Mol. Cells, Vol. 10, No. 6, pp. 705–711

Molecules and Cells

Proteome Analysis of Light-induced Proteins in *Synechocystis* sp. PCC 6803: Identification of Proteins Separated by 2D-PAGE Using N-terminal Sequencing and MALDI-TOF MS

Jong-Soon Choi*, Dae-Sup Kim[1], Jinsuk Lee, Soo-Jung Kim, Seung-Il Kim, Young Hwan Kim[2], Jongki Hong[2], Jong-Shin Yoo[2], Kyong Hoon Suh[1], and Young Mok Park

Biomolecule Research Team, Korea Basic Science Institute, Taejon 305-333, Korea;
[1] Department of Biochemistry & Bio-Med RRC, Paichai University, Taejon 302-735, Korea;
[2] Mass Spectrometry Analysis Team, Korea Basic Science Institute, Taejon 305-333, Korea.

(Received on July 28, 2000)

The cyanobacterium *Synechocystis* sp. PCC 6803 is an ideal model organism for the proteome study of light-induced gene expression because the whole genomic sequence has been determined. The soluble proteins extracted from light- and dark-cultured cells were separated by two-dimensional polyacrylamide gel electrophoresis. Light-induced protein spots electroblotted on a polyvinyldiene difluoride membrane were analyzed by N-terminal Edman sequence determination and followed by CyanoBase. The tryptic digests of some proteins were also confirmed by matrix-assisted laser desorption ionization/time-of-flight (MALDI-TOF) and MS-Fit search. Interestingly, eight proteins were related to photosynthesis and respiration (RbcS/L, CbbA, Gap2, AtpB, CpcB, PsbO, and PsbU). Four proteins (SodB, DnaK, GroEL2, and Tig) were involved in cellular processes and the functions of another two proteins (rehydrin and membrane protein) were unknown. The proteome analysis by N-terminal Edman sequencing and MALDI-TOF enabled us to characterize one-shot protein profiles expressed under different physiological conditions.

Keywords: 2D-PAGE; MALDI-TOF; Mass Finger Printing; N-terminal Sequencing.

Introduction

Unicellular cyanobacterium *Synechocystis* sp. PCC 6803 is a transformable organsim (Grigorieva and Shestakov, 1982) capable of facultative photoheterotrophic growth in glucose-supplemented media. Owing to the high transformation efficiency, *Synechocystis* sp. PCC 6803 has often served as an ideal model system for the study of photosynthesis (Vermaas *et al.*, 1987). Moreover, the entire genome sequence of *Synechocystis* sp. PCC 6803 has been determined and released via Cyanobase (http://www.kazusa.or.jp/cyano) in the Kazusa DNA Research Institute (Kaneko *et al.*, 1996). For understanding the cyanobacterium responding to external light, the entire *prote*in complement expressed by genome (proteome) should be carefully monitored by proteomic analysis techniques. To conduct the proteome study, one possible approach would be to identify the whole proteins expressed under different physiological conditions, which is based on 2D-PAGE (O'Farrell, 1975). Recently whole proteins from *Synechocystis* sp. PCC 6803 were resolved by 2D-PAGE and the N-terminal sequences of 234 protein spots were determined (Sazuka *et al.*, 1999). This useful information on cyanobacterial genome and proteins may help to accelerate the extensive molecular study of the cellular responses by environmental stresses, i.e., light, heat, salt and drought. Massively comparative proteomics is currently available owing to advanced technological skills, (1) high-resolution of all proteins on 2D-PAGE, (2) gel-image analysis of the complex expression pattern, and (3) rapid identification and quantitation of each protein component by mass spectrometry (Quadroni and James, 1999).

As an initial attempt, we established the proteomic techniques using N-terminal sequencing and MALDI-TOF MS to identify the light-induced proteins in

* To whom correspondence should be addressed.
Tel: 82-42-865-3428; Fax: 82-42-865-3419
E-mail: jschoi@comp.kbsi.re.kr

Abbreviations: 2D-PAGE, two-dimensional polyacrylamide gel electrophoresis; IEF, isoelectrofocusing; MALDI-TOF MS, matrix-assisted laser desorption ionization/time-of-flight mass spectrometry; PVDF, polyvinyldiene difluoride.

Synechocystis sp. PCC 6803. This attempt enabled us to identify new light-dependently-expressed proteins, which were not known by conventional methods. Here we describe the cyanobacterial proteome study by means of the gene–protein linkage database.

