Electronic Journal of Biotechnology ISSN: 0717-3458
Vol.10 No.4, Issue of October 15, 2007

Received September 21, 2006 / Accepted April 17, 2007
DOI: 10.2225/vol10-issue4-fulltext-2
RESEARCH ARTICLE

Experimental and bioinformatic approaches for analyzing and visualizing cyanobacterial nitrogen and hydrogen metabolism

Röbbe Wünschiers*#
Institute for Genetics
University of Cologne
Germany
Tel: 49 39482 760112
Fax: 49 39482 760199
E-mail: robbe.wunschiers@uni-koeln.de

Rikard Axelsson
Swedenborgsgatan 41 C 753 34
Uppsala, Sweden
Tel: 46 18 4891216
Fax: 46 18 4891022
E-mail: r.axelsson@q-med.com

Martin Vellguth
Institute for Genetics
University of Cologne
Germany
Zülpicher Str. 47
50674 Cologne, Germany
Tel: 49 221 4705661
Fax: 49 221 4705975
E-mail: martin.vellguth@uni-koeln.de

Peter Lindblad
Department Photochemistry and Molecular Science
The Ångström Laboratories
Uppsala University
Box 523, 751 20 Uppsala, Sweden
Tel: 46 18 471 2826
Fax: 46 18 471 6844
E-mail: peter.lindblad@fotomol.uu.se

Financial support: This work is part of the BMBF-funded Cologne University Bioinformatics Center (CUBIC) and was financially supported by grants to PL from the Swedish Research Council, the Swedish Energy Agency, and the EU/NEST Project SOLAR-H (contract 516510).

Keywords: bioinformatics, cyanobacteria, DNA-microarrays, gene expression, metabolic pathways, nitrogen availability, nitrogen fixation.

Present address: #SunGene GmbH & Co KGaA, Correnstrasse 3, 06466 Gatersleben, Germany.

Abbreviations: ATP: adenosine 5'-triphosphate
BSA: bovine serum albumin
CAB: chlorophyll a/b-binding
ELIP: early light-induced proteins
GC: guanine/cytosine nucleotides
HLIP: high light-induced stress proteins
KEGG: Kyoto encyclopedia of genes and genomes
ORF: open reading frame

Many cyanobacteria are capable of utilizing light energy for nitrogen fixation. As a by-product of this nitrogenase mediated catalysis, hydrogen gas is produced. Several approaches to increase hydrogen production from cyanobacteria exist. Usually, these approaches are non-targeted. Here we exemplify how DNA-microarray based gene-expression analysis and bioinformatic visualization techniques can be used to analyze nitrogen and hydrogen metabolism from the filamentous, heterocyst forming cyanobacterium *Nostoc* PCC 7120. We analyzed the expression of 1249 genes from major metabolic categories under nitrogen fixing and non-nitrogen fixing growth. Of the selected genes, 494 show a more than 2-fold expression difference in the two conditions analyzed. Under nitrogen-fixing conditions 465 genes, mainly involved in energy metabolism, photosynthesis, respiration and nitrogen-fixation, were found to be stronger expressed, whereas

*Corresponding author

only 29 genes showed a stronger expression under non-nitrogen fixing conditions. To help understanding probe hybridization, all expression data were correlated with potential target secondary structures and probe GC-content. For the first time the expression of high light-induced stress proteins (HLIP-family) is shown to be linked to the nitrogen availability.

Cyanobacteria play an increasing in biotechnology (Singh et al. 2005). One potential but not yet feasible application is the photobiological production of hydrogen gas (Wünschiers and Lindblad, 2003; Dutta et al. 2005; Prince and Kheshgi, 2005). Cyanobacterial hydrogen metabolism is closely linked to dinitrogen assimilation in nitrogen-fixation. All known nitrogenases evolve 1 to 7.5 mole hydrogen gas as they convert one mole dinitrogen to two mole ammonia. This apparent loss of energy is circumvented by uptake hydrogenases that allow the cyanobacteria to regain energy. Mutants which lack uptake hydrogenases and thus produce substantial amounts of hydrogen gas have successfully been engineered (Happe et al. 2000; Masukawa et al. 2002). The genetic and metabolic regulatory networking behind photobiological hydrogen production is only understood in fragments.

Heterocystous nitrogen-fixing cyanobacteria like *Nostoc* sp. strain PCC 7120 (*Nostoc* PCC 7120; formerly *Anabaena* PCC 7120) respond to the deprivation of combined nitrogen with morphological changes, *i.e.*, the formation of heterocysts. These specialized non-dividing cells develop more or less equidistantly along the filaments with a ratio of about one heterocyst per 10 vegetative cells. In comparison to vegetative cells, the heterocyst is larger and more rounded. It provides an environment with low oxygen partial pressure since it lacks oxygen-evolving photosystem II activity and has a higher respiration rate (Böhme, 1998). Furthermore, it is surrounded by a thick glycolipidic cell wall that reduces the diffusion of oxygen. Heterocysts

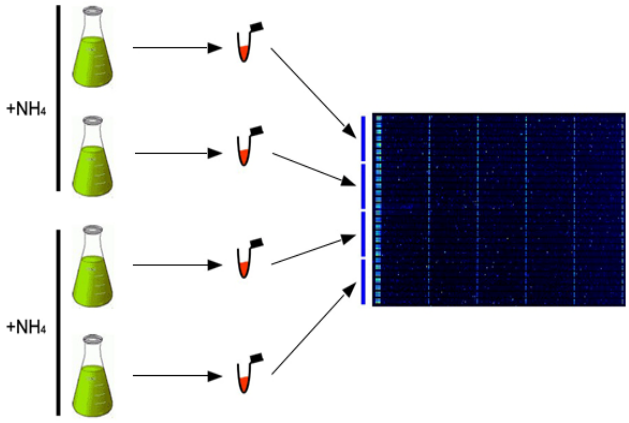

Figure 1. Experimental setup. *Nostoc* PCC 7120 batch cultures were grown under either nitrogen fixing (-NH_4) or non-nitrogen fixing conditions (+NH_4), harvested and subjected to total-RNA purification. After removal of rRNA by means of affinity chromatography, labeled and fragmented cRNA was hybridized to four individual arrays at single physical slides.

provide amino acids and receive carbohydrates from their neighboring vegetative cells. The differentiation process begins within a few hrs after combined-nitrogen deprivation, *i.e.* growth on dinitrogen as sole nitrogen source, and requires approximately 24 hrs to complete. Single components involved in this process have been described (Golden and Yoon, 1998; Adams, 2000; Wolk, 2000; Meeks and Elhai, 2002; Herrero et al. 2004).

A key protein in heterocyst formation is NtcA (Herrero et al. 2001; Herrero et al. 2004), which is the global nitrogen regulator that controls, *e.g.*, the expression of genes essential for heterocyst development such as the ABC-type transporter genes *devABC* and the regulator genes *hetR* and *patS*. Among many other factors that have been demonstrated to be important for heterocyst development are enzymes that are involved in the formation of

Table 1. Analyzed ORFs. Number of ORFs analyzed and present in each metabolic category of *Nostoc* PCC 7120.

Metabolis class		Number of ORFs analized
A	Amino acid biosynthesis	64 of 113 (57%)
B	Biosynthesis of cofactors, prosthetic groups and carriers	14 of 152 (9%)
C	Cell envelope	15 of 77 (19%)
D	Cellular processes	47 of 96 (49%)
E	Central intermediary metabolism	40 of 72 (56%)
F	DNA-replication, recombination, and repair	29 of 105 (28%)
G	Energy metabolism	20 of 100 (20%)
H	Fatty acid, phospholipid and sterol metabolism	2 of 40 (5%)
I	Hypothetical	567 of 3573 (16%)
J	Other categories	204 of 678 (30%)
K	Photosynthesis and respiration	98 of 157 (62%)
L	Purines, pirymidines, nucleosides and nucleotides	2 of 58 (3%)
M	Regulatory functions	36 of 360 (10%)
N	Transcription	31 of 41 (76%)
O	Translation	13 of 200 (7%)
P	Transport and binding proteins	67 of 313 (21%)
Total		1249 of 6135 (20%)

heterocyst-specific glycolipids and polysaccharides. The driving forces for these structural changes occurring during heterocyst formation are metabolic requirements: the dinitrogen fixing multi-enzyme complex nitrogenase requires a low oxygen partial pressure and a high supply of ATP and reduction equivalents (Dixon and Kahn, 2004). ATP may be generated by either cyclic photophosphorylation or oxidative phosphorylation while low-potential electrons may be generated from the degradation of carbohydrates produced during photosynthesis (Haselkorn and Buikema, 1992). Obviously, profound regulatory events coincide with growth on dinitrogen.

