Microbiology (2007), **153**, 608–618 DOI 10.1099/mic.0.29198-0

Proteomic analyses of the photoauto- and diazotrophically grown cyanobacterium *Nostoc* sp. PCC 73102

Liang Ran, Fang Huang,† Martin Ekman, Johan Klint and Birgitta Bergman

Department of Botany, Stockholm University, SE-106 91 Stockholm, Sweden

Correspondence
Birgitta Bergman
bergmanb@botan.su.se

Received 7 June 2006
Revised 4 October 2006
Accepted 23 October 2006

The filamentous cyanobacteria of the genus *Nostoc* are globally distributed, phenotypically complex organisms, capable of cellular differentiation and of forming symbiotic associations with a wide range of plants. To further our understanding of these processes and functions, the proteome of photoautotrophically and diazotrophically grown *Nostoc* sp. PCC 73102 (*N. punctiforme*) cells was examined. Extracted proteins were separated into membrane and soluble protein fractions and analysed using two-dimensional gel electrophoresis and matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS). The analysis led to the identification of 82 proteins that could be divided into 12 functional categories. Significantly, 65 of these proteins have not been previously documented in the *Nostoc* proteome. Many of the proteins identified were readily recognized as housekeeping proteins involved in carbon, nitrogen and energy metabolism, but a number of proteins related to stress, motility, secretion and post-translational modifications were also identified. Ten unclassified proteins were also detected, representing potential novel functions. These proteins were highly expressed, suggesting that they play key roles during photoautotrophic and diazotrophic growth. Nineteen of the proteins expressed under the growth conditions examined contained putative thioredoxin (Trx) targets, a motif that functions in redox regulation via redox equivalent mediators and is known to be significant in a wide range of biological processes. These observations contribute to our understanding of the complex *Nostoc* life cycle.

INTRODUCTION

Cyanobacteria are photosynthetic prokaryotic microbes, which make a substantial contribution to global CO_2 assimilation, O_2 production and N_2 fixation, and were evolutionary precursors of plant plastids. *Nostoc* is a nitrogen-fixing (diazotrophic) filamentous genus (Rippka *et al.*, 1979), with multiple physiological properties, occupying a wide range of terrestrial and limnic niches (Dodds *et al.*, 1995).

To survive in such varying environments (from cold arctic to tropical areas), the photosynthetic vegetative cells of *Nostoc* species have evolved elaborate developmental alternatives. Firstly, they may divide their vegetative cells into new photosynthetically competent cells; secondly, they may terminally differentiate a proportion of the vegetative cells (5–10 %) into highly specialized non-photosynthetic, but diazotrophic, heterocysts as a response to lack of combined nitrogen; thirdly, the vegetative cells may develop into transient spore-like cells, termed akinetes, able to withstand harsh environmental conditions; and finally, whole filaments may differentiate into motile hormogonia, functioning in short-distance dispersal and as 'infective units' in cyanobacteria–eukaryote symbioses (Adams & Duggan, 1999; Meeks & Elhai, 2002). The multiple differentiation strategies allow *Nostoc* species to use a photoautotrophic and diazotrophic mode of growth, making them uniquely independent of combined carbon and nitrogen. There has been significant research into the developmental cycles exhibited by this organism and many of the genes involved in N_2 fixation for example have been characterized (Adams & Duggan, 1999) but still many unknowns remain, such as the factors that govern the highly symbiosis-competent nature of members of this genus (Rai *et al.*, 2000), the regulatory elements central to their complex life cycle, and the factors that contribute to their global success in terrestrial systems (Dodds *et al.*, 1995).

†Present address: Research Center of Photosynthesis, Institute of Botany, Chinese Academy of Sciences, Beijing, China.

Abbreviations: AvAK, akinete marker protein; G6PD, glucose-6-phosphate dehydrogenase; OEP, outer-membrane efflux protein; OPP, oxidative pentose phosphate; PBS, phycobilisome; 6PGD, 6-phosphogluconate dehydrogenase; PRK, phosphoribulokinase; Trx, thioredoxin; vWA, von Willebrand factor type A.

The GenBank/EMBL/DDBJ accession numbers for the sequences reported in this paper are listed in the tables.

A supplementary table and figure are available with the online version of this paper.

 0002-9198 © 2007 SGM *Printed in Great Britain*

Whole-genome sequencing of *Nostoc punctiforme* (ATCC 29133; http://www.jgi.doe.gov), the only symbiotically competent cyanobacterium sequenced so far, revealed an exceptionally large prokaryotic genome approaching 10 Mb with 7432 predicted ORFs (Meeks *et al.*, 2001). The availability of genomic information is just beginning to pave the way for the use of proteomics as a powerful tool for large-scale comparison of *Nostoc* protein levels and for the identification of these proteins. To date, approximately 35 cytoplasmic proteins of *N. punctiforme* (ATCC 29133) have been identified (Hunsucker *et al.*, 2004), and 10 proteins in *N. punctiforme* (PCC 73102) are known to be differentially expressed during hormogonium differentiation (Klint *et al.*, 2006).

Here, proteins were analysed from the free-living, photoautotrophic and diazotrophic life stage of *Nostoc* sp. PCC 73102, which is the predominant stage of its complex life cycle and at which cyanobacterial–plant interactions are initiated. To increase the number of identified proteins, cellular and subcellular fractionation procedures were employed together with a quantitative strategy. Proteins were identified by MALDI-TOF MS and divided into 12 categories, based on predicted functions; their putative functions are discussed.

METHODS

Cyanobacterial strains and culture conditions. The cyanobacterium *Nostoc* sp. PCC 73102 (originally isolated from the coralloid roots of the cycad *Macrozamia* sp.; Rai *et al.*, 2002) and equivalent to *Nostoc* ATCC 29133 and *N. punctiforme*, was grown in nitrogen-depleted (diazotrophic conditions) $BG11_0$ medium (Stanier *et al.*, 1971) with constant light (18 μmol photons m^{-2} s^{-1}) shaken at 63 r.p.m. at 25 °C.

Protein extraction and fractionation. Two-week-old cyanobacterial cultures were harvested by centrifugation and the cell suspension was washed three times in fresh $BG11_0$ medium. The cell pellets were weighed and then ground in five times their volume of extraction buffer [40 mM Tris supplemented with a Complete Mini Protease Inhibitor Cocktail Tablet (Roche)] with a mortar and pestle in liquid nitrogen, followed by sonication (Bandelin Sonoplus, DPC Scandinavia) 10 times (70 % intensity) for 20 s each in an ice bath, with 40 s cooling breaks. The homogenates were centrifuged for 1 min at 13 600 ***g*** to remove unbroken cells and cell debris. Supernatants were pooled and centrifuged at 160 000 ***g*** for 40 min at 4 °C. These supernatants were designated the soluble protein fraction. The pellets were washed once in the same buffer, harvested by ultracentrifugation and designated the membrane fraction. The protein concentrations of the soluble and membrane protein fractions were determined using the DC Protein Assay kit (Bio-Rad). The proteins were solubilized in two-dimensional gel electrophoresis rehydration buffers: the soluble fraction proteins in 7 M urea, 2 M thiourea, 4 % (w/v) CHAPS, 30 mM DTT and 1 % (v/v) IPG buffer (GE Healthcare); and the membrane fraction proteins in 7 M urea, 2 M thiourea, 1 % (v/v) ASB-14, 4 % (w/v) CHAPS, 30 mM DTT and 2 % (v/v) IPG buffer, to a protein concentration of 2 mg ml^{-1}, and then stored at −80 °C until used.

