protein: May be involved in the formation or repair of [Fe-S] clusters present in iron-sulfur proteins (Potential).	Transformation/competence
51	G17m (P80241)	yflT (BG19020)	Unknown	No similarity
52	YqcE (P45940)	yqcE (BG11296)	Unknown; similar to phage-related protein	Phage-related functions
53	YlyA (Q45478)	ylyA (BG11795)	Unknown; similar to unknown proteins from *B. subtilis*	From *B. subtilis*
54	YwqH (P96720)	ywqH (BG12512)	Unknown	No similarity
55	YtzF (O32068)	ytzF (BG13940)	Unknown; similar to 16S pseudouridylate synthase	RNA modification
56	Ak2 (P08495)	lysC (BG10350)	Aspartokinase II (alpha and beta subunits)	Metabolism of amino acids and related molecules
57	Ak2 (P08495)	lysC (BG10350)	Aspartokinase II (alpha and beta subunits)	Metabolism of amino acids and related molecules
58	Ak2 (P08495)	lysC (BG10350)	Aspartokinase II (alpha and beta subunits)	Metabolism of amino acids and related molecules
59	BsaA (P52035)	bsaA (BG11530)	Putative glutathione peroxidase (stress response)	Adaptation to atypical conditions
60	FliH (P23449)	fliH (BG10242)	Flagellar assembly protein	Motility and chemotaxis
61	CssR (O32192)	cssR (BG14131)	Two-component response regulator involved in the control of cellular responses to protein secretion stress	Transcription regulation

(Table 5) contd….

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

(Table 5) contd….

Spot no.	Protein ID / no.	Gene name / no.	Description	Functional category
62	TrpA (P07601)	trpA (BG10291)	Tryptophan synthase (alpha subunit)	Metabolism of amino acids and related molecules
63	YflN (O34409)	yflN (BG12949)	Unknown; similar to unknown proteins	From *B. subtilis*
64	YkkA (P49854)	ykkA (BG11420)	Unknown; similar to unknown proteins	From *B. subtilis*
65	MntR (P54494)	mntR (BG11702)	Transcriptional regulator of manganese uptake (repression of mntH in high Mn(II) conditions, activation of mnt ABCD under low Mn(II) conditions)	Transcription regulation
66	RpoZ (O35011)	rpoZ (BSU15690)	Promotes RNA polymerase assembly	RNA synthesis
67	YqgO (P54494)	yqgO (BG11682)	Unknown	No similarity
68	OpbA (Q45460)	opuBA (BG1633)	Choline ABC transporter (ATP-binding protein) (high affinity transport of choline)	Transport/binding proteins and lipoproteins
69	MnaA (P39131)	mnaA (BSU35660)	Catalyzes the conversion of UDP-N-acetylglucosamine into UDP-N-acetylmannosamine, a precursor of the teichoic acid linkage unit.	Membrane bioenergetics (electron transport chain and ATP synthase)
70	CypC (O31440)	cypC (BG12729)	Fatty acid beta-hydroxylating cytochrome P450	Metabolism of lipids
71	PgdH (P35136)	serA (BG10509)	Phosphoglycerate dehydrogenase	Metabolism of amino acids and related molecules
72	Syc (Q06752)	cysS (BG10156)	Cysteinyl-tRNA synthetase	Aminoacyl-tRNA synthetases
73	AroD (P35146)	aroD (BG11522)	Shikimate 5-dehydrogenase	Metabolism of amino acids and related molecules
74	OdbB (P37941)	BFMBAB (BSU24040)	The branched-chain alpha-keto dehydrogenase complex catalyzes the overall conversion of alpha-keto acids to acyl-CoA and CO.	Metabolism of coenzymes and prosthetic groups
75	KraC (P31749)	kraC (-)	-	-
76	ComB (O06738)	comb (BSU10940)	Probable 2-phosphosulfolactate phosphatase	Metabolism of coenzymes and prosthetic groups
77	Odb2 (P37942)	BFMBB (BSU24030)	The branched-chain alpha-keto dehydrogenase complex catalyzes the overall conversion of alpha-keto acids to acyl-CoA and CO	Metabolism of coenzymes and prosthetic groups
78	YrkQ (P54444)	yrkQ (BG11783)	Unknown; similar to two-component sensor histidine kinase [YrkP]	Sensors (signal transduction)
79	YqcC (P45938)	yqcC (BG11294)	Unknown; similar to phage-related protein	Phage-related functions
80	BioI (P53554)	bioI (BG11528)	Cytochrome P450 enzyme	Metabolism of coenzymes and prosthetic groups
81	TarB (P53554)	tarB (-)	Putative CDP-glycerol: glycerophos-phate glycerophosphotransferase	Membrane bioenergetics (electron transport chain and ATP synthase)
82	PurA (P29726)	purA (BG10002)	Adenylosuccinate synthetase	Metabolism of nucleotides and nucleic acids
83	FliI (P23445)	fliI (BG10243)	Flagellar-specific ATP synthase	Motility and chemotaxis
84	Idi2 (P50740)	idi (BSU22870)	Catalyzes the 1,3-allylic rearrangement of the homoallylic substrate isopentenyl (IPP) to its allylic isomer, dimethylallyl diphosphate (DMAPP) (By similarity).	Metabolism of coenzymes and prosthetic groups
85	AroC (P31104)	aroC (BG10538)	3-Dehydroquinate dehydratase	Metabolism of amino acids and related molecules
86	Odo2 (P16263)	odhB (BG10273)	2-Oxoglutarate dehydrogenase (dihydrolipoamide transsuccinylase, E2 subunit)	TCA cycle
87	Gpr (P22322)	gpr (BG10438)	Spore protease (degradation of SASPs) (initiates degradation of small, acid-soluble spore proteins (SASPs) during the first minutes of germination)	Germination

(Table 5) contd….

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

(Table 5) contd….

Spot no.	Protein ID / no.	Gene name / no.	Description	Functional category
88	BltR (P39842)	bltR (BG10904)	Transcriptional regulator of the bltD operon	Transcription regulation
89	ScrK (O05510)	ydhR (BG12195)	Unknown; similar to fructokinase	Specific pathways
90	DinG (P54397)	dinG (BG11515)	ATP-dependent DNA helicase	DNA restriction/modification and repair
91	KduI (P50843)	kduI (BG11401)	5-Keto-4-deoxyuronate isomerase	Specific pathways
92	YqiK (P54527)	yqiK (BG11719)	Unknown; similar to glycerophospho-diester phosphodiesterase	Metabolism of lipids
93	AroE (P54374)	aroE (BG10294)	5-Enolpyruvoylshikimate-3-phosphate synthase	Metabolism of amino acids and related molecules
94	DhpS (P28822)	sul (BG10140)	Dihydropteroate synthase	Metabolism of coenzymes and prosthetic groups
95	Pgi (P80860)	pgi (BG12366)	Glucose-6-phosphate isomerase	Main glycolytic pathways
96	OxdC (O34714)	oxdC (BG14148)	Oxalate decarboxylase	Specific pathways
97	YqjM (P54550)	yqjM (BG11742)	Unknown; similar to NADH-dependent flavin oxidoreductase	Membrane bioenergetics (electron transport chain and ATP synthase)
98	ArgD (P36839)	argD (BG10194)	N-acetylornithine aminotransferase	Metabolism of amino acids and related molecules
99	Syr (P25499)	argS (BG11341)	Arginyl-tRNA synthetase	Aminoacyl-tRNA synthetases
100	HrcA (P25499)	hrcA (BG10662)	Transcriptional repressor of class I heat-shock genes	Transcription regulation
101	Ffp (Q9F4F7)	ffp (-)	4'-phosphopantetheinyl transferase ffp (Fengycin synthetase-activating enzyme)	Antibiotic production
102	PucG (O32148)	pucG (BSU32520)	Purine catabolism protein pucG (Could encode ureidoglycolase or L-alanine: glyoxylate aminotransferase, or perhaps both activities)	Metabolism of nucleotides and nucleic acids

proteins (SASP) could be potential biomarkers for offline identification of spores of *Bacillus* species by MS. The ensemble of SASP masses revealed that each MALDI spectrum allows genetically distinct species and strains to be identified and readily confirms engineering of specific genes (Hathout *et al*., 2003). Furthermore, it is well known that some antibiotics are produced by *Bacillus* bacteria, such as gramicidin, tyrocidine, bacitracin, mycobacillin, surfactin, bacilysin and subtilin (Mannanov and Sattarova, 2001).

The proteomic technology is also useful for drug discovery; for example, it has been used to elucidate the complex cellular response of *B. subtilis* to antimicrobial compounds belonging to classical and emerging antibiotic classes (Bandow *et al*., 2003). Two-dimensional gels were also used to establish a comprehensive database of cytoplasmic proteins with *p*Is covering a range of 4-7. These proteins were synthesized during treatment with antibiotics or agents known to cause generalized cell damage. The proteomic analysis using 2-D PAGE and MALDI-TOF MS allowed identification of 38 proteins cross-reacting with sera from *B. anthracis* immunized animals. Among them, five immunogens were preselected by the bioinformatics analysis (EA1, Sap, 2 novel SLH proteins and peroxiredoxin/AhpC) as vaccine candidates (Ariel *et al*., 2003). On the other hand, proteomic techniques used to investigate the influence of the two anthracyclin antibiotics daunomycin and adriamycin on the proteome of *B. subtilis* found that both compounds induce proteins related to DNA damage and oxidative stress as indicated by the induction of some members of the PerP and the DinR-regulon (Sender *et al*., 2004).

7. FUTURE OF *BACILLUS* SP. PROTEOME RESEARCH

Since the characterization of the yeast transcriptome (Velculescu *et al*., 1997), the transcriptome has become the new tool for the large scale analysis of biological processes. It is defined as all mRNA species of a cell under defined conditions and its profile strongly varies depending on different environmental conditions. For example, physical stress factors like a rapid increase in environmental temperature would initiate fast and extensive re-organization of gene expression, which would result in another set of specific transcriptome. Consequently, *B. subtilis* transcriptome analyses (transcriptomics) in conjunction with its proteome analysis (proteomics) will allow a more comprehensive global expression profile of the *B. subtilis* genome under defined set of conditions. This would result in comparable induction and repression patterns of almost all corresponding genes. Based on the completed genome sequences of some species of *Bacillus* such as *B. subtilis*, *B. halodurans* and *B. cereus*, the available nucleotide sequences of all the genes has made systematic analysis of the the *Bacillus* proteome

Table 6. Comparison of the Functional Systems of the *Bacillus* sp. in Protein Coding Genes by COG and SubtiList Databases. The Pathway and Functional Systems are Classified from COG Database Using COGNITOR Program

Functional systems	*B. subtilis*	*B. halodurans*	*B. stearothermophilus* TLS33*
Information storage and processing			
1. Translation, ribosomal structure and biogenesis	157	157	3 (*inf*B, *yhb*H, *ybb*B)
2. Transcription	289	281	2 (*nus*G, *rsf*A)
3. DNA replication, recombination and repair	133	238	2 (*top*A, *xtm*A)
Cellular processes			
1. Cell division and chromosome partitioning	34	34	1 (mrp)
2. Posttranslational modification, protein turnover, chaperones	98	99	2 (*msr*A, *yqh*T)
3. Cell envelope biogenesis, outer membrane	178	132	6 (slp, *spo*IIIAE, *pon*A, *tag*E, *yqi*Z, *rsf*A)
4. Cell motility and secretion	79	75	3 (*lep*A, *fli*T, *yvy*F)
5. Inorganic ion transport and metabolism	162	156	2 (*fhu*C, *opu*CA)
6. Signal transduction mechanisms	121	134	-
Metabolism			
1. Energy production and conversion	168	160	-
2. Carbohydrate transport and metabolism	289	281	3 (*suh*B, pgm, pgi)
3. Amino acid transport and metabolism	291	289	6 (*prs*A, *his*H, *cys*H, *lys*C, *lev*E, *ypw*A)
4. Nucleotide transport and metabolism	82	74	2 (*prs*A, *gua*B)
5. Coenzyme metabolism	106	107	5 (*fhu*C, *cys*H, *mob*B, *opu*CA)
6. Lipid metabolism	88	92	-
7. Secondary metabolites biosynthesis, transport and catabolism	88	67	-
Poorly characterized			
1. General function prediction only	348	336	3 (*com*FC, *ybb*B, *rsb*T)
2. Function unknown	307	319	5 (*ypb*G, *yce*B, *yqg*S, *yhb*H, *yla*L)
Not in COGs	1026	961	-

* Some identification

possible. DNA microarrays (DNA chips) have become one of the key high-throughput techniques for characterization of gene expression in the post-genomic era. Recently, 4096 putative open reading frames (ORFs) of *B. subtilis* genome were successfully spotted in duplicate onto aldehyde glass slides to produce *B. subtilis* microarray, which is available commercially (Nippon Gene Co., Ltd). Although transcriptomics can quantify thousands of defined mRNA species, the mRNA and the corresponding protein levels may not always be directly correlated. This is primarily because the rates of degradation of individual mRNAs and corresponding proteins are often different. Also, since proteins can be modified after translation, translation of a specific mRNA molecule can give rise to more than one protein. Nevertheless, transcriptomics and proteomics can complement each other and their combined usage would allow an extensive and global characterization of gene expression patterns in response to varying environmental conditions. The combination of approaches of transcriptome analysis using microarrays and proteome analysis using 2-D PAGE and MS has been increasingly used to study the *B. subtilis* in response to cold shock (Kaan *et al*., 2002), nitrogen and glucose starvations (Jarmar *et al*., 2004; Blencke *et al*., 2003), disulfide stress (Leichert *et al*., 2003), overproduction of an insoluble heterologous protein porA (Jurgen *et al*., 2001), amino acid availability (Mader *et al*., 2002b, in studying the DegS-DegU regulon (Mader *et al*., 2002a) and the correlation between scoC phenotype and gene expression (Caldwell *et al*., 2001), and in comparing protein synthesis patterns of wild-type and relA mutant cells (Eymann *et al*., 2002). Thus, this combination of transcriptome and proteome analysis becomes a new powerful approach to study genes and proteins under experimental conditions, revealing the flow of genetic information from genome to phenotypes. Another novel "ome", the translatome, that describes the members of the proteome weighted by their abundance, plays an important role in genomic and proteomic research. In contrast, the transcriptome refers to the population of mRNA transcripts (Greenbaum *et al*., 2001, 2002). Differences between the translatome and the transcriptome exist, since transcripts from different genes can give rise to different amounts of proteins due to difference in the rates of translation and protein degradation. Posttranslational modifications further affect the translatome. Although the variation between transcriptome and translatome is much smaller for global properties that are computed by averaging over the properties of many individual genes, it is sufficient to look at this smaller number of dominating proteins to characterize the whole population.