A powerful tool to study gene expression and its regulation is the DNA-microarray technique. Previously, we analyzed the expression of individual genes and operons of Nostoc spp. that are involved in nitrogen metabolism (Wünschiers and Lindblad, 2003). Here we describe an oligonucleotide based DNA-microarray expression analysis, where each gene is covered by up to 10 unique probes. Until now, only few oligonucleotide DNA-microarray based gene-expression analyses with heterocystous cyanobacteria have been reported (Ehira et al. 2005; Imashimizu et al. 2005; Ehira and Ohmori, 2006; Higo et al. 2006; Lechno-Yossef et al. 2006). Only two among these focus on the effect of nitrogen sources on gene expression. While Ehira and Ohmori (2006) follow the time course of nitrogen deprivation, ending at 24 hrs after removal of anorganic nitrogen from growth medium, Lechno-Yossef et al. (2006) looked only at the expression differences after 14 hrs of such a shift. In our experiment we look at the effect at continuous growth, *i.e.*, fully shifted and equilibrated cultures.

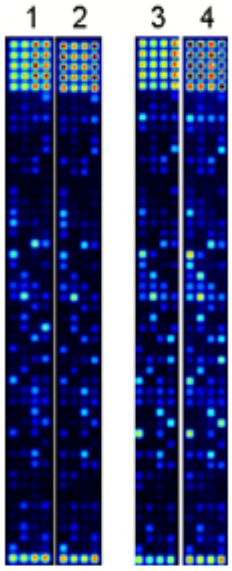

Figure 2. DNA-microarrays. Sections of four arrays used in the present analysis. The bright spots at the top and button of the arrays are controls used by the spot-finding software. (1) non-nitrogen fixing culture (2) non-nitrogen fixing culture, biological replicate (3) nitrogen fixing culture (4) nitrogen fixing culture, biological replicate.

Therefore we employed a novel, recently developed microarray technique where probe synthesis, hybridization, and signal detection take place in one device at strongly controlled physical conditions (Baum et al. 2003; Güimil et al. 2003). The expression data were used to (a) validate the technique employed and (b) obtain a global overview about the effect of growth on dinitrogen on *Nostoc* PCC 7120. For the latter, we set up a convenient data-processing

Table 2. Highly expressed ORFs under non-nitrogen fixing conditions. Annotated ORFs that show a more than 2-times higher expression level under non-nitrogen fixing than under nitrogen fixing conditions (the values are negative because nitrogen fixation was taken as reference). All ORFs annotated as putative, hypothetical, or unknown have been omitted. The key for metabolic classes can be found in Table 1.

ORF	Class	Subclass	Annotation	Expr. diff.	p-value
all0410	(A)	Aromatic amino acid family	tryptophan sysnthase beta subunit TrpB	-2.0	0.02
alr7354	(B)	Thioredoxin, glutaredoxin, and glutathione	glutathione S-transferase	-5.9	0.003
all0166	(E)	Polysaccharides and glycoproteins	alpha, alpha-trehalase	-6.9	0.003
asl0873	(J)	Adaptation and atipical conditions	CAB/ELIP/HLIP superfamilty	-4.4	0.1
asr3042	"	"	"	-11.3	0.08
asr3043	"	"	"	-7.9	0.02
asl0449	"	"	"	-2.3	0.04
all0168	"	Other	alpha-amylase	-9.5	0.001
all0275	"	"	glycerophosphoryl diester phospodiesterase	-3.4	0.1
all0167	"	"	maltooligosyltrehalose synthase	-2.3	0.02
all0875	"	"	probable alpha-glucanotransferase	-4.4	0.0005
all0412	"	"	putative zinc-binding oxidoreductase	-2.2	0.02
alr2405	(K)	Soluble electron carriers	flavodoxin	-7.6	0.003
all0258	"	"	plastocyanin precursor; PetE	-2.3	0.007
alr0702	(O)	Degradation of proteins, peptides, and glycopeptides	serine proteinase	-2.8	0.002
all2674	(P)	Protein modification and translation factors	ferrichrome-iron receptor	-2.1	0.09
all0322	"	"	sulfate-binding protein SbpA	-4.0	0.006

pipeline based on a MySQL database and a web-based graphical user interface. This front-end allows users to visualize gene-expression data on KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway maps.

MATERIALS AND METHODS

Strains and culture conditions

The cyanobacterium *Nostoc* sp. strain PCC 7120 (formerly *Anabaena* sp. strain PCC 7120) was grown on either dinitrogen (nitrogen fixing) or combined nitrogen (non-nitrogen fixing) in batch cultures. Non-nitrogen fixing conditions were obtained by growing cells in $BG11_0$ (Stanier et al. 1971) supplemented with 5 mM NH_4Cl and 10 mM HEPES (pH 7.5). Nitrogen-fixing conditions were obtained by growing cells in $BG11_0$. All cultures were grown in continuous white light (Thorn Polylux 4000 and Osram Warmton Warm Light 400 to 700 nm; 40 $\mu Em^{-2}s^{-1}$) at 26ºC with a magnetic stirrer being used to retain a homogeneous suspension. Cultures were harvested at mid-logarithmic growth phase.

RNA isolation

Total RNA was extracted from 100 ml cultures at A_{730nm} = 0.5. The cells were harvested by centrifugation at 6,000 x g for 10 min together with 50 ml crushed ice in a 250 ml centrifuge bottle. The cell pellet was frozen in liquid nitrogen and thawed on ice. The cells were suspended in resuspension buffer (0.3 M sucrose, 10 mM sodium acetate, pH 4.5), transferred to an Eppendorf tube and pelleted at 12,000 x g for 5 min. The pellet was suspended in 250 µl resuspension buffer with 75 µl 250 mM Na_2-EDTA, and the suspension was incubated on ice for 5 min. 375 µl lysis buffer (2% (w/v) SDS, 10 mM Na-acetate, pH 4.5) was added, followed by incubation at 65ºC for 3 min. 700 µl of 65ºC phenol was added to the lysed cells, followed by incubation at 65ºC for 3 min and then at -70ºC for 15 sec. The suspension was centrifuged at 12,000 x g for 5 min, the upper phase was collected and the hot phenol treatment was repeated twice, followed by an extraction with hot phenol:chloroform (1:1). 1/5 volume of 10 M LiCl and 2.5 volumes of 99.5% ethanol was added and the RNA was precipitated at -20ºC for 30 min. The pellet was washed with 80% ethanol, suspended in water and stored in aliquots at -70ºC. Quality and quantity of total RNA was analyzed with an Agilent 2100 Bioanalyser using the RNA 6000 Nano LabChip kit (Agilent Technologies, Boeblingen, Germany). If necessary, the RNA preparation was treated with 40 units of RNase-free DNase I (Amersham-Pharmacia Biotech, Uppsala, Sweden) according to the manufacturer's instructions. The average RNA yield per 100 ml culture at A_{730nm} = 0.5 was 50 µg.

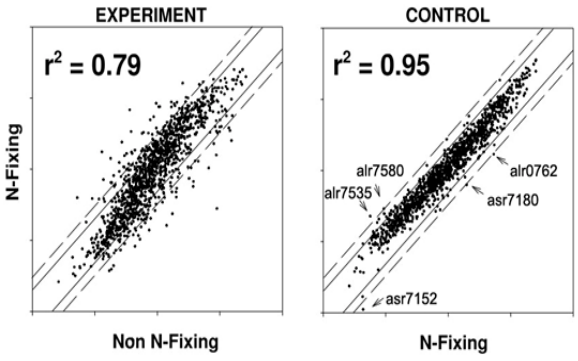

Figure 3. Data quality. For the double-logarithmic plots non-normalized raw data were used. The left plot shows the gene-expression differences between nitrogen fixing (ordinate) and non-nitrogen fixing, ammonia grown (abscissa) cultures of *Nostoc* PCC 7120. The right plot visualizes the gene-expression differences between two individual nitrogen fixing cultures. Significant outliers are marked by arrows. The straight and dashed lines represent 2- and 3-fold expression differences, respectively.

Preparation of biotin labeled, fragmented cRNA

From 10 µg of total RNA, low molecular weight RNA, *e.g.*, tRNA and 5S rRNA, were removed by size exclusion chromatography (MEGAclear kit, Ambion). To remove 16S and 23S rRNA, the MICROBExpress kit from Ambion was used. To increase the abundance of low expressed transcripts, the remaining RNA was linearly amplified by a modified Eberwine protocol (Eberwine et al. 1992) as follows. If not differently stated, all enzymes and chemicals were purchased from Invitrogene.

