CHARACTERIZATION OF CYANATE METABOLISM IN MARINE *SYNECHOCOCCUS* AND *PROCHLOROCOCCUS* spp.

Nina A. Kamennaya[1] and Anton F. Post[2*]

[1]Dept. of Plant and Environmental Sciences, Hebrew University of Jerusalem, Edmond J. Safra Campus, Jerusalem, 91904, Israel

[2]Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, MA 02543, USA

*Corresponding author: Anton F. Post

Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, 7 MBL St, Woods Hole, MA 02543

Phone: 508 289-7131, Fax: 508 457-4727, Email: apost@mbl.edu

Key Words: cyanase, cyanate, marine cyanobacteria

Running Head: Cyanase in marine unicellular cyanobacteria

ABSTRACT

Cyanobacteria of the genera *Synechococcus* and *Prochlorococcus* are the most abundant photosynthetic organism on Earth occupying a key position at the base of marine food webs. The *cynS* gene that encodes cyanase was identified among bacterial, fungi and plant sequences in public databases and the gene was particularly prevalent among cyanobacteria, including numerous *Prochlorococcus* and *Synechococcus* strains. Phylogenetic analysis of *cynS* sequences retrieved from the Global Ocean Survey database identified >60% as belonging to unicellular marine cyanobacteria, suggesting an important role for cyanase in their nitrogen metabolism. Here we showed that marine cyanobacteria have a functionally active cyanase, the transcriptional regulation of which varies among strains and reflects the genomic context of *cynS*. In *Prochlorococcus* sp. MED4, *cynS* was presumably transcribed as part of the *cynABDS* operon, implying cyanase involvement in cyanate utilization. In *Synechococcus* sp. WH8102, expression was not related to nitrogen stress responses and here cyanase presumably serves in the detoxification of cyanate resulting from intracellular urea and/or carbamoyl phosphate decomposition. Lastly, we report on a cyanase activity encoded by *cynH*, a novel gene found in marine cyanobacteria only. The presence of dual cyanase genes in genomes of seven marine *Synechococcus* strains and their respective roles in nitrogen metabolism remain to be clarified.

INTRODUCTION

Cyanase (EC 4.2.1.104) converts cyanate to carbon dioxide and ammonia in a bicarbonate dependent reaction:

$$NCO^- + HCO_3^- + 2H^+ \rightarrow 2CO_2 + NH_3$$

(19). The enzyme is encoded by *cynS* and is found in a wide range of organisms: Cyanobacteria, Proteobacteria (including enterobacteria), some Gram positive Bacteria, Fungi, and Plants. Transcriptional regulation and enzymatic activity were initially studied in *E. coli* strain B/1 (1). In response to cyanate addition, transcription was induced as a polycistronic message of *cynS* together with a cyanate transporter gene (1). Twenty years later, protein structure and subunit organization of *E. coli* CynS were determined at 1.65Å resolution (41). The cyanase monomer was found to be composed of two domains: an N-terminal domain with similarity to the DNA-binding α-helix bundle motif and an 'open fold' C-terminal domain with no structural homology to other proteins. The dimer structure of the cyanase subunit revealed intertwined C-terminal domains with five dimers forming a decameric cyanase holoenzyme. The proposed active site contains three conserved residues, Arg-96, Glu-99, Ser-122, so that five catalytic sites found in the active decamer form an inner ring around a hollow core (41).

Cyanase activity in cyanobacteria was first described for the freshwater *Synechococcus* sp. PCC6301 (UTEX 625) and cyanate decomposition did not require pre-exposure of cells to cyanate (27). Instead, the decomposition of exogenous cyanate by *Synechocystis* sp. PCC6803 and *Synechococcus elongatus* PCC7942 was found to be light dependent (12). Based on sequence similarity, *cynS* was identified in *Synechocystis* sp. PCC6803, *Synechococcus elongatus* PCC7942 (17), and freshwater *Synechococcus*

68 sp. PCC6301 (12). In the *Synechococcus* strains, *cynS* was transcribed as a part of an
69 operon together with *cynABD*, encoding a ABC-type cyanate transporter, while in
70 *Synechocystis* it was co-transcribed with four molybdenum-cofactor biosynthesis genes
71 (17). Transcription of the operon was negatively regulated by ammonium and required
72 the presence of NtcA, a global nitrogen (N) regulator of cyanobacteria (17). Comparative
73 genomics of marine cyanobacteria revealed *cynS* in the majority of *Synechococcus* (29,
74 34) and in some *Prochlorococcus* (13, 36). The physiological and ecological roles of
75 cyanase in marine cyanobacteria have not yet been elucidated. In the presence of a
76 specific transporter, cyanase may play a role in cyanate assimilation. Marine
77 cyanobacteria strains that possess the *cynABD* genes, encoding an ABC-type cyanate
78 transporter, grew at near maximal growth rates with cyanate as the sole N-source (20).
79 The CynABD complex was recently shown to also contribute to nitrite uptake in
80 *Synechococcus elongatus* PCC7942 (25). Conversely, transport systems for CO_2, HCO_3^-,
81 NO_3^-, NO_2^-, Cl^-, PO_4^{2-}, and SO_4^{2-} do not contribute to cyanate acquisition (12). The great
82 majority of cyanobacteria that contain *cynS* in fact lack the genes for cyanate acquisition,
83 suggesting a role for CynS in detoxification of internally generated cyanate, which
84 accumulates as a byproduct of the urea cycle or via the degradation of carbamoyl
85 phosphate (36). Here, we characterize transcriptional regulation of *cynS* and the resulting
86 cyanase activity in marine cyanobacteria. Furthermore, we report on a novel source of
87 cyanase activity associated with a conserved hypothetical gene in seven marine
88 *Synechococcus* strains. Based on this activity we have named it as *cynH* (cyanate
89 hydratase) and we refer to this gene as such throughout this report.
90

MATERIALS AND METHODS

Strains and media *Prochlorococcus* sp. MED4 was grown in the seawater based PRO99 medium (28) while *Synechococcus* spp. WH8102 and WH7803 were grown in artificial seawater medium (44), supplemented to a final concentration of 0.8 mmol L^{-1} ammonium chloride (NH_4Cl; J. T. Baker, Deventer), up to mid-log phase. They were maintained at 25±1 °C with gentle agitation at 80–90 rpm on a gyratory shaker, model G2 (New Brunswick Scientific Co., New Brunswick, NJ) with continuous illumination provided by "warm-white" fluorescence tubes at 20–25 μmol photons m^{-2} s^{-1}. For N nutrition experiments, NH_4Cl was replaced with 0.8 mmol L^{-1} of nitrate, 0.8 mmol L^{-1} of freshly prepared sodium cyanate (NaOCN; Aldrich), 0.4 mmol L^{-1} of freshly prepared urea (Amresco), or 0.8 mmol L^{-1} sodium chloride to produce a N-free medium. Cultures supplemented with fresh ammonium were used as a control.

For protein expression, we used HMS174 (Novagen, EMD Biosciences Inc.), Rosetta™ pLysS (Novagen, EMD Biosciences Inc.), and BL21-CodonPluS®-RIL (Stratagen, Agilent Technologies) *E. coli* strains. Cloning and propagation of recombinant plasmids was performed following protocols of the manufacturers. For a negative control of specificity of the cyanate activity, Maltose Binding Protein (MBP) fused NtcA and MBP itself were overexpressed in *E. coli* strains Rosetta™ pLysS and BL21-CodonPluS®-RIL respectively.

Sequence analysis Protein sequence data of *cynS*, *cynH* genes and its genomic context were obtained from the genome of *Synechococcus* sp. WH7803 available from GenBank, using the DNA sequence viewer and annotation tool Artemis (35). Further genomic data

for comparative genomic context study, alignments, and phylogenetic analyses was obtained from the "nr" public database at the NCBI (http://www.ncbi.nlm.nih.gov/Blast.cgi) and the All Metagenomic ORF Peptides subject database of the Community Cyberinfrastructure for Advanced Marine Microbial Ecology Research and Analysis (CAMERA) (http://camera.calit2.net/), by BLASTN and BLASTP searches. Multiple protein sequence alignments were performed using the Multiple Sequence Comparison by Log-Expectation (MUSCLE) program (9). Phylogenetic relationships were analyzed with MrBayes 3.1 (18), with nucleotide frequencies and parameters for the GTR + gamma invariant model estimated from the data. Two independent runs of four chains were run for 2 million generations and sampled every 100 generations; comparison of parameter estimates from the two runs indicated convergence (15). The secondary structure of CynS and CynH was determined by the Jpred Prediction Server (www.compbio.dundee.ac.uk/www-jpred/).

Transcriptional regulation Cyanobacterial cells (200 ml) grown with different N sources were harvested 3, 6, 9, 12, and 24 hours after medium replacement and centrifuged at 10,500 *g* for 6 min at 4 °C. Cell pellets were resuspended in 0.5 ml of TRI Reagent® (Ambion) and immediately frozen at -20 ºC. Further RNA isolation was performed according to standard procedure recommended by the manufacturer. Prior to the analysis, RNA samples were treated with the DNA-*free*™ kit (Ambion) to eliminate DNA contamination, and RNA purity was confirmed by PCR using the primer pairs listed in Table 1. Final nucleic acid concentrations were determined photometrically (NanoDrop). The ImProm-II™ Reverse Transcription System (Promega) kit was used for

two-step reverse transcription PCR analysis. For the first step, a standard reaction was applied with the gene-specific reverse primers listed in Table 1. Subsequent PCR reactions with specific primer sets were performed in a final volume of 50 µl containing 1–2 µl cDNA, 0.5 µmol L^{-1} of each primer, 0.25 mmol L^{-1} of each deoxyribonucleotide triphosphate, 1.25 units of Taq DNA polymerase (PEQLAB), and 10x PCR buffer containing 15 mmol L^{-1} $MgCl_2$. Using a PTC200 Thermo Cycler (MJ Research Inc.), the reaction mix was preincubated at 94 °C for 5 min followed by cycles of denaturation at 94 °C for 45 s, primer annealing for 30 s (see Table 1 for primer-specific temperatures), and elongation at 72 °C for 45 s. Samples of 6 µl were promptly collected after 20, 25 and 30 cycles. Equal volumes of PCR products from the three sets (20, 25, and 30 cycles) were run on 1.5% agarose gels and visualized with ethidium bromide (Sigma). Quantification of gene expression level was performed with ImageJ analysis software (rsbweb.nih.gov/ij/). Density of EtBr stained amplicons on 1% agarose gels bands was determined the three subset of samples, but always before saturation of the PCR amplification was reached (Fig. 1.) In the rare case of early saturation, cDNA was PCR-amplified for 16-18 cycles instead. Reference genes in RT-PCR analyses were PMM0615 of *Prochlorococcus* sp. MED4 (cell wall hydrolase/autolyin, COG0388) and *orf0250* of *Synechococcus* sp. WH7803 (Ycf48-like photosystem II stability/assembly factor, COG4447). These genes were chosen as they are part of the cyanobacterial core genomes and do not alter their transcript level as assessed by micro-array (Zer, Singer and Post, unpubl. data). For *Synechococcus* sp. WH8102 the reference was the 16S rRNA gene. Transcript levels of the genes of interest determined were normalized to those of the reference genes. The regression lines describing gene expression over time were

compared while testing the hypotheses of coincidence, parallelism and equality of intercepts (as described in "Master of Applied Statistics" by Pia Veldt Larsen (http://statmaster.sdu.dk/courses/st111/module09/)). The p-value was calculated from the F value (www.graphpad.com/quickcalcs/) obtained with the equation

$$F = \frac{(Ss(Time) + Ss(Time * Type))/2}{Ms(\mathrm{Re}sidual)}$$

from the MANOVA-derived variables.