Furthermore, through advanced analysis of its genome, transcriptome and proteome, fundamental discoveries dealing with the informational architecture of the *Bacillus* sp. chromosome, as well as with the elucidation of its pathway-level regulation of gene and protein expression, have been

Copyright © by Chiang Mai University
All rights reserved

Table 7. Comparison of Protein Analysis of *B. Subtilis*, *B. Halodurans*, *B. Cereus* ATCC14579 and *B. Anthracis* (Strain Ames) by Proteome Analysis@EBI Database Based on High-Level Terms of the Gene Ontology (GO) Data that has been Assigned to InterPro Entries and Shown the General Statistics for the Number of Proteins in the Proteome

GO Classification	*B. subtilis*	*B. halodurans*	*B. cereus* ATCC14579	*B. anthracis* (Ames)
Number of proteins in proteome	4105	4007	5239	5312
Molecular_function	2515 (61.2%)	2454 (61.2%)	3051 (58.2%)	2554 (48.0%)
Nucleic acid binding	497 (12.1%)	514 (12.8%)	573 (10.9%)	479 (9.0%)
Transcription regulator activity	232 (5.6%)	243 (6.0%)	262 (5.0%)	233 (4.3%)
Chaperone activity	35 (0.8%)	14 (0.3%)	16 (0.3%)	16 (0.3%)
Motor activity	15 (0.3%)	13 (0.3%)	13 (0.2%)	10 (0.1%)
Catalytic activity	1658 (40.3%)	1641 (40.9%)	2118 (40.4%)	1642 (30.9%)
Enzyme regulator activity	9 (0.2%)	2 (0.0%)	3 (0.0%)	-
Structural molecule activity	74 (1.8%)	72 (1.7%)	76 (1.4%)	80 (1.5%)
Transporter activity	446 (10.8%)	408 (10.1%)	477 (9.1%)	519 (9.7%)
Binding	1089 (26.5%)	1094 (27.3%)	1204 (22.9%)	882 (16.6%)
Signal transducer activity	90 (2.1%)	116 (2.8%)	124 (2.3%)	118 (2.2%)
Molecular_function unknown	124 (3.0%)	117 (2.9%)	127 (2.4%)	111 (2.0%)
Biological_process	**2412 (58.7%)**	**2194 (54.7%)**	**2493 (47.5%)**	**2137 (40.2%)**
Metabolism	1746 (42.5%)	1621 (40.4%)	1826 (34.8%)	1548 (29.1%)
Transport	527 (12.8%)	427 (11.7%)	536 (10.2%)	475 (8.9%)
Death	8 (0.1%)	2 (0.0%)	10 (0.1%)	3 (0.0%)
Cell motility	21 (0.5%)	18 (0.4%)	21 (0.4%)	11 (0.2%)
Response to stress	116 (2.8%)	83 (2.0%)	94 (1.7%)	74 (1.3%)
Cell cycle	39 (0.9%)	30 (0.7%)	25 (0.4%)	57 (1.0%)
Cell communication	112 (2.7%)	144 (3.5%)	151 (2.8%)	139 (2.6%)
Development	263 (6.4%)	93 (2.3%)	98 (1.8%)	34 (0.650
Physiological process	2402 (58.5%)	2183 (54.4%)	2485 (47.4%)	2091 (39.3%)
Cellular component	**1351 (32.9%)**	**960 (23.9%)**	**1161 (22.1%)**	**1038 (19.5%)**
Extracellular	13 (0.3%)	15 (0.3%)	24 (0.4%)	19 (0.3%)
Cell	1322 (32.2%)	926 (23.1%)	1112 (21.2%)	982 (18.4%)
External encapsulating structure	94 (2.2%)	52 (1.2%)	73 (1.3%)	26 (0.4%)
Unlocalized protein complex	28 (0.6%)	27 (0.6%)	34 (0.6%)	30 (0.5%)
Cellular component unknown	2 (0.0%)	-	-	-

achieved. The possibility of performing a complete metabolic manipulation of the secretory and signal transduction pathways of *Bacillus* is important biotechnological discovery.

8. SUMMARY

Functional and structural analysis of proteins in *Bacillus* sp. is important for understanding the proteome of this bacterium under defined conditions, which provides significant biological information such as secretory network, signal transduction network, protein interaction and other mechanisms. Proteomic analysis using high throughput technologies is one of the most important post-genomic approaches to understand gene function; however, it also has a limitation in terms of resolution of analysis and sensitivity of detection. Thus, the combination of more than two technology platforms of transcriptome and proteome analysis in conjunction with bioinformatics is becoming increasingly useful for interpretation of comparative genomic, functional genomic and proteomic results from many *Bacillus* species. This provides better understanding of a complementary view into cellular responses to experimental conditions. The impact of the *Bacillus* genome and proteome has philosophically revolutionized the way basic knowledge is translated into applied microbiology and biotechnology, making this bacterium the workhorse of post-genomic microbiology.

ACKNOWLEDGEMENTS

We gratefully thank Academia Sinica and the National Science Council, Taiwan for providing funding for this research. We also thank the core facilities for proteomic research, Academia Sinica, Taiwan for a part of proteomic analysis.

ABBREVIATIONS

2-D PAGE	=	Two-dimensional polyacrylamide gel electrophoresis
BAA	=	*Bacillus amyloliquefaciens*
BACTH	=	Bacterial two hybrid system
BLA	=	*Bacillus licheniformis*
CAP	=	Cold acclimatization protein

Copyright © by Chiang Mai University All rights reserved

CE	=	Capillary electrophoresis
CIP	=	Cold induced protein
CMM	=	Chemically modified mutant
COG	=	Clusters of orthologous groups of proteins
DIGE	=	Difference gel electrophoresis
DPA	=	Dipicolonic acid
ER	=	Endoplasmic reticulum
ESI	=	Electrospray ionization
FAD	=	Flavin adenine dinucleotide
FMN	=	Flavin mononucleotide
FT-IR	=	Fourier-transform infrared spectroscopy
HPLC	=	High performance liquid chromatography
ICAT	=	Isotope coded affinity tag
IMM	=	Indian meal moth
LIF	=	Laser-induced fluorescence
MALDI-TOF	=	Matrix-assisted laser desorption/ionization time-of-flight
MS	=	Mass spectrometry
ORF	=	Open reading frame
PAcM	=	Poly (N-acryloylmorpholine)
PSBD	=	Peripheral subunit-binding domain
PTM	=	Post translational modification
PyMS	=	Pyrolysis mass spectrometry
SASP	=	Small acid soluble proteins
SCOP	=	Structural classification of proteins
TyrRS	=	Tyrosyl-tRNA synthetase
UC	=	Uricase

REFERENCES

Abee, T. and Wouters, J.A. (1999). Microbial stress response in minimal processing. *Int. J. Food Microbiol.* **50**: 65-91.

Akbar, S., Kang, C.M., Gaidenko, T.A. and Price, C.W. (1997). Modulator protein RsbR regulates environmental signaling in the general stress pathway of *Bacillus subtilis*. *Mol. Microbiol.* **24**: 567-78.

Alper, S., Duncan, L. and Losick, R. (1994). An adenosine nucleotide switch controlling the activity of a cell type-specific transcription factor in *B. subtilis*. *Cell* **77**: 195-205.

Antelmann, H., Bernhardt, J., Schmid, R., Mach, H., Volker, U. and Hecker, M. (1997a). First steps from a two-dimensional protein index towards a response-regulation map for *Bacillus subtilis*. *Electrophoresis* **18**: 1451-63.

Antelmann, H., Engelmann, S., Schmid, R., Sorokin, A., Lapidus, A. And Hecker, M. (1997b). Expression of a stress- and starvation-induced *dps*/*pex*B-homologous gene is controlled by the alternative sigma factor ^{B} in *Bacillus subtilis*. *J. Bacteriol.* **179**: 7251-6.

Antelmann, H., Scharf, C. and Hecker, M. (2000). Phosphate starvation-inducible proteins of *Bacillus subtilis*: proteomics and transcriptional analysis. *J. Bacteriol*. **182**: 4478-90.

Antelmann, H., Tjalsma, H., Voigt, B., Ohlmeier, S., Bron, S., van Dijl, J.M. and Hecker, M. (2001). A proteomic view on genome-based signal peptide predictions. *Genome Res*. **9**: 1484-1502.

Antelmann, H., Yamamoto, H., Sekiguchi, J. and Hecker, M. (2002). Stabilization of cell wall proteins in *Bacillus subtilis*: a proteomic approach. *Proteomics* **2**: 591-602.

Ariel, N., Zvi, A., Makarova, K.S., Chitlaru, T., Elhanany, E., Velen, B., Cohen, S., Friedlander, A.M.*, et al.* (2003). Genome-based bioinformatics selection of chromosomal *Bacillus anthracis* putative vaccine candidates coupled with proteomic identification of surface-associated antigens. *Infect. Immun.* **71**: 4563-79.

Atalla, A. and Schumann, W. (2003). The pst operon of *Bacillus subtilis* is specifically induced by alkali stress. *J. Bacteriol.* **185**: 5019-22.

Bader, G.D., Betel, D. and Hogue, C.W. (2003). BIND: the Biomolecular Interaction Network Database. *Nucleic Acids Res.* **31**: 248-50.

Bairoch, A. and Apweiler, R. (2002). The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* **28**: 45-8

Bandow, J.E., Brotz, H., Leichert, L.I.O., Labischinski, H. and Hecker, M. (2003). Proteomic approach to understanding antibiotic action. *Antimicrob. Agents Chemo.* **47**: 948-55.

Bateman, A., Birney, E., Cerruli, L., Durbin, R., Etwiller, L., Eddy, S.R., Griffiths-Jones, S., Howe, K.L., *et al.* (2002). The Pfam protein families database. *Nucleic Acids Res.* **30**: 276-80.

Batey, R.T. and Williamson, J.R. (1996). Interaction of the *Bacillus stearothermophilus* ribosomal protein S15 with 16S rRNA: II. specificity determinants of RNA-protein recognition. *J. Mol. Biol.* **261**: 550-67.

Bearson, B.L., Wilson, L. and Foster, J.W. (1998). A low pH-inducible, PhoPQ-dependent acid tolerance response protects *Salmonella typhimurium* against inorganic acid stress. *J. Bacteriol.* **180**: 2409-17.

Beckering, C.L., Steil, L., Weber, M.H.W., Volker, U. and Marahiel, M.A. (2002). Genome-wide transcriptional analysis of the cold shock response in *Bacillus subtilis*. *J. Bacteriol.* **184**: 6395-402.

Berman, H.M., Battistuz, T., Bhat, T.N., Bluhm, W.F., Bourne, P.E., Burkhardt, K., Feng, Z., Gilliland, G.L., *et al.* (2002). The protein data bank. *Acta Crystallogr. D* **58**: 899-907.

Bernhardt, J., Volker, U., Volker, A., Antelmann, H., Schmid, R., Mach, H. and Hecker, M. (1997). Specific and general stress proteins in *Bacillus subtilis* – a two-dimensional protein electrophoresis study. *Microbiol.* **143**: 999-1017.

Bernhardt, J., Buttner, K., Scharf, C. and Hecker, M. (1999). Dual channel imaging of two-dimensional electropherograms in *Bacillus subtilis*. *Electrophoresis* **20**: 2225-40.

Bernhardt, J., Weibezahn, J., Scharf, C. and Hecker, M. (2003). *Bacillus subtilis* during feast and famine: visualization of the overall regulation of protein synthesis during glucose starvation by proteome analysis. *Genome Res.* **13**: 224-37.

Blackman, S.A., Smith, T.J. and Foster, S.J. (1998). The role of autolysins during vegetative growth of *Bacillus subtilis* 168. *Microbiol.* **144**: 73-82.

Blenckc, H.M., Homuth, G., Luwig, H., Mader, U., Hecker, M. and Stulke, J. (2003). Transcriptional profiling of gene expression in response to glucose in *Bacillus subtilis*: regulation of the central metabolic pathways. *Metab. Eng.* **5**: 133-49.

Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M.-C., Estreicher, A., Gasteiger, E., Martin, M.J., Michoud, K., *et al.* (2003). The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.* **31**: 365-70.

Bolhuis, A., Sorokin, A., Azevedo, V., Ehrlich, S.D., Braun, P.G., de Jong, A., Venema, G., Bron, S*., et al*. (1996). *Bacillus subtilis* can modulate its capacity and specificity for protein secretion through temporally controlled expression of the *sip*S gene for signal peptidase I. *Mol. Microbiol.* **22**: 605-18.