First strand synthesis. The pelleted RNA from the previous mRNA-enrichment steps was resuspended in 4.25 µl water and mixed with 1 µl of T7 random hexamers (0.5 µg/µl; 5'-GGC CAG TGA ATT GTA ATA CGA CTC ACT ATA GGG AGG CGG NNN NNN-3'). Following incubation at 70ºC for 10 min, 4ºC for 2 min and 23ºC for 5 min, 3.75 µl reaction mix (2 µl 5x first strand synthesis buffer (1 µl 0.1 M DTT, 0.5 µl 10 mM dNTP mix, 0.25 µl 40 U RNase OUT) and 200 U Superscript II polymerase) was added to the RNA/primer mix. First strand synthesis reaction was performed with the following temperature scheme: 37ºC for 20 min, 42ºC for 20 min, 50ºC for 15 min, 55ºC for 10 min and 65ºC for 15 min. After adding 0.5 µl RNase H the reaction mix was incubated for another 30 min at 37ºC and 2 min at 95ºC.

Second strand synthesis. The product of the first strand synthesis was mixed with 43.8 µl water and 15 µl 5x second strand synthesis buffer (20 U DNA-polymerase I, 1.5 µl 10 mM dNTP and 1 U RnaseH) and incubated for 2 hrs at 16ºC. After addition of 10 U T4 DNA-polymerase the reaction mix was first incubated at 16ºC for 15 min and then at 70ºC for 10 min.

Table 3. Photosynthesis and respiration. Highly expressed ORFs involved in photosynthesis or respiration. Expression levels are at least 30x above background (300).

ORF	Submetabolism	Annotation	+N	-N	p-value
all0138	PS II	CP47 protein	13326	9658	0.3
all0259	PS II	cytochrome c550	9311	13906	0.4
asr3845	PS II	cytochrome b559 alpha subunit	12527	10674	0.3
asr3846	PS II	cytochrome b559 beta subunit	7606	10848	0.5
alr3421	Cyt. b6/f	Plastoquinol plastocyanin reductase, cytochrome b6; PetB	16697	12112	0.03
alr3422	Cyt. b6/f	Plastoquinol plastocyanin reductase, aposytochrome subunit 4; PetD	9381	13981	0.03
all0109	PS I	Subunit III precursor; PsaF	19723	23286	0.007
all0107	PS I	Subunit XI; PsaL	18416	25334	0.04
all4121	Electron Carriers	ferredoxin-NADP(+) reductase	5205	12290	0.03
all4148	Electron Carriers	ferredoxin I	10984	14541	0.3
alr0021	Phycobilisome	allophycocyanin alpha subunit; ApcA	28417	24927	0.03
alr0022	Phycobilisome	allophycocyanin beta subunit; ApcB	25066	19930	0.2
alr0023	Phycobilisome	core linker protein Lc7.8; ApcC	13165	16575	0.2
alr0020	Phycobilisome	core membrane linker protein; ApcE	7285	18364	0.003
alr0529	Phycobilisome	phycosyanin alpha chain; CpcA	16927	18047	0.04
alr0528	Phycobilisome	phycosyanin beta chain; CpcB	20430	15579	0.4
alr0530	Phycobilisome	phycosyanin associated rod linker protein CpcC	13993	11404	0.1
alr0532	Phycobilisome	phycocyanobilin lyase alpha subunit; CpcE	11497	18861	0.001
alr0524	Phycobilisome	phycoerythrocyanin alpha chain; PecA	11660	13985	0.2
alr0523	Phycobilisome	phycoerythrocyanin beta chain; PecB	22153	16828	0.1
alr0525	Phycobilisome	phycoerythrocyanin associated rod linker protein PecC	4337	12756	0.01
all1842	NADH	NADH dehydrogenase	2762	15611	0.004
all3840	NADH	Chain J	9877	14945	0.004
alr0225	NADH	Subunit 6; NdhG	8989	10480	0.002
alr0226	NADH	Subunit 4L; NdhE	9310	12526	0.004
alr3956	NADH	Subunit 5	7635	12323	0.02
all0006	ATP Synthase	Subunit delta; AtpD	10175	10859	0.1
all0010	ATP Synthase	Subunit a; AlpI	15528	17750	0.03
all0011	ATP Synthase	Subunit I; AtpI	13805	13505	0.07

Isolation of ds-cDNA. Double stranded cDNA was isolated from the product of second strand synthesis according to standard procedures (Maniatis et al. 1982).

***In vitro* transcription**. The pelleted ds-cDNA was resuspended in 1.5 µl water. The MEGAscript T7 kit (Ambion) was used for *in vitro* transcription. In addition to the standard nucleotides, 3.75 µl 10 mM Bio-16-CTP (NEN) and 3.75 µl 75 mM Bio-11-UTP (Roche) were added to the reaction mix. This led to the formation of biotinylated cRNA.

cRNA-isolation. The RNeasy kit (Qiagen) was applied for cRNA-isolation. All steps were performed according to the manufacturer's instructions.

cRNA-fragmentation. For cRNA-fragmentation 15 µg cRNA was resuspended in 2.5 µl water and 2.5 µl 2 x fragmentation buffer (5 x stock: 200 mM Tris, 150 mM Mg-acetate, 500 mM K-acetate, pH 8.1). The reaction mix was incubated for 5 min at 94ºC. The fragmentation reaction was performed immediately prior to hybridization and checked by alkaline agarose electrophoresis.

Oligonucleotide probe selection

A unique *Nostoc* PCC 7120 probe set (as many 25-mer probes per open reading frame (ORF) as possible) was calculated based on the full genome sequence (retrieved online from CyanoBase: http://www.kazusa.or.jp/cyanobase/Anabaena/index.html) using a combination of sequence uniqueness criteria and rules for selection of oligonucleotides likely to hybridize with high specificity and sensitivity. The selection criteria were as described in Lockhart et al. (1996) with modifications for the longer probes used here (25-mers instead of 20-mers). If available, 10 unique probes per ORF were used in the experiments.

DNA-microarray production and *in situ* oligonucleotide synthesis

Light-activated *in situ* oligonucleotide synthesis was performed as described by Singh-Gasson et al. (1999) using a digital micromirror device, which is part of the geniom one device (febit biotech GmbH, Heidelberg/Germany). The synthesis was performed within the geniom one device on an activated three-dimensional reaction carrier consisting of a glass-silica-glass sandwich (DNA-processor). Four individually accessible microchannels (referred to as arrays), etched into the silica layer of the

DNA-processor, were connected to the microfluidic system of the geniom device. Using standard DNA-synthesis reagents and 3'-phosphoramidites with a photolabile protecting group (Hasan et al. 1997; Beier and Hoheisel, 2000), oligonucleotides were synthesized in parallel in all four translucent arrays of one reaction carrier. Prior to synthesis, the glass surface was activated by coating with a silane-bound spacer.

Hybridization

Non-competitive hybridizations were performed with 7.5 µg fragmented cRNA (see above) in a final volume of 10 µl. The hybridization solution contained 100 mM MES (pH 6.6), 0.9 M NaCl, 20 mM EDTA, 0.01% (v/v) Tween 20, 0.1 mg/ml sonicated herring sperm DNA, and 0.5 mg/ml BSA. RNA-samples were heated in the hybridization solution to 95ºC for 3 min followed by 45ºC for 3 min before being placed in an array which had been prehybridized for 15 min with 1% (w/v) BSA in hybridization solution at room temperature. Hybridizations were carried out at 45ºC for 16 hrs. After removing the hybridization solutions, arrays were first washed with non-stringent buffer (0.005% (v/v) Triton X-100 in 6 x SSPE) for 20 min at 25ºC and subsequently with stringent buffer (0.005% (v/v) Triton X-100 in 0.5 x SSPE) for 20 min at 45ºC. After washing, the hybridized RNA was fluorescence-stained by incubating with 10 µg/ml streptavidin-phycoerythrin and 2 µg/µl BSA in 6 x SSPE at 25 for 1ºC 5 min. Unbound streptavidin-phycoerythrin was removed by washing with non-stringent buffer for 20 min at 25ºC. Two biological replicates of each condition have been hybridized.

Detection and data processing

The CCD-camera based fluorescence detection system, equipped with a Cy3 filter set, integrated into the geniom one automate was used. 36 pixels per spot were available for data analysis. Processing of raw data, including background correction, array to array normalization and determination of gene-expression levels, as well as calculation of expression differences were performed as described before (Zhou and Abagyan, 2002). All steps were carried out using the PROP-algorithm of the geniom application software which is based on the MOID-algorithm described by Zhou and Abagyan (2002). Background correction is based on probes with no corresponding mRNA-target and the average of the lowest 5% expressed genes. Data normalization is based on iteratively correcting the raw data on these genes. Significance levels for differentially expressed ORFs have been calculated using a two-sided t-test.