Cloning of cyanase genes Genomic DNA of *Synechococcus* sp. WH7803 was extracted by phenol-chloroform as described in (31). The complete coding sequence of *Synechococcus* sp. WH7803 *cynS* (424 bp) and *cynH* (201-bp) were PCR amplified with primer combinations cSEcoRI2-F (5'-AGAAAGGGGAATTCATGAGTTTCGCCGATC-3'), cSPstI-R (5'-CACACGATTCAAGCTGCAGTTACCATTTTTTGTAAGGAAGG-3') and scSEcoRI-F (5'-AGTTCGTGGAATTCCATGAGTGCTCTTTTCCGTTCC-3'), scSPstI-R (5'-GCCCGAGGGCTGCAGTTACGGGGAGTCGAGATAGG-3') respectively. The forward/reverse primers contain *EcoRI* and *PstI* restriction sites respectively to facilitate synthesis of the MBP fusion construct. PCR reactions (50 µl) were performed with Phusion high-fidelity DNA polymerase (Finnzymes) and 1.2 ng μl^{-1} DNA template. PCR reactions were run over 30 cycles of denaturation (98 °C, 30 s), annealing (58 °C for *cynS* and 66 °C for *cynH*, 20 sec) and elongation (72 °C, 20 s) followed by a final 5 min of elongation. PCR products were purified on 1.2% TAE-buffered agarose gels and eluted with the Wizard® SV Gel and PCR Clean-Up System. Both the amplicons and pMBP1 vector (kindly provided by Dr P. Sheffield, University of Virginia) were digested with *EcoRI* and *PstI*. Purified amplicons (150 ng) were ligated

183 downstream of *malE* on the pMBP1 vector in a 1:3 molar ratio. In-frame assembly of
184 fusion constructs was verified from DNA sequence analysis following transformation
185 into a suitable *E. coli* host strain.
186
187 **Over-expression of cyanase genes** Following transformation into *E.coli*, expression of
188 the CynS-MBP and CynH-MBP fusion proteins was tested in several strains suitable for
189 protein overexpression. In order to determine optimal conditions for protein expression,
190 we tested different IPTG concentrations, temperatures and incubation times.
191 Accumulation of recombinant protein was then confirmed by SDS-PAGE analysis using
192 whole cell lysate of IPTG-induced cells. The cells were centrifuged at 20,000 *g* for 1 min
193 at 4 ºC, and the cell pellet was kept overnight at -20 ºC to ease further lysis. The next day,
194 the cell pellet was re-suspended in column buffer (20 mmol L^{-1} Tris-HCl pH 7.5, 200
195 mmol L^{-1} NaCl; 1 mmol L^{-1} EDTA) containing 100 μg ml^{-1} lysozyme, 1 mmol L^{-1} PMSF,
196 0.1 mg ml^{-1} DNase, and 0.1 mmol L^{-1} $MgSO_4$, incubated on ice for 30 min, and then
197 sonicated. Membrane debris was removed by centrifugation at 20,000 *g* for 15 min at 4
198 ºC. Both pellet and soluble fractions were analyzed for recombinant protein accumulation
199 using SDS-PAGE. Experimental conditions for expression and accumulation of
200 recombinant proteins in the soluble fraction of specific *E. coli* strains were as follows:
201 MBP-CynS was expressed in *E.coli* strain HMS174 incubated for 4 h at 37 ºC after
202 addition of 0.4 mM IPTG; MBP-CynH was expressed in *E.coli* strain Rosetta™ pLysS
203 incubated for 4 h at 22 ºC after addition of 0.1 mM IPTG; MBP-NtcA was expressed in
204 *E.coli* strain Rosetta™ pLysS incubated for overnight at 17 ºC after addition of 0.2 mM
205 IPTG; MBP was expressed in *E.coli* strain BL21-CodonPluS®-RIL incubated for 3 h at

206 37 ºC after addition of 0.3 mM IPTG. All *E.coli* strains were supplemented with 100 μg
207 ml^{-1} ampicillin, and the Rosetta™ pLysS strain was supplemented with 34 μg ml^{-1}
208 chloramphenicol.

209 **SDS-PAGE and western blotting** The identity of the induced protein was confirmed
210 by western blotting, where cell lysate from a non-induced culture, alongside pellet and
211 eluted fractions from an induced culture, were separated on 14.5% acrylamide gels by
212 SDS-PAGE. Proteins were transferred to a polyvinylidene difluoride (PVDF) membrane
213 at 4 °C using transfer buffer containing 54 mmol L^{-1} Tris, 384 mmol L^{-1} glycine, and
214 20% (v/v) methanol. After blocking the membrane with 5% (w/v) skimmed milk powder
215 (Difco) dissolved in TBS (10 mmol L^{-1} Tris-HCl pH 7.5, 250 mmol L^{-1} NaCl) with
216 0.06% (v/v) Tween-20, it was washed three times with TBS buffer containing 0.01%
217 Tween-20 (TTBS). Primary antibody (maltose binding protein antibody [R29.6], ab65,
218 Abcam) diluted 1:1,000 was added and incubated overnight at 4 °C, then washed with
219 TTBS buffer and exposed to a blocking buffer containing peroxidase-conjugated anti-
220 mouse-IgG secondary antibody (Jackson ImmunoResearch Laboratories Inc.) diluted
221 1:5,000. Chemiluminescence detection was carried out with the EZ-ECL Enhanced
222 Chemiluminescence Detection Kit for HRP (Biological Industries Ltd.) using a LAS-
223 3,000 Image Analyzer (Fujifilm).
224
225 **Purification of recombinant protein** Purification of CynS-MBP, CynH-MBP
226 fusion proteins and MBP was performed using amylose resins (E8021) according to New
227 England BioLabs protocols. In brief, cell pellets were resuspended in 1/10 culture volume
228 of phosphate buffered saline (PBS, 0.05 mmol L^{-1}, pH 7.6) complemented with 0.2 mg

229 ml^{-1} lysozyme, 0.05 mg ml^{-1} DNase, 10 mmol L^{-1} $MgCl_2$, 1 mmol L^{-1} DMSF, and 1:200
230 protease cocktail inhibitor, and cells were disrupted by sonication. After centrifugation at
231 10,000 *g* for 15 min, the supernatant was mixed with the amylose resin, and after binding
232 for 1 hour at 4 °C, the column was washed gently with PBS buffer. Recombinant protein
233 was eluted following the addition of 10 mmol L^{-1} free maltose in PBS buffer.
234
235 **Cyanase activity assay** *In vitro* cyanase activity of the *Synechococcus* sp. WH7803
236 associated with the CynS and CynH fusion proteins was measured as described
237 previously (1) with small modifications. In brief, 20 μg recombinant protein aliquots
238 were used for activity measurements using two controls. The first control was
239 supplemented with all reaction mixture ingredients except cyanate and tested for
240 background ammonia levels occasionally introduced with the recombinant protein
241 mixture. The second control consisted of the reaction mixture without protein aliquot and
242 tested spontaneous decomposition of cyanate. The reactions were initiated by addition of
243 sodium cyanate (NaOCN, Aldrich), and terminated by addition of 225 μl Nessler reagent
244 (K_2HgI_4, Aldrich) diluted 1:3 with DDW. The temperature of the reaction solution was
245 adjusted to the desired value prior to cyanate addition. Reactions were performed in 48-
246 well plates (Nunc™) and analyzed on a Microplate Reader Synergy2 (BioTek
247 Instruments Inc.) within 5 min after termination of the reaction. The specificity of CynS
248 and CynH activity was shown in parallel control experiments containing MBP by itself or
249 MBP-NtcA, proteins that lack enzymatic activity. Cyanase inhibition reactions were
250 performed following 200 μmol L^{-1} Na-azide additions. All control reactions were
251 performed at 26 °C. One unit of cyanase activity was defined as the amount of enzyme,