Two-dimensional (2D) gel electrophoresis. Protein extracts containing 200 μg protein were loaded onto rehydrated 18 cm immobiline gel strips (pH 4–7; GE Healthcare) by cup loading. The isoelectric focusing of the proteins was performed on an IPGphore (GE Healthcare) and the focusing time was adjusted to a total of 80 000 Vh. The strips were equilibrated in the equilibration buffer described by Nouwens *et al.* (2000) for 30 min then positioned on top of an SDS–polyacrylamide gel (10 % polyacrylamide) and sealed with 0.5 % (w/v) agarose. The second dimension was carried out in a Protean II xi 2-D cell (Bio-Rad) at 20 mA for 20 min and 40 mA for 5.5 h. The gels were stained with the fluorescent dye SYPRO Ruby (Molecular Probes) as described by the manufacturer. A laser-scanning instrument (Typhoon 8600, GE Healthcare) was used for digitizing protein spot maps of the gels.

Gel analysis, in-gel digestion and MALDI-TOF analysis. The scanned gel images were further analysed, using the PDQuest software (Bio-Rad) for visualization of protein spot distribution and annotation. Reproducible spots (clearly detected at least three times) were selected for identification. In-gel digestion was performed manually according to Fulda *et al.* (2000), followed by analyses in a MALDI-TOF MS instrument (Voyager-DE STR mass spectrometer, Applied Biosystems). The peptide mass fingerprints obtained were internally calibrated in MoverZ software (http://www.genomicsolutionscanada.com) using known autolytic trypsin peaks.

Database search. The proteins were identified by comparing peptide mass fingerprints to the NCBInr database using the Mascot search engine (http://www.matrixscience.com). The search parameters allowed for oxidation of methionines, carbamidomethylation of cysteines, one mis-cleavage of trypsin, and 30 p.p.m. mass accuracy. The proteins were successfully identified, based on the first-ranking result and Mascot scores >74, which indicates that the hits were significant. Some hits with Mascot scores <74, for which more than half of the peptides used in the search matched, and hits with the correct M_r and pI, were also accepted in this study.

Bioinformatic analyses. Signal peptides and their cleavage sites were predicted using the SignalP program (www.cbs.dtu.dk/services/SignalP-3.0). The prediction of transmembrane helices in identified proteins was performed using the TMHMM program (www.cbs.dtu.dk/services/TMHMM). Putative thioredoxin (Trx) targets were predicted based on homology (BLAST) to sequences of known Trx-linked proteins. To determine the number of conserved cysteines in cyanobacterial Trx targets, homologues (see Supplementary Table S1, available with the online version of this paper) from cyanobacteria, plants or green alga were compared and analysed by CLUSTALW (Lemaire *et al.*, 2004; Lindahl & Florencio, 2003) (Supplementary Fig. S1). For each alignment, at least one of the selected proteins was previously identified as a Trx target.

RESULTS

Protein fractionation, 2D electrophoresis and protein identification

To achieve optimal cell protein coverage, total protein extracts from vegetative filaments (containing vegetative cells and heterocysts) of *Nostoc* sp. PCC 73102 were first fractionated into soluble and membrane proteins. After separation of the cyanobacterial proteins using 2D gel electrophoresis, gels were stained with SYPRO Ruby, which is more sensitive than Coomassie brilliant blue (Berggren *et al.*, 2000). About 180 of the most abundant and reproducible (clearly detected on at least three gels) protein spots were selected for identification. As expected, a considerably higher number of protein spots, distributed over the entire gel, were detected in the soluble fraction as

compared to the membrane fraction (composed of thylakoids, and inner and outer cell membranes) of the vegetative cells (Fig. 1a, b). The latter proteins dominated in the lower pH and higher-M_r region. Of the 180 protein spots, about 130 were successfully identified using peptide mass fingerprints (MALDI-TOF MS) and by searching public databases. The protein spots identified corresponded to 82 proteins, 60 soluble (Fig. 1a) and 22 membrane (Fig. 1b) fraction proteins from the whole-filament extractions (dominated by vegetative cells; typically about 92–95 % of the total cell population under N_2-fixing conditions). Of the proteins identified, 65 were detected in the *Nostoc* proteome for the first time. Some proteins, particularly in the membrane fraction, appeared as multiple spots, possibly due to post-translational modifications. However, the nature of the modifications was not discernible from the MALDI-TOF spectra, but in most cases resulted in a shift in pI rather than in M_r.

Four vegetative membrane proteins Mp7, 9, 10 and 11 (see Table 2) were identified as being identical to proteins previously described as up-shifted in heterocysts (annotated

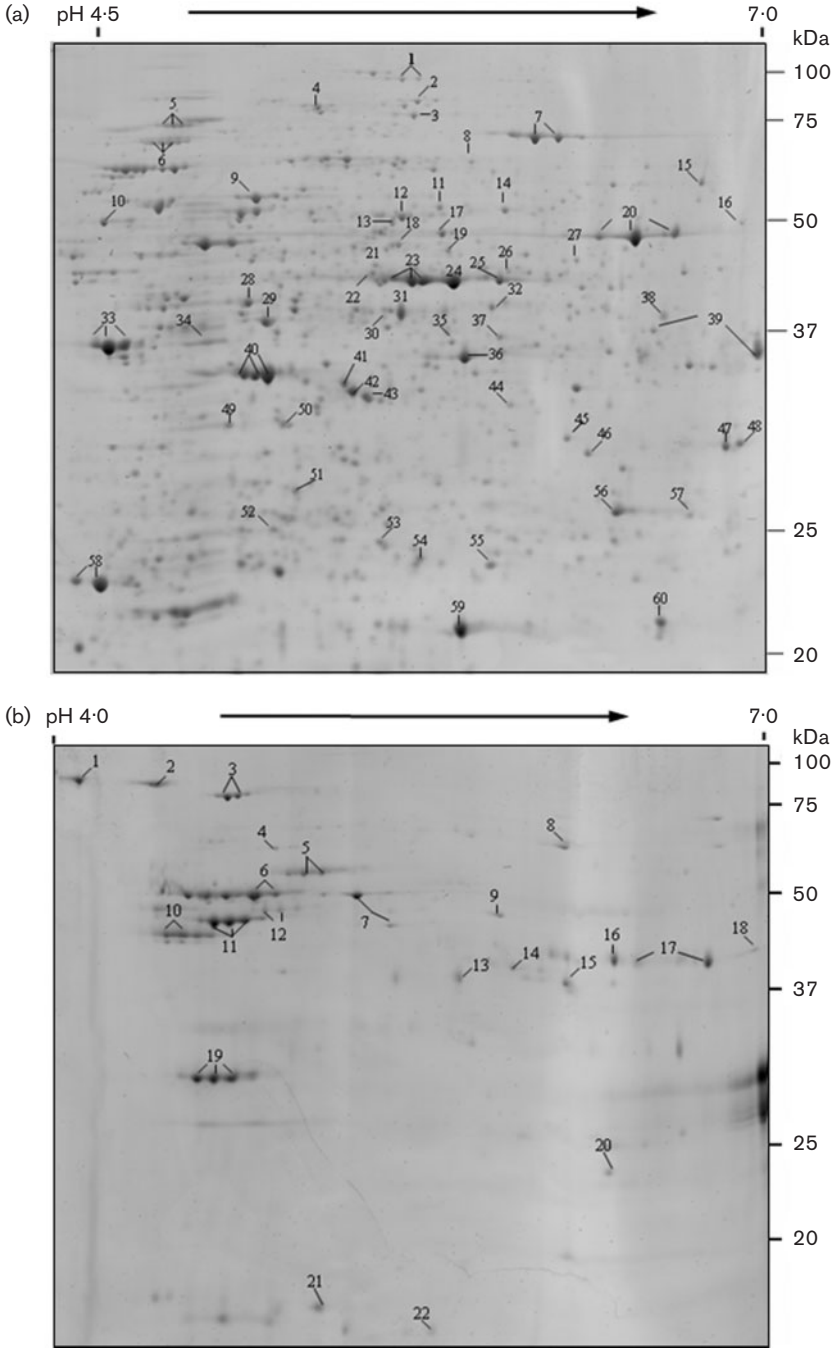

Fig. 1. 2D protein patterns of *Nostoc* sp. PCC 73102 from (a) the cellular soluble fraction and (b) the membrane fraction, of photoautotrophically and diazotrophically grown cells. The proteins were separated using a linear pH 4–7 immobilized gradient and 10 % SDS–PAGE. The gels were stained with SYPRO Ruby fluorescent dye.

as h3, h4, h2 and h6 respectively in Ekman *et al.*, 2006). The heterocyst-specific NifK (Sp11), one of the larger subunits of the nitrogenase complex, was identified in the soluble fraction of vegetative filaments (Fig. 1a).

Categorization of the proteins

The 82 proteins identified in the proteome (Fig. 1a, b) were divided into 12 categories as illustrated in Tables 1 and 2 and Fig. 2. These were: unknown proteins (10), and proteins involved in carbohydrate metabolism (11), energy metabolism (10), amino acid biosynthesis (9), protein synthesis (3), post-translational modification of proteins (12), nucleotide metabolism (3), lipid metabolism (2), reductases (3), cell motility and secretion and cell envelope biogenesis (11), inorganic nutrient transport and metabolism (7), and signal transduction (1). These categories are based on those selected for the automated annotation of the *Nostoc* sp. ATCC 29133 genome, and recognized as being the most relevant to the organism's lifestyle (Meeks *et al.*, 2001). Consistent with the smaller number of proteins (22 from the membrane and 60 from the soluble fraction) there was a corresponding reduction in the number of membrane-associated functional categories (4) compared with those observed for the soluble proteins (12). Moreover, the functions of the membrane proteins were primarily related to membrane biosynthesis, transport across membranes, surface components and thylakoid photosystems. This also verifies the purity of the membrane fractions.

Signal peptide and transmembrane helix prediction

To gain deeper insights into the membrane proteins identified, a search for signal peptides was performed based on the SignalP program for Gram-negative bacteria. As shown in Table 2, 13 of the 22 membrane proteins were predicted to possess a signal peptide (Nielsen *et al.*, 1997), and N-terminal transmembrane helices were predicted in six of these. This suggests that these are either integral membrane proteins or located on the outer side of the membrane surface (von Heijne, 1988). However, as the transmembrane regions were mostly located within the signal peptide, which is likely to be cleaved off during translocation, these proteins may represent membrane-associated proteins. Moreover, the 20 kDa subunit of NADH : ubiquinone oxidoreductase (Table 2) was predicted to have one transmembrane helix although no signal peptide was detected.

Putative thioredoxin (Trx) targets

Thioredoxin (Trx) targets among the proteins found were identified by sequence similarity to known Trx targets detected in cyanobacteria, plants and algae that had been characterized previously using Trx affinity chromatography coupled to proteomic analysis. The numbers of cysteines and conserved cysteines in the cyanobacterial proteins were first determined by multiple sequence alignments (Table 3). Nineteen of the cyanobacterial proteins expressed under the growth conditions examined were predicted to be putative Trx targets (Table 3). These included one reductase (Sp51) and proteins involved in carbohydrate and energy metabolism (Sp7, 25, 28, 31, 36, 52, 20 and Mp5,); amino acid biosynthesis (Sp8, 29, 30, 43 and 46); translation (Sp23); protein folding and degradation (Sp6, 9 and 10); and fatty acid biosynthesis (Sp3). All of the predicted Trx target proteins contained at least one cysteine residue, and 16 of the proteins contained conserved cysteines as in reported Trx targets (Table 3).

DISCUSSION

The *Nostoc* sp. PCC 73102 proteins identified in the present study represent a proteomic 'snap-shot', illustrating the most prevalent proteins in the life cycle of *Nostoc* during its most common life stage: a combination of photoautotrophy and diazotrophy. Of the 82 proteins identified, 65 were reproducibly identified in *Nostoc* sp. PCC 73102 for the first time and the data considerably extend those of a previous proteomic study using *Nostoc* sp. ATCC 29133 (*N. punctiforme*) (Hunsucker *et al.*, 2004). This was achieved by introducing cell fractionation to increase the number of proteins identified, improving sample preparation, and introducing SYPRO Ruby staining to improve protein spot visualization (and allow quantitative comparisons). Protein categorization revealed the involvement of a number of specific metabolic and cellular pathways in *Nostoc* under these growth conditions.

Interestingly, our data indicated that growth under free-living conditions in light under combined N-depletion may be stressful. This is however not a surprising result, since two incompatible processes, aerobic photosynthesis and anaerobic nitrogen fixation, occur within single filaments. For instance, the known stress-related enzyme superoxide dismutase (Sp59) (Kliebenstein *et al.*, 1998) was strongly expressed, together with three reductases (Sp42, 44 and 51), which also may have roles in stress defence and repair systems. Additionally, four molecular chaperones (Sp1, 6, 9 and 10) known to play roles in stabilizing and refolding proteins during exposure of cells to stress, for instance in the GroEL system (Bukau & Horwich, 1998; Fink, 1999), were highly expressed (Fig. 1a). The proteins represented by spots Sp49 and 60, representing peptidyl-prolyl *cis-trans* isomerases (PPIase), are additional protein-folding catalysts related to stress (Lodish *et al.*, 1999). The genes encoding the Sp53 (an ATP-dependent Clp protease) and Sp32 (an uncharacterized protein containing a vWA domain) proteins were located in a conserved gene cluster. The Clp protease removes oxidatively damaged proteins in response to environmental stress (Adam & Clarke, 2002; Agrawal *et al.*, 2002; Zheng *et al.*, 2002), while the vWA domain is found in cell adhesion proteins and promotes protein–protein interactions via a highly conserved metal ion-dependent adhesion site (Whittaker & Hynes, 2002; Lee *et al.*, 1995). These proteins may function in reducing or

Table 1. Proteins identified in the cellular soluble fraction (Sp) of *Nostoc* sp. PCC 73102, corresponding to the proteins marked in Fig. 1(a)

Spot no.*	Accession no.	Gene product‡	Mowse score	Matched peptide/total	Putative M_r/pI
Unknown proteins					
Sp24†	ZP_00106617	Hypothetical protein Npun02007358	124	7/7	43.8/5.57
Sp32†	ZP_00108612	Uncharacterized protein containing a vWA domain	171	9/9	45.6/5.64
Sp33	ZP_00108348	AvAK-like protein	225	13/15	34.3/4.72
Sp39†	ZP_00109274	Hypothetical protein Npun02004266	124	11/31	42.1/6.38
Sp40†	ZP_00106350	Uncharacterized conserved protein	120	7/8	39.2/5.11
Sp47	ZP_00112085	Hypothetical protein Npun02000353	112	7/9	34.2/6.25
Carbohydrate metabolism					
Sp2†	ZP_00111620	Phosphoketolase	86	6/7	79.4/5.50
Sp7	ZP_00106110	Transketolase	157	10/10	72.9/5.77
Sp12	ZP_00107727	Glucose-6-phosphate isomerase	116	8/11	57.9/5.46
Sp18†	ZP_00108646	Malic enzyme	156	10/12	49.6/5.43
Sp19†	ZP_00109278	Phosphomannomutase	166	12/17	52.2/5.55
Sp25†	ZP_00111860	6-phosphogluconate dehydrogenase	81	6/11	52.4/5.62
Sp28	ZP_00107735	Fructose-1,6-bisphosphatase	89	7/13	37.2/4.99
Sp31†	ZP_00109191	Phosphoribulokinase (PRK)	115	7/9	38.7/5.41
Sp36†	ZP_00110670	Fructose/tagatose bisphosphate aldolase	178	11/13	38.9/5.47
Sp41	ZP_00107110	Transaldolase	99	6/8	36.3/5.35
Sp52†	ZP_00109456	Ribose 5-phosphate isomerase	96	6/11	24.9/5.03
Energy metabolism (photosynthesis, respiratory electron transport, ATP synthesis)					
Sp20	ZP_00108159	Rubisco, large subunit	139	10/14	53.4/6.25
Sp38†	ZP_00110492	Carbon dioxide concentrating mechanism/carboxysome shell protein	79	4/7	10.9/6.10
Sp48†	ZP_00109192	Ferredoxin–NADP reductase (FNR)	87	6/9	48.4/6.78
Sp56†	ZP_00108261	Phycobilisome linker polypeptide	136	8/11	29.3/6.64
Amino acid biosynthesis					
Sp8†	ZP_00111241	Dihydroxyacid dehydratase/phosphogluconate dehydratase	103	7/8	59.7/5.58
Sp14†	ZP_00112058	Phosphoglycerate dehydrogenase	258	17/20	56.0/5.69
Sp21†	ZP_00106298	Aspartate/tyrosine/aromatic aminotransferase	138	10/16	45.4/5.34
Sp26†	ZP_00109678	Asp–$tRNA^{Asn}$/Glu–$tRNA^{Gln}$ amidotransferase A subunit	78	5/6	52.6/5.73
Sp29†	ZP_00110976	Isocitrate/isopropylmalate dehydrogenase	96	7/13	38.8/5.15
Sp30†	ZP_00110649	Argininosuccinate synthase	91	6/9	44.1/5.31
Sp43†	ZP_00106143	Ketol–acid reductoisomerase	115	6/6	36.0/5.31
Sp45	ZP_00107756	Cysteine synthase	164	9/10	34.0/5.84
Sp46†	ZP_00112380	Cysteine synthase	138	8/9	33.8/5.91
Protein synthesis					
Sp22†	ZP_00105922	$tRNA^{Ser}$ synthetase	140	9/11	47.8/5.37
Sp23	ZP_00107088	GTPases translation elongation factors	96	7/12	44.7/5.46
Sp34†	ZP_00110655	Translation elongation factor Ts	135	9/15	34.4/4.97
Post-translational modification, protein turnover, chaperones					
Sp1	ZP_00110397	ATPases with chaperone activity, ATP-binding subunit	94	10/11	91.2/5.50
Sp5	ZP_00107203	Dipeptidyl aminopeptidases/acylaminoacyl-peptidases	107	9/15	70.4/4.85
Sp6†	ZP_00107038	Molecular chaperone	208	16/23	68.1/4.84
Sp9†	ZP_00107939	Chaperonin GroEL (HSP60 family)	200	15/21	57.8/5.17
Sp10†	ZP_00110155	Chaperonin GroEL (HSP60 family)	243	18/27	58.9/4.90
Sp16†	ZP_00110931	Predicted Zn-dependent peptidases	54	4/6	60.1/6.75
Sp17†	ZP_00105816	Subtilisin-like serine proteases	92	6/9	37.7/6.09
Sp49†	ZP_00106954	Peptidyl–prolyl *cis–trans* isomerase (rotamase)	86	6/11	38.8/5.01
Sp60†	ZP_00111066	Peptidyl–prolyl *cis–trans* isomerase (rotamase)	83	5/9	21.0/6.07
Sp53†	ZP_00108611	Protease subunit of ATP-dependent Clp proteases	93	5/6	24.3/5.61
Sp54†	ZP_00345050	Putative intracellular protease/amidase	103	5/5	24.3/5.42
Sp57†	ZP_00110671	Predicted amidohydrolase	83	5/8	30.0/6.25

Table 1. cont.

Spot no.*	Accession no.	Gene product‡	Mowse score	Matched peptide/total	Putative M_r/pI
Nucleotide metabolism					
Sp4†	ZP_00106148	Polyribonucleotide nucleotidyltransferase (polynucleotide phosphorylase)	108	7/7	78.3/5.21
Sp13†	ZP_00109890	AICAR transformylase/IMP cyclohydrolase PurH	94	6/7	54.2/5.21
Sp35†	ZP_00110674	IMP dehydrogenase/GMP reductase	90	6/10	40.4/5.56
Lipid metabolism					
Sp3†	ZP_00108549	Acyl-coenzyme A synthetases/AMP-(fatty) acid ligases	171	11/11	73.5/5.44
Sp27†	ZP_00111289	Biotin carboxylase	107	6/6	49.6/5.81
Reductases					
Sp42	ZP_00107451	Predicted oxidoreductases	76	5/8	35.8/5.30
Sp44†	ZP_00107560	Putative NADP-dependent oxidoreductases	85	5/6	37.6/5.73
Sp51†	ZP_00109876	Peroxiredoxin	113	7/11	27.2/5.09
Cell motility and secretion, cell envelope biogenesis					
Sp15†	ZP_00110813	Large exoproteins involved in haem utilization or adhesion	89	6/7	67.1/6.42
Sp55†	ZP_00111758	Secreted and surface protein containing fasciclin-like repeats	112	7/11	28.7/8.92
Inorganic nutrients (transport and metabolism)					
Sp11†	ZP_00112339	Nitrogenase molybdenum–iron protein, alpha and beta chains	75	5/6	57.8/5.53
Sp37†	ZP_00107658	Oxyanion-translocating ATPase	110	7/10	40.0/5.68
Sp58†	ZP_00111829	Rhodanese-related sulfurtransferase	105	7/16	21.8/4.77
Sp59	ZP_00108516	Superoxide dismutase	90	7/9	22.4/5.67
Signal transduction					
Sp50†	ZP_00108081	Phage shock protein A (IM30); suppresses σ^{54}-dependent transcription	113	7/9	28.7/5.08

*Spot number refers to the number of the protein given in Fig. 1

†Proteins identified in the proteome of *Nostoc punctiforme* for the first time.

‡Gene product names are those given in the NCBI database (http://www.ncbi.nlm.nih.gov/) with four exceptions: Sp31, 33, 48, and 56 (see Results and Discussion).

repairing misfolded proteins, facilitate the refolding of proteins, and prevent protein aggregation.

NifK (Sp11; one of the larger subunits of dinitrogenase) was detected in total extracts from vegetative filaments (Fig. 1a). This shows that the proteomic procedure allows detection of proteins present only in 5–10 % of the cells (heterocysts). The nitrogen fixed in heterocysts is efficiently scavenged into amino acids to avoid inhibition by the ammonium ions synthesized after fixation, and indeed, 24 proteins (Table 1) were identified as being involved in amino acid and protein biosynthesis or in protein modification. In fact, this was one of the largest categories based on the number of proteins identified under N_2-fixing conditions (Fig. 2).

Proteins involved in energy and carbohydrate metabolism together represented the largest fraction of the proteome. Both photosynthesis and diazotrophy are associated with high energy demands, the latter requiring cell differentiation, synthesis of the nitrogenase complex and 12–16 ATPs per N_2 molecule fixed. Many of the components of the Calvin cycle were identified (Sp2, 7, 12, 28, 30, 36 and 52) and most of them were highly abundant (Sp7 and 36; Fig. 1). The oxidative pentose phosphate (OPP) pathway enzyme 6-phosphogluconate dehydrogenase (6PGD, Sp25) was highly abundant, while glucose-6-phosphate dehydrogenase (G6PD) (Summers *et al.*, 1995; Hagen & Meeks, 2001) was not identified. The latter finding corroborates earlier data showing that light inhibits the activity of G6PD in *Anabaena* sp. PCC 7120 (Gleason, 1996). In contrast, G6PD protein levels in *Nostoc* symbionts are up-shifted under the dark heterotrophic conditions offered in plant symbiosis (Ekman *et al.*, 2006). This result is expected since the OPP pathway is a major route for providing reductants (NADPH) to nitrogenase in heterocysts (Winkenbach & Wolk, 1973; Bothe & Neuer 1988). This illustrates the metabolic flexibility of the genus and the findings are supported by our data.

Two proteins (Mp19 and Sp56), annotated as hypothetical proteins in the *Nostoc* sp. ATCC 29133 (*N. punctiforme*) genome, are listed in the energy metabolism category. Mp19 contains a conserved domain of MSP (manganese-stabilizing protein) and shows high similarity (95 %) to an MSP protein (P13907) in *Nostoc* sp. PCC 7120, a 33 kDa protein associated with photosytem II in plants and cyanobacteria (Borthakur & Haselkorn, 1989). We propose that Sp56 (Table 1) is equivalent to a phycobilisome (PBS) rod–core

Table 2. Proteins identified in the membrane fraction (Mp) of *Nostoc* sp. PCC 73102 and prediction of putative signal peptides and transmembrane helices

Spot no.*	Accession no.	Gene product‡	Mowse score	Matched peptides/total	M_r/pI	Position of transmenbrane helices	Putative cleavage site	
							Position	Sequence
Unknown proteins								
Mp1†	ZP_00108704	Uncharacterized conserved protein	97	8/11	81.6/4.19	–	23	VTT-AD
Mp6†	ZP_00105969	Hypothetical protein Npun02008175	118	11/23	67.9/5.12	–	32	VQA-LP
Mp10	ZP_00106112	Hypothetical protein Npun02008102	111	9/18	61.6/4.86	21–23	45	AIA-AP
Mp14†	ZP_00111668	Hypothetical protein Npun02001042	62	5/10	51.0/5.77			
Cell motility and secretion, cell envelope biogenesis (outer-membrane proteins)								
Mp2†	ZP_00110047	Outer-membrane receptor proteins, mostly Fe transport	98	9/17	87.1/4.56	–		
Mp3†	ZP_00107621	Outer-membrane protein/protective antigen OMA87	229	19/24	91.5/5.17	–		
Mp4†	ZP_00111842	Haemolysin activation/secretion protein	175	12/14	63.9/5.10	–	35	AQA-VD
Mp7	ZP_00107013	Outer-membrane protein	136	13/30	78.9/9.43	–	31	AWA-GT
Mp11	ZP_00108185	Flagellin and related hook-associated proteins	188	14/21	60.2/4.83	12–34	35	ALA-AE
Mp18†	ZP_00109958	Outer-membrane protein	141	9/11	51.5/6.67	5–27	28	ASA-VT
Cell motility and secretion, cell envelope biogenesis (related to surface structure)								
Mp8†	ZP_00111132	Type II secretory pathway, component HofQ	100	8/12	83.3/9.41	7–26	27	VWA-QI
Mp9	ZP_00110600	Periplasmic protein involved in polysaccharide export	95	7/12	51.2/7.88	–	34	SLA-QG
Mp22†	ZP_00106426	Methyl-accepting chemotaxis protein	67	4/7	21.3/9.14	5–27	25	SDA-PA
Energy metabolism (photosynthesis, respiratory electron transport, ATP synthase)								
Mp5†	ZP_00111405	F_0F_1-type ATP synthase, alpha subunit	150	12/20	54.5/5.02	–		
Mp12†	ZP_00107336	F_0F_1-type ATP synthase, beta subunit	171	12/17	51.8/4.98	–		
Mp21†	ZP_00111407	F_0F_1-type ATP synthase, subunit b	92	6/13	20.4/5.20	–		
Mp15†	ZP_00111098	NADH : ubiquinone oxidoreductase 49 kDa subunit 7	138	10/14	45.4/5.91	–		
Mp20†	ZP_00110637	NADH : ubiquinone oxidoreductase 20 kDa subunit and related Fe–S oxidoreductases	91	5/6	27.6/6.75	46–68	–	–
Mp19†	ZP_00111456	Manganese-stabilizing protein/photosystem II polypeptide	148	10/21	30.2/4.92	5–27	25	SDA-PA
Inorganic nutrients (transport and metabolism)								
Mp13†	ZP_00107283	ABC-type phosphate transport system, periplasmic component	89	6/10	41.2/7.63	–	31	ACG-GQ
Mp17†	ZP_00105675	ABC-type nitrate/sulfonate/bicarbonate transport systems, periplasmic components	97	7/13	29.4/5.62	–	12	AGG-LD
Mp16†	ZP_00109838	ABC-type branched-chain amino acid transport systems, periplasmic component	163	10/12	45.2/6.57	–		

*Spot number refers to the number of the protein given in Fig. 1

†Proteins identified in the proteome of *Nostoc punctiforme* for the first time.