Probe secondary structure and GC-content analysis

In order to correlate probe secondary structures and GC-contents with their expression value, the secondary

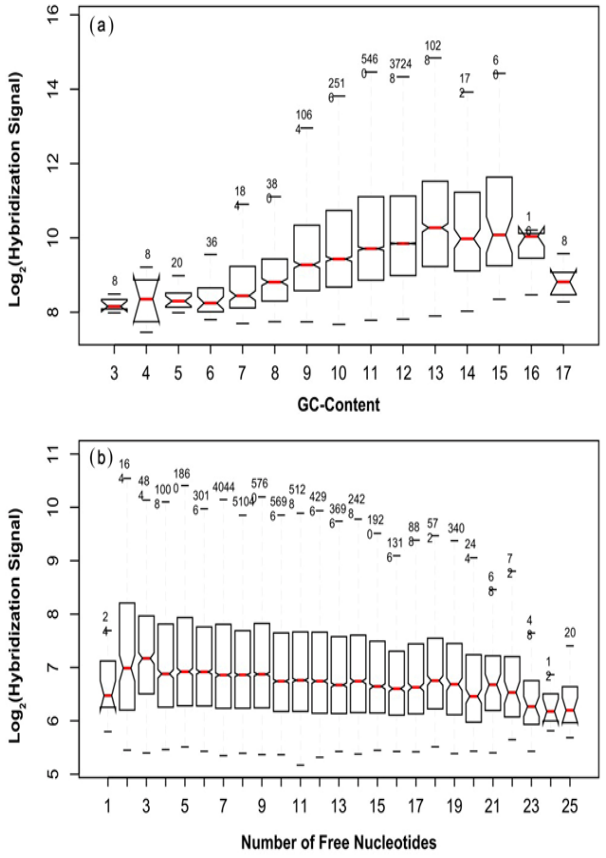

Figure 4. Hybridization effects. The notches around the medians in both box plots (a and b) indicate the 95% confidence interval that the median from one box differs from the median of another box, *i.e.*, if the notches do not overlap the corresponding medians are significantly different. The upper and lower box indicate the second and third quartile, respectively. The plot whiskers extending out from the box to the extreme values. The number above the upper whisker state the number of observations for the corresponding box.
(a) The logarithmized intensity of the hybridization signal is plotted against the GC-content of the 25-mere probe sequence 1.
(b)The logarithmized intensity of the hybridization signal is plotted against the number free secondary structure features (bulges and loops) in the target sequence.

structure of each ORF was calculated using the Vienna RNA Package Version 1.4 (Hofacker et al. 1994). The probes were aligned to their corresponding ORF and the potential probe structure extracted. The number of hybridized (stem duplexes) versus free (loops and bulges) probe nucleotides as well as the probe's GC-content were used in further analysis.

HyDaBa database

All gene-expression data obtained are saved in the Hydrogenase Database (HyDaBa). This relational database allows cross-linking of the expression data with the annotated genome data from NCBI and Cyanobase and pathway maps available from KEGG. The latter is achieved in real-time via a SOAP-interface. HyDaBa is based on a Apache Webserver, MySQL database and a front-end

Table 4. Heterocysts. ORFs included in the present study that are specific to heterocysts and show significant expression differences. Expr.: expression difference between non-nitrogen and nitrogen fixing condition.

ORF	Description	Expr.	p-value
all0521	two-component response regulator, heterocyst pattern formation protein PatA	2.4	0.09
all0813	heterocyst-specific glycolipids-directing protein HglK	4.6	0.001
all1430	heterocyst ferredoxin; FdxH	12.9	0.01
all1431	HesB protein	9.7	0.03
all1432	HesA protein	8.6	0.01
all1730	similar to HetF protein	3.4	0.01
all2512	transcriptional regulator; PatB	2.5	0.06
all5346	heterocyst specific ABC-transporter, membrane spanning subunit DevC homolog	8.9	0.4
all5347	heterocyst specific ABC-transporter, membrane fusion protein DevB homolog	3.0	0.007
alr1603	putative heterocyst to vegetative cell conection protein (fraH)	2.4	0.07
alr2339	heterocyst differentiation protein HetR	2.1	0.02
alr2817	heterocyst differentiation protein HetC	5.8	0.009
alr2818	heterocyst differentiation protein HetP	2.6	0.03
alr2834	glycosyltransferase; hepC	18.5	0.2
alr2835	heterocyst specific ABC-transporter hepA	5.7	0.01
alr3234	similar to heterocyst formation protein HetP	6.6	0.2
alr3648	heterocyst specific ABC-transporter, membrane spanning subunit DevC homolog	2.7	0.1
alr3649	heterocyst specific ABC-transporter, ATP-binding subunit DevA homolog	3.7	0.1
alr3698	heterocyst envelope polysaccharide sysnthesis protein HepB	6.5	0.08
alr3710	heterocyst specific ABC-transporter, membrane fusion protein DevB	3.3	0.2
alr3711	heterocyst specific ABC-transporter, membrane spanning subunit DevC	12.4	0.002
alr3712	heterocyst specific ABC-transporter, ATP-binding subunit DevA	13.1	0.005
alr4281	heterocyst specific ABC-transporter, membrane spanning subunit DevC homolog	3.5	0.1
alr4812	heterocyst differentiation related protein PatN	2.3	0.03
alr5355	heterocyst glycolipid synthase; HglC	4.1	0.03
alr5358	ketoacyl reductase; HetN	4.0	0.05

programmed in PHP. All data are publicly accessible via this web interface that can be accessed using “guest” and “hydaba06” as user name and password, respectively.

RESULTS AND DISCUSSION

In the present study we analyzed the expression of 1249 selected genes from 16 metabolic categories (ca. 20% of the complete genome) of *Nostoc* PCC 7120 cultures under nitrogen fixing and non-nitrogen fixing conditions (Table 1). Therefore we applied a DNA-microarray based approach (Figure 1).

Preparation of the DNA-processor

Oligonucleotide synthesis, hybridization with target cRNA, and signal detection were performed with one single device, named geniom one (febit biotech GmbH, Heidelberg/Germany, Baum et al. (2003)). 25-mer oligonucleotide probes were synthesized *in situ* on the DNA-microarray surface. In order to obtain a broad picture of gene-expression differences between nitrogen fixing and non-nitrogen fixing *Nostoc* PCC 7120 cultures, 500 manually and 749 randomly selected target genes from all major metabolic categories were analyzed (Table 1; Kaneko et al. (2001)). This selection was based on the genome sequence and annotation available from the CyanoBase consortium. In order to ensure reproducibility of the microarray analysis, up to 10 unique 25-mer oligonucleotide probes per target ORF were distributed randomly over the DNA-processor. Due to their small size, 132 ORFs were represented by fewer than 10 unique probes. Of theses, 78 represent unknown and 15 hypothetical proteins, respectively. Of the remaining only 5 ORFs were represented by less than 4 unique probes.

General data analysis

Figure 2 shows a section of four arrays used in this analysis. Due to *in situ* probe synthesis with a digital micromirror device both the spot morphology and topology are extremely homogeneous. The use of one physical surface for all arrays and the fixed placement of the slide during all processing steps results in very low experimental variation. In order to visualize the signal-to-noise ratio the fluorescence-intensity ratio of either two RNA-samples from different growth conditions or from two RNA-samples from the same growth condition are plotted in double-logarithmic scales (Figure 3). It can be clearly seen that the comparison of two different metabolic states scatters much broader than the self-to-self comparison. The variance of the data is displayed by their respective Pearson correlation value r. Five ORFs in the self-to-self comparison show an unexpectedly large variation. In three cases (asr7152, alr7535 and alr7580) this can be explained