252 required to catalyze the formation of 1 μmol of ammonia per minute. Temperature
253 optima for the fusion proteins were estimated by performing enzymatic reactions at five
254 different temperatures in the range of 4-50 °C. Protein concentrations were determined
255 with the Bradford assay (4).
256
257 **RESULTS**
258 **Phylogenetic analysis** Cyanase, encoded by *cynS*, is a well characterized enzyme
259 in *E. coli.* Orthologs to *cynS* are commonly found in a wide range of microorganisms,
260 including marine unicellular cyanobacteria. In an attempt to better define cyanase
261 evolution in cyanobacteria, we performed alignments and phylogenetic analyses of both
262 genomic and environmental CynS sequences. Among 12 marine *Synechococcus*
263 genomes, 11 were found to possess one or more *cynS* orthologs. The only *Synechococcus*
264 lacking *cynS* was WH5701, a strain representative of halotolerant (43), estuarine
265 *Synechococcus*, ancestral to marine *Synechococcus*. Among 12 *Prochlorococcus*
266 genomes only three carried a *cynS* ortholog (20). Fig. 2 shows an alignment of translated
267 *cynS* sequences of *Synechococcus* and *Prochlorococcus* along with cyanase of *E. coli.*
268 With an overall 37-44% identity, the alignment revealed a higher degree of sequence
269 conservation for the C-terminal region as compared to the N-terminal region. Moreover,
270 amino acid residues that are proposed to contribute to the catalytic activity of the *E. coli*
271 protein (41) were fully conserved in all *Synechococcus* and *Prochlorococcus* CynS.
272 Using CynS from *Synechococcus* sp. WH7803 and *Roseovarius* sp. 217 as queries
273 in the BLAST searches (http://blast.ncbi.nlm.nih.gov/Blast.cgi), we retrieved 107 full
274 length CynS sequences from genomes of bacterial isolates and strains, 9 fungal CynS

sequences, and a single Archaeal sequence. In addition, 147 translated open reading frames (ORF) derived from the Global Ocean Survey (30) were retrieved using the CAMERA BLAST Wizard Tool. Following a preliminary alignment we discarded truncated and incoherent sequences. Subsequently, 90 translated environmental sequences >120 amino acids in length, from surface waters of the Atlantic, Pacific, Indian and Southern Oceans, were aligned with 29 CynS proteins of known origin. The Archaeal CynS was chosen as an outgroup for the construction of a phylogenetic tree. Tree topology suggests that cyanobacterial *cynS* evolved from a common ancestor near the base of the bacterial radiation. Out of a total of 90 GOS-derived cyanase sequences, 56 clustered with cyanobacterial CynS and they were affiliated with known *Synechococcus* and *Prochlorococcus* CynS (Fig. 3). Of these 77% clustered with *Synechococcus* and the remaining 23% with *Prochlorococcus*. Branching patterns for *cynS* closely mirror those of 16S and *ntcA* topologies (31, 36) suggesting that cyanobacterial *cynS* distributions resulted from vertical evolution with limited (if at all) contribution of early lateral gene transfer events. Interestingly, *cynS* has so far not been identified among any of the marine diazotrophic cyanobacteria.

Transcriptional regulation of *cynS* Consistent with the fact that cyanate may serve as an N-source, cyanate acquisition genes form an integral part of the regulon controlled by NtcA, a global N-stress regulator in marine cyanobacteria (36). However NtcA control over *cynS* transcription has not been clearly established. Whole genome microarray analyses for *Prochlorococcus* sp. MED4 (40) showed elevated transcript levels for genes encoding urea (*urtA*) and cyanate (*cynA*) transporters in N-deprived cells.

298 However, despite the *cynABD* and *cynS* organization as an NtcA controlled operon, the
299 authors reported that *cynS* was not differentially expressed (40). In *Synechococcus* sp.
300 WH8102 *cynABD* and *cynS* are separated by 1,932 bp with two putative open reading
301 frames (SYNW2488 and SYNW2489) between them. A putative NtcA binding site was
302 found upstream of *cynA* but not of *cynS* (Fig. 4), suggesting that *cynS* transcription may
303 be uncoupled from N stress responses in strain WH8102. Likewise, the promoter region
304 of *cynS* in *Synechococcus* sp. WH7803 lacked an NtcA binding motif. Here we aimed at
305 confirmation of the microarray result for *ntcA, cynA* and *cynS* in N-starved
306 *Prochlorococcus* sp. MED4 by RT-qPCR. We further expanded transcription studies of
307 these genes to cells supplemented with different N-sources (Fig. 5). We further monitored
308 *cynA* and *cynS* responses in two *Synechococcus* strains, in an effort to tease apart their
309 contribution to N-scavenging (uptake and conversion of cyanate to ammonium and CO_2)
310 and detoxification (cyanate conversion only) pathways.
311 In order to monitor the N-status of *Prochlorococcus* sp. MED4 during the
312 experiment, *ntcA* transcript levels were determined alongside those of *cynA* and *cynS*.
313 *ntcA* transcript levels, of cells maintained under different N regimes, increased with time
314 before reaching a maximum after six (cyanate, urea) or nine hours (nitrate, no N) of
315 incubation, as compared to basal levels in the presence of ammonium. We found as a
316 general trend that during the first 9 hours the transcription patterns of *cynA* and *cynS*
317 followed that of *ntcA* in *Prochlorococcus* sp. MED4, indicating that their transcription
318 occurred in response to changes in N-source and/or its availability (Fig. 5 A-C). After 9
319 hours the transcript had reached steady state levels or decreased in all treatments. In
320 *Synechococcus* sp. WH8102, the transcription pattern of *cynA* (Fig. 5D) was similar to

that in *Prochlorococcus* sp. MED4 (Fig. 5A). However, increase of *cynS* transcript was minor and no clear pattern could be discerned in any of the treatments (Fig. 5E). In *Synechococcus* sp. WH7803, basal transcription of *cynS* was determined in cells grown with ammonium. After 3 hours, *cynS* transcript accumulated in cells grown with nitrate, urea, or in N-free medium. However, in the cyanate-grown culture, transcript levels decreased after 6 hours and returned to levels seen in ammonium-grown cells. After 12 hours *cynS* remained strongly transcribed in N-starved cells only (Fig. 5F). Fig. 6 illustrates the parallel pattern of *cynA* and *cynS* transcript accumulation observed for *Prochlorococcus* sp. MED4 (panel A). In contrast a significantly different pattern ($p<0.05$) was observed in *Synechococcus* sp. WH8102, suggesting that the *cynA* and *cynS* respond to different N signals and controls in this strain.

Genomic context of *cynS* The genomic context of *cynS* was different in different marine cyanobacterial strains. In *Prochlorococcus* strains NATL1A and NATL2A, *cynS* was positioned among conserved hypothetical genes. In *Prochlorococcus* sp. MED4, *cynS* was positioned immediately downstream of *cynABD* (Fig. 4) and it was probably transcribed as part of a polycistronic message (Fig. 6) as in *E. coli*. In marine *Synechococcus* genomes *cynS* is confined to a 60 Kb region that contains the major N-acquisition genes (36). It is typically found downstream of *nirA* (assimilatory nitrite reductase) and *focA* (formate/nitrite transporter) genes and flanked by four genes with a fully conserved localization on these genomes: *trpD*, *proP*, *ppk*, and *rpoD* (respectively encoding for glycosyl transferase family protein, an unidentified permease, polyphosphate kinase, and an alternative RNA polymerase sigma factor, Fig. 4). Interestingly, *Synechococcus* sp. WH7805 carries an open reading frame near *cynS* which

345 was identified as a cyanate hydratase (= cyanase) in the automated annotation pipeline
346 (Fig. 4). The predicted amino acid sequence appeared unique and orthologs were found in
347 7 marine *Synechococcus* genomes as well as on a clone GRIST19 from a metagenomic
348 library obtained from the Atlantic Ocean (GenBank accession EU795157).
349 *Synechococcus* CC9311 was found to possess two copies of *cynH*. Sequence comparison
350 revealed no significant similarity between known CynS sequences and the short protein
351 encoded by *cynH*. Moreover, CynH could not be assigned to a functional protein family,
352 since no known structural domains were identified in its amino acid sequence by
353 SUPERFAMILY 1.73 (45) and Phyre Server 0.2 (23). In the sections below we describe
354 experiments pertaining to cyanase activity associated with overexpressed fusion
355 constructs of WH7803 *cynS* and *cynH*.
356
357 **CynS and CynH over-expression** Over-expression of the CynS-MBP and CynH-MBP
358 fusion proteins in different *E. coli* hosts following IPTG-induction was found to be
359 optimal in *E. coli* strains HMS174 and Rosetta™ pLysS respectively. The recombinant
360 proteins were purified on amylose resin and protein presence in the elution fraction was
361 confirmed by SDS-PAGE (Fig. 7). Following elution purified fusion proteins appeared as
362 a ~50 kDa peptide for CynS-MBP while apparent molecular weight of CynH-MBP was
363 slightly smaller, around ~48 kDa. Both were in close approximation of estimated
364 molecular weights for *Synechococcus* sp. WH7803 CynS (14.3 kDa) and CynH (6.2 kDa)
365 fused to MBP (42.1 kDa). Purified fusion proteins appeared as a single band following
366 the second elution off the maltose resin (Fig. 7A, B). We confirmed identity of the
367 purified protein by immunoblotting with monoclonal antibodies against MBP (not

368 shown). Levels of fusion protein were below detection for both constructs in crude lysate
369 of non-induced cells, but were readily identified in lysate of IPTG-induced cells. A
370 distinct crossreactivity with the α-MBP antibody was obtained in both supernatant and
371 pellet fractions, implying that the recombinant protein was in part directed to inclusion
372 bodies. In the elution fraction for CynH-MBP, a single band indicated the presence of the
373 fusion protein in a stable configuration. For CynS however, two bands were detected,
374 presumably representing CynS-MBP and a product resulting from spontaneous cleavage
375 of the fusion protein.

376 Aliquots of the second eluted fraction for both fusion constructs were
377 subsequently tested for cyanase activity. Measured as ammonium liberated from cyanate,
378 cyanase activity of CynS was maximal (5.56 U mg^{-1}) at 26 °C, and rapidly dropped to
379 50% of this maximum activity at both higher and lower temperatures (Fig. 8). The protein
380 became rapidly inactive at higher temperatures and only residual activity (0.09 U mg^{-1})
381 was detected at 50 °C (Fig. 8). Furthermore, we report here on a distinct cyanase activity
382 associated with the gene product of *cynH*, identified as a cyanate hydratase in genome
383 annotations, a characterization that so far lacked experimental evidence. In general,
384 cyanate dependent ammonium accumulation rates were similar to those obtained with
385 CynS: highest cyanase activities (5.87 U mg^{-1}) were measured at 26 °C and activities
386 declined at both higher and lower temperatures to 0.80 U mg^{-1} at 4 °C and 0.43 U mg^{-1} at
387 55 °C. Cyanase activities were derived specifically from either CynS or CynH, as
388 ammonium failed to accumulate when MBP or overexpressed NtcA-MBP (an N
389 regulatory protein) were added to the reaction mix (data not shown). No ammonium

390 accumulation was observed for CynS nor CynH in the presence of the cyanase inhibitor
391 Na-azide (Fig. 8).
392
393 **DISCUSSION**
394 Cyanase serves several functions, the most pronounced being the detoxification of
395 cyanate generated in various metabolic pathways (8, 38). Besides detoxification,
396 microorganisms employ cyanase in the assimilation of cyanate from the environment. *E.*
397 *coli* transports cyanate into the cell via a cyanate permease encoded by *cynX* (2, 37).
398 Cyanobacteria utilize cyanate following its acquisition via an ABC-type transport system
399 (12, 20, 27). It has been proposed that cyanate and urea play an important role in the N
400 cycle of marine oligotrophic environments (20). In this study we characterized the
401 evolution, marine distribution, and transcriptional regulation of *cynS* (cyanase) and the
402 activity associated with its gene product. We further report on a novel cyanase encoded
403 by *cynH* in marine *Synechococcus*.
404 CynS tree topologies show cyanobacteria as a well defined branch emerging at the
405 base of the bacterial lineage. Cyanase was found in *Synechocystis* sp. PCC6803 (22),
406 *Synechococcus elongatus* PCC7942 (42), the filamentous diazotrophs *Anabaena* sp.
407 ATCC29413 and *Nostoc* sp. PCC7120 (11), *Synechococcus* spp. PCC7002 and PCC7335
408 from brackish, estuarine waters (32), toxic bloom-forming *Microcystis* (21) and members
409 of the unicellular marine *Synechococcus* and *Prochlorococcus* (this study). >60% of the
410 GOS-derived cyanase sequences were identified as cyanobacterial. This includes ten
411 clones from a hypersaline lagoon that clustered with *Synechococcus* sp. RS9917, a
412 euryhaline ecotype (7, 14).