Britz-McKibbin, P., Markuszewski, M.J., Iyanagi, T., Matsuda, K., Nishioka, T. and Terabe, S. (2003). Picomolar analysis of flavins in biological samples by dynamic pH junction-sweeping capillary electrophoresis with laser-induced fluorescence detection. *Anal. Biochem.* **313**: 89-96.

Brody, M. and Price, C.W. (1998). *Bacillus licheniformis sig*B operon encoding the general stress transcription factor ^{B}. *Gene* **212**: 111-8.

Bron, S., Bolhuis, A., Tjalsma, H., Holsappel, S., Venema, G. and van Dijl, J.M. (1998). Protein secretion and possible roles for multiple signal peptidases for precursor processing in *Bacilli*. *J. Biotechnol.* **64**: 3-13.

Bruno, J.G. and Kiel, J.L. (1999). *In vitro* selection of DNA aptamers to anthrax spores with electrochemiluminescence detection. *Biosens. Bioelectron.* **14**: 457-64.

Buttner, K., Bernhardt, J., Scharf, C., Schmis, R., Mader, U., Eymann, C., Antelmann, H., Volker, A*., et al.* (2001). A comprehensive two-dimensional map of cytosolic proteins of *Bacillus subtilis*. *Electrophoresis* **22**: 2908-35.

Caldwell, R., Sapolsky, R., Weyler, W., Maile, R.R., Causey, S.C. and Ferrari, E. (2001). Correlation between *Bacillus subtilis* scoC phenotype

and gene expression determined using microarrays for transcriptome analysis. *J. Bacteriol.* **183**: 7329-40.

Candas, M., Loseva, O., Oppert, B., Kosaraju, P. and Bulla, L.A. (2003). Insect resistance to *Bacillus thuringiensis*. *Mol. Cell. Proteomics* **2**: 19-28.

Chalmers, M.J. and Gaskell, S.J. (2000). Advances in mass spectrometry for proteome analysis. *Curr. Opin. Biotechnol.* **11**: 384-90.

Chandonia, J.M., Walker, N.S., Lo Conte, L., Koehl, P., Levitt, M. and Brenner, S.E. (2002). ASTRAL compendium enhancements. *Nucleic Acids Res.* **30**: 260-3.

Choudhary, J.S., Blackstock, W.P., Creasy, D.M. and Cottrell, J.S. (2001). Matching peptide mass spectra to EST and genomic DNA databases. *Trends Biotechnol.* **19**: S17-22.

Coppée, J.Y., Auger, S., Turlin, E., Sekowska, A., Caer, J.P.L., Labas, V., Vagner, V., Danchin, A. and Martin-Verstraete, I. (2001). Sulfur-limitation-regulated proteins in *Bacillus subtilis*: a two-dimensional gel electrophoresis study. *Microbiol.* **147**: 1631-40.

Core, L. and Perego, M. (2003). TPR-mediated interaction of RapC with ComA inhibites response regulator-DNA binding for competence development in *Bacillus subtilis*. *Mol. Microbiol.* **49**: 1509-22.

de Wildt, R.M., Mundy, C.R., Gorick, B.D. and Tomlinson, I.M. (2000). Antibody arrays for high-throughput screening of antibody-antigen interactions. *Nat. Biotechnol.* **18**: 989-94.

Delehanty, J.B. and Ligler, F.S. (2002). A microarray immunoassay for simultaneous detection of proteins and bacteria. *Anal. Chem.* **74**: 5681-7.

Delumeau, O., Lewis, R.J. and Yudkin, M.D. (2002). Protein-protein interactions that regulate the energy stress activation of ^{B} in *Bacillus subtilis*. *J. Bacteriol.* **184**: 5583-9.

Deutzmann, R. (2004). Structural characterization of proteins and peptides. *Methods Mol. Med.* **94**: 269-97.

Driessen, A.J.M. (1994). How proteins cross the bacterial cytoplasmic membrane. *J. Membr. Biol.* **142**: 145-59.

Elhanany, E., Barak, R., Fisher, M., Kobiler, D. And Altboum, Z. (2001). Detection of specific *Bacillus anthracis* spore biomarkers by matrix-assited laser desorption/ionization time-of-flight mass spectrometry. *Rapid Comm. Mass Spectrom.* **15**: 2110-16.

English, R.D., Warscheid, B., Fenselau, C. and Cotter, R.J. (2003). *Bacillus* spore identification *via* proteolytic peptide mapping with a miniaturized MALDI-TOF mass spectrometer. *Anal. Chem.* **75**: 6886-93.

Eymann, S., Homuth, G., Scharf, C. and Hecker, M. (2002). *Bacillus subtilis* functional genomics: global characterization of the stringent response by proteome and transcriptome analysis. *J. Bacteriol.* **184**: 2500-20.

Fleming, K., Müller, A., MacCallum, R.M. and Sternberg, M.J.E. (2004). 3D-GENOMICS: a database to compare structural and functional annotations of proteins between sequenced genomes. *Nucl. Acids Res.* **32**: D245-50.

Foster, J.W. and Hall, H.K. (1991). Inducible pH homeostasis and the acid tolerance response of *Salmonella typhimurium*. *J. Bacteriol.* **173**: 5129-35.

Francis, K.P., Mayr, R., von Stetten, F., Stewart, G.S. and Scherer, S. (1998). Discrimination of psychrotrophic and mesophilic strains of the *Bacillus cereus* group by PCR targeting of major cold shock protein genes. *Appl. Environ. Microbiol.* **64**:3525-9.

Gaidenko, T.A. and Price, C.W. (1998). General stress transcriptional factor ^{B} and sporulation sigma factor ^{H} each contribute to survival of *Bacillus subtilis* under extreme conditions. *J. Bacteriol.* **180**: 3730-3.

Galvao-Botton, L.M., Katsuyama, A.M., Guzzo, C.R., Almeida, F.C., Farah, C.S. and Valente, A.P. (2003). High-throughput screening of structural proteomics targets using NMR. *FEBS Lett.* **552**: 207-13.

Gevaert, K. and Vanderkerckhove, J. (2000). Protein identification methods in proteomics. *Electrophoresis* **21**: 1145-54.

Gilmour, R., Messner, P., Guffanti, A.A., Kent, R., Scheberl, A., Kendrick, N. and Krulwich, T.A. (2000). Two-dimensional gel electrophoresis analyses of pH-dependent protein expression in facultatively alkaliphilic *Bacillus pseudofirmus* OF4 lead to characterization of an S-laayer protein with a role in alkaliphily. *J. Bacteriol.* **182**: 5969-81.

Godovac-Zimmermann, J. and Brown, L.R. (2001). Perspectives for mass spectrometry and functional proteomics. *Mass Spectrom. Rev.* **20**: 1-57.

Gohar, M., Okstad, O.A., Gilois, N., Sanchis, V., Kolsto, A.B. and Lereclus, D. (2002). Two-dimensional electrophoresis analysis of the extracellular proteome of *Bacillus cereus* reveals the importance of the PlcR regulon. *Proteomics* **2**: 784-91.

Goodacre, R., Shann, B., Gilbert, R.J., Timmins, E.M., McGovern, A.C., Alsberg, B.K., Kell, D.B. and Logan, N.A. (2000). Detection of the dipicolinic acid biomarker in *Bacillus* spores using curie-point pyrolysis mass spectrometry and fourier transform infrared spectroscopy. *Anal. Chem.* **72**: 119-27.

Graumann, P. and Marahiel, M.A. (1996). Some like it cold: response of microorganisms to cold shock. *Arch. Microbiol.* **166**: 293-300.

Graumann, P.L. and Marahiel, M.A. (1999). Cold shock response in *Bacillus subtilis*. *J. Mol. Microbiol. Biotechnol.* **1**: 203-9.

Graumann, P., Wendrich, T.M., Weber, M.H.W., Schroder, K. and Marahiel, M.A. (1997). A family of cold shock proteins in *Bacillus subtilis* is essential for cellular growth and for efficient protein synthesis at optimal and low temperatures. *Mol. Microbiol.* **25**: 741-56.

Graumann, P., Schroder, K., Schmid, R. and Marahiel, M.A. (1996). Cold shock stress-induced proteins in *Bacillus subtilis*. *J. Bacteriol.* **178**: 4611-9.

Greenbaum, D., Jansen, R., Gerstein, M. (2002). Analysis of mRNA expression and protein abundance data: an approach for the comparison of the enrichment of features in the cellular population of proteins and transcripts. *Bioinformatics* **18**: 585-96.

Greenbaum, D., Luscombe, N.M., Jansen, R., Qian, J. and Gerstein, M. (2001). Interrelating different types of genomic data, from proteome to secretome: 'oming in on function. *Genome Res.* **11**: 1463-8.

Grow, A.E., Wood, L.L., Claycomb, J.L. and Thompson, P.A. (2003). New biochip technology for label-free detection of pathogens and their toxins. *J. Microbiol. Meth.* **53**: 221-33.

Gygi, S.P., Rist, B., Gerber, S.A., Turecek, F., Gelb, M.H. and Aebersold, R. (1999). Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17**: 994-9.

Haab, B.B., Dunham, M.J. and Brown, P.O. (2001). Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. *Genome Biol.* **2**: 0004.1-0004.13.

Haldenwang, W.G., Losick, R. (1979). A modified RNA polymerase transcribes a cloned gene under sporulation control in *Bacillus subtilis*. *Nature* **282**: 256-60.

Hamamoto, T., Hashimoto, M., Hino, M., Kitada, M., Seto, Y., Kudo, T. and Horikoshi, K. (1994). Characterization of a gene responsible for the Na^+/H^+ antiporter system of alkalophilic *Bacillus* species strain C-125. *Mol. Microbiol.* **14**: 939-46.

Hathout, Y., Demirev, P.A., Ho, Y.P., Bundy, J.L., Ryzhov, V., Sapp, L., Stutler, J., Jackman, J., *et al.* (1999). Identification of *Bacillu*s spores by matrix-assisted laser desorption ionization-mass spectrometry. *Appl. Envir. Microbiol.* **65**: 4313-9.

Hathout, Y., Setlow, B., Cabrera-Martinez, R.M., Fenselau, C. and Setlow, P. (2003). Small, acid-soluble proteins as biomarkers in mass spectrometry analysis of *Bacillus* spores. *Appl. Envir. Microbiol.* **69**: 1100-7.

Hecker, M. (2003). A proteomic view of cell physiology of *Bacillus subtilis*--bringing the genome sequence to life. *Adv. Biochem. Eng. Biotechnol.* **83**: 57-92.

Hecker, M., Schumann, W. and Volker, U. (1996). Heat-shock and general stress response in *Bacillus subtilis*. *Mol. Microbiol.* **19**: 417-28.

Hecker, M. and Volker, U. (1998). Non-specific, general and multiple stress resistance of growth-restricted *Bacillus subtilis* cells by the expression of the ^{B} regulon. *Mol. Microbiol.* **29**: 1129-36.

Hecker, M. and Volker, U. (2001). General stress response of *Bacillus subtilis* and other bacteria. *Adv. Microb. Physiol.* **44**: 35-91.

Helmann, J.D., Wu, M.F.W., Gaballa, A., Kobel, P.A., Morshedi, M.M., Fawcett, P. and Paddon, C. (2003). The global trnsciptional response of *Bacillus subtilis* to peroxide stress is coordinated by three transcription factors. *J. Bacteriol.* **185**: 243-53.

Hirose, I., Sano, K., Shioda, I., Kumano, M., Nakamura, K. and Yamane, K. (2000). Proteome analysis of *Bacillus subtilis* extracellular proteins: a two-dimensional protein electrophoretic study. *Microbiol.* **146**: 65-75.

Hoffmann, T., Schütz, A., Brosius, M., Volker, A., Volker, U. and Bremer, E. (2002). High-salinity-induced iron limitation in *Bacillus subtilis*. *J. Bacteriol.* **184**: 718-27.

Howorka, S., Sara, M., Wang, Y., Kuen, B., Sleytr, U.B., Lubitz, W. and Bayley, H. (2000). Surface-accessible residues in the monomeric and assembled forms of a bacterial surface layer protein. *J. Biol. Chem.* **275**: 37876-86.

Hountondji, C., Beauvallet, C., Pernollet, J.C. and Blanquet, S. (2000). Enzyme-induced covalent modification of methionyl-tRNA synthetase from *Bacillus stearothermophilus* by methionyl-adenylate: identification of the labeled amino acid residues by matrix-assisted laser desorption-ionization mass spectrometry. *J. Protein Chem.* **19**: 563-8.

Huang, R.P., Huang, R., Fan, Y. and Lin, Y. (2001). Simultaneous detection of multiple cytokines from conditioned media and patient's sera by an antibody-based protein array system. *Anal. Biochem.* **294**: 55-62.

Hulett, F.M. (1996). The signal-transduction network for Pho regulation in *Bacillus sutilis*. *Mol. Microbiol.* **19**: 933-9.

Ippel, J.H., Pouvreau, L., Kroef, T., Gruppen, H., Versteeg, G., can den Putten, P., Struik, P.C. and van Mierlo, C.P. (2004). *In vivo* uniform (15).N-isotope labeling of plants: using the greenhouse for structural proteomics. *Proteomics* **4**: 226-34.

Ivanova, N., Sorokin, A., Anderson, I., Galleron, N., Candelon, B., Kapatral, V., Bhattacharyya, A., Reznik, G*., et al.* (2003). Genome sequenceof Bacillus cereus and comparative analysis with *Bacillus anthracis*. *Nature* **423**: 87-91.