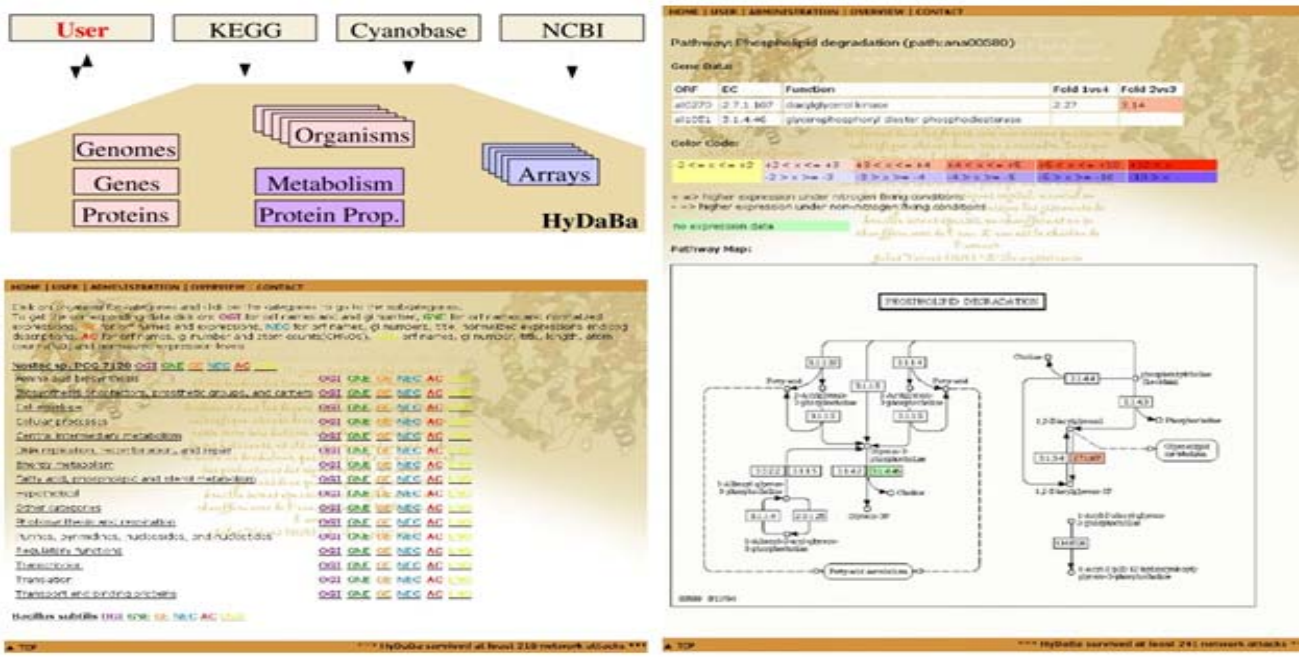

Figure 5. HyDaBa database. All experimental gene-expression data are saved in a relational database (http://www.hydaba.uni-koeln.de). This database allows cross-linking of the gene-expression data with genome data from NCBI and Cyanobase and metabolic charts available from KEGG. HyDaBa is based on a Apache Webserver, a MySQL database and a front-end programmed in PHP. All data are publicly accessible via this web interface. Upper left: database outline; Lower left: screen shot of metabolic categories; Right: screen shot of expression data overlayed on metabolic charts.

by their low, close to threshold fluorescence-signal intensity.

GC-content and probe secondary structure analysis

It is often observed in oligonucleotide-based DNA-microarray experiments that probes directed against one single transcript show large hybridization level variations. The reason for this fluctuation remains still unknown and is one major reason for the necessity to calculate the expression value for each transcript from several unique probes (Kuo et al. 2002; Pozhitkov et al. 2006). We analyzed the influence of both GC-content and secondary structure formations on the hybridization signal.

Figure 4a shows the correlation between the number of guanine and cytosine nucleotides (GC-content) in the 25-mere probes and the corresponding hybridization signal. There is no probe with less than 3 or more than 17 GCs. In the range between 7 and 13 GCs there is a clear linear correlation between GC-content and hybridization signals. To many (more than 50%) or to few (less than 25%) GCs in the probe result in non-linear behavior. As can be seen from the number of observations given in the plot, less than 0.7% (328 out of 48208) of all unique probes are affected by this non-linear behavior. At the current stage it is not possible to draw conclusions from the extremes on both sides of Figure 4a because they are only represented by few data. In contrary to the GC-content, the predicted secondary structure (stems, loops and bulges) of the transcript has only little influence on the hybridization signal (Figure 4b). This was expected because (a) the target was chopped into smaller fragments prior to hybridization, and (b) the hybridization conditions are set such that no secondary structures should form in either the probe or the target.

The effect of individual probe hybridization signals is usually ignored in oligonucleotide-based DNA-microarray experiments. Instead, the average over all probes is used for each transcript. However, these effects have to be taken into account if only few probes are available for particular transcripts. Furthermore, we can conclude that a large portion of the hybridization signal variation is intrinsic to the probe sequence and can not be explained by currently known DNA-duplex formation physico-chemistry.

Data processing and visualization

DNA-microarray experiments involve accumulation and management of large amounts of data. Apart from the experimental data, information from open access knowledge databases and sequence analysis are collected. To provide optimal accessibility to all data we set up a MySQL database on an Apache driven Internet server. The database holds both raw and processed data. Besides data management the database allows cross-connectivity of expression data with annotations from NCBI database, Cyanobase and KEGG (Figure 5) In order to access and query the database (which has been coined HyDaBa) a PHP-based and Internet-accessible front-end has been developed. This front-end helps to query the data, guides

Table 5. Respiratory terminal oxidases. Expression differences for respiratory terminal oxidases under nitrogen fixing conditions.

ORF	Description	Expr.	p-value
COX 1			
alr0950	cytochrome c oxidase subunit II (coxB)	less than ± 2	0.2
alr0951	cytochrome c oxidase subunit I (coxA)	less than ± 2	0.04
alr0952	cytochrome c oxidase subunit III (coxC)	less than ± 2	0.01
COX 2			
alr2514	cytochrome c oxidase subunit II (coxB)	4.0	0.07
alr2515	cytochrome c oxidase subunit I (coxA)	7.0	0.03
alr2516	cytochrome c oxidase subunit III (coxC)	31.8	0.002
COX 3			
alr2731	cytochrome c oxidase subunit II (coxB)	less than ± 2	0.7
alr2732	cytochrome c oxidase subunit I (coxA)	3.3	0.02
alr2734	cytochrome c oxidase subunit III (coxC)	4.7	0.04

the user to define and store new queries, allows data up-and download, and can be easily extended due to its modular setup with template pages. The most important feature of HyDaBa constitutes the mapping of gene-expression data onto metabolic charts from the KEGG database (Figure 5). Technically, this has been achieved by using a SOAP interface (Kawashima et al. 2003). Equally important is the possibility to query for all data available for a given ORF. HyDaBa can be accessed at http://www.hydaba.uni-koeln.de/ using "guest" and "hydaba06" as user name and password, respectively.

Global differences in gene expression upon growth on dinitrogen

Growth on dinitrogen as sole nitrogen source acts like a positive transcriptional switch in *Nostoc* PCC 7120. There is a much larger fraction of genes stronger expressed under nitrogen fixing than under non-nitrogen fixing conditions and only a minority of genes shows a decreased expression level. Only 17 annotated and 12 hypothetical ORFs exhibit a significantly higher expression under non-nitrogen fixing conditions (Table 2) whereas 281 annotated and 184 hypothetical ORFs are more strongly expressed under nitrogen fixing conditions. In Figure 6, these gene-expression differences are clustered according to the participation of the corresponding ORF in specific metabolic categories. The strongest expressed genes participate in photosynthesis and respiration (K). Closer analysis reveals that 21 of the 29 strongest expressed genes in this group belong to photosynthesis, 11 of which are structural proteins of phycobilisomes (Table 3). These findings clearly illustrate the extensive energy demand for nitrogen fixation. The cell expands its light harvesting complexes in order to direct more light energy to the photosystems and produce both more ATP and NADPH. The stronger expression of proteins involved in respiration underlines previous findings that the respiration rate is increased under nitrogen fixing conditions in cyanobacteria (Murry and Wolk, 1989). It is believed that this process supports the removal of oxygen which otherwise would inactivate the nitrogenase enzyme complex (Fay, 1992).