413 Cyanobacteria likely acquired cyanase during the very early stages of their
414 evolution. Tree topologies of cyanobacterial CynS matched phylogenies based on 16S
415 rRNA, ITS and *ntcA* (14, 24, 31, 33, 36), and branching was supported by strong
416 posterior probabilities. Based on these observations, we suggest that *cynS* was common in
417 ancestral cyanobacteria and *cynS* was lost from many modern cyanobacteria. Our tree
418 topology suggests that the importance of lateral gene transfer (LGT) of *cynS* was minor,
419 however it might occur among related species. Thus, the estuarine *Synechococcus* sp.
420 PCC7002 carries two *cynS* orthologs that share 74% identity at the amino acid level
421 (*Synechococcus* PCC7002a and b in Fig. 3). *Synechococcus* PCC7002a, encoded by a
422 stand-alone *cynS* gene, clustered with CynS of the endosymbiont *Acaryochloris marina*
423 MBIC11017, which is ancestral to *Synechococcus* sp. PCC7002 (3, 39). The
424 *Synechococcus* PCC7002b homolog is most closely related to CynS of *Synechococcus* sp.
425 PCC7335, and partakes in an NtcA-regulated *cynABDS* operon, very similar to our
426 observations for *Prochlorococcus* MED4. Thus, the cyanase gene is involved in lateral
427 gene transfer, suggesting that different CynS may carry out distinct functions in the
428 cyanobacterial cell. We propose that CynS by itself may provide the cell with means to
429 detoxify internally generated cyanate, whereas the *cynABDS* operon encodes the
430 utilization of external cyanate. The presence of two cyanase homologs on a single
431 genome suggests that both functions play distinct roles in the N-metabolism and N-
432 assimilation of cyanobacteria.

433 In an attempt to set apart the cyanate detoxification and utilization functions of
434 cyanase, we studied transcriptional regulation of *cynS* and *cynA* alongside that of the N-
435 regulatory gene *ntcA*. In *Prochlorococcus* sp. MED4, in contrast to *Synechococcus* sp.

436 WH8102, *cynABDS* genes showed coordinated expression in response to N deprivation
437 (Fig. 6). This is in agreement with the gene arrangement in the strains examined (Fig. 4).
438 In *Synechococcus* sp. WH7803, which lacks the transporter genes, N-independent
439 regulation of *cynS* was observed. This implies that *cynS* in *Prochlorococcus* sp. MED4
440 takes part in utilization of external cyanate, while the presence of *cynS* in *Synechococcus*
441 genomes indicates a requirement to detoxify an intracellular cyanate. There are several
442 possible sources for the intracellular cyanate in cyanobacteria. A substrate of the urea
443 cycle, carbamoyl phosphate is known to decompose spontaneously to cyanate and
444 phosphate (2). Urea undergoes spontaneous transformation to cyanate by an isomeric
445 change (6, 26). The origin of urea in *Synechococcus* and *Prochlorococcus* cells is
446 unclear, since they lack the *arg* gene product that facilitates urea hydrolysis (36).
447 However, marine cyanobacteria may convert excess arginine to spermidine by sequential
448 action of arginine decarboxylase (EC 4.1.1.19), agmatine ureaohydrolase (EC 3.5.3.11),
449 and spermidine synthase (EC 2.5.1.16). Hence, despite a lack of arginase (EC 3.5.3.1),
450 the toxic cyanate can transform from urea produced by agmatine ureaohydrolase.
451 In an attempt to confirm cyanase activity by CynS from unicellular marine
452 cyanobacteria, the *cynS* was cloned as a fusion construct with MBP (CynS-MBP),
453 overexpressed in an *E. coli* background and subsequently purified on amylose resin.
454 Using enzyme assays, we clearly identified CynS as a functionally active cyanase.
455 Similarly to cyanases described in other studies (2, 11, 41), the CynS-MBP construct
456 showed bicarbonate-dependent cyanate degrading activity that was inhibited by Na-azide,
457 the latter preventing binding of substrate to the holoenzyme (41). Cyanase activity was
458 also confirmed for a short peptide, product of an ORF that was tentatively annotated as

cyanate hydratase, and we propose to rename this ORF as *cynH*. The identification of an additional cyanase raises questions about its origin and physiological importance. The *cynH* gene was found in seven marine *Synechococcus* genomes as well as on the metagenomic clone GRIST19. Although located on the same genomic region, *cynS* is separated from *cynH* by 210-244 bp and is transcribed in the opposite orientation (Fig. 4). It is unlikely that *cynH* resulted from (partial) gene duplication as its amino acid sequence does not align with any part of *cynS*. The two genes are thus paralogs. Nine *cynH* sequences shared a high degree of identity in the C-terminal half (Fig. 9B) similar to CynS, suggesting that the catalytic domain of CynH is confined to this region. The catalytic site of CynS contains Arg (R), Glu (E), and Ser (S) residues (41) in a configuration conserved across members of the bacteria, fungi and plants (11). Interestingly, we identified three fully conserved amino acid residues in CynH, identical to those of CynS, and their configuration is reminiscent of the active site of CynS. Moreover, secondary structure predictions indicated the presence of an α-helix followed by a short β-sheet in the C-terminal domain of CynH in agreement with the secondary structure of the C-terminal domain of CynS (41) (Fig. 9A, B). No reliable prediction for tertiary structure of CynH is available due to the low similarity to any known protein and a lack of defined motifs. The dual role the *cynS* and *cynH* genes in marine *Synechococcus* remains to be clarified.

ACKNOWLEDGEMENTS

Plasmids with recombinant MBP-NtcA constructs were kindly provided by Dr. H. Zer and B. Rihtman. The Niedersachsen State Fund at the Hebrew University, the Israel

Science Foundation (grant 135/05) and the NATO Science for Peace program (grant SfP 98216) all provided financial support.

REFERENCES

1. **Anderson, P. M.** 1980. Purification and properties of the inducible enzyme cyanase. Biochemistry **19:**2882-2888.
2. **Anderson, P. M., Y. C. Sung, and J. A. Fuchs.** 1990. The cyanase operon and cyanate metabolism. FEMS Microbiol. Rev. **87:**247-252.
3. **Blank, C. E., and P. Sánchez-Baracaldo.** 2010. Timing of morphological and ecological innovations in the cyanobacteria - a key to understanding the rise in atmospheric oxygen. Geobiology **8:**1-23.
4. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248-254.
5. **Chin, C. C., P. M. Anderson, and F. Wold.** 1983. The amino acid sequence of *Escherichia coli* cyanase. J. Biol. Chem. **258:**276-282.
6. **Dirnhuber, P., and F. Schütz.** 1948. The isomeric transformation of urea into ammonium cyanate in aqueous solutions. Biochem. J. **42:**628-632.
7. **Dufresne, A., M. Ostrowski, D. J. Scanlan, L. Garczarek, S. Mazard, B. P. Palenik, I. T. Paulsen, N. T. de Marsac, P. Wincker, C. Dossat, S. Ferriera, J. Johnson, A. F. Post, W. R. Hess, and F. Partensky.** 2008. Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. Genome Biol. **9:**R90.
8. **Ebbs, S.** 2004. Biological degradation of cyanide compounds. Curr. Opin. Biotechnol. **15:**231-236.
9. **Edgar, R. C.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32:**1792-1797.
10. **Efron, B., and R. J. Tibshirani.** 1993. An Introduction to the Bootstrap. Chapman and Hall, New York.

11. **Elleuche, S., and S. Pöggeler.** 2008. A cyanase is transcriptionally regulated by arginine and involved in cyanate decomposition in *Sordaria macrospora*. Fungal Genet. Biol. **45:**1458-1469.
12. **Espie, G. S., F. Jalali, T. Tong, N. J. Zacal, and A. K.-C. So.** 2007. Involvement of the *cynABDS* operon and the CO_2-concentrating mechanism in the light-dependent transport and metabolism of cyanate by cyanobacteria. J. Bacteriol. **189:**1013-1024.
13. **Fuhrman, J.** 2003. Genome sequences from the sea. Nature **424:**1001-1002.
14. **Fuller, N. J., D. Marie, F. Partensky, D. Vaulot, A. F. Post, and D. J. Scanlan.** 2003. Clade-specific 16S ribosomal DNA oligonucleotides reveal the predominance of a single marine *Synechococcus* clade throughout a stratified water column in the Red Sea. Appl. Environ. Microbiol. **69:**2430-2443.
15. **Gelman, A., and D. Rubin.** 1992. Inference from iterative simulation using multiple sequences. Stat. Sci. **7:**457-511.
16. **Graybill, F. A., and H. K. Iyer.** 1994. Regression Analysis: Concepts and Applications, First eddition ed. Duxbury Pr, Florence, KY.
17. **Harano, Y., I. Suzuki, S.-i. Maeda, T. Kaneko, S. Tabata, and T. Omata.** 1997. Identification and nitrogen regulation of the cyanase gene from the cyanobacteria *Synechocystis* sp. strain PCC 6803 and *Synechococcus* sp. strain PCC 7942. J. Bacteriol. **179:**5744-5750.
18. **Huelsenbeck, J. P., and F. Ronquist.** 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics **17:**754-755.
19. **Johnson, W. V., and P. M. Anderson.** 1987. Bicarbonate is a recycling substrate for cyanase. J. Biol. Chem. **262:**9021-9025.
20. **Kamennaya, N. A., M. Chernihovsky, and A. F. Post.** 2008. The cyanate utilization capacity of marine unicellular cyanobacteria. Limnol. Oceanogr. **53:**2485-2494.
21. **Kaneko, T., N. Nakajima, S. Okamoto, I. Suzuki, Y. Tanabe, M. Tamaoki, Y. Nakamura, F. Kasai, A. Watanabe, K. Kawashima, Y. Kishida, A. Ono, Y. Shimizu, C. Takahashi, C. Minami, T. Fujishiro, M. Kohara, M. Katoh, N. Nakazaki, S. Nakayama, M. Yamada, S. Tabata, and M. M. Watanabe.** 2007.