Materials and Methods

Cyanobacterial strain and culture condition Cells of *Synechocystis* sp. PCC 6803 were inoculated in a BG11 medium (Elhai, 1990), supplemented with 10 mM glucose. The cells were grown axenically at 28°C under continuous illumination of white fluorescence light adjusted to 50 μmol/m^2/s. Cell suspensions grown up to the mid log phase (OD at 730 nm = 0.8 ∼ 1.2) were kept in darkness for 1 d. Thereafter, the cells were divided into two groups: one group was kept further for a subsequent 2 d in darkness and the other group was illuminated at 100 μmol/m^2/s for 12 h.

Preparation of soluble proteins Cell pellets were harvested and resuspended with cell fractionation buffer containing 50 mM Tris-maleate, 20 mM $MgCl_2$, pH 7.5, and 1 mM phenylmethylsulfonyl fluoride. The cells were disrupted at 4°C using a French pressure cell (SLM Aminco) at 20,000 lb/in^2. Cell debris and insoluble protein particulates were removed by centrifugation at 800 × *g* for 10 min. The supernatant was subsequently fractionated by centrifugation (Beckman TLX Ultracentrifuge) at 45,000 × *g* for 1 h. The supernatant was used as soluble proteins and precipitated with 10% (w/v) trichloroacetic acid and washed with 50 mM Tris-maleate buffer, pH 7.5.

One-dimensional and two-dimensional electrophoresis The soluble proteins of *Synechocystis* sp. PCC 6803 were dissolved in the sample buffer described by Laemmli (1970) or in the isoelectrofocusing (IEF) lysis buffer for two-dimensional electrophoresis (O'Farrell, 1975). Sodium dodecyl sulfate (SDS)-PAGE of whole proteins was performed on a 12% acrylamide slab gel using an electrophoresis kit (BioRad Mini-Protean II cell). The gels were stained with Coomassie Brilliant Blue R250 or silver nitrate prior to drying.

For IEF, the soluble fraction was mixed with lysis buffer containing 9.5 M urea, 4% NP40 or 2% CHAPS, 2% β-mercaptoethanol, and 2% ampholyte carrier (pH range, 3–10). The protein sample of 50 μg was focused in the direction of anode (20 mM NaOH) to cathode (10 mM phosphoric acid) for 12,000 Vhr in IEF gel [5% acrylamide, 9 M urea, 2% CHAPS, 4% ampholyte carrier (pH range, 3–10:5–8:2.5–5, 2:1:1, v/v/v)]. IEF was performed in the vertical gel electrophoresis system (Hoefer Mighty Small SE245). After focusing, the gel was fully equilibrated for at least 1 h with IEF equilibrium buffer containing 12.5 mM Tris-HCl, pH 6.8, 2% SDS, 10% glycerol, and 2% β-mercaptoethanol. For two-dimensional electrophoresis, the IEF rod gel was placed on a 12% SDS-polyacrylamide gel (10 × 8 cm) and electrophoresed as described previously.

N-terminal Edman sequencing of protein spots electroblotted on PVDF The soluble proteins resolved on 2D-PAGE were electroblotted on a PVDF membrane using a semidry blotting apparatus (BioRad Trans-Blot SD) at 1.25 mA/cm^2 for 60 min. The blotted membrane was stained with Coomassie Brilliant Blue R250 and washed several times with 50% methanol. The stained single protein spots were excised and applied to an Automated Peptide Sequencer (Perkin–Elmer Procise 491) in accordance with a previous report (Kim and Ha, 1997). The protein sequence determined up to 12 amino acid residues and was searched and identified using Cyano-Base (http://www.kazusa.or.jp/cyano).

Mass spectrometry MALDI-MS spectra were recorded on the Voyager-DE STR (Biosystems PerSeptive) TOF-MS in the positive reflectron mode at 20-kV accelerating voltage and 75% grid voltage. The MS was equipped with a nitrogen laser operating at 337 nm. Successful mass spectra were obtained by accumulating 150–200 laser shots. Samples for MALDI-MS were prepared by adding 1 μl of digested peptide mixture diluted in deionized water (20 μl) to 9 μl of 53 mM α-cyano-4-hydroxy cinnamic acid in 0.1% trifluoroacetic acid/acetonitrile (1:1, v/v). Then 1 μl of the mixture was applied to a stainless-steel plate and allowed to dry at room temperature before the target was introduced into the spectrometer. The spectra were calibrated externally using a peptide standard mix spotted immediately next to the sample. The peak lists were transferred into the MS-Fit algorithm (http://prospector.ucsf.edu) for database searching. A mass tolerance of 70 ppm was used.