Heterocyst-related genes

As a key global regulator, NtcA plays an important role in the expression of many genes involved in heterocyst differentiation and nitrogen assimilation. For the unicellular, non-differentiating cyanobacterium *Synechococcus* PCC 7942 it has been shown that the binding affinity of NtcA to its target DNA-sequence is elevated by 2-oxoglutarate (Vázquez-Bermúdez et al. 2002; Vázquez-Bermúdez et al. 2003). Thus, 2-oxoglutarate exerts a direct role on NtcA-mediated transcription activation. Furthermore, it plays a central role in sensing the nitrogen status, or rather the C/N-balance. Although synthesized in the heterocyt, 2-oxoglutarate cannot serve as

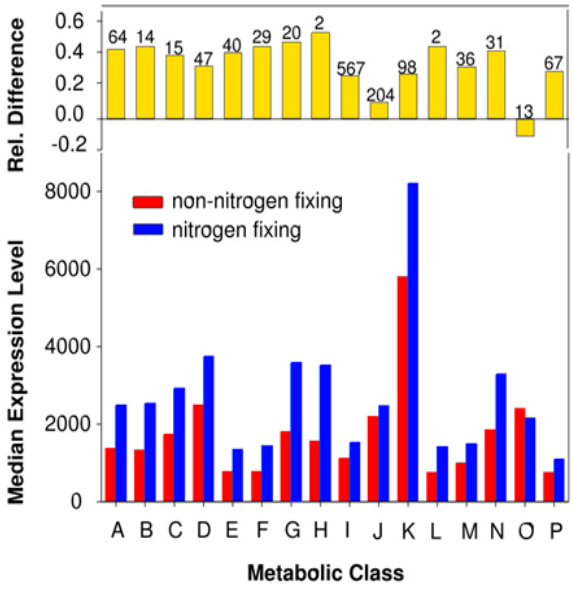

Figure 6. Global expression differences. Measured gene-expression differences in major metabolic categories. The lower plot shows the median expression level of the analyzed ORFs in the corresponding metabolic class. The upper plot shows the corresponding relative change (maximum range: -1 to 1). The numbers above the bars represent the number of ORFs analyzed in the corresponding category. The legend to the letters denoting metabolic classes can be found in Table 1.

a substrate for glutamate synthesis in heterocysts because the necessary ferredoxin glutamine 2-oxoglutarate amidotransferase is not expressed in heterocysts (Martín-Figueroa et al. 2000). This conversion is done in the vegetative cells. Thus, glutamate is imported, while 2-oxoglutarate and glutamine are exported to vegetative cells (Figure 7). We found key enzymes catalyzing the synthesis of 2-oxoglutarate, aconitase hydratase (2.8-times; EC 4.2.1.3) and isocitrate dehydrogenase (4.7-times; EC 1.1.1.42), respectively, being stronger expressed under nitrogen fixing condition (Figure 7). The *hetC* gene (alr2817), which encodes a putative ABC-transporter that is essential for heterocyst formation, has been shown to be a direct target of the transcriptional regulator NtcA (Muro-Pastor et al. 1999). Indeed we see a 5.8-time stronger expression under nitrogen fixing conditions. Table 4 shows expression differences for all known ORFs involved in heterocyst formation included in this study.

Nitrogen metabolism-related genes

The conversion of dinitrogen to ammonia, catalyzed by the nitrogenase enzyme complex, is only the first step in a series of reactions that make nitrogen available to the cell. The nitrogenase enzyme complex provides two products that are metabolized, hydrogen gas and ammonia, respectively. The former is taken up by an uptake hydrogenase while the latter is incorporated to glutamate by the glutamine synthase yielding glutamine. Figure 7 gives an overview over the main pathways and enzyme complexes involved in nitrogen fixation. The nitrogenase consists of three subunits, the molybdenum-iron protein alpha chain (NifD, all1454), the molybdenum-iron protein beta chain (NifK, all1440), and the iron protein (NifH, all1455). We found *nifK* and *nifH* to be more than 10-times more strongly expressed under nitrogen fixing conditions. One ORF (*nifH2*, alr0874) of yet unknown function that is paralogous (92% similarity, 86% identity) to the iron protein (*nifH*, all1455) was found to be expressed at a very lower level and is only slightly stronger expressed under nitrogen fixing conditions (Figure 7). Thus, the gene product of *nifH2* is probably not involved in the nitrogen fixation reaction. The glutamine synthase (glutamate-ammonia ligase) is encoded by *glnA* (alr2328). Northern blot studies in *Nostoc* PCC 7120 have shown that the *glnA* transcript is present in both nitrogen-fixing as well as non-nitrogen fixing cultures, but more abundant in the latter (Orr and Haselkorn, 1982). This is in accordance with our results.

Other genes

Phycobilisomes. Phycobilisomes are the major light-harvesting complexes of cyanobacteria. These are multiprotein assemblies that are functionally associated with photosystem II and constitute up to 50% of the total cellular protein. It has been shown previously that phycobilisomes serve as a nitrogen storage. Upon nitrogen starvation they can be completely degraded within two days. Phycobilisome degradation is thought to provide substrates for protein biosynthesis. As discussed above, we observed strong expression of major components involved in photosynthesis under nitrogen fixing conditions. Accordingly, all proteins included in this analysis and that constitute phycobilisomes are around 3-times more strongly expressed under nitrogen-fixing conditions.

Respiratory terminal oxidases. *Nostoc* PCC 7120 possesses three cytochrome c oxidase gene clusters, cox1 (alr0950, alr0951, alr0952), cox2 (alr2514, alr2515, alr2516), and cox3 (alr2729, alr2730, alr2731, alr2732, alr2734), respectively (Valladares et al. 2003). While cox1 and cox2 are homologous to aa-type cytochrome c oxidases (Schmetterer, 1994; Jones and Haselkorn, 2002), cox3 is most similar to alternative respiratory terminal oxidases (Valladares et al. 2003). The expression of cox2 and cox3 has been reported to be restricted to heterocysts (Valladares et al. 2003). In accordance to this result we see no difference in the expression of cox1, while cox2 and cox3 are more strongly expressed under nitrogen fixing conditions (Table 5).

Adaptations and atypical conditions. One group of genes, members of which are up to 11-times more strongly expressed under non-nitrogen fixing conditions belong to the high light-induced proteins (HLIP-family; Table 6) (Montané and Kloppstech, 2000; Heddad and Adamska, 2002). These proteins belong to the CAB/ELIP/HLIP-superfamily and are evolutionary related to each other (Heddad and Adamska, 2002). While CAB (chlorophyll a/b-binding) proteins are major constitutions of the light harvesting complexes, ELIPs (early light-induced proteins) and HLIPs are taking over photo-protective functions. The HLIP-family in pro-and eukaryotic photosynthetic

Table 6. HLIP-family. Expression differences for the HLIP-family involved in adaptations and atypical conditions. Negative values indicate stronger expression under non-nitrogen fixing conditions.

ORF	Description	Expr.-diff	p-value
asl0449	CAB/ELIP/HLIP-related protein	-2.3	0.04
asl0514	CAB/ELIP/HLIP-superfamily	less than ± 2	0.3
asl0873	CAB/ELIP/HLIP-superfamily	-4.4	0.2
asl2354	CAB/ELIP/HLIP-related protein	less than ± 2	0.5
asr3042	CAB/ELIP/HLIP-superfamily of proteins	-11.3	0.09
asr3043	CAB/ELIP/HLIP-superfamily of proteins	-7.9	0.02
asl3726	CAB/ELIP/HLIP-superfamily	less than ± 2	0.2
asr5262	CAB/ELIP/HLIP-superfamily of proteins	less than ± 2	0.05

organisms consists of more than 100 different stress proteins which have one membrane spanning alpha helix. They accumulate only transiently in photosynthetic membranes in response to light stress and have photoprotective functions. At the amino acid level, members of the HLIP-family are closely related to light-harvesting CAB antenna proteins of photosystem I and II, present in higher plants and some algae. Despite this similarity it is believed that HLIP-proteins fulfill their photoprotective role by either transient binding of free chlorophyll molecules or by participating in energy dissipation (Montané and Kloppstech, 2000). Photooxidative stress in not necessarily connected to high light fluxes but can also be caused by nutrient deprivation that ultimately lead to oversaturation of the photosynthetic electron transport chain. At this point one can only speculate why the HLIP-family is stronger expressed under non-nitrogen fixing conditions. Since the same light and temperature settings were employed for both growing conditions, one argument would be that more light is required (and thus "consumed") under nitrogen fixing conditions. Indeed the nitrogenase enzyme complex has immense energy and reducing power demands. In concert with the higher availability of nitrogen upon growth on dinitrogen (see above) we conclude that this growth condition frees *Nostoc* PCC 7120 from stress. To our knowledge this is the first report about differential expression of members of the HLIP-family in cyanobacteria upon combined-nitrogen deprivation. Analysis of the location of the HLIP-family members shown in Table 6 reveals no link to known ORFs that are involved in nitrogen metabolism. Only asl0449 is located immediately downstream of the allophycocyanin alpha subunit (*apcA*, all0450) and thus demonstrates its potential functional relation to photosynthesis. Future work might help to uncover the function of these proteins in pro-and eukaroytes.