Complete genomic structure of the bloom-forming toxic cyanobacterium *Microcystis aeruginosa* NIES-843. DNA Res. **14:**247-256.

22. **Kaneko, T., and S. Tabata.** 1997. Complete genome structure of the unicellular cyanobacterium *Synechocystis* sp. PCC6803. Plant Cell Physiol. **38:**1171-1176.

23. **Kelley, L. A., and M. J. E. Sternberg.** 2009. Protein structure prediction on the web: a case study using the Phyre server. Nature Protocols **4:**363-371.

24. **Lavin, P., P. Gómez, B. González, and O. Ulloa.** 2008. Diversity of the marine picocyanobacteria *Prochlorococcus* and *Synechococcus* assessed by terminal restriction fragment length polymorphisms of 16S-23S rRNA internal transcribed spacer sequences. Rev. Chil. Hist. Nat. **81:**515-531.

25. **Maeda, S.-i., and T. Omata.** 2009. Nitrite transport activity of the ABC-type cyanate transporter of the cyanobacterium *Synechococcus elongatus*. J. Bacteriol. **191:**3265-3272.

26. **Marier, J. R., and D. Rose.** 1964. Determination of cyanate, and a study of its accumulation in aqueous solutions of urea. Anal. Biochem. **7:**304-314.

27. **Miller, A. G., and G. S. Espie.** 1994. Photosynthetic metabolism of cyanate by the cyanobacterium *Synechococcus* UTEX 625. Arch. Microbiol. **162:**151-157.

28. **Moore, L. R., A. F. Post, G. Rocap, and S. W. Chisholm.** 2002. Utilization of different nitrogen sources by the marine cyanobacteria. Limnol. Oceanogr. **47:**989-996.

29. **Palenik, B., B. Brahamsha, F. W. Larimer, M. Land, L. Hauser, P. Chain, J. Lamerdin, W. Regala, E. E. Allen, J. McCarren, I. Paulsen, A. Dufresne, F. Partensky, E. A. Webb, and J. Waterbury.** 2003. The genome of a motile marine *Synechococcus*. Nature **424**.

30. **Parthasarathy, H., E. Hill, and C. MacCallum.** 2007. Global ocean sampling collection. PLoS biol. **5:**e83.

31. **Penno, S., D. Lindell, and A. F. Post.** 2006. Diversity of *Synechococcus* and *Prochlorococcus* populations determined from DNA sequences of the N-regulatory gene *ntcA*. Environ. Microbiol. **8:**1200-1211.

32. **Rippka, R., J. Deruelles, J. B. Waterbury, M. Herdman, and R. Y. Stanier.** 1979. Generic assignments, strain histories and properties of pure cultures of cyanobacteria. J. Gen. Microbiol. **111:**1-61.

33. **Rocap, G., D. L. Distel, J. B. Waterbury, and S. W. Chisholm.** 2002. Resolution of *Prochlorococcus* and *Synechococcus* ecotypes by using 16S-23S ribosomal DNA internal transcribed spacer sequences. Appl. Environ. Microbiol. **68:**1180-1191.

34. **Rocap, G., F. W. Larimer, J. Lamerdin, S. Malfatti, P. Chain, N. A. Ahlgren, A. Arellano, M. Coleman, L. Hauser, W. R. Hess, Z. I. Johnson, M. Land, D. Lindell, A. F. Post, W. Regala, M. Shah, S. L. Shaw, C. Steglich, M. B. Sullivan, C. S. Ting, A. Tolonen, E. A. Webb, E. R. Zinser, and S. W. Chisholm.** 2003. Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. Nature **424:**1042-1047.

35. **Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M.-A. Rajandream, and B. Barrell.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944-945.

36. **Scanlan, D. J., M. Ostrowski, S. Mazard, A. Dufresne, L. Garczarek, W. R. Hess, A. F. Post, M. Hagemann, I. Paulsen, and F. Partensky.** 2009. Ecological genomics of marine picocyanobacteria. Microbiol. Mol. Biol. Rev. **73:**249-299.

37. **Sung, Y. C., and J. A. Fuchs.** 1988. Characterization of the cyn Operon in *Escherichia coli* K12. J. Biol. Chem. **263:**14769-14775.

38. **Sung, Y. C., and J. A. Fuchs.** 1989. Identification and characterization of a cyanate permease in *Escherichia coli* K-12. J. Bacteriol. **171:**4674-4678.

39. **Swingley, W. D., M. Chen, P. C. Cheung, A. L. Conrad, L. C. Dejesa, J. Hao, B. M. Honchak, L. E. Karbach, A. Kurdoglu, S. Lahiri, S. D. Mastrian, H. Miyashita, L. Page, P. Ramakrishna, S. Satoh, W. M. Sattley, Y. Shimada, H. L. Taylor, T. Tomo, T. Tsuchiya, Z. T. Wang, J. Raymond, M. Mimuro, R. E. Blankenship, and J. W. Touchman.** 2008. Niche adaptation and genome expansion in the chlorophyll *d*-producing cyanobacterium *Acaryochloris marina*. Proc. Natl. Acad. Sci. U. S. A. **105:**2005-2010.

40. **Tolonen, A. C., J. Aach, D. Lindell, Z. I. Johnson, T. Rector, R. Steen, G. M. Church, and S. W. Chisholm.** 2006. Global gene expression of *Prochlorococcus* ecotypes in response to changes in nitrogen availability. Mol. Syst. Biol. **2:**1-15.

41. **Walsh, M. A., Z. Otwinowski, A. Perrakis, P. M. Anderson, and A. Joachimiak.** 2000. Structure of cyanase reveals that a novel dimeric and decameric arrangement of subunits is required for formation of the enzyme active site. Structure **8:**505-514.

42. **Waterbury, J. B., and R. Rippka.** 1989. Subsection 1: Order Croococcales. Wettsten 1924, emend. Rippka et al., 1979, p. 1728-1746. *In* J. T. Staley, M. P. Bryant, N. Pfenning, and a. J. G. Holt (ed.), Bergey's Manual of Systematic Bacteriology, vol. 3. Williams and Wilkins, Baltimore, MD.

43. **Waterbury, J. B., S. W. Watson, F. W. Valois, and D. G. Franks.** 1986. Biologial and ecological characterization of the marine unicellular cyanobacterium *Synechococcus*, p. 71-120. *In* T. Platt and W. Li (ed.), Photosynthetic Picoplankton, vol. 214. Department of Fisheries and Oceans, Ottawa.

44. **Waterbury, J. B., and J. M. Willey.** 1988. Isolation and growth of marine planktonic cyanobacteria, p. 100-105, Methods Enzymol., vol. Volume 167. Academic Press.

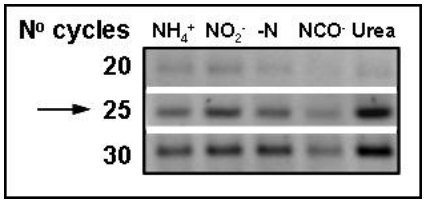
45. **Wilson, D., R. Pethica, Y. Zhou, C. Talbot, C. Vogel, M. Madera, C. Chothia, and J. Gough.** 2009. SUPERFAMILY - sophisticated comparative genomics, data mining, visualization and phylogeny. Nucleic Acids Res. **37:**380-386.

TABLES

Table 1. Primer (F, forward; R, reverse) sequences used in the RT-PCR amplification of *cynA*, *cynS*, *ntcA*, and PMM0615 products from *Prochlorococcus* sp. MED4; *cynS*, *ntcA*, and *orf0252* – from *Synechococcus* sp. WH7803; and *cynA*, *cynS*, and 16S rRNA – from *Synechococcus* sp. WH8102, with their annealing temperatures and expected size of PCR products.

Target gene	Primer	Sequence (5'–3')	Ta (°C)	size (bp)
Prochlorococcus sp.				
cynA	RTcA-Med4	GGAGGTAGCTAAGGCTATTT	52.5	198
	RTcA-Med4	CCTCCTAGATCCCATCTTAT		
cynS	RTcS-Med4	CCTACGGATCCTCTTATCTA	52	163
	RTcS-Med4	CTAGAACCCTATCTCCCTTT		
ntcA	RTnA-Med4	AGAGGAGCAGTAAGGTTATC	51	116
	RTnA-Med4	TCAGACCTATGTCCTGTTAG		
PMM0615	RTorf1206	CCCTGAACTTTATAGACACC	52	202
	RTorf1206	GACTTTGTCTTCTCCCATAG		
Synechococcus sp.				
cynA	cA-8102 F	GCCTCTATTCACTCTAGTTCCC	53	828
	cA-8102 R	GCGAATTATGCAACAAATCCTA		
cynS	cS-8102 F	AGGTTTGGGTTGCATCTTTG	52.5	235
	cS-8102 R	TCTCCGAAATGCTCCTGAAT		
16S rRNA	16S-8102	CATCATGCCCCTTACATCCT	56	103
	16S-8102	AACTGAGCCACGGTTTATGG		
Synechococcus sp.				
cynS	RTcS-2 F	GGCCACAGCATCAGCGGAGG	63	283
	RTcS-2 R	GGTGATCTTCACGCGATCGCC		
orf0252	RTorf0252	GTGCCCGGTATGTCTTCCTT	59	150
	RTorf0252	ATTGGCGTCTGTGTTGAGGT		

633 **FIGURES**



634

635 Fig. 1. Typical example of RT-PCR amplification of *cynS* cDNA for *Synechococcus* sp.
636 WH7803. 50 µl samples were run for 30 cycles, and 6 µl subsamples were collected at 5
637 cycle intervals. PCR products from each cycle set were run on a 1.5% agarose gel and
638 visualized by staining with ethidium bromide. Band density was determined for the set in
639 which subsamples had clearly not reached saturation phase (arrow), assuming that they
640 most closely resembled the phase of exponential amplification.