‡Gene product names are those given in the NCBI database (http://www.ncbi.nlm.nih.gov/) with one exception: Mp19 (see Results and Discussion).

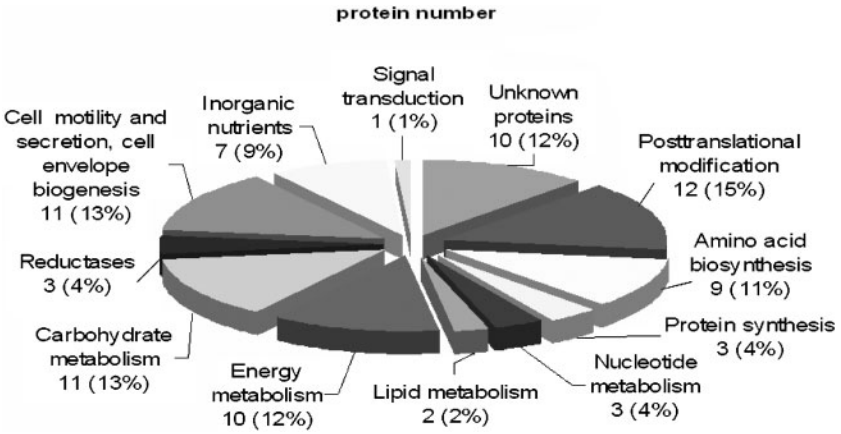

Fig. 2. Categories and relative proportion of the proteins identified in the proteome of *Nostoc* sp. PCC 73102. The 82 proteins identified from the two fractions (Sp and Mp) were divided into categories, based on their expected functions. The number of proteins identified in each category is given, and also expressed as percentage of the total number of identified proteins (in parentheses).

linker polypeptide due to the high homology to such proteins (67 % identity to NP440438 in *Synechocystis* sp. PCC 6803), and that it contains a PBS-linker domain as does protein Sp48, a ferredoxin : $NADP^+$ reductase (FNR).

The unknown protein category was prominent. Some of its protein members were abundant, such as Mp6, Mp10, Sp33 and Sp40 (Fig. 1), suggesting that they may have vital functions. SLH (S-layer homology) domains are present in

Table 3. Putative Trx targets identified in soluble (Sp) and membrane-bound (Mp) proteins of *Nostoc* sp. PCC 73102

Spot no.*	Gene product	Cys†	References‡
Carbohydrate metabolism			
Sp7	Transketolase	5/4/4	[1, 2]
Sp25	6-phosphogluconate dehydrogenase	4/4/2	[2]
Sp28	Fructose-1,6-bisphosphatase	5/5/0	[1]
Sp31	PRK	4/4/4	[1, 3]
Sp36	Fructose/tagatose bisphosphate aldolase	3/3/1	[3, 4]
Sp52	Ribose 5-phosphate isomerase	1/0/0	[1, 3]
Energy metabolism			
Sp20	Rubisco large subunit	7/6/6	[1, 3, 5]
Mp5	F_0F_1-type ATP synthase, alpha subunit	1/1/1	[3, 4]
Amino acid biosynthesis			
Sp8	Dihydroxyacid dehydratase/phosphogluconate dehydratase	6/6/4	[1, 3]
Sp29	Isocitrate/isopropylmalate dehydrogenase	3/3/2	[3]
Sp30	Argininosuccinate synthase	2/2/0	[5]
Sp43	Ketol-acid reductoisomerase	2/2/2	[3]
Sp46	Cysteine synthase	1/1/0	[2, 4]
Protein synthesis, modification and chaperones			
Sp23	GTPases translation elongation factors	1/1/1	[2, 3, 5]
Sp6	Molecular chaperone	4/4/2	[2, 3, 6]
Sp9	Chaperonin GroEL	1/1/1	[3–5]
Sp1	ATPases with chaperone activity, ATP-binding subunit	1/1/1	[2]
Lipid metabolism			
Sp3	Acyl-coenzyme A synthetases/ AMP-(fatty) acid ligases	7/7/6	[3]
Reductase			
Sp51	Peroxiredoxin	4/4/2	[1, 4]

*Spot number refers to the number of the protein given in Fig. 1 and Tables 1 and 2.

†Cys: total number of cysteines in the protein/number of conserved cysteines in cyanobacteria/number of conserved cysteines in plant and/or green alga.

‡References: [1] Marchand *et al.* (2004); [2] Balmer *et al.* (2003); [3] Lemaire *et al.* (2004); [4] Balmer *et al.* (2004); [5] Lindahl & Florencio (2003); [6] Wong *et al.* (2004).

two of the abundant proteins (Mp6 and Mp10). The latter, Mp10, also shows high similarity (78 % identity) to a component in the phosphotransferase system of *Anabaena* ATCC 29413 (ZP_00158087). The gene downstream of Mp10 is predicted to encode a signal transduction histidine kinase and the gene upstream of Mp6 is predicted to encode a cAMP-binding protein. Therefore, these two proteins may be involved in signal transduction and/or adhesion, which makes them candidate components for intracellular communication in cyanobacteria. The high homology (92 % similarity) between the hypothetical protein Sp39 and the ALAase, an actin-like ATPase in *Anabaena variabilis*, suggests a role in morphogenesis in *Nostoc* PCC 73102.

The protein Sp33 was identified as an 'uncharacterized protein conserved in bacteria' showing similarity to the 'akinete marker' protein (AvaK) in *A. variabilis*, particularly in the conserved N-terminus (93 % similarity) (Zhou & Wolk, 2002). AvaK is proposed to be exclusive to akinetes (Argueta *et al.*, 2004), but as Sp33 was strongly expressed in the actively growing vegetative cells in *Nostoc*, other functions related to this protein must be considered. The PRC domain of Sp33 is widespread and ancient, and has been associated with a variety of biological processes, ranging from RNA processing to photosynthesis (Anantharaman & Aravind, 2002).

Signal peptides were identified in several of the membrane category proteins, which confirms their membrane location. This category was primarily related to cell motility, secretion and cell envelope biogenesis (Table 2). Six proteins were identified as outer-membrane proteins, and Mp7 and Mp18 were predicted to contain two conserved domains: the OEP (outer-membrane efflux protein) domain and the TolC protein domain. The OEP family forms trimeric channels allowing export of a variety of substrates in bacteria (Johnson & Church, 1999), whereas TolC proteins are outer-membrane, multifunctional proteins producing peptide antibiotics (Delgado *et al.*, 1999) and maintaining membrane integrity (Bernadac *et al.*, 1998). Located upstream of Mp18 are two genes encoding protein components of the ABC-type antimicrobial peptide transport system, and we predict that Mp7 and Mp18 may function in outer-membrane secretion systems and antimicrobial signalling. As the outer-membrane proteins Mp3 and Mp4 both contain a bacteria surface antigen (Bac surface Ag) domain, a family including protective surface antigens (Ruffolo & Adler, 1996), they are likely to serve a protective function in *Nostoc* sp. PCC 73102. Mp4 also contains a domain homologous to FhaC (homolysin activation/secretion protein) known to be a membrane transporter with a role in intracellular trafficking and secretion (Guedin *et al.*, 2000). Due to the unique symbiotic competence of *Nostoc* sp. PCC 73102, we propose that Mp9, a polysaccharide biosynthesis/export protein, and Sp19 phosphomannomutase, are related to lipopolysaccharide biosynthesis (Maroda & Valvano, 1993), and possibly of symbiotic significance.

Trxs are small, multifunctional, widely distributed redox-active proteins operating through a reversible reduction of intra- or intermolecular disulfides (Balmer *et al.*, 2004; Lindahl & Florencio, 2003). Recently, numerous putative Trx targets have been identified in plant chloroplasts (Balmer *et al.*, 2003, 2004; Marchand *et al.*, 2004), in green algae (Lemaire *et al.*, 2004) and in cyanobacteria (Lindahl & Florencio, 2003). About 200 proteins appear to be linked to Trx in plants (Buchanan & Balmer, 2005). Since 19 of the proteins identified here in *Nostoc* were predicted Trx targets (Table 3), Trx-linked processes also appear to be essential in *Nostoc*. Trx targets were detected in fructose-1,6-bisphosphatase, translation elongation factors, the Rubisco large subunit, chaperones, ATPase and peroxiredoxins of both plants and cyanobacteria (Table 3). All putative Trx targets identified here contained at least one cysteine, and 16 of them contained cysteines in conserved positions (Table 3). The presence of single cysteines in some of the proteins (Sp23, 46, 9, 1 and Mp5) suggests a regulatory function, although biochemical evidence is required to verify the regulation.

In summary, the protein profiling and expression data presented here provide an expanded view of the proteins expressed and involved in growth under photoautotrophic and diazotrophic conditions and illustrate the complexity of the proteome and lifestyle of *Nostoc* sp. PCC 73102 (*N. punctiforme*).

ACKNOWLEDGEMENTS

The Swedish Research Council and the Swedish Research Council for Environment, Agricultural Science and Spatial Planning are gratefully acknowledged for financial support.

REFERENCES

Adam, Z. & Clarke, A. K. (2002). Cutting edge of chloroplast proteolysis. *Trends Plant Sci* **7**, 451–456.

Adams, D. G. & Duggan, P. S. (1999). Heterocyst and akinete differentiation in cyanobacteria. *New Phytol* **144**, 3–33.

Agrawal, G. K., Rakwal, R., Yonekura, M., Kubo, A. & Saji, H. (2002). Proteome analysis of differentially displayed proteins as a tool for investigating ozone stress in rice (*Oryza sativa* L.) seedlings. *Proteomics* **2**, 947–959.

Anantharaman, V. & Aravind, L. (2002). The PRC-barrel: a widespread, conserved domain shared by photosynthetic reaction center subunits and proteins of RNA metabolism. *Genome Biol* **3**, research0061, 1–9.

Argueta, C., Yuksek, K. & Summers, M. (2004). Construction and use of GFP reporter vectors for analysis of cell-type-specific gene expression in *Nostoc punctiforme*. *J Microbiol Methods* **59**, 181–188.

Balmer, Y., Koller, A., del Val, G., Manieri, W., Schurmann, P. & Buchanan, B. B. (2003). Proteomics gives insight into the regulatory function of chloroplast thioredoxins. *Proc Natl Acad Sci U S A* **100**, 370–375.

Balmer, Y., Venswl, W. H., Tanaka, C. K., Hurkman, W. J., Gelhaye, E., Rouhier, N., Jacquot, J. P., Manieri, W., Schurmann, P. & other authors (2004). Thioredoxin links redox to the regulation of

fundamental processes of plant mitochondria. *Proc Natl Acad Sci U S A* **101**, 2642–2647.

Berggren, K., Chernokalskaya, E., Steinberg, T. H., Kemper, C., Lopez, M. F., Diwu, Z., Haugland, R. P. & 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.

Bernadac, A., Gavioli, M., Lazzaroni, J. C., Raina, S. & Lloubes, R. (1998). *Escherichia coli tol-pal* mutants form outer membrane vesicles. *J Bacteriol* **180**, 4872–4878.

Borthakur, D. & Haselkorn, R. (1989). Nucleotide sequence of the gene encoding the 33 kDa water oxidizing polypeptide in *Anabaena* sp. strain PCC 7120 and its expression in *Escherichia coli*. *Plant Mol Biol* **13**, 427–439.

Bothe, H. & Neuer, G. (1988). Electron donation to nitrogenase in heterocysts. *Methods Enzymol* **167**, 496–501.

Buchanan, B. B. & Balmer, Y. (2005). Redox regulation: a broadening horizon. *Annu Rev Plant Biol* **56**, 187–220.

Bukau, B. & Horwich, A. L. (1998). The Hsp70 and Hsp60 chaperone machines. *Cell* **92**, 351–366.

Delgado, M. A., Solbiati, J. O., Chiuchiolo, M. J., Farias, R. N. & Salomon, R. A. (1999). *Escherichia coli* outer membrane protein TolC is involved in production of the peptide antibiotic microcin J25. *J Bacteriol* **181**, 1968–1970.

Dodds, K., Gudder, D. A. & Mollenhauer, D. (1995). The ecology of *Nostoc*. *J Phycol* **31**, 2–18.

Ekman, M., Tollbäck, P., Klint, J. & Bergman, B. (2006). Protein expression profiles in an endosymbiotic cyanobacterium revealed by a proteomic approach. *Mol Plant Microbe Interact* **19**, 1251–1261.

Fink, A. L. (1999). Chaperone-mediated protein folding. *Physiol Rev* **79**, 425–449.

Fulda, S., Huang, F., Nilsson, F., Hagemann, M. & Norling, B. (2000). Proteomics of *Synechocystis* sp. PCC 6803: identification of periplasmic proteins in cells grown at low and high salt concentrations. *Eur J Biochem* **267**, 5900–5907.

Gleason, F. K. (1996). Glucose-6-phosphate dehydrogenase from the cyanobacterium, *Anabaena* sp. PCC 7120: purification and kinetics of redox modulation. *Arch Biochem Biophys* **334**, 277–283.

Guedin, S., Willery, E., Tommassen, J., Fort, E., Drobecq, H., Locht, C. & Jacob-Dubuisson, F. (2000). Novel topological features of FhaC, the outer membrane transporter involved in the secretion of the *Bordetella pertussis* filamentous hemagglutinin. *J Biol Chem* **275**, 30202–30210.

Hagen, K. D. & Meeks, J. C. (2001). The unique cyanobacterial protein OpcA is an allosteric effector of glucose-6-phosphate dehydrogenase in *Nostoc punctiforme* ATCC 29133. *J Biol Chem* **276**, 11477–11486.

Hunsucker, S. W., Klage, K., Slaughter, S. M., Potts, M. & Helm, R. F. (2004). A preliminary investigation of the *Nostoc punctiforme* proteome. *Biochem Biophys Res Commun* **317**, 1121–1127.

Johnson, J. M. & Church, G. M. (1999). Alignment and structure prediction of divergent protein families: periplasmic and outer membrane proteins of bacterial efflux pumps. *J Mol Biol* **287**, 695–715.

Kliebenstein, D. J., Monde, R. A. & Last, R. L. (1998). Superoxide dismutase in Arabidopsis: an eclectic enzyme family with disparate regulation and protein localization. *Plant Physiol* **118**, 637–650.

Klint, J., Ran, L., Rasmussen, U. & Bergman, B. (2006). Identification of developmentally regulated proteins in cyanobacterial hormogonia using a proteomic approach. *Symbiosis* **41**, 87–95.

Lee, J. O., Rieu, P., Arnaout, M. A. & Liddington, R. (1995). Crystal structure of the A domain from the alpha subunit of integrin CR3 (CD11b/CD18). *Cell* **80**, 631–638.

Lemaire, S. D., Guillon, B., Le Marechal, P., Keryer, E., Miginiac-Maslow, M. & Decottigniesv, P. (2004). New thioredoxin targets in the unicellular photosynthetic eukaryote *Chlamydomonas reinhardtii*. *Proc Natl Acad Sci U S A* **101**, 7475–7480.

Lindahl, M. & Florencio, F. J. (2003). Thioredoxin-linked processes in cyanobacteria are as numerous as in chloroplasts, but targets are different. *Proc Natl Acad Sci U S A* **100**, 16107–16112.

Lodish, H., Berk, A., Zipursky, S. L., Matsudaira, P., Baltimore, D. & Darnell, J. (1999). *Molecular Cell Biology*, pp. 708–709. New York: W. H. Freeman.

Marchand, C., Le Marechal, P., Meyer, Y., Miginiac-Maslow, M., Issakidis-Bourguet, E. & Decottignies, P. (2004). New targets of *Arabidopsis* thioredoxins revealed by proteomic analysis. *Proteomics* **4**, 2696–2706.

Maroda, C. L. & Valvano, M. A. (1993). Identification, expression, and DNA sequence of the GDP-mannose biosynthesis genes encoded by the O7 rfb gene cluster of strain VW187 (*Escherichia coli* O7 : K1). *J Bacteriol* **175**, 148–158.

Meeks, J. C. & Elhai, J. (2002). Regulation of cellular differentiation in filamentous cyanobacteria in free-living and plant-associated symbiotic growth states. *Microbiol Mol Biol Rev* **66**, 94–121.

Meeks, J. C., Elhai, J., Thiel, T., Potts, M., Larimer, F., Lamerdin, J., Predki, P. & Atlas, R. (2001). An overview of the genome of *Nostoc punctiforme*, a multicellular, symbiotic cyanobacterium. *Photosynthesis Research* **70**, 85–106.

Nielsen, H., Engelbrecht, J., Brunak, S. & von Heijne, G. A. (1997). Neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Int J Neural Syst* **8**, 581–599.

Nouwens, A. S., Cordwell, S. J., Larsen, M. R., Molloy, M. P., Gillings, M., Willcox, M. D. & Walsh, B. J. (2000). Complementing genomics with proteomics: the membrane subproteome of *Pseudomonas aeruginosa* PAO1. *Electrophoresis* **21**, 3797–3809.

Rai, A. N., Söderbäck, E. & Bergman, B. (2000). Cyanobacterium-plant symbioses. *New Phytol* **147**, 449–481.

Rai, A. N., Bergman, B. & Rasmussen, U. (2002). *Cyanobacterial-Plant Symbiosis*. Dordrecht, the Netherlands: Kluwer.

Rippka, R., Deruells, J., Waterbury, J. B., Herdman, M. & Stanier, R. (1979). Generic assignment, strain histories and properties of pure cultures of cyanobacteria. *J Gen Microbiol* **111**, 1–61.

Ruffolo, C. G. & Adler, B. (1996). Cloning, sequencing, expression, and protective capacity of the oma87 gene encoding the *Pasteurella multocida* 87-kilodalton outer membrane antigen. *Infect Immun* **64**, 3161–3167.

Stanier, R. Y., Kunisawa, R., Mandel, M. & Cohen-Blazire, G. (1971). Purification properties of unicellular blue-green algae (order Chlorococcales). *Bacteriol Rev* **35**, 171–205.

Summers, M. L., Wallis, J. G., Campbell, E. L. & Meeks, J. C. (1995). Genetic evidence of a major role for glucose-6-phosphate dehydrogenase in nitrogen fixation and dark growth of the cyanobacterium *Nostoc* sp. strain ATCC 29133. *J Bacteriol* **177**, 6184–6194.

von Heijne, G. (1988). Transcending the impenetrable: how proteins come to terms with membranes. *Biochim Biophys Acta* **947**, 307–333.

Whittaker, C. A. & Hynes, R. O. (2002). Distribution and evolution of von Willebrand/integrin A domains: widely dispersed domains with roles in cell adhesion and elsewhere. *Mol Biol Cell* **13**, 3369–3387.

Winkenbach, F. & Wolk, C. P. (1973). Activities of enzymes of the oxidative and the reductive pentose phosphate pathways in heterocysts of a blue-green alga. *Plant Physiol* **52**, 480–483.

Wong, J. H., Cai, N., Balmer, Y., Tanaka, C. K., Vensel, W. H., Hurkman, W. J. & Buchanan, B. B. (2004). Thioredoxin targets of developing wheat seeds identified by complementary proteomic approaches. *Phytochemistry* **65**, 1629–1640.

Zheng, B., Halperin, T., Hruskova-Heidingsfeldova, O., Adam, Z. & Clarke, A. K. (2002). Characterization of chloroplast Clp proteins in *Arabidopsis*: localization, tissue specificity and stress responses. *Physiol Plant* **114**, 92–101.

Zhou, R. & Wolk, C. P. (2002). Identification of an akinete marker gene in *Anabaena variabilis*. *J Bacteriol* **184**, 2529–2532.

Edited by: M. Hecker