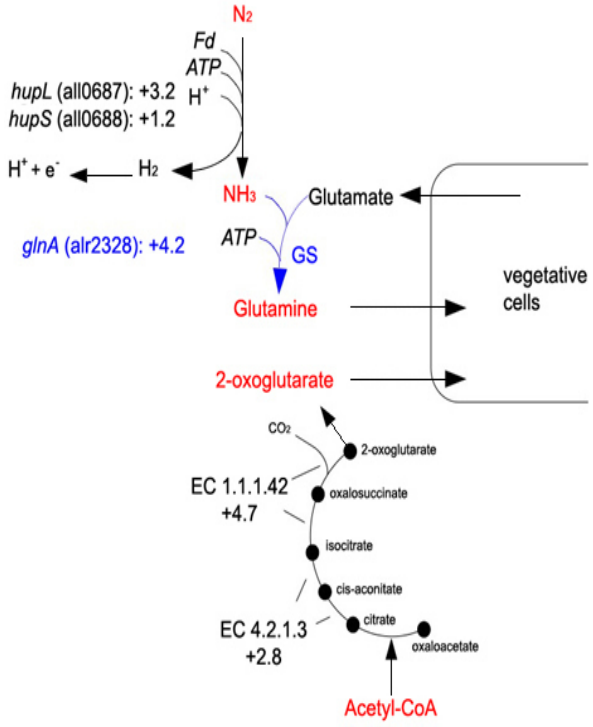

Figure 7. Nitrogen metabolism in *Anabaena* PCC 7120 heterocysts. The main components involved in nitrogen fixation and assimilation include nitrogenase (*nif*-genes) and glutamine synthase (GS, *glnA*). 2-oxoglutarate is synthezised from oxaloacetate and acetyl-CoA by the incomplete TCA-cycle, of which the aconitate hydratase (EC 4.2.1.3) and isocitrate dehydrogenase (EC 1.1.1.42) are shown. Together with glutamine, 2-oxoglutarate is exported to vegetative cells while glutamate is imported.

ACKNOWLEDGMENTS

The authors like to thank Dr. M. Baum and N. Rittner from febit for technical assistance with the microarray experiments and H. Eckes for initial work on the database front-end. R.W. likes to thank Prof. D. Tautz for his continuous support.

REFERENCES

ADAMS, David G. Heterocyst formation in cyanobacteria. *Current Opinion in Microbiology*, December 2000, vol. 3, no. 6, p. 618-624.

BAUM, Michael; BIELAU, Simone; RITTNER, Nicole; SCHMID, Kathrin; EGGELBUSCH, Kathrin; DAHMS, Michael; SCHLAUERSBACH, Andrea; TAHEDL, Harald; BEIER, Markus; GÜIMIL, Ramon; SCHEFFLER, Matthias; HERMANN, Carsten; FUNK, Jörg-Michael; WIXMERTEN, Anke; REBSCHER, Hans; HÖNIG, Matthias; ANDREAE, Claas; BÜCHNER, Daniel; MOSCHEL, Erich; GLATHE, Andreas; JÄGER, Evelyn; THOM, Marc; GREIL, Andreas; BESTVATER, Felix; OBERMEIER, Frank; BURGMAIER, Josef; THOME, Klaus; WEICHERT, Sigrid; HEIN, Silke; BINNEWIES, Tim; FOITZIK, Volker; MÜLLER, Manfred; STÄHLER, Cord Friedrich and STÄHLER, Peer Friedrich. Validation of a novel, fully integrated and flexible microarray benchtop facility for gene expression profiling. *Nucleic Acids Research*, December 2003, vol. 31, no. 23, p. e151.

BEIER, Markus and HOHEISEL, Jörg D. Production by quantitative photolithographic synthesis of individually quality checked DNA microarrays. *Nucleic Acids Research*, February 2000, vol. 28, no. 4, p. e11.

BÖHME, Herbert. Regulation of nitrogen fixation in heterocyst-forming cyanobacteria. *Trends in Plant Science*, September 1998, vol. 3, no. 9, p. 346-351.

DIXON, Ray and KAHN, Daniel. Genetic regulation of biological nitrogen fixation. *Nature Reviews Microbiology*, August 2004, vol. 2, no. 8, p. 621-631.

DUTTA, Debajyoti; DE, Debojyoti; CHAUDHURI, Surabhi and BHATTACHARYA, Sanjoy K. Hydrogen production by Cyanobacteria. *Microbial Cell Factories*, December 2005, vol. 4, no. 1, p. 36.

EBERWINE, James; YEH, Hermes; MIYASHIRO, Kevin; CAO, Yanxiang; NAIR, Suresh; FINNELL, Richard; ZETTEL, Martha and COLEMAN, Paul. Analysis of gene expression in single live neurons. *Proceedings of the National Academy of Sciences of the United States of America*, April 1992, vol. 89, no. 7, p. 3010-3014.

EHIRA, Shigeki; OHMORI, Masayuki and SATO, Naoki. Identification of Low-temperature-regulated ORFs in the cyanobacterium *Anabaena* sp. strain PCC 7120: distinguishing the effects of low temperature from the effects of photosystem II excitation pressure. *Plant and Cell Physiology*, August 2005, vol. 46, no. 8, p. 1237-1245.

EHIRA, Shigeki and OHMORI, Masayuki. NrrA, a nitrogen-responsive response regulator facilitates heterocyst development in the cyanobacterium *Anabaena* sp. strain PCC 7120. Molecular Microbiology, March 2006, vol. 59, no. 6, p. 1692-1703.

FAY, Peter. Oxygen relations of nitrogen fixation in cyanobacteria. *Microbiological Reviews*, June 1992, vol. 56, no. 2, p. 340-373.

GOLDEN, James W. and YOON, Ho-Sung. Heterocyst formation in *Anabaena*. *Current Opinions in Microbiology*, December 1998, vol. 1, no. 6, p. 623-629.

GÜIMIL, Ramon; BEIER, Markus; SCHEFFLER, Matthias; REBSCHER, Hans; FUNK, Jörg-Michael; WIXMERTEN, Anke; BAUM, Michael; HERMANN, Carsten; TAHEDL, Harald; MOSCHEL, Erich; OBERMEIER, Frank; SOMMER, I.; BÜCHNER, Daniel; VIEHWEGER, R.; BURGMAIER, Josef; STÄHLER, Cord Friedrich; MÜLLER, Manfred and STÄHLER, Peer Friedrich. Geniom® technology - the benchtop array facility. *Nucleosides, Nucleotides and Nucleic Acids*, October 2003, vol. 22, no. 5-8, p. 1721-1723.

HAPPE, Thomas; SCHÜTZ, Kathrin and BÖHME, Herbert. Transcriptional and mutational analysis of the uptake hydrogenase of the filamentous cyanobacterium *Anabaena variabilis* ATCC 29413. *Journal of Bacteriology*, March 2000, vol. 182, no. 6, p. 1624-1631.

HASAN, Ahmad; STENGELE, Klaus-Peter; GIEGRICH, Heiner; CORNWELL, Paul; ISHAM, Kenneth R.; SACHLEBEN, Richard A.; PFLEIDERER, Wolfgang and FOOTE, Robert S. Photolabile protecting groups for nucleotides: synthesis and photodeprotection rates. *Tetrahedron*, March 1997, vol. 53, no. 12, p. 4247-4264.

HASELKORN, Robert and BUIKEMA, William J. Nitrogen fixation in cyanobacteria. In: STACEY, Gary; BURRIS, Robert H. and EVANS, Harold J. *eds. Biological Nitrogen Fixation*. Chapman & Hall, Inc. NY, USA, 1992, p. 166-190.

HEDDAD, Mounia and ADAMSKA, Iwona. The evolution of light stress proteins in photosynthetic organisms. *Comparative and Functional Genomics*, December 2002, vol. 3, no. 6, p. 504-510.

HERRERO, Antonia; MURO-PASTOR, Alicia M. and FLORES, Enrique. Nitrogen control in cyanobacteria. *Journal of Bacteriology*, January 2001, vol. 183, no. 2, p. 411-425.

HERRERO, Antonia; MURO-PASTOR, Alicia M.; VALLADARES, Ana and FLORES, Enrique. Cellular differentiation and the NtcA transcription factor in filamentous cyanobacteria. *FEMS Microbiology Reviews*, October 2004, vol. 28, no. 4, p. 469-487.

HIGO, Akiyoshi; KATOH, Hiroshi; OHMORI, Kazuko; IKEUCHI, Masahiko and OHMORI, Masayuki. The role of a gene cluster for trehalose metabolism in dehydration tolerance of the filamentous cyanobacterium *Anabaena* sp. PCC 7120. *Microbiology*, April 2006, vol. 152, no. 4, p. 979-987.

HOFACKER, Ivo L.; FONTANA, Walter; STADLER, Peter F.; BONHOEFFER, L. Sebastian; TACKER, Manfred and SCHUSTER, Peter. Fast folding and comparison of RNA secondary structures. *Monatshefte für Chemie*, February 1994, vol. 125, no. 2, p. 167-188.