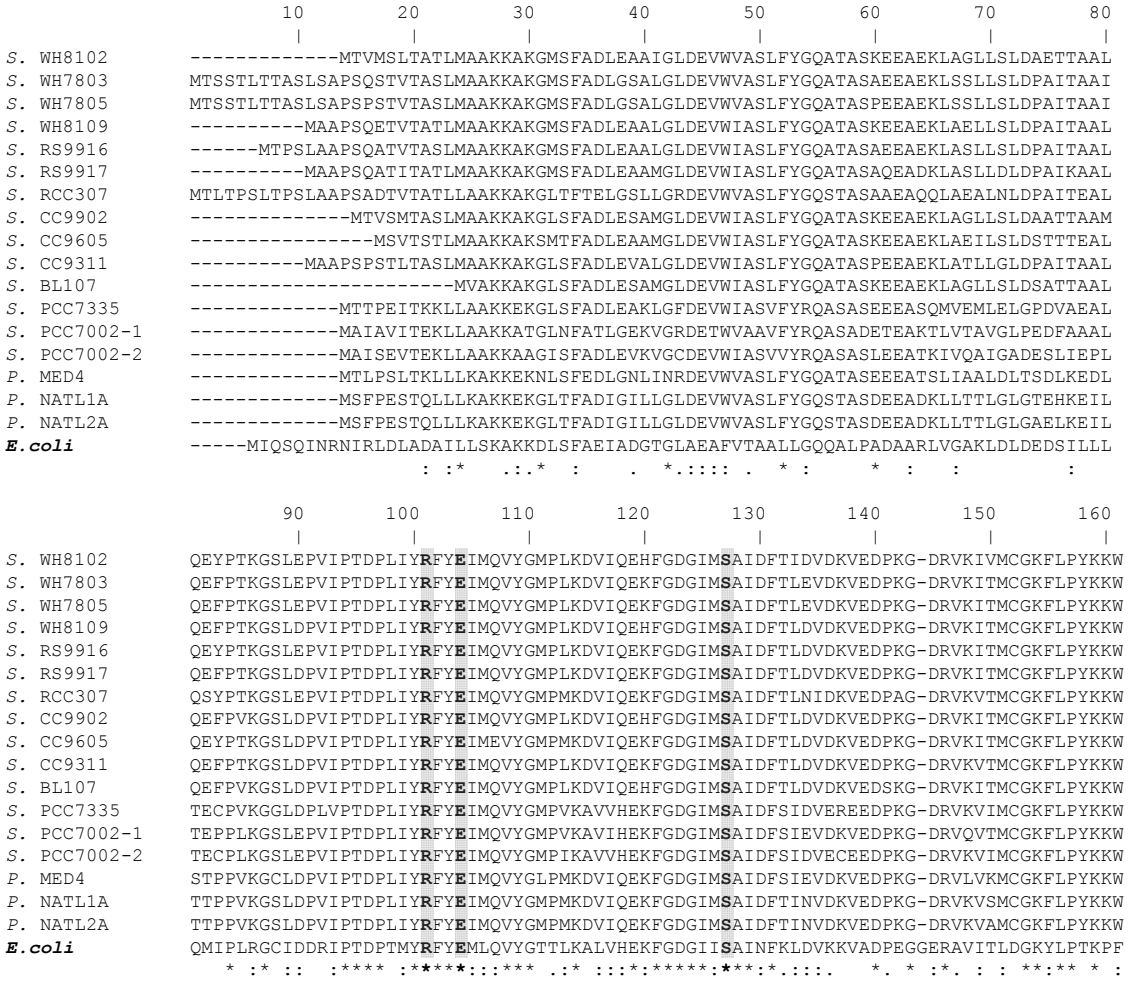
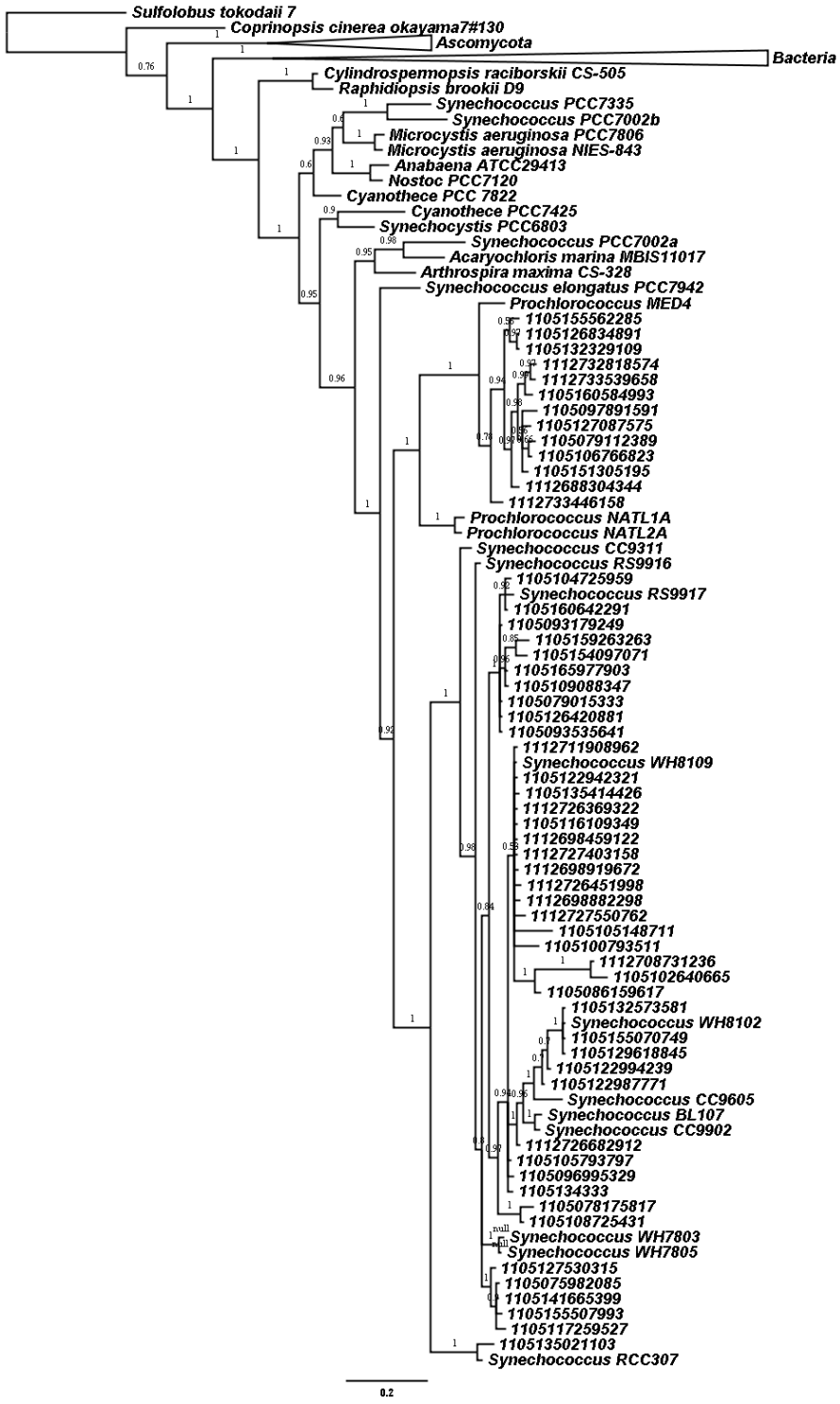


Fig. 2. Alignment of the amino acid sequences for seventeen putative cyanase genes in marine *Prochlorococcus* (P.) and *Synechococcus* (S.) strains along with that of the characterized *E. coli* cyanase. Fully conserved residues are labeled with "*", conserved replacements – with ":" and functional similarity – with ".". The proposed active-site residues Arg-96, Glu-99, Ser-122 are shown in bold against a gray background. GenBank accession numbers for cyanase sequences in alignment are: *Synechococcus* spp. WH8102 (NP_898579), WH7803 (YP_001226218), WH7805 (ZP_01124911), WH8109 (ZP_05789360), RS9916 (ZP_01471501), RS9917 (ZP_01079240), RCC307 (YP_001228741), CC9902 (YP_378288), CC9605 (YP_382939), CC9311 (YP_732088), BL107 (ZP_01469110), PCC7335 (ZP_05037924), PCC7001 (ZP_05043889,

652 YP_001733904); *Prochlorococcus* spp. MED4 (NP_892492), NATL1A

653 (YP_001013899), NATL2A (YP_292581); *E. coli* strain K12 (NP_414874) (5).

654



655

Fig. 3. Tree topology resulting from Bayesian phylogenetic analysis of translated environmental *cynS* sequences (135 aa) derived from the Global Ocean Survey database, along with CynS sequences retrieved from GenBank. Detailed relationships among cyanobacterial clades are shown while bacterial and fungal branches are collapsed for the purpose of presentation. The scale bar provides a distance measure for 2 substitutions per 100 nucleotides between sequences. Posterior probabilities are given at the nodes using a 0–1 scale. The denomination "JCVI PEP" has been omitted from all environmental sequences in order improve presentation.