Results and Discussion

One-dimensional and two-dimensional electrophoresis of soluble proteins The soluble proteins in *Synechocystis* sp. were resolved on 2D-PAGE according to their characteristic isoelectric point (p*I*) and molecular mass by IEF and SDS-PAGE, respectively. We found that most of the proteins were prevalently distributed in the pH 4–8 range, while a few proteins were observed in the extreme pH range. The new 2D technique using the immobilized pH gradient (IPG) has been improved to generate a 2D map accessible to high resolution and reproducibility (Görg *et al*., 1985). However, there still remained problems in the identification of some proteins on the basis of the IPG strip gels, i.e., the limitation of electrophoresis by a fixed pH gradient, expense, a time constraint, and rare proteins missing on 2D-PAGE (Sazuka and Ohara, 1997). In contrast, the application of IEF and SDS-PAGE to identify specific protein spots provides some benefits, such as low-cost performance on a small scale and flexibility of analysis to determine proteins of interest in combination with N-terminal microsequencing and MALDI-TOF MS. In the present study, we chose IEF characterizing pH 4–7 in a narrow range and pH 3–10 in a wide range instead of the IPG gel for subsequent experiments.

Identification of light-induced proteins by N-terminal sequencing The comparison of the 2D gel pattern of

the soluble proteins extracted from light- and dark-cultured cells is shown in Fig. 1. Around 50–60 proteins could be detected on the PVDF membranes transferred after 2D-PAGE. We selected 18 spots of light-induced proteins to be subjected to N-terminal Edman sequencing. Among the 18 spots, N-terminal sequences of 17 proteins encoded by the *Synechocystis* sp. PCC 6803 genome successfully determined up to 12 amino acid residues. However, we could not obtain N-terminal sequences from spot 3, which seemed to be caused by N-terminal blocking. Protein spot 4 was shown to be a protein mixture containing two proteins, CpcB and SodB. Actually about 20 and 40% of all the protein spots were reported to be protein mixtures in procaryotes and eucaryotes, respectively (Link *et al.*, 1997).

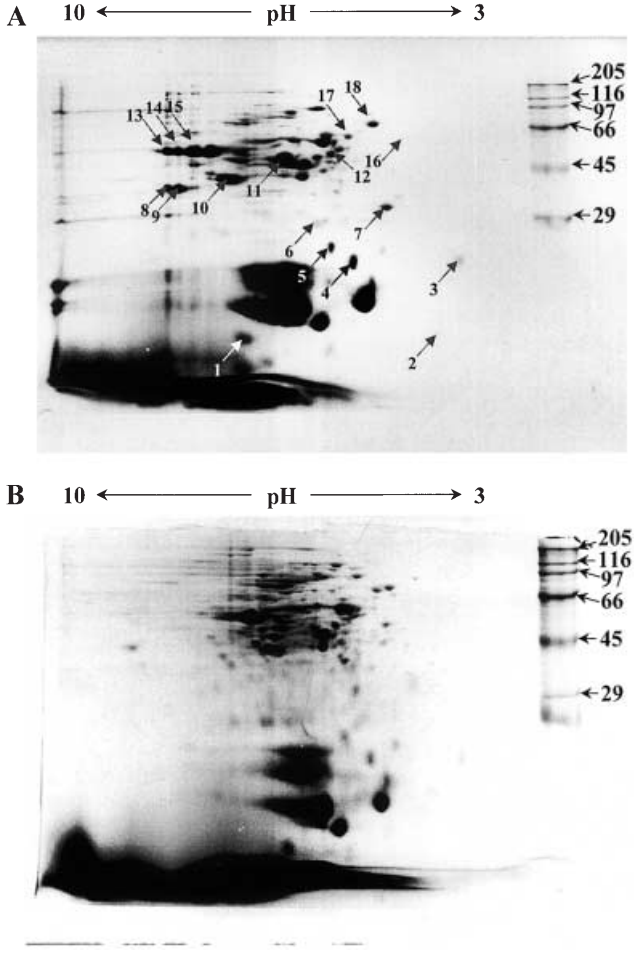

Fig. 1. Light-induced protein expression profile of *Synechocystis* sp. PCC 6803. Whole soluble proteins were extracted from *Synechocystis* sp. cultured under light (**A**) and dark (**B**) conditions. The equivalent amount (50 μg) of proteins was separated on 2D-PAGE and stained with Coomassie Brilliant Blue R250. The arrows in the left panel represent the upregulated protein spots by light treatment. The short bars on each right side indicate the location of the molecular size marker (kDa).

As seen from the 2D gel, protein spots 8 and 9 were identified as same the N-terminal sequences of Gap2. Spots 13, 14, and 15 were proven to be isoforms of RbcL. Taken together, we considered the isoform spots with different p*I* values as one of the posttranslational modifications such as acetylation, methylation, and phosphorylation of amino acid residues. Among them, phosphorylation or methylation, is possibly one of the most frequent events found in protein modification (Müller *et al.*, 1999; Yan *et al.*, 1999).

The light-induced proteins identified here are listed in Table 1. The 11 protein spots among the 14 proteins identified have been also determined by N-terminal microsequencing in the Cyano2Dbase (Sazuka *et al.*, 1999). However, three proteins (RbcL, Tig, and DnaK), estimated as being over 60 kDa in our analysis system, were not observed in Cyano2Dbase despite the higher sensitivity of the 2D technique using IPG gel. Spots 1 (RbcS), 2 (PsbU), 4-1 (CpcB), 7 (PsbO), and 12 (AtpB), formerly known as peripheral membrane proteins, were identified in the soluble fraction, suggesting that those proteins were associated with thylakoid membrane peripherally (Wang *et al.*, 2000), and were extracted during sample preparation. Table 2 shows the function-based categorization of light-induced proteins that was performed in this study. As expected, many proteins (11 spots out of 17) were shown to be components of photosynthetic machinery and the respiration system. Another four proteins (Tig, GroEL2, DnaK and SodB) were involved in cellular processes, such as protein and peptide secretion, chaperones, and detoxification. The transcriptional level of molecular chaperones, *gro*ESL and *cpn*60, is known to be significantly increased in *Synechocystis* sp. by light (Glatz *et al.*, 1997). Especially, DnaK protein is one of the abundant soluble proteins, in which the *dna*K gene expression is reported to obey upregulation with light on and downregulation with light off, i.e., in a fashion of circadian rhythm (Aoki *et al.*, 1995). Light-induced chaperones including subunits of chloroplast and cyanobacteria protein complex were involved in the assembly of Rubisco for the photosynthetic processes (Chitnis and Nelson, 1991). In a previous study, two downregulated proteins which correspond to chaperonin 2 and rehydrin were detected after exposing *Synechocystis* sp. to short illumination for 4 h (Sazuka and Ohara, 1997); however, as the exposure time increased to 12 h, the downregulated proteins increased compared with the dark control. The result suggests that DnaK and rehydrin protect the cells from continuous illumination and assist the functional assembly of photosynthetic machinery proteins.

One characteristic feature of N-terminal Edman sequencing offers the actual protein structures after the translation of corresponding genes. As shown in Fig. 1, four proteins (RbcS, CpcB, sll1621, and Tig) identified by N-terminal sequencing revealed Met in the first

residue as expected by translation codons. However, ten of the proteins lack Met at the N-terminal ends, which suggest Met to be processed. Among these proteins, PsbU and PsbO were generated by posttranslational cleavage of 36 and 28 amino acid residues from each premature protein. The N-terminal residue of the remaining eight proteins was shown as the penultimate amino acid encoded by the codon adjacent to the putative translation initiation codon. The cleavage of Met is likely to depend on the methionyl aminopeptidases (MAPs) which preferentially cleave at the N-terminus of the second amino acid residues containing the small and uncharged side groups, such as Gly, Ala, Ser, Thr, Pro, and Val (Li and Chang, 1995; Sazuka *et al*., 1999). MAP is a ubiquitous metallic enzyme that removes the initiator methionine from nascent proteins. MAP has been characterized widely in *Escherichia coli* (Ben-Bassat *et al*., 1987), *Salmonella typhymurium* (Movva *et al*., 1990), *Bacillus subtilis* (Nakamura *et al*., 1990), *Saccharomyces cerevisiae* (Chang *et al*., 1992), and human (Arfin *et al*., 1995). MAP homologs of *Synechocystis* sp. can be found on CyanoBase as putative genes, i.e., sll0555, slr0786 and slr0918.

Identification of light-induced proteins by mass finger printing Two example proteins, which were identified

Table 1. N-terminal sequence of light-induced 18 protein spots on 2D-PAGE.

Spot no.	kDa/p*I*	N-terminal sequence	Induction level	ORF no.	Predicted protein
1	9/6.9	MKTLPKERRYET	+ + +	slr0012	Ribulose bisphosphate carboxylase small subunit (RbcS)
2	12/5.4	ELNAVDAKLTXD	+ +	sll1194	Photosystem II 12kDa extrinsic protein (PsbU)
3	18/5,0	N-blocked	+		
4-1	23/6.4	MFDVFTRVVSQA	+	sll1577	Phycocyanin b subunit (CpcB)*
4-2	23/6.4	AYALPNLPYDYT	+	slr1516	Superoxide dismutase (SodB)*
5	25/6.5	MTPERVPSVVFK	+ +	sll1621	Membrane protein
6	27/6.6	ALQLGDVVPDFT	+	slr1198	Rehydrin
7	33/6.2	VDKSQLTYDDIV	+ +	sll0427	Photosystem II-Mn stabilizing polypeptide (PsbO)
8	40/7.3	TRVAINGFXRIG	+ +	sll1342	Glyceraldehyde 3-phosphate dehydrogenase isoform (Gap2)
9	40/7.4	TRVAINGFXRIG	+ +	sll1342	Glyceraldehyde 3-phosphate dehydrogenase isoform (Gap2)
10	43/6.9	ALVPMRLLLDHA	+ +	sll0018	Fructose 1,6-bisphosphate aldolase (CbbA)
11	50/6.8	MDXTXX	+ + +		Mixed proteins
12	55/6.5	VAVKEATNVGKI	+	slr1329	ATP synthase b subunit
13	60/7.2	VQAKAGFKAGVQ	+ +	slr0009	Ribulose bisphosphate carboxylase large subunit (RbcL)
14	60/7.3	VQAKAGFKAGVQ	+ +	slr0009	Ribulose bisphosphate carboxylase large subunit (RbcL)
15	60/7.4	VQAKAGFKAGVQ	+ + +	slr0009	Ribulose bisphosphate carboxylase large subunit (RbcL)
16	62/5.9	MKVTQEKLPDSQ	+	sll0533	Trigger factor (Tig)
17	64/6.5	SKLISFKDESPR	+	sll0416	60 kD chaperonin 2 (GroEL2)
18	68/6.3	GKVVGIDLGTTN	+ +	sll0170	DnaK protein

*N-terminal sequences of spots 4-1 and 4-2 were codetected. By serial chromatogram analysis of Edman sequencing, spot 4-1 was prevalently observed as a major component.

Table 2. Function-based classification of light-induced proteins.

Protein category	Spot no. on 2D-PAGE	Gene product
Photosynthesis & respiration	65%	
Photosystem II	2, 7	PsbU, PsbO
Phycobilisome	4	CpcB
CO_2 fixation	1, 8, 9, 10, 13, 14, 15	RbcS, Gap2, CbbA, RbcL
ATP synthesis	12	AtpB
Cellular process	24%	
Protein & peptide secretion	16	Tig
Chaperones	17, 18	GroEL2, DnaK
Detoxification	4	SodB
Other categories	12%	
Others	5, 6	membrane protein, rehydrin

successfully by N-terminal Edman sequencing followed by CyanoBase, were selected and applied to reconfirm the identification by the mass finger printing technique. One fingerprint example, corresponding to spot 4, is shown in Fig. 2. For eight of 17 fragments of input data, the protein containing correct tryptic digests within tolerated mass values was displayed as a possible candidate with the highest molecular weight search score. Given the datum from the analysis was considered to be CpcB of *Synechocystis* sp. PCC 6803 and it could be clearly distinguished from the second- and third-rank candidates. The fact that the corresponding spots were determined by N-terminal Edman sequencing is in good agreement with the protein predicted by mass finger printing. Tryptic fragment 1 was observed as m/z 915.4551 and 931.4337 as shown in Fig. 2, suggesting that both peaks contained a significant footprint such as the oxidation of Met (Sutton *et al.*, 1995) showing isopeaks of mass difference (m/z 16). Actually spot 4 was codetected as CpcB (major peak) and SodB (minor peak) in N-terminal sequencing but was identified as only CpcB in mass fingerprinting. This suggests that the preferential detection in tryptic digests of the mixture might be caused by the relative abundance of CpcB. Another example is shown in Fig. 3. Seven major peaks of tryptic digests were matched with CbbA within a mass tolerance of 70 ppm.

In this study we have attempted proteome analysis using a method in combination with N-terminal Edman sequencing and MALDI-MS analysis of tryptic digests based on 2D-PAGE. The MALDI-MS approach was already successfully applied to the proteomics of *E. coli* and higher plant chloroplasts (Henzel *et al.*, 1993; Peltier *et al.*, 2000). Both techniques are mutually complementary to get useful information on expressed proteins. MALDI-MS analysis provides the rapid iden-

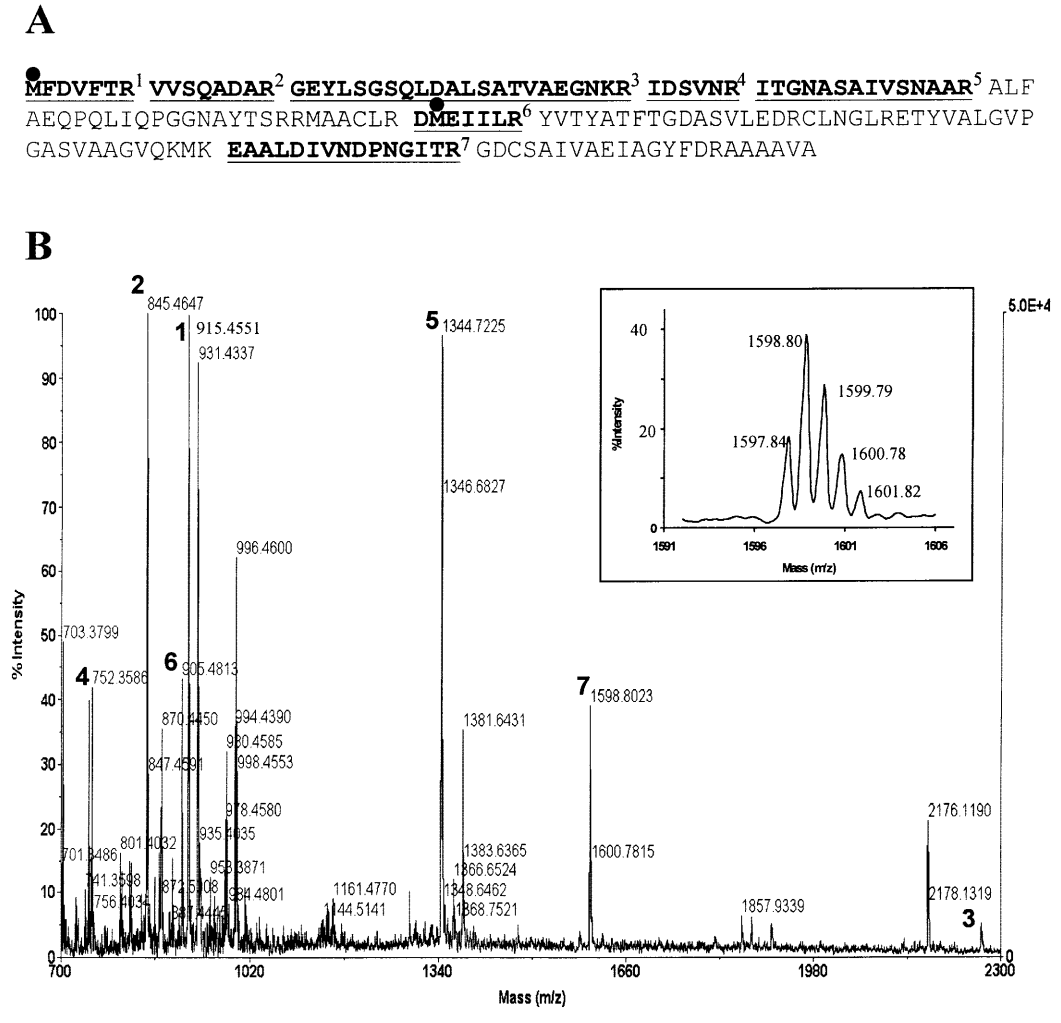

Fig. 2. MALDI-TOF MS peptide map of spot 4. **A.** Amino acid sequence of CpcB (Q54 714). Seven peptides identified by peptide mass fingerprinting are indicated in the protein sequence (bold and underlined). The dot point means the modifying site at the Met residue, such as oxidation during sample preparation. **B.** The MALDI-TOF MS spectrum of $[M + H^+]$ molecular ions from the peptides generated by tryptic digestion. The numbers in bold over the peaks mean the matched tryptic fragments within a mass tolerance of 20 ppm based on the amino acid sequence. The insert shows the isotopic peaks in the range $1591 < m/z < 1606$.

A

MALVPMRLLLDHAAENGYGIPAFNVNNMEQIISIMQAADETDSPVILQASRGAR **SYAGENFL**
R[1] HLVLGAVETYPHIPIAMHQDHGNSPATCYSAIRNGFTSVMMDGSLEADAK **TPASFEYNVN**[2]
VTAEVVK VAHSVGASVEGELGCLGSLETGQGEAEDGHGFEGK **LDHSQLLTDPEEAVEFVNK**
TQVDALAVAIGTSHGAYK[3] FTR **KPTGEVLAISR**[4] **IEEIHR**[5] LLPNTHLVMHGSSSVPQEWID
MINEFGGAIPETYGVPVEEIQKGIKSGVRKVNIDTDNR **LAITAAFR**[6] EAAAKDPKNFDPRHF
LKPSIKYMK **QVCADR**[7] YQQFWTAGNASKIKQLTLDDYAAKYAKGELTATSRTSVAV

B

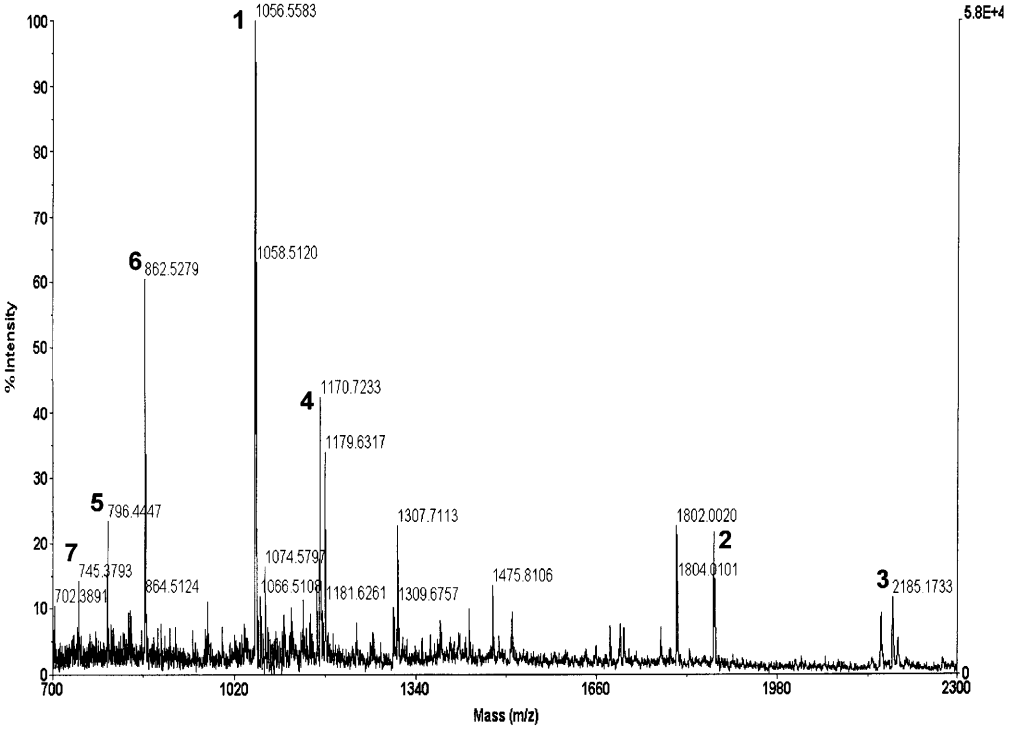

Fig. 3. MALDI-TOF MS peptide map of spot 10. **A.** Amino acid sequences of CbbA (Q55664). The actual N-terminal amino acid identified by Edman sequencing was Ala, at the second position of the protein. The down arrow (↓) indicates the probable cleavage site by cyanobacterial methionine aminopeptidase. Seven peptides identified by peptide mass fingerprinting are indicated as bold and underlined. **B.** The MALDI-TOF MS spectrum of [M + H^+] molecular ions from the peptides generated by tryptic digestion. Arabic numbers over the peaks mean the matched tryptic fragments based on the amino acid sequence.

tification of proteins on a large scale, while the Edman sequencing technique gives additional information, such as the removal of leader sequences. The proteomic approach using two methods is likely to provide reliable one-shot protein expression profiles under different physiological conditions of organisms.

Acknowledgments This work was supported by a grant (1999-2-203-003-5) from the Korea Science and Engineering Foundation through the Korea Basic Science Institute.

References

Aoki, S., Kondo, T., and Ishiura, M. (1995) Circadian expression of the *dna*K gene in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *J. Bacteriol.* **177**, 5606–5611.

Arfin, S. M., Kendall, R. L., Hall, L., Weaver, L. H., Stewart, A. E., Matthews, B. W., and Bradshaw, R. A. (1995) Eucaryotic methionyl aminopeptidase: two classes of cobalt–dependent enzymes. *Proc. Natl. Acad. Sci. USA* **92**, 7714–7718.

Ben-Bassat, A., Bauer, K., Chang, S. -Y., Myambo, K., Boosman, A., and Chang, S. (1987) Processing of the initiation methionine from proteins: properties of the *Escherichia coli* methionine aminopeptidase and its gene structure. *J. Bacteriol.* **169**, 751–757.

Chang, Y. -H., Teichert, U., and Smith, J. A. (1992) Molecular cloning, sequencing, deletion, and overexpression of a methionine aminopeptidase gene from *Saccharomyces cerevisiae*. *J. Biol. Chem.* **267**, 8007–8011.

Chitnis, P. R. and Nelson, N. (1991) Molecular cloning of the genes encoding two chaperone proteins of the cyanobacterium *Synechocystis* sp. PCC 6803. *J. Biol. Chem.* **266**, 58–65.

Elhai, J. T., Thiel, T., and Pakrasi, H. B. (1990) DNA transfer into cyanobacteria. *Plant Mol. Biol. Man.* **A12**, 1–23.

Glatz, A., Horvath, I., Varvasovszki, V., Kovacs, E., Török, Z., and Vigh, L. (1997) Chaperonin genes of the *Synechocystis* PCC 6803 are differentially regulated under light-dark transition during heat stress. *Biochem. Biophys. Res. Commun.* **239**, 291–297.

Görg, A., Postel, W., Günther, S., and Weser, J. (1985) Improved horizontal two-dimensional electrophoresis with hybrid isoelectric focusing in immobilized pH gradients in the first dimension and laying-on transfer to the second dimension. *Electrophoresis* **6**, 599–604.

Grigorieva, G. and Shestakov, S. (1982) Transformation in the cyanobacterium *Synechocystis* sp. PCC 6803. *FEMS Microbiol. Lett.* **13**, 367–370.

Henzel, W. J., Billeci, T. M., Stults, J. T., Wong, S. C., Grimley, C., and Watanabe, C. (1993) Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence database. *Proc. Natl. Acad. Sci. USA* **90**, 5011–5015.

Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirisawa, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Yamada, M., Yasuda, M., and Tabata, S. (1996) Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. PCC 6803 II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* **3**, 109–136.

Kim, S. I. and Ha, K. -S. (1997) Peptide mapping and amino acid sequencing of two catechol 1,2-dioxygenases (CD I_1 and CD I_2) from *Acinetobacter lwoffii* K24. *Mol. Cells* **7**, 635–640.

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.

Li, X. and Chang, Y. -H. (1995) Amino-terminal protein processing in *Saccharomyces cerevisiae* is an essential function that requires two distinct methionine aminopeptidases. *Proc. Natl. Acad. Sci. USA* **92**, 12357–12361.

Link, A. J., Hays, L. G., Carmack, E. B., and Yates III, J. R. (1997) Identifying the major proteome components of *Haemophilus influenzae* type-strain NCTC 8143. *Electrophoresis* **18**, 1314–1334.

Movva, N. R., Semon, D., Meyer, C., Kawashima, E., Wingfield, P., Miller, J. L., and Miller, C. G. (1990) Cloning and nucleotide sequence of the *Salmonella typhimurium pcp*M gene. *Mol. Gen. Genet.* **223**, 345–348.

Müller, D. R., Schindler, P., Coulot, M., Voshol, H., and van Oostrum, J. (1999) Mass spectrometric characterization of stathmin isoforms separated by 2D PAGE. *J. Mass Spectrom.* **34**, 336–345.

Nakamura, K., Nakamura, A., Takamatsu, M., Yoshikawa, H., and Yamane, K. (1990) Cloning and characterization of a *Bacillus subtilis* gene homologous to *E. coli* SccY. *J. Biochem.* **107**, 603–607.

O'Farrell, P. (1975) High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **250**, 4007–4021.

Peltier, J. -B., Friso, G., Kalume, D. E., Roepstorff, P., Nilsson, F., Adamska, I., and Wijk, K. J. (2000) Proteomics of the chloroplast: systematic identification and targeting analysis of lumenal and peripheral thylakoid proteins. *Plant Cell* **12**, 319–341.

Quadroni, M. and James, P. (1999) Proteomics and automation. *Electrophoresis* **20**, 664–677.

Sazuka, T., Yamaguchi, M., and Ohara, O. (1999) Cyano2Dbase updated: linkage of 234 protein spots to corresponding genes through N-terminal microsequencing. *Electrophoresis* **20**, 2160–2171.

Sutton, C. W., Pemberton, K. S., Cottrell, J. S., Corbett, J. M., Wheeler, C. H., Dunn, M. J., and Pappin, D. J. (1995) Identification of myocardial proteins from two-dimensional gels by peptide mass fingerprinting. *Electrophoresis* **16**, 308–316.

Vermaas, W., Williams, J., and Arntzen, C. (1987) Sequencing and modification of *psb*B, the gene encoding the CP-47 protein of photosystem II in the cyanobacterium *Synechocystis* 6803. *Plant Mol. Biol.* **8**, 317–326.

Yan, J. X., Sanchez, J. -C., Binz, P. A., Williams, K. L., and Hochstrasser, D. F. (1999) Method for identification and quantitative analysis of protein lysine methylation using matrix–assisted laser desorption ionization time-of-flight mass spectrometry and amino acid analysis. *Electrophoresis* **20**, 749–754.