IMASHIMIZU, Masahiko; YOSHIMURA, Hidehisa; KATOH, Hiroshi; EHIRA, Shigeki and OHMORI, Masayuki. NaCl enhances cellular cAMP and upregulates genes related to heterocyst development in the cyanobacterium, *Anabaena* sp. strain PCC 7120. *FEMS Microbiology Letters*, November 2005, vol. 252, no. 1, p. 97-103.

JONES, Kathryn M. and HASELKORN, Robert. Newly identified cytochrome *c* oxidase operon in the nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120 specifically induced in heterocysts. *Journal of Bacteriology*, May 2002, vol. 184, no. 9, p. 2491-2499.

KANEKO, Takakazu; NAKAMURA, Yasukazu; WOLK, C. Peter; KURITZ, Tanya; SASAMOTO, Shigemi; WATANABE, Akiko; IRIGUCHI, Mayumi; ISHIKAWA, Atsuko; KAWASHIMA, Kumiko; KIMURA, Takaharu; KISHIDA, Yoshie; KOHARA, Mitsuyo; MATSUMOTO, Midori; MATSUNO, Ai; MURAKI, Akiko; NAKAZAKI, Naomi; SHIMPO, Sayaka; SUGIMOTO, Masako;

TAKAZAWA, Masaki; YAMADA, Manabu; YASUDA, Miho and TABATA, Satoshi. Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120. *DNA Research*, October 2001, vol. 8, no. 5, p. 205-213.

KAWASHIMA, Shuichi; KATAYAMA, Toshiaki; SATO, Yoko and KANEHISA, Minoru. KEGG API: A web service using SOAP/WSDL to access the KEGG system. *Genome Informatics*, 2003, vol. 14, p. 673-674.

KUO, Winston Patrick; JENSSEN, Tor-Kristian; BUTTE, Atul J.; OHNO-MACHADO, Lucila and KOHANE, Isaac S. Analysis of matched mRNA measurements from two different microarray technologies. *Bioinformatics*, March 2002, vol. 18, no. 3, p. 405-412.

LECHNO-YOSSEF, Sigal; FAN, Qing; EHIRA, Shigeki; SATO, Naoki and WOLK, C. Peter. Mutations in four regulatory genes have interrelated effects on heterocyst maturation in *Anabaena* sp. strain PCC 7120. *Journal of Bacteriology*, November 2006, vol. 188, no. 21, p. 7387-7395.

LOCKHART, David J.; DONG, Helin; BYRNE, Michael C.; FOLLETTIE, Maximillian T.; GALLO, Michael V.; CHEE, Mark S.; MITTMANN, Michael; WANG, Chunwei; KOBAYASHI, Michiko; NORTON, Heidi and BROWN, Eugene L. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nature Biotechnology*, December 1996, vol. 14, no. 13, p. 1675-1680.

MANIATIS, Tom; FRITCH, Edward F. and SAMBROOK, Joseph. Synthesis and cloning of cDNA. In: SAMBROOK, Joseph and RUSSELL, David W. *Molecular cloning*. Cold Spring Harbor Laboratory Press, 1982, p. 212-246.

MARTÍN-FIGUEROA, Eugenio; NAVARRO, Francisco and FLORENCIO, Francisco J. The GS-GOGAT pathway is not operative in the heterocysts. Cloning and expression of *glsF* gene from the cyanobacterium *Anabaena* sp. PCC 7120. *FEBS Letters*, July 2000, vol. 476, no. 3, p. 282-286.

MASUKAWA, H.; MOCHIMARU, M. and SAKURAI, H. Disruption of the uptake hydrogenase gene, but not of the bidirectional hydrogenase gene, leads to enhanced photobiological hydrogen production by the nitrogen-fixing cyanobacterium *Anabaena* sp. PCC 7120. *Applied Microbiology and Biotechnology*, April 2002, vol. 58, no. 5, p. 618-624.

MEEKS, John C. and ELHAI, Jeff. Regulation of cellular differentiation in filamentous cyanobacteria in free-living and plant-associated symbiotic growth states. *Microbiology and Molecular Biology Reviews*, March 2002, vol. 66, no. 1, p. 94-121.

MONTANÉ, Marie-Hélène and KLOPPSTECH, Klaus. The family of light-harvesting-related proteins (LHCs, ELIPs, HLIPs): was the harvesting of light their primary function? *Gene*, November 2000, vol. 258, no. 1-2, p. 1-8.

MURO-PASTOR, Alicia M.; VALLADARES, Ana; FLORES, Enrique and HERRERO, Antonia. The *hetC* gene is a direct target of the NtcA transcriptional regulator in cyanobacterial heterocyst development. *Journal of Bacteriology*, November 1999, vol. 181, no. 21, p. 6664-6669.

MURRY, Marcia A. and WOLK, C. Peter. Evidence that the barrier to the penetration of oxygen into heterocysts depends upon two layers of the cell envelope. *Archives of Microbiology*, May 1989, vol. 151, no. 6, p. 469-474.

ORR, J. and HASELKORN, R. Regulation of glutamine synthetase activity and synthesis in free-living and symbiotic *Anabaena* spp. *Journal of Bacteriology*, November 1982, vol. 152, no. 2, p. 626-635.

POZHITKOV, Alex; NOBLE, Peter A.; DOMAZET-LOSO, Tomislav; NOLTE, Arne W.; SONNENBERG, Rainer; STAEHLER, Peer; BEIER, Markus and TAUTZ, Diethard. Tests of rRNA hybridization to microarrays suggest that hybridization characteristics of oligonucleotide probes for species discrimination cannot be predicted. *Nucleic Acids Research*, May 2006, vol. 34, no. 9, p. e66.

PRINCE, Roger C. and KHESHGI, Haroon S. The photobiological production of hydrogen: potential efficiency and effectiveness as a renewable fuel. *Critical Reviews in Microbiology*, February 2005, vol. 31, no. 1, p. 19-31.

SCHMETTERER, Georg. Cyanobacterial respiration. In: BRYANT, D.A. ed. *The Molecular Biology of Cyanobacteria*. Kluwer Academic Publisher, NL, 1994, p. 409-435.

SINGH, S.; KATE, B.N. and BANERJEE, U.C. Bioactive compounds from cyanobacteria and microalgae: an overview. *Critical Reviews in Biotechnology*, July 2005, vol. 25, no. 3, p. 73-95.

SINGH-GASSON, Sangeet; GREEN, Roland D.; YUE, Yongjian; NELSON, Clark; BLATTNER, Fred; SUSSMAN, Michael R. and CERRINA, Franco. Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array. *Nature Biotechnology*, October 1999, vol. 17, no. 10, p. 974-978.

STANIER, R.Y.; KUNISAWA, R.; MANDEL, M. and COHEN-BAZIRE, G. Purification and properties of unicellular blue-green algae (order Chroococcales).

Bacteriological Reviews, June 1971, vol. 35, no. 2, p. 171-205.

VALLADARES, Ana; HERRERO, Antonia; PILS, Dietmar; SCHMETTERER, Georg and FLORES, Enrique. Cytochrome c oxidase genes required for nitrogenase activity and diazotrophic growth in *Anabaena* sp. PCC 7120. *Molecular Microbiology*, March 2003, vol. 47, no. 5, p. 1239-1249.

VÁZQUEZ-BERMÚDEZ, María Félix; HERRERO, Antonia and FLORES, Enrique. 2-Oxoglutarate increases the binding affinity of the NtcA (nitrogen control) transcription factor for the *Synechococcus glnA* promoter. *FEBS Letters*, February 2002, vol. 512, no. 1-3, p. 71-74.

VÁZQUEZ-BERMÚDEZ, María Félix; HERRERO, Antonia and FLORES, Enrique. Carbon supply and 2-oxoglutarate effects on expression of nitrate reductase and nitrogen-regulated genes in *Synechococcus* sp. strain PCC 7942. *FEMS Microbiology Letters*, April 2003, vol. 221, no. 2, p. 155-159.

WOLK, C. Peter. Heterocyst formation in *Anabaena*. In: BRUN, Yves and SHIMKETS, Lawrence J. eds. *Prokaryotic Development*. American Society for Microbiology, USA, 2000, p. 83-104.

WÜNSCHIERS, Röbbe and LINDBLAD, Peter. Light-dependent hydrogen metabolism and generation by Cyanobacteria. In: NALVA, Hari Singh ed. *Handbook of Photochemistry and Photobiology*. American Scientific Publishers, North Lewis Way, CA, USA, 2003, vol. 4, p. 295-328.

ZHOU, Yingyao and ABAGYAN, Ruben. Match-only integral distribution (MOID) algorithm for high-density oligonucleotide array analysis. *BMC Bioinformatics*, January 2002, vol. 3, p. 3.