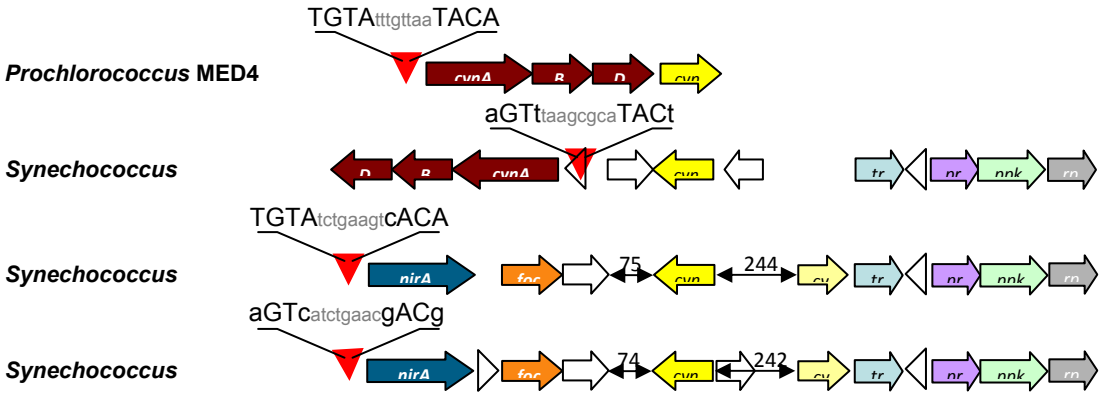
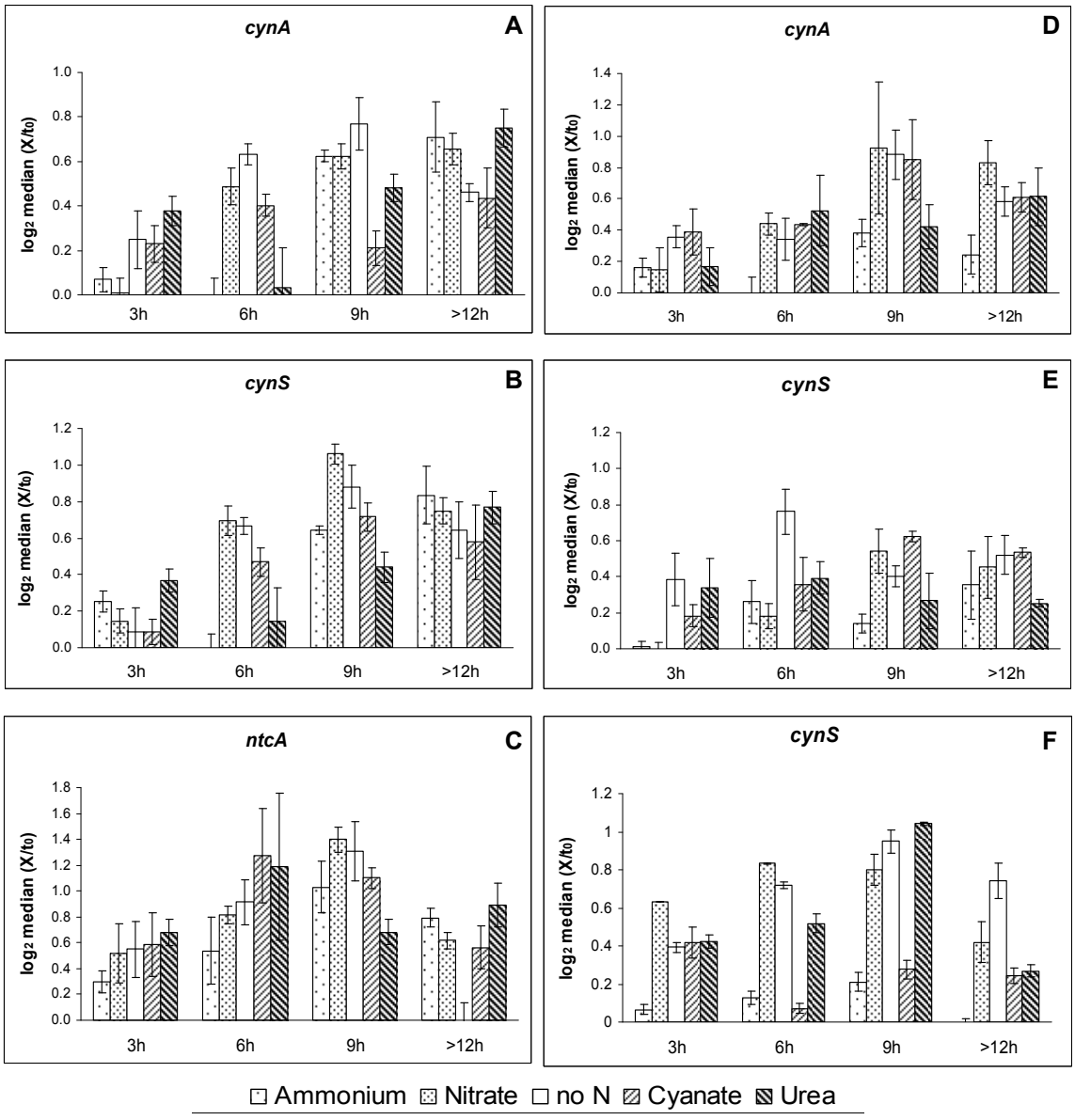


Fig. 4. Schematic representation of the genome regions that contain the cyanate utilization genes in *Prochlorococcus* sp. MED4 and *Synechococcus* spp. WH8102, WH7803, and WH7805. Vertical red arrows indicate location of putative NtcA binding sites shown above the arrows, numbers indicate the nucleotide distance between two adjacent genes. Horizontal arrows indicate *cynABD* genes encoding for the cyanate transporter (brown), *nirA* (COG0155, dark blue); *focA* (COG2116, orange); *cynS* (bright yellow), *cynH* (pale yellow), *trpD* (COG0547E, blue), *proP* (COG2814G, purple), *ppk* (COG0855P, green), *rpoD* (COG0568K, grey), and ORFs encoding unidentified proteins (white).



674

675 Fig. 5. Transcript accumulation of *cynA* (A, D), *ntcA* (C) and *cynS* (B, E, F), in
676 *Prochlorococcus* sp. MED4 (A-C), *Synechococcus* sp. WH8102 (D, E), and
677 *Synechococcus* sp. WH7803 (F) cells grown on ammonium, alternative N-sources or
678 deprived of combined N, for 3, 6, 9, and >12 hours after resuspension in fresh medium.
679 Data are $\log_2$ of median values of 3 replicates, normalized to their initial transcription
680 level, ±25-75 percentiles.

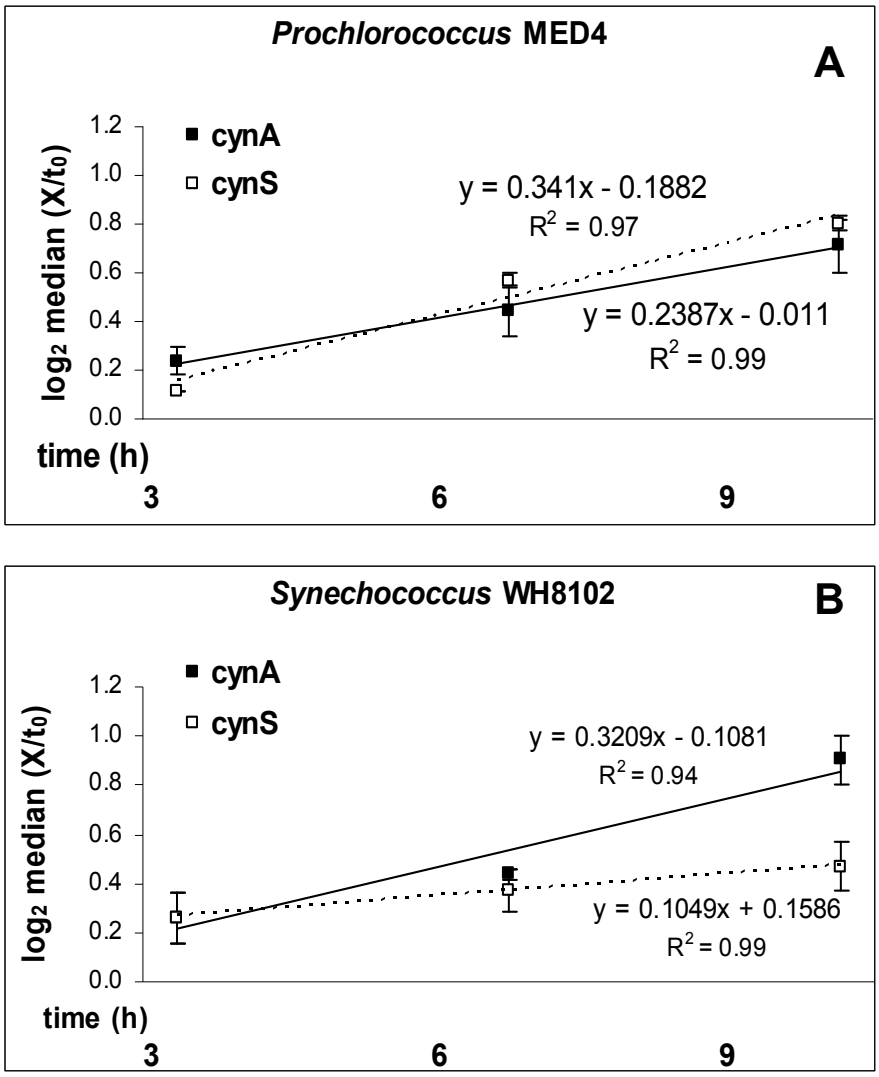
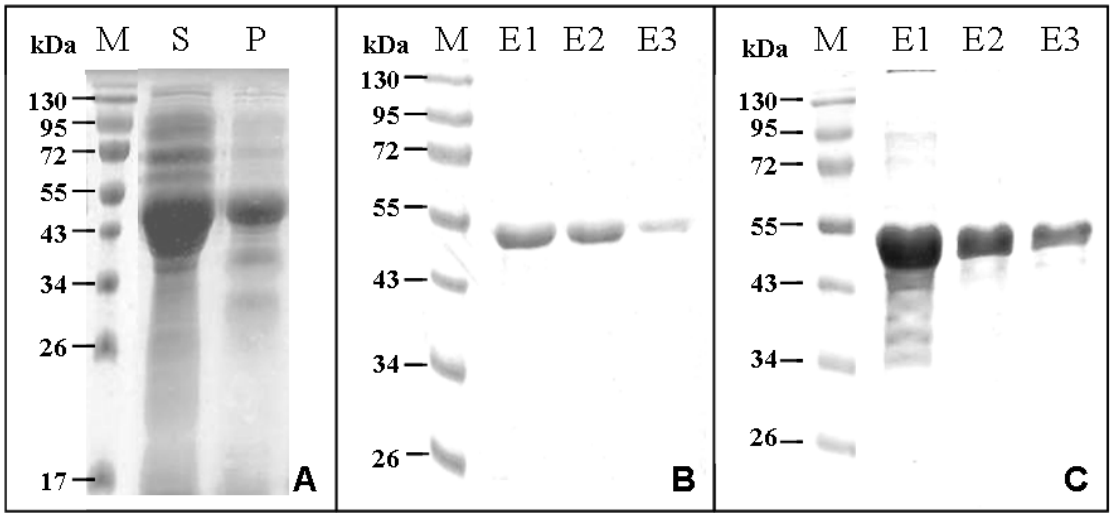


Fig. 6. Median log normalized values of *cynA* and *cynS* transcription for *Prochlorococcus* sp. MED4 (A) and *Synechococcus* sp. WH8102 (B) cultures grown on alternative N-sources or with no N, for 3, 6, and 9 hours after medium replacement. The error bars represent ±25-75 percentiles. Linear regression values and R-square values for *cynA* (continuous) and *cynS* (dashed) accumulation are presented above and below the respective trend line.