Jarmer, H., Berka, R., Knudsen, S. and Saxild, H.H. (2002).. Transcriptome analysis documents induced competence of *Bacillus subtilis* during nitrogen limiting conditions. *FEMS Microbiol. Lett.* **206**: 197-200.

Jongbloed, J.D.H., Antelmann, H., Hecker, M., Nijland, R., Bron, S., Airaksinen, U., Prices, F., Quax, W.J*., et al.* (2002). Selective contribution of the Twin-arginine translocation pathway to protein secretion in *Bacillus subtilis*. *J. Biol. Chem.* **277**: 44068-78.

Jucovic, M. and Hartley, R.W. (1996). Protein-protein interaction: a genetic selection for compensating mutations at the barnase-barstar interface. *Proc. Natl. Acad. Sci. USA* **93**: 2343-7.

Jung, H.I., Cooper, A. and Perham, R.N. (2003). Interactions of the peripheral subunit-binding domain of the dihydrolipoyl acetyltransferase component in the assembly of the pyruvate dehydrogenase multienzyme complex of *Bacillus stearothermophilus*. *Eur. J. Biochem.* **270**: 4488-96.

Jung, J.W. and Lee, W. (2004). Structure-based functional discovery of proteins: structural proteomics. *J. Biochem. Mol. Biol.* **37**: 28-34.

Jungblut, P. and Thiede, B. (1997). Protein identification from 2-DE gels by MALDI mass spectrometry. *Mass Sepc. Rev.* **16**: 145-62.

Jurgen, B., Hanschke, R., Sarvas, M., Hecker, M. and Schweder, T. (2001). Proteome and transcriptome based analysis of *Bacillus subtilis* cells overproducing an insoluble heterologous protein. *Appl. Microbiol. Biotechnol.* **55**: 326-32.

Kaan, T., Homuth, G., Mader, U., Bandow, J. and Schweder, T. (2002). Genome-wide transcriptional profiling of the *Bacillus sutilis* cold-shock response. *Microbiol.* **148**: 3441-55.

Kai, K., Noguchi, S. and Sone, N. (1997). Over-expression and post-translational modification of thermophilic *Bacillus cytochrome* c-551 in *Bacillus subtilis*. *J. Ferment. Bioeng.* **84**: 190-4.

Kang, C.M., Brody, M.S., Akbar, S., Yang, X. and Price, C.W. (1996). Homologous pairs of regulatory proteins control activity of *Bacillus subtili*s transcription factor sigma (b) in response to environmental stress. *J. Bacteriol.* **178**: 3846-53.

Karimova, G., Ullmann, A. and Ladant, D. (2001). Protein-protein interaction between *Bacillus stearothermophilis* tyrosyl-tRNA synthtase subdomains revealed by a bacterial two-hybrid system. *J. Mol. Micro. Biotechnol.* **3**: 73-82.

Khajeh, K., Naderi-Manesh, H., Ranjbar, B., Moosavi-Movahedi, A.A. and Nemat-Gorgani, M. (2001). Chemical modification of lysine residues in *Bacillus* -amylases: effect on activity and stability. *Enzyme Micro. Technol.* **28**: 543-9.

Kontinen, V.P. and Sarvas, M. (1993). The PrsA lipoprotein is essential for secretion in *Bacillus subtili*s and sets a limit for high-level secretion. *Mol. Microbiol.* **8**: 727-37.

Krulwich, T.A. (1995). Alkaliphiles: 'basic' molecular problems of pH tolerance and bioenergetics. *Mol. Microbiol.* **15**: 403-10.

Krulwich, T.A. (1998). Energies of alkaliphilic *Bacillus* species: physiology and molecules. *Adv. Mcrob. Physiol.* **40**: 401-438.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessieres, P., *et al.* (1997). The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* **390**: 249-56.

Lai, E.M., Phadke, N.D., Kachman, M.T., Giorno, R., Vazquez, S., Vazquez, J.A., Maddock, J.R. and Driks, A. (2003). Proteomic analysis of the spore coats of *Bacillus subtilis* and *Bacillus anthracis*. *J. Bacteriol.* **185**: 1443-54.

Lee, J.K., Movahedi, S., Harding, S.E. and Waites, W.M. (2003). The effect of acid shock on sporulating *Bacillus subtilis* cells. *J. Appl. Microbiol.* **94**: 184-90.

Lefkovits, I. (2003). Functional and structural proteomics: a critical appraisal. *J. Chromatogr. B* **787**: 1-10.

Leichert, L.I.O., Scharf, C. and Hecker, M. (2003). Global characterization of disulfide stress in *Bacillus subtilis*. *J. Bacteriol.* **185**: 1967-75.

Lentz, O., Li, Q.-S., Schwaneberg, U., Lutz-Wahl, S., Fischer, P. and Schmid, R. D. (2001). Modification of the fatty acid specificity of cytochrome P450 BM-3 from *Bacillus megaterium* by directed evolution: a validated assay. *J. Mol. Catal. B-Enzym.* **629**: 1-11.

Leskela, S., Wahlstrom, E., Kontinen, V.P. and Sarvas, M. (1999). Lipid modification of prelipoproteins is dispensable for growth but essential for efficient protein secretion in *Bacillus subtilis*: characterization of the Lgt gene. *Mol. Microbiol.* **31**: 1075-85.

Lessard, I.A., Domingo, G.J., Borges, A. and Perham, R.N. (1998). Expression of genes encoding the E2 and E3 components of the *Bacillus stearothermophilus* pyruvate dehydrogenase complex and the stoichiometry of subunit interaction in assembly *in vitro*. *Eur. J. Biochem.* **258**: 491-501.

Link, A.J., Eng, J., Schieltz, D.M., Carmack, E., Mize, G.J., Morris, D.R., Garvik, B.M. and Yates, J.R. (1999). Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* **17**: 676-82.

Liu, C.M and Tzeng Y.M. (2001). Quantitative analysis of parasporal crystal proten from *Bacillus thuringiensis* by capillary electrophoresis. *J. Food Drug Anal.* **9**: 79-83.

Lloyd, R.C., Davis, B.G. and Jones, J.B. (2000). Site-selective glycosylation of subtilisin Bacillus lentus causes dramatic increases in esterase activity. *Bioorg. Med. Chem.* **8**: 1537-44.

Lo Conte, L., Brenner, S.E., Hubbard, T.J., Chothia, C. and Murzin, A.G. (2002). SCOP database in 2002: refinements accommodate structural genomics. *Nucleic Acids Res.* **30**: 264-7.

Lopez, C.S., Heras, H., Garda, H., Ruzal, S., Sanchez-Rivas, C. and Rivas, E. (2000). Biochemical and biophysical studies of *Bacillus subtilis* envelopes under hyperosmotic stress. *Int. J. Food Microbiol.* **55**: 137-42.

Lopez, C.S., Garda, H.A. and Rivas, E.A. (2002). The effect of osmotic stress on the biophysical behavior of the *Bacillus subtilis* membrane studied by dynamic and steady-state fluorescence anisotrophy. *Ach. Biochem. Biophys.* **408**: 220-8.

Lucet, I., Feucht, A., Yudkin, M.D. and Errington, J. (2000). Direct interaction between the cell division protein FtsZ and the cell differentiation protein SpoIIE. *EMBO J.* **19**: 1467-75.

Luwig, H. and Stülke, J. (2001). The *Bacillus subtilis* catabolite control protein CcpA exerts all its regulatory functions by DNA-binding. *FEMS Microbiol. Lett.* **203**: 125-9.

Madec, E., Stensballe, A., Kjellstrom, S., Cladiere, L., Obuchowski, M., Jensen, O.N. and Seror, S.J. (2003). Mass spectrometry and site-directed mutagenesis identify several autophosphorylated residues required for the activity of PrkC, a Ser/Thr kinase from *Bacillus subtilis*. *J. Mol. Biol.* **330**: 459-72.

Mader, U., Antelmann, H., Buder, T., Dahl, M.K., Hecker, M. and Homuth, G. (2002a). *Bacillus subtilis* functional genomics: genome-wide analysis of the DegS-DegU regulon by transcriptomics and proteomics. *Mol. Genet. Genomics* **268**: 455-67.

Mader, U., Homuth, G., Scharf, C., Buttner, K., Bode, R. and Hecker, M. (2002b). Transcriptome and proteome analysis of *Bacillus subtilis* gene expression modulated by amino acid availability. *J. Bacteriol.* **184**: 4288-95.

Mann, M., Hendrickson, R.C. and Pandey, A. (2001). Analysis of proteins and proteomes by mass spectrometry. *Annu. Rev. Biochem.* **70**: 437-73.

Mannanov, R.N. and Sattarova, R.K. (2001). Antibiotics produced by *Bacillus* bacteria. *Chem. Nat. Compounds* **37**: 117-123.

Markuszewski, M.J., Britz-McKibbin, P., Terabe, S., Matsuda, K. and Nishioka, T. (2003). Determination of pyridine and adenine nucleotide metabolites in *Bacillus subtilis* cell extract by sweeping borate complexation capillary electrophoresis. *J. Chromatogr. A* **989**: 293-301.

Marx, R., Stein, T., Entian, K.D. and Glaser, S.J. (2001). Structure of the *Bacillus subtilis* peptide antibiotic subtilosin A determined by ^{1}H-NMR and matrix assisted laser desorption/ionization time-of-flight mass spectrometry. *J. Prot. Chem.* **20**: 501-6.

Matsumoto, K., Davis, B.G. and Jones J.B. (2002). Chemically modified "polar patch" of subtilisin in peptide synthesis with remarkably broad substrate acceptance: designing combinatorial biocatalysts. *Chemistry* **8**: 4129-37.

Moody, M.D., Van Arsdell, S.W., Murphy, K.P., Orencole, S.F. and Burns, C. (2001). Array-based ELISAs for high-throughput analysis of human cytokines. *Biotechniques* **31**: 186-94.

Mortz, E., Vorm, O., Mann, M. and Roepstorff, P. (1994). Identification of proteins in polyacrylamide gels by mass spectrometric peptide mapping combined with database search. *Biol. Mass Spectrom.* **23**: 249-61.

Mostertz, J., Scharf, C., Hecker, M. and Homuth, G. (2004). Transcriptome and proteome analysis of *Bacillus subtilis* gene expression in response to superoxide and peroxide stress. *Microbiol.* **150**: 497-512.

Moszer, I., Glaser, P. and Danchin, A. (1995). SubtiList: a relational database for the *Bacillus subtilis* genome. *Microbiology* **141**: 261-8.

Muller, J.P. and Wagner, M. (1999). Localisation of the cell wall-associated phosphodiesterase PhoD of *Bacillus subtilis*. *FEMS Microbiol Lett.* **180**: 287-96.

Murzin, A.G., Brenner, S.E., Hubbard, T. and Chothia, C. (1995). SCOP: a structural classification of protein database for the investigation of sequences and structures. *J. Mol. Biol.* **247**: 536-40.

Nakamura, K., Nishiguchi, M., Honda, K. and Yamane, K. (1994). The *Bacillus subtilis* SRP54 homologue, Ffh, has an intrinsic GTPase activity and forms a ribonucleoprotein complex with small cytoplasmic RNA *in vivo*. *Biochem. Biophys. Res. Commun.* **199**: 1394-9.

O'Driscoll, B., Gahan, C.G.M. and Hill, C. (1997). Two-dimensional polyacrylamide gel electrophoresis analysis of the acid tolerance response in *Listeria monocytogenes* LO28. *Appl. Environ. Microbiol.* **63**: 2679-85.

Ohlmeier, S., Scharf, C. and Hecker, M. (2000). Alkaline proteins of *Bacillus subtilis*: first steps towards a two-dimensional alkaline master gel. *Electrophoresis* **21**: 3701-9.

Oosthuizen, M.C., Steyn, B., Lindsay, D., Brozel, V.S. and von Holy, A. (2001). Novel method for the proteomic investigation of a dairy-associated *Bacillus cereus* biofilm. *FEMS Microbiol. Lett.* **194**: 47-51.

Oosthuizen, M.C., Steyn, B., Theron, J., Cosette, P., Lindsay, von Holy, A. and Brozel, V.S. (2002). Proteomic analysis reveals differential protein expression by *Bacillus cereus* biofilm formation. *Appl. Environ. Microbiol.* **68**: 2770-80.

Periago, P.M., van Schaik, W., Abee, T. and Wouters, J.A. (2002). Identification of proteins involved in the heat stress response of *Bacillus cereus* ATCC14579. *Appl. Environ. Microbiol.* **68**: 3486-95.

Petersohn, A., Brigulla, M., Haas, S., Hoheisel, J.D., Volker, U. and Hecker, M. (2001). Global analysis of the general stress response of *Bacillus subtilis*. *J. Bacteriol.* **183**: 5617-31.

Petit-Glatron, M.F., Grajcar, L., Munz, A. and Chambert, R. (1993). The contribution of the cell wall to a transmembrane calcium gradient could play a key role in *Bacillus subtilis* protein secretion. *Mol. Microbiol.* **9**: 1097-106.

Phizicky, E.M. and Fields, S. (1995). Protein-protein interactions: methods for detection and analysis. *Microbiol. Rev.* **59**: 94-123.

Pragai, Z., Allenby, N.E.E., O'Connor, N., Dubrac, S., Rapoport, G., Msadek, T. and Harwood, C.R. (2004). Transcriptional regulation of the *phoPR* operon in *Bacillus subtilis*. *J. Bacteriol.* **186**: 1182-90.

Price, C.W. (2002). *Bacillus subtilis* and its closest relatives: from genes to cells. In General stress response, 1st ed. (eds. A.L. Sonenshein, J.A. Hoch and R. Losick, ASM Press, Washington, D.C., pp 369-84.

Price, C.W., Fawcett, P., Ceremonie, H., Su, N., Murphy, C.K. and Toungman, P. (2001). Genome-wide analysis of the general stress response in *Bacillus subtilis*. *Mol. Microbiol.* **41**: 757-74.

Pugsley, T. (1993). The complete general secretory pathway in Gram-negative bacteria. *Microbiol. Rev.* **57**: 50-108.

Rajagopal, P., Waygood, E.B., Reizer, J., Saier, M.H. Jr. and Klevit, R.E. (1997). Demonstration of protein-protein interaction specificity by NMR chemical shift mapping. *Protein Sci.* **6**: 2624-7.

Read, T.D., Peterson, S.N., Tourasse, N., Baillie, L.W., Paulsen, I.T., Nelson, K.E., Tettelin, H., Fouts, D.E., *et al.* (2003). The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. *Nature* **423**: 81-6.

Ryzhov, V., Hathout, Y. and Fenselau, C. (2000). Rapid characterization of spores of *Bacilus cereus* group bacteria by matrix-assisted laser desorption-ioninzation time-of flight mass spectrometry. *Appl. Envir. Microbiol.* **66**: 3828-34.

Sali, A., Glaeser, R., Earnest, T. and Baumeister, W. (2003). From words to literature in structural proteomics. *Nature* **422**: 216-25.

Schiavon, O., Caliceti, P., Ferruti, P. and Veronese, F.M. (2000). Therapeutic proteins: a comparison of chemical and biological properties of uricase conjugated to linear or branched poly (ethylene glycol). and poly (N-acryloylmorpholine). *Farmaco.* **55**: 264-9.

Schreier, H.J. and Rostkowski, C.A. (1995). *Bacillus subtilis* glnR mutants defective in regulation. *Gene* **161**: 51-6.

Schumann, W. (2000). Function and regulation of temperature-inducible bacterial proteins on the cellular metabolism. *Adv. Biochem. Eng.* **67**: 1-33.

Sender, U., Bandow, J., Engelmann, S., Lindequist, U. And Hecker, M. (2004). Proteomic signatures for daunomycin and adriamycin in *Bacillu subtilis*. *Pharmazie*. **59**: 65-70.

Sinchaikul, S., Sookkheo, B., Phutrakul, S., Pan, F.M. and Chen, S.T. (2001a). Optimization of a thermostable lipase from *Bacillus stearothermophilus* P1: overexpression, purification, and characterization. *Prot. Express. Pure.* **22**: 388398.

Sinchaikul, S., Sookkheo, B., Phutrakul, S., Wu, Y.T., Pan, F.M. and Chen, S.T. (2001b). Structural modeling and characterization of a thermostable lipase from *Bacillus stearothermophilus* P1. *Biochem. Biophys. Res. Com.* **283**: 868-75.

Sinchaikul, S., Sookkheo, B., Pan, F.M., Phutrakul, S. and Chen, S.T. (2002a). Proteomic study of cold shock protein in *Bacillus stearothermophilus* P1: comparison of temperature downshifts. *Proteomics* **2**: 1316-24.

Sinchaikul, S., Sookkheo, B., Topanuruk, S., Juan, H.F., Phutrakul, S. and Chen, S.T. (2002b). Bioinformatics, functional genomics, and proteomics study of *Bacillus* sp. *J. Chromatogr. B* **771**: 261-87.

Sinchaikul, S., Tyndall, J.D.A., Fothergill-Gilmore, L.A., Phutrakul, S., Chen, S.T. and Walkinshaw, M.D. (2002c). Expression, purification, crystallization and preliminary crystallographic analysis of a thermostable lipase from *Bacillus stearothermophilus* P1. *Acta Cryst. D* **58**: 182-5.

Soga, T., Ohashi, Y., Ueno, Y., Naraoka, H., Tomita, M. and Nishioka, T. (2003). Quantitative metabolome analysis using capillary electrophoresis mass spectrometry. *J. Proteome Res.* **2**: 488-94.

Sookkheo, B., Sinchaikul, S., Phutrakul, S. and Chen, S.T. (2000). Purification and characterization of the highly thermostable proteases from *Bacillus stearothermophilus* TLS33. *Protein Expr. Purif.* **20**: 142-51.

Sookkheo, B., Sinchaikul, S., Thannan, H., Thongprasong, O., Phutrakul, S. and Chen, S.T. (2002). Proteomic analysis of a thermostable superoxide dismutase from *Bacillus stearothermophilus* TLS33. *Proteomics* **2**: 1311-5.

Steil, L., Hoffmann, T., Budde, I., Volker, U. and Bremer, E. (2003). Genome-wide transcriptional profiling analysis of adaptation of *Bacillus subtilis* to high-salinity. *J. Bacteriol.* **185**: 6358-70.

Stevens, T.J. and Arkin, I.T. (2000). Do more complex organisms have a greater proportion of membrane proteins in their genomes? *Proteins* **39**: 417-20.

Stülke, J. and Hillen, W. (2000). Regulation of carbon catabolism in *Bacillus* species. *Annu. Rev. Microbiol.* **54**: 849-80.

Takami, H., Nakasone, K., Takaki, Y., Maeno, G., Sasaki, R., Masui, N., Fuji, F., Hirama, C., *et al.* (2000). Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*. *Nucleic Acids Res.* **28**: 4317-31.

Tatusov, R.L., Koonin, E.V. and Lipman, D.J. (1997). A genomic perspective on protein families. *Science* **278**: 631-7.

Tatusov, R.L., Natale, D.A., Garkavtsev, I.V., Tatusova, T.A., Shankavaram, U.T., Rao, B.S., Kiryutin, B., Galperin, M.Y., *et al.* (2001). The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res.* **29**: 22-8.

Thackray, P.D. and Moir, A. (2003). SigM, an extracytoplasmic function sigma factor of *Bacillus subtilis*, is activated in response to cell wall antibiotics, ethanol, heat, acid, and superoxide stress. *J. Bacteriol.* **185**: 3491-8.

Tjalsma, H., Noback, M.A., Bron, S., Venema, G., Yamane, K. and van Dijl, J.M. (1997). *Bacillus subtilis* contains four closely related type I signal peptidases with overlapping substrate specificities. *J. Biol. Chem.* **272**: 26983-92.

Tjalsma, H., Kontinen, V.P., Prágai, Z., Wu, H., Meima, R., Venema, G., Bron, S., Sarvas, M. and van Dijl, J.M. (1999). The role of lipoprotein processing by signal peptidase II in the gram-positive eubacterium *Bacillus subtilis*. Signal peptidase II is required for the efficient secretion of -amylase a non-lipoprotein. *J. Biol. Chem.* **274**: 1698-707.

Tjalsma, H., Bolhuis, A., Jongbloed, J.D.H., Bron, S. and van Dijl, J.M. (2000). Signal peptide-dependent protein transport in *Bacillus subtilis*: a genome-based survey of the secretome. *Microbiol. Mol. Biol. Rev.* **64**: 515-47.

Tobisch, S., Zühlke, D., Berngardt, J., Stülke, J. and Hecker, M. (1999). Role of CcpA in regulation of the central pathways of carbon catabolism in *Bacillus subtilis*. *J. Bacteriol.* **181**: 6996-7004.

Topanurak, S., Sinchaikul, S., Sookkheo, B., Tseng, A., Phutrakul, S. and Chen, S.T. (2005). Functional proteomics and correlated signaling pathway of thermophilic bacterium *Bacillus stearothermophilus* TLS33 under cold shock stress. *Proteomics* (in press).

Tyndall, J.D.A., Sinchaikul, S., Fothergill-Gilmore, L.A., Taylor, P. and Walkinshaw, M.D. (2002). Crystal structure of a thermostable lipase from *Bacillus stearothermophilus* P1. *J. Mol. Biol.* **323**: 859-69.

Unlu, M., Morgan, M.E. and Minden, J.S. (1997). Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* **18**: 2071-7.

van Dijl, J.M., Braun, P.G., Robinson, C., Quax, W.J., Antelmann, H., Hecker, M., Muller, J., Tjalsma, H., Bron, S. and Jongbloed, J.D.H.

(2002). Functional genomic analysis of the *Bacillus subtilis* Tat pathway for protein secretion. *J. Biotechnol.* **98**: 243-54.

van Dijl, J.M., de Jong, A., Vehmaanpera, J., Venema, G. and Bron, S. (1992). Signal peptidase I in *Bacillus subtilis*: patterns of conserved amino acids in prokaryotic and eukaryotic type I signal peptidases. *EMBO J.* **11**: 2819-28.

van Schaik, W., Tempelaars, M.H., Wouters, J.A., de Vos, W.M. and Abee, T. (2004). The alternative sigma factor ^{B} of *Bacillus cereus*: response to stress and role in heat adaptation. *J. Bacteriol.* **186**: 316-25.

Velculescu, V.E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M.A., Bassett, D.E., Hieter, P., Vogelstein, B.*, et al.* (1997). Characterization of the yeast transcriptome. *Cell* **88**: 243-51.

Voelker, U., Voelker A. and Haldenwang, W.G. (1996). The yeast two-hybrid system detects interactions between *Bacillus subtilis* ^{B} regulators. *J. Bacterial.* **178**: 7020-3.

von Heijne, G. (1990). The signal peptide. *J. Membr. Biol.* **115**: 195-201.

Wallis, N.G. and Perham, R.N. (1994). Structural dependence of post-translational modification and reductive acetylation of the lipoyl domain of the pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **236**: 209-16.

Weber, M.H.W. and Marahiel, M.A. (2003). Bacterial cold shock response. *Science Prog.* **86**: 9-75.

Wiegert, T., Homuth, G., Versteeg, S. and Schumann, W. (2001). Alkaline shock induces the *Bacillus subtilis* ^{W} regulon. *Mol. Microbiol.* **41**: 59-71.

Williams, B.H., Hathout, Y. and Fenselau, C. (2002). Structural characterization of lipopeptide biomarkers isolated from *Bacillus globigii*. *J. Mass Spectrom.* **37**: 259-64.

Winters, M.S. and Day, R.A. (2003). Detecting protein-protein interactions in the intact cell of *Bacillus subtilis* (ATCC6633). *J. Bacteriol.* **185**: 4268-75.

Wray, L.V. Jr., Zalieckas, J.M. and Fisher, S.H. (2001). *Bacillus subtilis* glutamine synthetase controls gene expression through a protein-protein interaction with transcription factor TnrA. *Cell* **107**: 427-35.

Yoshida, K., Nobayashi, K., Miwa, Y., Kang, C.M., Matsunaga, M., Yamaguchi, H., Tojo, S., Yamamoto, M., Nishi, R., Ogasawara, N., Nakayama, T. and Fujita, Y. (2001). Combined transcriptome and proteome analysis as a powerful approach to study genes under glucose repression in *Bacillus subtilis*. *Nucl. Acids Res.* **29**: 683-92.

Received: March 29, 2004 **Accepted: January 11, 2005**

CHIANG MAI UNIVERSITY 1964

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Vol. 6 (1). 2005

Thai J. Biotechnol., March 2005, p.6-15

Differential Gene Expression in Proteome Level of the Thermophillic Bacterium *Bacillus stearothermophilus* TLS33 in Environmental Cold Stresses

Supachai Topanurak[1,2], Supachok Sinchaikul[1], Suree Phutrakul[2], Boonyaras Sookkheo[3] and Shui-Tein Chen[1,4,*]

[1]*Institute of Biological Chemistry and the Genomic Research Center, Academia Sinica, Taipei, 11529, Taiwan;*
[2]*Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai, 50200, Thailand;*
[3]*Department of Chemistry, Faculty of Science, Rangsit University, Pathum Thani, 12000, Thailand;*
[4]*Institute of Biological Sciences, National Taiwan University, Taipei, 106, Taiwan*

Proteomic analysis is a powerful method to study the biological functions of thermophile Bacillus stearothermophilus TLS33 in the environmental cold stress. In this study, we investigated the proteins that involved in signaling pathway of sporulation in this bacterium. Form two-dimensional electrophoresis (2-DE), image analysis and MALDI-TOF mass spectrometry including bioinformatics aided us to investigate the change of protein and its function in lower optimum temperature conditions those are 37°C and 25°C. It found that eight cold shock proteins showed markedly differential expression in different temperatures such as glucosyltransferase, anti-sigma B (σ^B) factor, Mrp protein homolog, dihydroorthase, hydroadenosine phosphosulfate reductase and prespore specific transcriptional activator RsfA. Furthermore, bioinformatics based on database searches was used to detail the biological functions of differentiated proteins including cold shock-induced proteins, in deeply to gene homology and their functions in cell protection and adaptation that lead to understand the signaling pathway of sporulation in this bacterium.