690



691

692 Fig. 7. IPTG-induced expression of CynS and its accumulation in soluble (S) and
693 particulate (P) fractions of *E. coli* cell lysates (panel A). Purification of CynS (B) and
694 CynH (C) recombinant proteins using amylose resin-based affinity chromatography. M,
695 molecular weight markers; E1-E3, elution fractions.

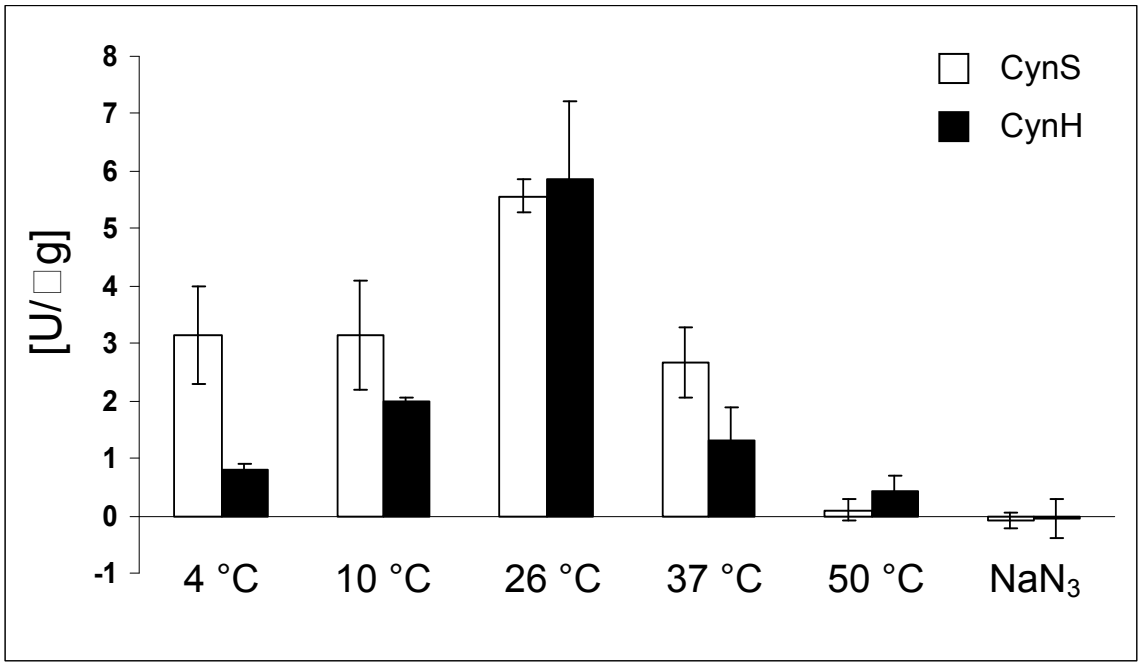


Fig. 8. Characterization of cyanase activity of CynS and CynH fusion constructs at different temperatures and their sensitivity to Na-azide (200 μmol L^{-1}) addition. Freshly prepared sodium cyanate (2 mmol L^{-1}) was added to the reaction mix (50 mmol L^{-1} PBS buffer (pH 7.6), 3 mmol L^{-1} sodium bicarbonate) complemented with aliquots of recombinant protein (20 μg) and incubated at the desired temperature for 10 min. Ammonium accumulation was determined with Nessler reagent. Activities shown are averages from at least nine replicates from three independent experiments.

A

```
S. WH8102      73  KGSLEPVIPTDPLIYRFYEIMQVYGMPLKDVIQEHFGDGIMSAIDFTIDV  122
S. WH7803      86  KGSLEPVIPTDPLIYRFYEIMQVYGMPLKDVIQEKFGDGIMSAIDFTLEV  135
S. WH7805      86  KGSLEPVIPTDPLIYRFYEIMQVYGMPLKDVIQEKFGDGIMSAIDFTLEV  135
S. WH8109      76  KGSLDPVIPTDPLIYRFYEIMQVYGMPLKDVIQEHFGDGIMSAIDFTLDV  125
S. RS9916      80  KGSLEPVIPTDPLIYRFYEIMQVYGMPLKDVIQEKFGDGIMSAIDFTLDV  129
S. RS9917      76  KGSLDPVIPTDPLIYRFYEIMQVYGMPLKDVIQEKFGDGIMSAIDFTLDV  125
S. RCC307      86  KGSLEPVIPTDPLIYRFYEIMQVYGMPMKDVIQEKFGDGIMSAIDFTLNI  135
S. CC9902      72  KGSLDPVIPTDPLIYRFYEIMQVYGMPLKDVIQEHFGDGIMSAIDFTLDV  121
S. CC9605     142  KGSLDPVIPTDPLIYRFYEIMEVYGMPMKDVIQEKFGDGIMSAIDFTLDV  191
S. CC9311      76  KGSLDPVIPTDPLIYRFYEIMQVYGMPLKDVIQEKFGDGIMSAIDFTLDV  125
S. BL107       69  KGSLDPVIPTDPLIYRFYEIMQVYGMPLKDVIQEHFGDGIMSAIDFTLDV  112
S. PCC7335     73  KGGLDPLVPTDPLIYRFYEIMQVYGMPVKAVVHEKFGDGIMSAIDFSIDV  122
S. PCC7002-1   73  KGSLEPVIPTDPLIYRFYEIMQVYGMPVKAVIHEKFGDGIMSAIDFSIEV  122
S. PCC7002-2   73  KGSLEPVIPTDPLIYRFYEIMQVYGMPIKAVVHEKFGDGIMSAIDFSIDV  122
P. MED4        73  KGCLDPVIPTDPLIYRFYEIMQVYGLPMKDVIQEKFGDGIMSAIDFSIEV  122
P. NATL1A      73  KGSLDPVIPTDPLIYRFYEIMQVYGMPMKDVIQEKFGDGIMSAIDFTINV  122
P. NATL2A      73  KGSLDPVIPTDPLIYRFYEIMQVYGMPMKDVIQEKFGDGIMSAIDFTINV  122
E. coli        81  RGCIDDRIPTDPTMYRFYEMLQVYGTTLKALVHEKFGDGIISAINFKLDV  130
                   :*     :****::*****:::***   * :  * *****:*** *   :
                   CCCCCCCCCCCCCHHHHHHHHHHHHHHHHHHHHHHHHHCCCCEEEEEECCCEE
```

B

```
S. WH7805              MSTLFRSFSSLFFSQQLESS------GASSLLLERLYYAEGRQNPDHPLHGSFAGLSYFDSP
S. RS9916              MIQLFARRLTRSATPSSAFTANN--TSFPSLIVERLYYAEGRHHPDHPLHGSHTGLSRW
S. CC9902              MSQLLRSLSSLLGSQKSDNA------TLTSLVLERLYYADGRHHPEHPNHGSFQGLTTFS
S. CC9311-1          MTNMKLFSVIAAITLMTTPAQASIG--MKAPNLLMEKLYYAEGRQHPDHPLHGSFSGLCCGTDP
S. CC9311-2  MFLSAIVNQFNQIQNQILKIINSFLTIDSLPVTPTVVSQTFPNLLIERLYYAEGRQHPFHPLHGSYAGLCRT
S. WH7803              MSALFRSFSSLLFSQQLESS------GASNLVLERLYYADGRQNPDHPLHGSFTGLTYLDSP
S. WH8109              MSQFFQSVATLLQAQFLPPT------TAPQLMLERLYYAEGRHHPQHPRHGSFEGLSRLSRP
S. BL107             KQNMSQLLRSLSSLLGSQQTNNA------TLTSLVLERLYYADGRHHPEHPKHGSFQGLSTVT
GRIST19                MSQLLRSLSFLLGSQQADNA------TLTSLVLERLYYADGRHHPEHPKHGSVQGLSTDA
                          ::              :        ..*::*:****:**::* ** ***. **
                          CCCHHHHHHHHHHHHHHHHC      CHHHHHHHHHHHHHCCCCCCCCCCCCCEEEEEC
```

Fig. 9. Amino acid alignments for catalytic site region of cyanases for 17 CynS sequences (panel A) and 9 full-length CynH sequences (panel B) found in marine *Synechococcus* and in a single metagenomic clone GRIST19. Identical residues are presented in bold and labeled (*), conserved residue replacements are labeled (:) and functionally similar residues as (.). The proposed active-site residues Arg (R), Glu (E), and Ser (S) are shown in bold against a gray background. The bottom line presents the consensus secondary structure predicted by the Jpred prediction server that identifies randomly coiled region (*C*), α-helix (*H*), and β-sheet (*E*) motifs. In panel B, the aligned sequences are: cyanate hydratase (ZP_01124909) in *Synechococcus* sp. WH7805; RS9916_37357

753 (ZP_01471502) in *S.* RS9916; Syncc9902_2288 (YP_378289) in *S.* CC9902; sync_2840
754 and sync_2903 (YP_732028 and YP_732090) in *S.* CC9311; SynWH7803_2496
755 (YP_001226219) is *S.* WH7803; SH8109_0550 (ZP_05789530) in *S.* WH8109; non-
756 annotated (reverse strand 1870903-1871064) in *S.* BL107 (NZ_AATZ00000000); and
757 metagenomic clone GRIST19 (EU795157).