Key words: *Bacillus stearothermophilus,* thermophile, cold shock, signaling pathway, sporulation

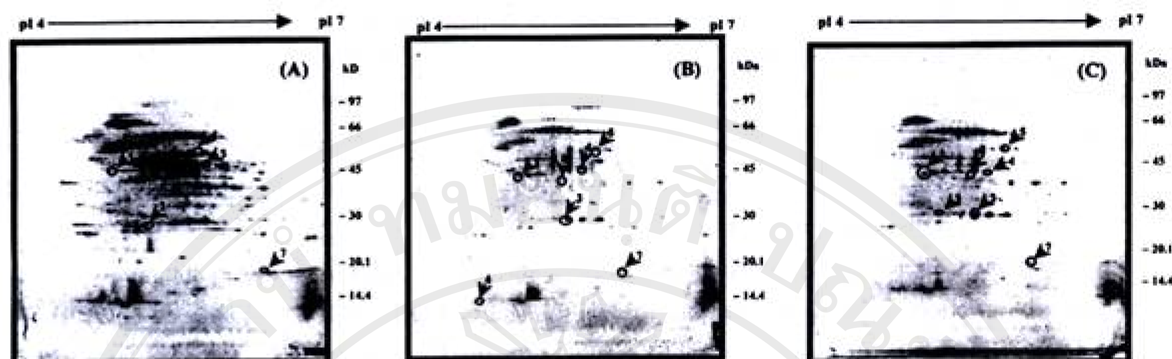
ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Introduction

Since thermophile organisms growing optimally at temperatures above 60-70°C have several unique physiological and metabolic characteristics, these may be successfully exploited within the biotechnology industry a range of applications. These may include biotransformations, biocatalysis, bioremediation or bioactive compound production, for example, whereby whole thermo-

* *Corresponding author: Dr. Shui-Tein Chen*
Institute of Biological Chemistry (RM707), Academia Sinica
128 Yen Chiu Yuan Rd., Sec II, Nankang, Taipei, 11529, Taiwan Tel: 886-2-27855981 ext.7071 Fax: 886-2-27883473
E-mail: bcchen@gate.sinica.edu.tw

Figure 1. The proteome patterns of *Bacillus stearothermophilus* TLS33 under environmental cold shock stresses. Label: A, proteome pattern at 65°C; B, proteome pattern at 37°C; C, proteome pattern at 25°C.



1,000 V, 1,000 Vh; 3,500 V, 10,500 Vh; 8,000 V, 60,000 Vh; then maintained at 100 V as last step until removal of strip. After 1-D IEF finished, IPG strips were equilibrated for 15 min in 3 mL of equilibration buffer I containing 50 mM Tris-HCl pH 8.8, 6 M urea, 30 % v/v glycerol, 2 % w/v SDS, 2 % w/v DTT and a trace of bromophenol blue, and then subsequently alkylated for 15 min in 3 mL of equilibration buffer II containing 50 mM Tris-HCl pH 8.8, 6 M urea, 30 % v/v glycerol, 2 % w/v SDS, 2.5 % w/v iodoacetamide and a trace of bromophenol blue. For the second dimension of SDS-PAGE, liner gradient of 10-20 % polyacrylamide gels (185×200×1.5 mm) was used to obtain an optimal separation. The electrophoresis was carried out on 45 mA/gel for approximately 5 h by using PROTEAN II xi Multi-Cells (Bio-Rad) and the 2-D gels were stained by Sypro Ruby staining method (Berggern *et al.*, 2000).

Image analysis

Computerized 2-D gel analysis (spot detection, spot editing, pattern matching, up- and down-regulations) was performed with the ImageMaster 2D elite software package (Amersham Biosciences, Uppsala, Sweden) using the high image quality of TIF format. On gel-to-gel comparison, the 2-D image of 65°C sample was set as the reference gel image. Before matching the images, the background subtraction was performed by using lowest-on-boundary method. The reference gel image was matched to another gel images and the matching was manually edited to ensure correct spot matches and to ensure more consistent determination of spot volume. The quantification of each spot was expressed as percent volume where %V = spot volume/ΣVolumes of all spots resolved in the gels. In addition, Proteomeweaver software (Definiens, Germany) was also used to construct 3-D viewing. In this comparison, the 2-D gel images were obtained from triplicate experiments.

In-gel tryptic digestion

The protein spots detected on 2-D images were manually excised from the gels to be small pieces. The gel pieces were dehydrated with acetronitrile for 10 min, vacuum dried and digested with a total of 20 ng of trypsin (Promega, Madison WI, USA) in 25 mM ammonium bicarbonate, pH 8.5 at 37°C for 16 hr. Following digestion, tryptic peptides were twice extracted with 50 % ACN and 5 % TFA for 15 min each coupled with moderate sonication. The extracted solutions were pooled and evaporated to dryness under vacuum. For MALDI-MS analysis, the dry peptide samples were redissolved in 0.1 % TFA and purified by C18 Zip-Tip™ (Millipore, Billerica, MA, USA) according to the manufacturer's instruction manual where necessary.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

MALDI-TOF mass spectrometry analysis

This method was done by according to Lee *et al.*, 2003. Briefly, tryptic peptides from 2-DE protein spots in 0.1 % TFA were subjected onto MALDI peptide mass fingerprinting (PMF) using a MALDI-TOF mass spectrometer (M@LDI™; Micromass Machester, UK) operated in reflectron positive ion mode. Samples were spotted onto 96-well format. MALDI target plate using a saturated matrix solution of α-cyano-4-hydroxycinnamic acid (CHCA) in 60 % ACN/1 % TFA, The instrument was externally calibrated with standard peptide mixtures and further adjusted with the lock mass feature using adenocorticotropic hormone (ACTH) as the near-point calibrant. Mass spectra were acquired for the mass range of 900-3000 Da and automatically processed by the MASCOT software (http://www.matrixsciences.com) for PMF searches against the Swiss-Prot database. The search paprameters allowed for one missed cleavage, oxidation of methionine and carbamidomethylaton of cysteines. Positive identification of proteins required at least five matching peptide masses with 50 ppm or better mass accuracy.

Results and Discussion

Two-dimensional electrophoresis and image analysis

Since a number of the proteome of *Bacillus* sp. located in a narrow p*I* range of 4-7 and MW range of 5-100 kDa (Graumann *et al.*, 1996; Graumann *et al.*, 1997) and the most protein spots of *B. stearothermophilus* TLS33 cell extracts also located in a narrow p*I* range of 4-7 (Sinchaikul *et al.*, 2002), we thus chose a narrow range of p*I* 4-7 for 2-DE analysis. The proteome patterns of *B. stearothermophilus* TLS33 cell extracts of each temperature condition are shown Figure 1. The protein spots on 2-D gel images were analyzed by ImageMaster 2D elite software in order to comparison of differential protein spots. We found that the synthesis of major intracellular proteins decreased directly to the cold shock response and only eight major cold shock-induced proteins had markedly changed under cold stress. Moreover, the image analysis could classify the protein expression to three groups. The first group, the protein was expressed in lower temperature and expressed in 25°C rather than 37°C and 65°C. The second group, the protein was expressed in only 37°C but could not be expressed in 65°C and 25°C. The third group, the protein was expressed in all temperatures or at least it was expressed in 65°C. Thus, these finding on 2-DE technique is useful for investigating the bacterial adaptation under cold shock stress and demonstrated the differential protein expressions in lower temperatures.

Protein identification and differential protein synthesis

The tryptic peptide digests were excised from 2-D gels and analyzed by MALDI-TOF mass spectrometry. The obtained peptide mass fingerprints (PMF) were searched on Swiss-Prot database with MASCOT software (http://www.matrixsciences. com). Eight of identified cold-shock induced proteins were glucosyltransferase, anti-sigma B (σ^B) factor (RSbT), Mrp protein homolog, dihydroorthase, hypothetical transcriptional regulator in the FeuA-SigW intergenic region, RibT protein, phosphoadenosine phosphosulfate reductase and prespore specific transcriptional activator RsfA. In addition, we also used Subtilist database (http://genolist.pasteur.fr/SubtiLsit) supplementing with EMBL/GenBank/DDBJ database to search the biological functions of the identified proteins including cold-shocked proteins. The protein identification of major cold shocked proteins including accession numbers, description and functional category is shown in Table 1. Based on database search, most of major cold shock-induced proteins from *B. stearothermophilus* TLS33 related to the cellular process and metabolism, which indicated the adaptation or maintenance of this bacterium under cold shock stress.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Table 1. Protein identification of the cold shock-induced proteins that identified by MALDI-TOF MS and searched on MASCOT software (http://www.matrixsciences.com).

Spot number	Protein name	Accession number	MW (Da)/pI	Number peptide match
1	Probable poly (glycerol-phosphate) alpha-glucosyltransferase (Teichoic acid biosynthesis protein) (TagE)	P13484	48700/4.44	62 %
2	Anti-sigma B factor (RsbT)	P42411	34803/5.14	29 %
3	MRP protein homolog (Mrp)	P50863	28628/5.55	39 %
4	Dihydroorotase (Dhoase) (PyrC)	P25995	39144/5.66	46 %
5	Hypothetical transcriptional regulator ybbB in feuA-sigW intergenic region (ORF3) (YbbB)	P40408	57121 / 4.55	50 %
6	RIBT protein (RibT)	P17622	18139/4.60	35 %
7	Phosphoadenosine phosphosulfate reductase (PAPS reductase, Thioredoxin dependent) (PADOPS reductase) (3'-Phosphoadenylsulfate reductase) (PAPS sulfotransferase) (Cyh1)	P94498	22031/5.95	32 %
8	Prespore specific transcriptional activator (RsfA)	P39650	58758/5.34	33 %

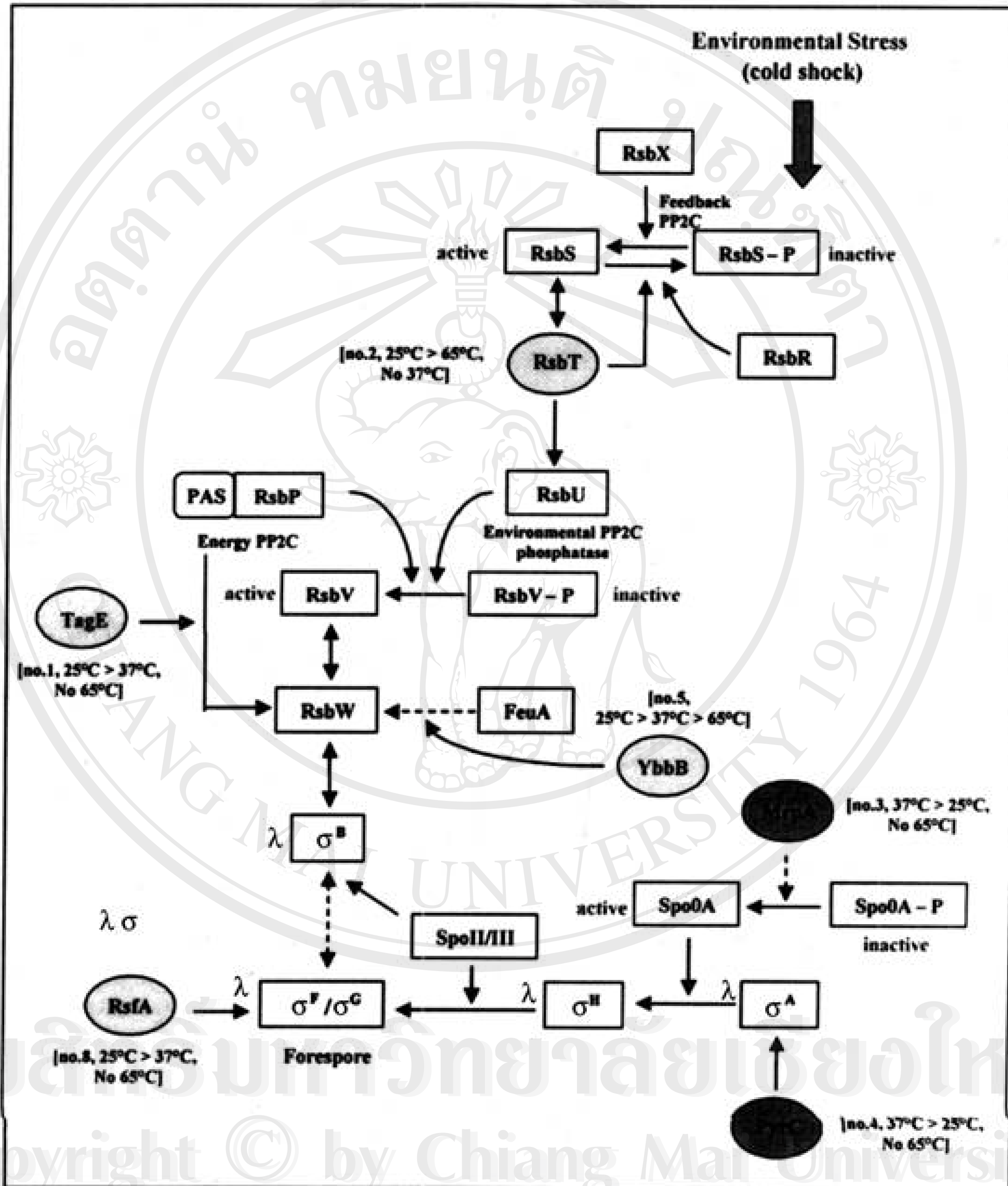
Correlation of cold shock-induced proteins in signaling pathway of sporulation

The sporulation is generally induced since the bacterium encountered the stress such as harmful or starvation condition (Movahedi *et al.*, 2002; Weber *et al.*, 2001). In this study, we attempted to investigate the proteins that involved in sporulation after bacterial cold shock. The proteomic analysis supplementing with bioinformatics can aid us to understand the cold shock-induced proteins correlating to theirs biological functions. However, we found that only six proteins of eight different cold shock-induced proteins correlated the signaling pathways of sporulation. It has been reported that TagE, YbbB, RsfA, RsbT, MrpA and PyrC correlated to the σ^B and σ^G factors which were involved "Forespore" stage of the sporulation process (Price *et al.*, 2000). In this report, we interpreted the correlated cold shock-induced proteins in the signaling pathway of sporulation by dividing to up- and down-stream of signaling pathways.

In the up-stream of signaling pathway, the proteins were expressed in lower temperature and induced the forespore stage of the sporulation process. The up-steam proteins in this study are RsbT, RsfA, TagE and YbbB. For RST protein, several previous studies suggested that RSbT generally played an important role for the coordinated expression of the σ^B factor that controlled early sporulation of vegetative cell cycle of *Bacillus* sp. (Benson *et al.*, 1992; Benson *et al.*, 1993; Boylan *et al.*, 1993; Volker *et al.*, 1996; Akbar *et al.*, 2001) The mechanism of this protein after cold shock stress can be assumed that RsbT from the up-stream module formed a protein complex with RsbS from the up-stream to module to phosphorylated RsbS and it also formed the complex with RsbU PP2C-type phosphatase, to activate the stress specific of RsbV in the down-stream module which caused stimulating its phosphatase activity towards RsbV-V (Figure. 3). The product of this reaction, RsbV, liberated σ^B from its inactive complex with RsbW and induced the expression of the general stress proteins (Dufour *et al.*, 1996; Voelker *et al.*, 1995; Kang *et al.*, 1996; Voelker *et al.*, 1995; Smirnova *et al.*, 1998). For TagE protein, the up-regulation of TagE activated the σ^F/σ^G in the forespore via RsbW/σ^B route and it alternately activated the PAS-RsbP gene to

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

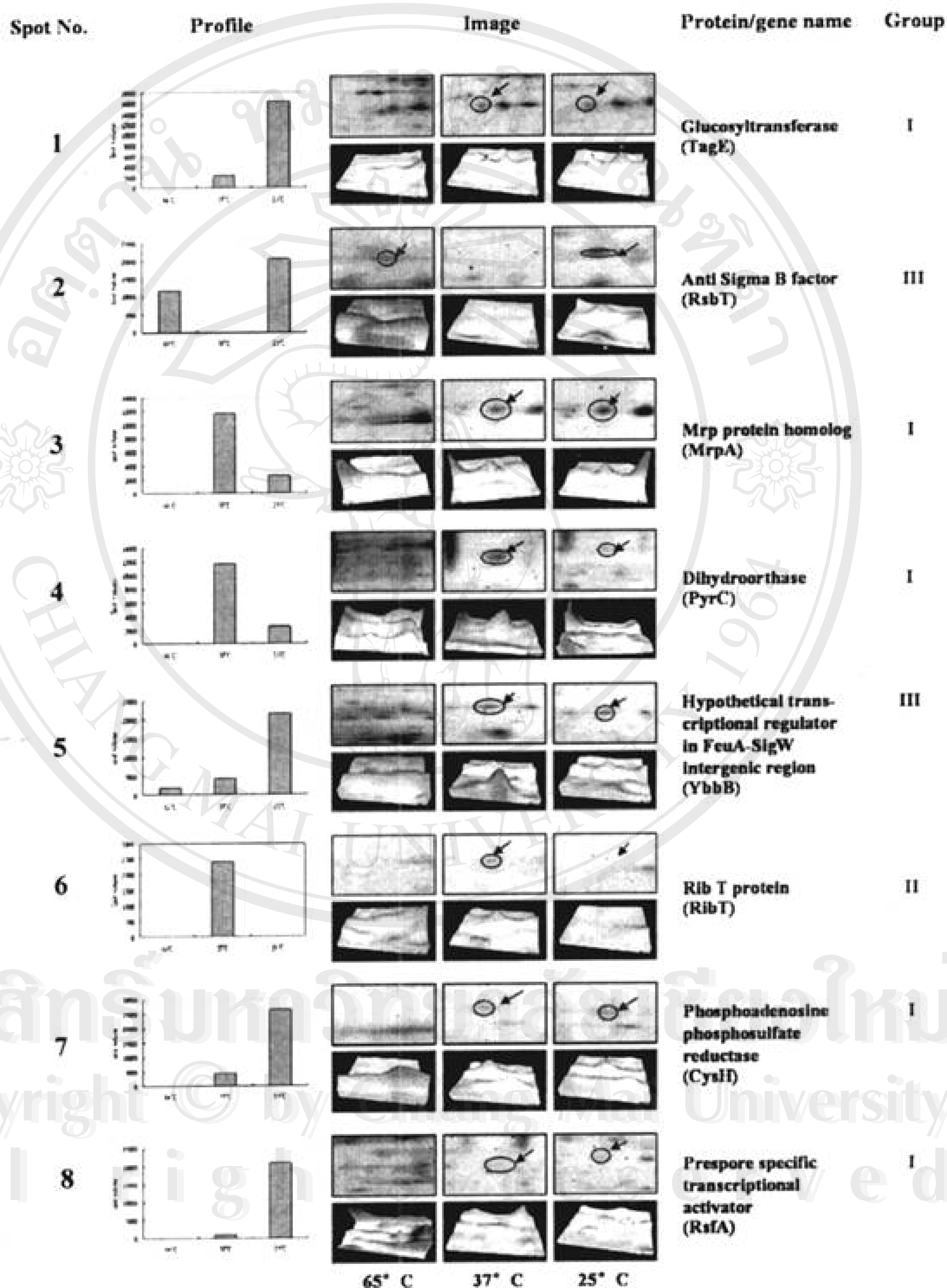
Figure 3. Summary of cold shock-induced proteins from *Bacillus stearothermophilus* TLS33 under cold shock stress which correlated to the signaling pathway of sporulation.



ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

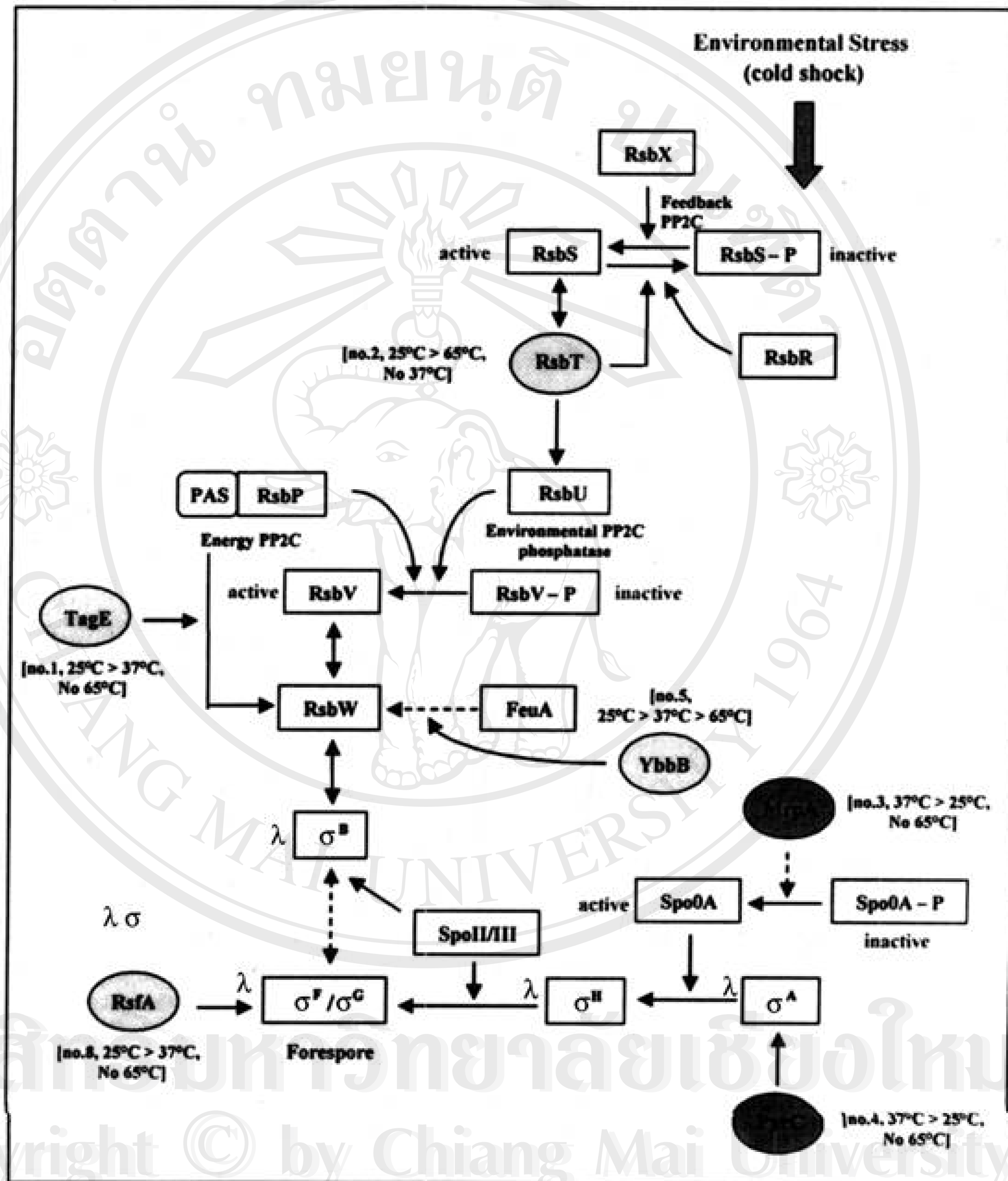
regulate the RsbW (Figure 3.). It has been found that this protein was expressed in only lower temperature at 37°C and 25°C whereas no expression at 65°C. This finding is similar to other reports that TagE could be functioned in transcription by PhoP~P under phosphate starvation conditions whereas its stress regulon was under control of the alternative transcription factor σ^B (Volker *et al.*, 1996). For YbbB protein, it was up-regulated proteins and can effect to the activation of FeuA-RsbW pathway (Figure 3). This led to process the σ^B production, in which

Figure 2. 2-D and 3-D viewings of eight cold shock-induced proteins of *Bacillus stearothermophilus* TLS33. Three groups of cold shock-induced proteins are classified by ImageMaster 2D elite software. (I, protein expression at 37°C and 25°C; II, protein expression at 37°C; III, protein expression at 65°C and 25°C)



ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Figure 3. Summary of cold shock-induced proteins from *Bacillus stearothermophilus* TLS33 under cold shock stress which correlated to the signaling pathway of sporulation.



ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

regulate the RsbW (Figure 3.). It has been found that this protein was expressed in only lower temperature at 37°C and 25°C whereas no expression at 65°C. This finding is similar to other reports that TagE could be functioned in transcription by PhoP~P under phosphate starvation conditions whereas its stress regulon was under control of the alternative transcription factor σ^B (Volker *et al.*, 1996). For YbbB protein, it was up-regulated proteins and can effect to the activation of FeuA-RsbW pathway (Figure 3). This led to process the σ^B production, in which

subsequently converted to σ^F or σ^G in the forespore. For RsfA protein, this protein was activated while the bacterium was cold shocked at 37°C and 25°C. The RsfA also directly activated σ^F/σ^G in the forespore (Figure 3.). Likewise, it had been found that the activation of σ^F in the forespore could lead to transcription of SpoIIR and SoIIQ immediately after asymmetric division of several genes that could be disrupted without preventing the formaton of stress spores and σ^G. Encoded SpoIIIG activated transcription in engulfed forespore (Price *et al.*, 2001; Wu *et al.*, 2000).

In the down-stream of signaling pathway, the proteins can be induced in the forespore stage when itself was down-regulated expression in lower temperature 37°C and 25°C. For MrpA and PyrC proteins, there was no protein expression in both of 37°C and 25°C. This finding was also found that MRP protein homolog functioned multi-resistance and pH homeostasis as Na^+/H^+ antiporter that affected post-translational regulation control of σ^H in the early sporulation of cell cycle (Ito *et al.*, 1999). For PyrC protein, it functioned in the regulation of pyrimidine biosynthetic operon by transcriptional attenuation and the control of gene expresson by mRNA-binding protein. According to the correlated signaling pathway of sporulation, the activation and deactivation of cold shock-induced proteins encoded genes involving to the σ^F/σ^G production in the forespore. It can be concluded that the forespore is a proceeding event of sporulation of the bacterium when the bacterium encountered the cold stress. For RibT and CysH proteins, they were expressed at 37°C and 25°C, but they did not involved in the signaling pathway of sporulation. In actually, RibT protein has a function as a reductase that involved in riboflavin or vitamin B_2 biosynthesis and reduction metabolism (Takami *et al.*, 2000). This has also been supported by other reports from RNA expression analysis using an antisense *B. subtilis* genome array (Lee *et al.*, 2001) and from investigation of the riboflavin operon in *B. subtilis* (Grundy *et al.*, 1998; Mansila *et al.*, 2000). For cysH (phosphoadenosine phosphosulfate reductase), it was expressed in the low temperature but its function was also not correlated to the signaling pathway of sporulation. However, the biological functions and mechanism of RibT and CysH under cold shock stress were not well known.

Conclusion

The cold shock response of the thermophile *B. stearothermophilus* TLS33 was analyzed by the proteomic approach and investigated each protein function by bioinformatics. We found that the eight cold shock-induced proteins related to sigma factor which is a core protein in sporulation stage. Furthermore, we found that six cold shock-induced proteins were correlated to signal transduction pathway of sporulation. These findings lead us to understanding the biological event of this bacterium under cold stresses. However, the biology in gene level and other mechanisms should be further studied to complete understanding of this bacterium that will be elucidated in future.

Acknowledgements

This report is part of Royal Golden Jubilee Ph.D. project of Mr. Supachai Topanurak supported by the Thailand Research Fund (TRF), Thailand. Also, this work supported by the Graudate School of Chiang Mai University. We gratefully acknowledge the support of the National Research Program for Genomic Medicine of National Science Council, Taiwan (NSC 91-3112-13-001-002) and the support of Academia Sinica, Taiwan for a grant of the *Bacillus* research. We also thank the core facilities for proteomic research, Academia Sinica, Taiwan for proteomics facilities.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

References

1. Akbar, S., Gaidenko, T.A., Kang, C.M., O'Reilly, M., Devine, K.M. and Price, C.W. (2001) New family of regulators in the environmental signaling pathway which activates the general stress transcription factor σ^B of *Bacillus subtilis*. J. Bacteriol. 183, 1329-1338.
2. Benson, A.K. and Haldenwang, W.G. (1992) Characterization of a regulatory network that controls sigma B expression in *Bacillus subtilis*. J. Bacteriol. 174, 749-757.
3. Benson, A.K. and Haldenwang W.G. (1993) *Bacillus subtilis* σ^B is regulated by a binding protein (RsbW) that blocks its association with core RNA polymerase. Proc. Natl. Acad. Sci. USA. 90, 2330-2334.
4. Berggren, K., Chernokalskaya, E., Steinberg, T.H., Kemper, C., Lopez, M.F., Diwu, Z., Haugland, R.P. and Patton, W.F. (2000) Background-free, high sensitivity staining of proteins in one- and two-dimensional sodium dodecyl sulfate-polyacrylamide gels using a luminescent ruthenium complex. Electrophoresis. 21, 2509-2521.
5. Boylan, S.A., Redfield, A.R., Brody, M.S. and Price, C.W. (1993) Stress-induced activation of the sigma B transcription factor of Bacillus subtilis. J. Bacteriol. 175, 7931-7937.
6. Büttner, K., Bernhardt, J., Scharf, C., Schmid, R., Mäder U., Eymann, C., Antelmann, H., Völker, A., Völker, U. and Hecker, M. (2001) A comprehensive two-dimensional map of cytosolic proteins of *Bacillus subtilis*. Electrophoresis. 22, 2908-2935.
7. Bustard, M.T., Burgess, J.G., Meeyoo V. and Wright, P.C. (2000) Novel opportunities for marine hyperthermophiles in emerging biotechnology and engineering industries. J. Chem. Tech. Biotech. 75, 1095-1109.
8. Dufour, A., Voelker, U., Voelker, A. and Haldenwang, W.G. (1996) Relative levels and fractionation properties of *Bacillus subtilis* sigma(B) and its regulators during balanced growth and stress. J. Bacteriol. 178, 3701-3709.
9. Graumann, P.L. and Marahiel, M.A. (1996) Some like it cold: response of microorganisms to cold. Arch. Microbiol. 166, 293-300.
10. Graumann, P., Schroder, K., Schmid, R. and Marahiel, M.A. (1996) Cold shock stress-induced proteins in *Bacillus subtilis*. J. Bacteriol. 178, 4611-4619.
11. Graumann, P., Wendrich, T. M., Weber, M. H. W., Svhroder, K. and Mirahiel, M. A. (1997) A family of cold shock proteins in *Bacillus subtilis* is essential for cellular growth and for efficient protein synthesis at optimal and low temperatures. Mol. Microbiol. 25, 741-756.
12. Grundy, F.J. and Henkin, T.M. (1998) The S box regulon: a new global transcription termination control system for methionine and cysteine biosynthesis genes in Gram-positive bacteria. Mol. Microbiol. 30, 737-749.
13. Hughes, K.T., Mathee, K. (1998) The anti-sigma factors. Annu. Rev. Microbiol., 52, 231-286.
14. Ito, M., Guffanti, A.A., Oudega, B. and Krulwich, T.A. (1999) Mrp, a multigene, multifunctional locus in *Bacillus subtilis* with roles in resistance to cholate and to Na^+ and in pH homeostasis. J. Bacteriol. 181, 2394-2402.
15. Kang, C.M., Brody, M.S., Akbar, S., Yang, X. and Price, C.W. (1996) Homologous pairs of regulatory proteins control activity of Bacillus subtilis transcription factor sigma(b) in response to environmental stress. J. Bacteriol. 178, 3846-3853.
16. Kang, C.M., Vijay, K. and Price, C.W. (1998) *Bacillus subtilis* contains multiple Fur homologues: identification of the iron uptake (Fur) and peroxide regulon (PerR) repressors. Mol. Microbiol. 30, 189-196.
17. La Teana, A., Brandi, A., Falconi, M., Spurio, R., Pon, C. L. and Gualerzi, C.O. (1991) Identification of a cold shock transcriptional enhancer of the *Escherichia coli* gene encoding nucleoid protein H-NS22. Proc. Natl. Acad. Sci. USA. 88, 10907-10911.
18. Lee, C.L., Hsiao, H.H., Lin, C. W., Wu, S.P., Huang, S.Y., Wu, C.Y., Wang, A.H. J. Khoo, K.H. (2003) Strategic shotgun proteomics approach for efficient construction of an expression map of targeted protein families in hepatoma cell lines. Proteomics. 3, 2472-2486.
19. Lee, J.M., Zhang, S., Saha, S., Anna, S.S., Jiang, C. and Perkins, J. (2001) RNA expression analysis using an antisense *Bacillus subtilis* genome array. J. Bacteriol. 183, 7371-7380.
20. Mansilla, M.C., Albanesi, D. and de Mendoza, D. (2000) Transcriptional control of the sulfur-regulated cysH operon, containing genes involved in L-cysteine biosynthesis in Bacillus subtilis. J. Bacteriol. 182, 5885-5892.
21. Movahedi S. and Waites W. (2002) Cold shock response in sporulating *Bacillus subtilis* and its effect on spore heat resistance. J. Bacteriol. 184, 5275-5281.
22. Price, C. W., Fawcett, P., Ceremonie, H., Su, N., Murphy, C.K. and Youngman, P. (2001) Genome-wide analysis of the general stress response in Bacillus subtilis. Mol. Microbiol. 41, 757-774.
23. Scott, J. M., Mitchell, T. and Haldenwang, W.G. (2000) Stress triggers a process that limits activation of the *Bacillus subtilis*

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

stress transcription factor B. J. Bacteriol. 182, 1452-1456.

24. Sinchaikul, S., Sookkheo, B., Phutrakul, S., Pan, F.M. and Chen, S.T. (2001) Optimization of a thermostable lipase from *Bacillus stearothermophilus* P1: overexpression, purification and characterization. Protein Expres. Purif. 22, 388-398.
25. Sinchaikul, S., Sookkheo, B., Topanurak, S., Juan, H.F., Phutrakul, S. and Chen, S.T. (2002) Bioinformatics, functional genomics, and proteomics study of *Bacillus* sp. J. Chromatogr. B 771, 261-287.
26. Shevchenko, A., Jensen, O.N., Podtelejnikov, A.V., Sagliocco, F., Wilm, M., Vorm O., Mortensen, P., Shevchenko, A., Boucherie, H. and Mann, M. (1996) Linking genome and proteome by mass spectrometry: Large-scale identification of yeast proteins from two dimensional gels. Proc. Natl. Acad. Sci. USA. 93, 14440-14445.
27. Smirnova, N., Scott, J., Voelker, U. and Haldenwang, W.G. (1998) Isolation and characterization of *Bacillus subtilis* sigB operon mutations that suppress the loss of the negative regulator rsbX. J. Bacteriol. 180, 3671-3680.
28. Sookkheo, B., Sinchaikul, S., Phutrakul, S. and Chen, S.T. (2000) Purification and characterization of the highly thermostable proteases from *Bacillus stearothermophilus* TLS33. Protein Expres. Purif. 20, 142-151.
29. Sookkheo, B., Sinchaikul, S., Thannan, H., Thongprasong, O., Phutrakul, S. and Chen, S.T. (2002) Proteomic analysis of a thermostable superoxide dismutase from *Bacillus stearothermophilus* TLS33. Proteomics. 2, 1311-1315.
30. Takami, H., Nakasone, K., Takaki, Y., Maeno, G., Sasaki, R., Masui, N., Fuji, F., Hirama, C., Nakamura, Y., Ogasawara, N., Kuhara, S. and Horikoshi, K. (2000) Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with *Bacillus subtilis*. Nucleic Acids Res. 28, 4317-4331.
31. Thieringer, H.A., Jones, P.G. and Inouye, M. (1998) Cold shock and adaptation. BioEssays 20, 49-57.
32. Volker, U., Englemann, S., Maul, B., Reithdorf, S., Volker, A., Schmid, R., Mach, H. and Hecker, M. (1994) Analysis of the induction of general stress proteins of *Bacillus subtilis*. Microbiol. 140, 741-752.
33. Voelker, U., Dulfour, A. and Haldenwang, W.G. (1995) The *Bacillus subtilis* rsbU gene product is necessary for RsbX-dependent regulation of sigma B. J. Bacteriol. 177, 114-122.
34. Volker, U., Voelker, A. and Haldenwang, W.G. (1996) Reactivation of the *Bacillus subtilis* anti-sigma B antagonist, RsbV, by stress- or starvation-induced phosphatase activities. J. Bacteriol. 178, 5456-5463.
35. Voelker, U., Voelker, A., Maul, B., Hecker, M., Dufour, A. and Haldenwang, W. G. (1995) Separate mechanisms activate sigma B of *Bacillus subtilis* in response to environmental and metabolic stresses. J. Bacteriol., 177, 3771-3780.
36. Weber, H.W.M., Volkov A.A., Fricker, I., Maraheil, M.A. and Grumann P.L. (2001) Localization of cold shock protein to cytosolic spaces surrounding nucleotides in *Bacillus subtilis* depends on active transcrption. J. Bacteriol. 180, 6435-6443.
37. Welker, C., Bohm, G., Schurig, H. and Jaenicke, R. (1999) Cloning, overexpression, purification, and physicochemical characterization of a cold shock protein homolog from the hyperthermophilic bacterium *Thermotoga maritime*. Prot. Sci. 8, 394-403.
38. Wu, L.J. and Errington, J. (2000) Identification and characterization of a new prespore-specific regulatory gene, rsfA, of *Bacillus subtilis*. J. Bacteriol. 182, 418-424.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

VITA

Name: Mr. Supachai Topanurak

Date of Birth : March 10, 1978

Place of Birth: Wiangpapao District, Chiang Rai.

Education: 1999 B. Sc. (Biochemistry and Biochemical technology) Faculty of Science, Chiang Mai University, Thailand

2006 Ph. D. (Biotechnology) Faculty of Science, Chiang Mai University, Thailand

Poster Presentation

1. May 29th -31st , **2002 "Proteomic analysis of cold stress induced proteins in thermophile *Bacillus stearothermophilus* TLS 33"** (PACON2002): International Conference & Exhibition on Pure and Applied Chemistry, Bangkok, Thailand.
2. May 14th-17th, **2004 "Proteomics viewed on stress response of thermophilic bacterium *Bacillus stearothermophilus* TLS 33**" 3rd International Proteomics Conference (IPC'03) 1st Taiwan Proteomics Conference Joint Meeting of AOHUPO 2004, Taipei, Taiwan.
3. November 22nd-26th, **2004, "Protemic analysis of *Bacillus stearothermophilus* TLS33 and post translational modification of its peroxiredoxin"** 17th FAOBMB-IUBMB Symposium, Thailand.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

4. May 28th-June 1, 2006 **"Proteomic Analysis of Stress Response of *Bacillus stearothermophilus* TLS33"** (2006), 54th ASMS conference on Mass spectrometry, Seattle, United States.

Oral Presentation.

February 3rd-6th **,** 2004 " **Differential gene expression in proteomic level of the thermophilic bacterium *Bacillus stearothermophilus* TLS33 in environmental cold stress**". The 15th Annual Meeting of Thai society for Biotechnology Sustainable Development of SMEs Through Biotechnology and The JSPS-NRCT Symposium on the Forefront of Bioinformatics Application, ,Chiang Mai, Thailand.

Publication

1. **Topanurak, S**., Sinchaikul, S., Sookkheo, B., Phutrakul, S., Chen, S. T., Functional proteomics and correlated signaling pathway of the thermophilic bacterium *Bacillus stearothermophilus* TLS33 under cold-shock stress. *Proteomics*, 2005, 5, 4456-4471.
2. **Topanurak, S**., Sinchaikul, S., Phutrakul S., Sookkheo, B., Chen, S. T., Proteomics viewed on stress response of thermophilic bacterium *Bacillus stearothermophilus* TLS 33. *Proteomics*, 2005, 5, 3722-3730.
3. Sinchaikul, S., Sookkheo, B., **Topanurak, S.**, Jaun, H.F., Phutrakul, S., Chen, S. T., Bioinformatics, functional genomics and proteomics study of *Bacillus* sp. *J. Chromatogr. B*., 2002, 771, 261-287.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

4. Sinchaikul, S., Sookkheo, B., **Topanurak, S**., Pan, F. M., Phutrakul, S., Chen, S. T., Functional and Structural Analysis of *Bacillus* proteome. *Current Proteomics*, 2005, 2, 109-145.

5. **Topanurak, S.,** Sinchaikul, S., Phutrakul, S., Sookkheo, B., Chen, S.T., Differential Gene Expression in Proteome Level of the Thermophillic Bacterium *Bacillus stearothermophilus* TLS33 in Environmental Cold stress. *Thai J. of Biotechnol*., 2005, 6, 6-15

Scholarship

2000-2006 The Royal Golden Jubilee Ph. D. project, The Thailand Research Fund, Bangkok, Thailand.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved