



# HOKKAIDO UNIVERSITY

Title	Identification of a gene essential for protoporphyrinogen IX oxidase activity in the cyanobacterium <i>Synechocystis</i> sp. PCC6803
Author(s)	Kato, K.; Tanaka, R.; Sano, S. et al.
Citation	Proceedings of the National Academy of Sciences, 107(38), 16649-16654 <a href="https://doi.org/10.1073/pnas.1000771107">https://doi.org/10.1073/pnas.1000771107</a>
Issue Date	2010-09-21
Doc URL	<a href="https://hdl.handle.net/2115/48275">https://hdl.handle.net/2115/48275</a>
Type	journal article
File Information	PNAS2011_tanaka_SI.pdf, Supporting Information



SI Text

## **SI Materials and methods**

### **Expression of an *slr1790* homologue from *Rhodobacter sphaeroides* in *E. coli***

The *Rs-slr1790* gene was amplified by PCR from *Rhodobacter sphaeroides* DNA (a kind donation from Prof. Shinji Masuda, Tokyo Institute of Technology). The amplicon was cloned into pMAL-c2E (New England Biolab), which encodes maltose-binding protein (MBP). The construct was introduced into the *E. coli* JM109 strain. The *E. coli* cells were inoculated with Luria broth (2 L) supplemented with ampicillin at 37°C and when the optical density of the culture reached 0.2 at 600 nm, IPTG was added to the culture at a final concentration of 0.4 mM. The cells were further grown for 4 hrs at 37°C and were harvested by centrifugation at 5,000g for 10 min. The cells were resuspended in 15 ml of a buffered solution containing 50 mM Tris-Cl (pH 8.0), 50 mM NaCl, 0.04% Tween20, protease inhibitor cocktail (P8465, Sigma). Cells were subsequently disrupted by sonication and a soluble fraction was collected after centrifugation at 40,000 x g for 20 min.

### **Purification of Rs-slr1790 protein with MBP**

Recombinant Rs-slr1790 protein was purified with amylose resin (Amylose Resin High Flow,

E8022S, New England Biolab), by passing supernatant (16 ml) over the resin at 4°C at a flow rate of 180 ul/min. The column was then washed with 10 ml buffer containing 50 mM Tris-Cl (pH 8.0), 50 mM NaCl, and 0.04% Tween20 at a flow rate of 240 ul/min. The Rs-slr1790 protein was eluted with 50 mM Tris-Cl (pH 8.0) containing 50 mM NaCl, 0.04% Tween20, and 20 mM maltose, at a flow rate of 120 ul/min.

### **Preparation of Protogen IX**

The Protox assay was carried out according to the method of Jacobs and Jacobs (1). Proto IX solution (200 uM Proto IX) was prepared by dissolving 3 mg of Proto IX (Sigma) in 25 ml of 20% ethanol and the solution was stirred for 30 min in dark conditions. The following procedure for preparation of Protogen IX was performed in the dark in a closed box under a stream of nitrogen gas.

An 800 ul Proto IX solution was placed into a 15-ml screw test tube and flushed with purified nitrogen gas for 5 min. Sodium amalgam (0.5 g, Sigma) was added to Proto IX solution and incubated for 2 min with vigorous shaking. The lid of the tube was opened one time to relieve the gas pressure that was generated by the sodium amalgam. After releasing the pressure, the lid was again closed and the incubation was continued for additional 3 min. , Protogen IX solution was

subsequently passed through a regenerated cellulose membrane filter (54513-RC-100, SUN-SRi). The pH of the filtrated Protogen IX solution (600 ul) was adjusted to pH 8.0 with 2M MOPS (ca. 300 ul). The Protogen IX was immediately used for the Protox assay. The actual concentration of Protogen IX solution was determined after completion of the assay. The solution was allowed to be oxidized by air over night, and was then mixed with 2.7N HCl. The Proto IX concentration was quantified by absorbance at 408 nm using a calibration curve that was generated with a serial dilution of the standard Proto IX pigment.

### **Protox assay**

Two hundred microliters of a reaction mixture containing a protein aliquot (equivalent to 20 ug protein), 50 mM Tris-Cl (pH 8.0), 50 mM NaCl , 0.04% Tween20, was preincubated for 30 min with 40 mM glucose, 5U glucose oxidase (from *Aspergillus niger*, Wako Chemicals) and 10U catalase (from bovine liver, Sigma) to reduce the oxygen concentration within the tube. The reaction mixture was dispensed into multiple wells of a 96-well plate and mineral oil was layered on the top of the reaction mixture in each well. The reaction was initiated by adding 1/10 volume of Protogen IX to the reaction mixture. This amount of Protogen IX corresponded to a final concentration of

approximately 10  $\mu$ M. Negative control reaction mixtures containing the MBP tag conjugated with the  $\alpha$  subunit of  $\beta$ -galactosidase, or a mixture lacking protein were prepared to estimate non-enzymatic oxidation of Protogen IX during the assay. The reaction was initiated by adding Protogen IX to the mixture and the reaction proceeded at room temperature up to 30 min. The plate was set in a fluoremetric photometer (Infinite M200 fluorometric photometer, Tecan Männedorf, Switzerland) equipped with the 405 nm (20 nm bandpass) and 633 nm (20 nm bandpass) filters for excitation and emission, respectively. The Proto IX contents were measured by fluorescence using a calibration curve constructed with a serial dilution of the standard Proto IX.

1. Jacobs JM & Jacobs NJ (1999) in *Current Protocols in Toxicology*, eds Costa LG, Hodgson E, Lawrence DA, & Reed DJ (John Wiley & Sons, Inc., Hoboken, NJ).

## Legends to supporting information

### Fig. S1 Disruption of the *slr1790* gene in the WT and AT strains

(A) Strategy for disruption of the *slr1790* in the genome of the *Synechocystis* WT and AT strains.

(B) Genomic DNA extracted from WT (lane 1), AT (lane 2), WTSK (lane 3) and ATSK (lane 4)

was amplified using the following primers: *slr1790* km *EcoRI* fr.

(5'-GGGGAATTCTGCTTGCATCAATATGGTGGC-3') and *slr1790* km *HindIII* rev.

(5'-GGGAAGCTTACCCTGGAGATCCACTGGTT-3') to check the segregation of the mutant genome.

Fig. S2. Multiple sequence alignment of HemJ (*slr1790* homologues) and NuoM (the M subunit of NADH oxidoreductase complex I). Four HemJ sequences from *Synechocystis* sp. PCC6803 (P72793), *Gloeobacter violaceus* PCC 7421 (NP\_925986), *Flavobacterium* sp. MED217 (ZP\_01059605) and *Rhodobacter sphaeroides* ATCC 17025 (YP\_001168861), and five *nuoM* sequences from *Bombyx mori* (BAB84656), *Drosophila melanogaster* (CAB91059), *Homo sapiens* (ABX40271), *Synechocystis* sp. PCC6803 (K05575) and *Escherichia coli* (CAA48372) were first aligned by the ClustalW program (1) and the alignment was then

manually adjusted. Black boxes with white letters show conserved residues among at least six sequences. Gray boxes indicate similar residues among at least six organisms. Arrows indicate the positions of essential residues for the activity of NuoM (2).

Fig. S3. Multiple sequence alignment of HemJ (slr1790 homologues) proteins among cyanobacteria. The sequences from thirteen cyanobacterial genomes were aligned by the ClustalW program (1). Bars above the alignment indicate the presumable trans-membrane regions that were predicted by the Phobius server (3). The accession numbers for the sequences are as follows: *Synechocystis* sp. PCC6803 (P72793), *Crocospaera watsonii* WH 8501 (ZP\_00515168), *Acaryochloris marina* MBIC11017 (YP\_001517117), *Nostoc punctiforme* PCC 73102 (ZP\_00112084), *Synechococcus elongatus* PCC 7942 (YP\_399867), *Gloeobacter violaceus* PCC 7421 (NP\_925986), *Synechococcus* sp. WH 7805 (ZP\_01124055), *Synechococcus* sp. WH 5701 (ZP\_01085493), *Synechococcus* sp. RCC307 (YP\_001227460), *Synechococcus* sp. WH 8102 (NP\_897336), *Prochlorococcus marinus* str. NATL2A (YP\_291524), *Prochlorococcus marinus* str. MIT 9312 (YP\_397415), *Prochlorococcus* (1)str. AS9601 (YP\_001009371).

Fig. S4. Deduced phylogenetic relationships of *slr1790* homologues

The evolutionary history was deduced using the Neighbor-Joining method (4). Evolutionary distances were computed using the JTT matrix-based method (5) and are in the units of the number of amino acid substitutions per site. There were a total of 248 positions in the final dataset. Phylogenetic analyses were conducted in MEGA4 (6). The bootstrap values from 500 calculations were indicated in percentage on the nodes for each clade unless the values were less than 80%. An asterisk indicates the position of the sole  $\delta$ -proteobacterium *Plesiocystis pacifica* SIR-1 in this phylogenetic tree. Accession numbers of the sequences are as follows:

*Acaryochloris marina* MBIC11017 (YP\_001517117), *Agrobacterium tumefaciens* str. C58

(NP\_355760), *Algoriphagus* sp. PR1 (ZP\_01719224), *Alkalilimnicola ehrlichei* MLHE-1

(YP\_741275) *Azoarcus* sp. EbN1 (YP\_157987), *Azotobacter vinelandii* AvOP (ZP\_00416918),

*Beggiatoa* sp. PS (ZP\_02003052), *Bordetella bronchiseptica* RB50 (NP\_890225),

*Bradyrhizobium japonicum* USDA 110 (NP\_767276), *Burkholderia xenovorans* LB400

(YP\_560405), *Coxiella burnetii* RSA 493 (NP\_820857), *Croceibacter atlanticus* HTCC2559

(ZP\_00949685), *Crocospaera watsonii* WH 8501 (ZP\_00515168), *Flavobacteriales bacterium*

HTCC2170 (ZP\_01105552), *Flavobacterium* sp. MED217 (ZP\_01059605), *Fulvimarina pelagi*  
HTCC2506 (ZP\_01438752), *Gloeobacter violaceus* PCC 7421 (NP\_925986), *Helicobacter*  
*pylori* HPAG1 (YP\_628170), *Legionella pneumophila subsp. pneumophila* str. Philadelphia 1  
(YP\_095565), *Maricaulis maris* MCS10 (YP\_758190), *Marinobacter* sp. ELB17  
(ZP\_01737473), *Mesorhizobium loti* MAFF303099 (NP\_105347), *Methylobacillus flagellatus*  
KT (YP\_544585), *Methylobacterium* sp. 4-46 (ZP\_01850302), *Methylophilales bacterium*  
HTCC2181 (ZP\_01552597), *Nitrococcus mobilis* Nb-231 (ZP\_01128952), *Nitrosomonas*  
*eutropha* C91 (YP\_747168), *Nitrospira multiformis* ATCC 25196 (YP\_411401), *Nostoc*  
*punctiforme* PCC 73102 (ZP\_00112084), *Novosphingobium aromaticivorans* DSM 12444  
(YP\_495402), *Oceanospirillum* sp. MED92 (ZP\_01166467), *Parvibaculum lavamentivorans*  
DS-1 (YP\_001412536), *Pedobacter* sp. BAL39 (ZP\_01884751), *Plesiocystis pacifica* SIR-1  
(ZP\_01905889), *Prochlorococcus marinus* str. AS9601 (YP\_001009371), *Prochlorococcus*  
*marinus* str. MIT 9313 (NP\_894557), *Prochlorococcus marinus* str. NATL2A (YP\_291524),  
*Pseudomonas syringae* pv. *phaseolicola* 1448A (YP\_272981), *Psychrobacter arcticus* 273-4  
(YP\_265144), *Rhizobium etli* CFN 42 (YP\_471595), *Rhodopseudomonas palustris* CGA009  
(NP\_945650), *Rhodospirillum rubrum* ATCC 11170 (YP\_428699), *Rickettsia bellii* OSU 85-389

(YP\_001496923), *Robiginitalea biformata* HTCC2501 (ZP\_01119876), *Roseobacter* sp.  
AzwK-3b (ZP\_01901421), *Roseobacter* sp. CCS2 (ZP\_01750153), *Roseovarius nubinhibens*  
ISM (ZP\_00959511), *Roseovarius* sp. 217 (ZP\_01038603), *Sulfurimonas denitrificans* DSM  
1251 (YP\_392907), *Synechococcus elongatus* PCC 7942 (YP\_399867), *Synechococcus* sp.  
RCC307 (YP\_001227460), *Synechococcus* sp. WH 5701 (ZP\_01085493), *Synechococcus* sp.  
WH 7805 (ZP\_01124055), *Synechococcus* sp. WH 8102 (NP\_897336), *Synechocystis* sp.  
PCC6803 (P72793), *Wolbachia* endosymbiont of *Drosophila melanogaster* (NP\_966207),  
*Wolinella succinogenes* DSM 1740 (NP\_907397)

### Supplementary Table Legends

Table S1. Distribution of *slr1790*, *HemY* and *HemG* homologues among cyanobacteria

+ indicates the presence of the homologues protein with an E value lower than  $10e^{-5}$  using a BLAST search; - indicates the absence of a protein with the E value lower than  $10e^{-5}$ ; Accession numbers for each query sequences are as follows: *slr1790* (*Synechocystis* sp. PCC 6803: P72793), *HemY* (*Homo sapiens*: P50336.1), *HemG* (*E. coli*: P27863). Note: Eight *Prochlorococcus* and eight *Synechococcus* species that are not shown in the list contain only

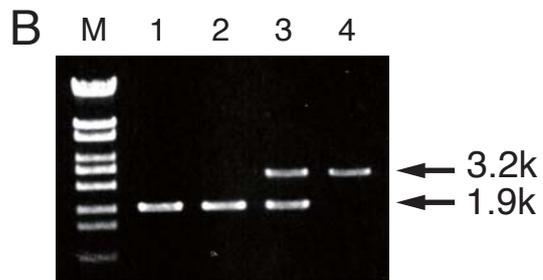
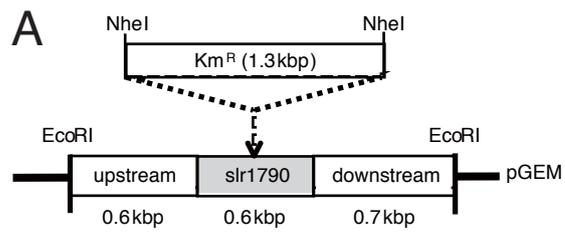
*slr1790* homologues.

Table S2. Distribution of *slr1790*, HemG and HemY homologues within Bacteroidetes/Chlorobi.

+ indicates the presence of the homologues protein with an E value lower than  $10e^{-5}$  using a BLAST search; - indicates the absence of a protein with the E value lower than  $10e^{-5}$ ; Accession numbers for each query sequences are as follows: *slr1790* (*Synechocystis* sp. PCC 6803: P72793), *HemY* (*Homo sapiens*: P50336.1), *HemG* (*E. coli*: P27863).

1. Chenna R, *et al.* (2003) Multiple sequence alignment with the Clustal series of programs. *Nucl Acid Res* 31:3497-3500.
2. Torres-Bacete J, Nakamaru-Ogiso E, Matsuno-Yagi A, & Yagi T (2007) Characterization of the NuoM (ND4) subunit in *Escherichia coli* NDH-1: CONSERVED CHARGED RESIDUES ESSENTIAL FOR ENERGY-COUPLED ACTIVITIES. *J Biol Chem* 282:36914-36922.
3. Käll L, Krogh A, & Sonnhammer EL (2007) Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. *Nucl Acid Res* 35:W429-432.
4. Saitou N & Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406-425.
5. Jones DT, Taylor WR, & Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* 8:275-282.
6. Tamura K, Dudley J, Nei M, & Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596-1599.





Synechocystis\_hemJ 1 -----MAYYWF-KAFHLIGIVVWFAG-----  
 Gloeobacter hemJ 1 -----MAYLWF-KAFHIVGFVTWFAG-----  
 Flavobacterium hemJ 1 -----MEAYYLYF-KSLHLIFIVTWFAG-----  
 Rhodobacter hemJ 1 -----MQMPNRFEEITMGTF LADYLLWT-KSLHVISVLAWMAG-----  
 Bombyx nuom 72 MIMASESLFKLN--FYFNFFLFNVI FLLIMLYLTFSVLNLLL FYFFFEASLIPTLLLIIGWGY----QPERIQAGMYLLFYTLFGSLPLLMG---IIYIF  
 Drosophila nuom 72 MLLASEMINKHN--NYKNLFLLNII IILLLLLILTFSMSL FMFYLF FESSLIPTLFLILGWGY----QPERLQAGLYLLFYTLVLSLPLMIG---IFYVM  
 Homo nuom 76 TIMASQRHLSSEPLSRKKLYLSMLISLQISLIMTFTATELIMFYIFFETTLIPTLAIITRWGN----QPERLNAGTYFLFYTLVGSPLLLIA---LIYTH  
 Synechocystis nuom 96 ITTLATMAAWPVTLKPK-LFYFLMLLMYGGQIAVFAVQDILLFFLVWELELVPVYLILSIWG----GKKRLYAATKFILYTAGGSLFILLAGLTLAFYG  
 Escherichia nuom 97 LGVLAVLC SWKEIEKYQGFHLLNLMWILGGVIGVFLAIDMFLFFFWFEMMLVPMYFLIALWGHKASDGKTRITAAATKFFIYTQASGLVMLIAILALVVFH



Synechocystis\_hemJ 20 -----LFYLVRLFV--YHAEADQEP--AKTILKKOYELMEKRLYNIIT-----  
 Gloeobacter hemJ 20 -----LFYLVRLFI--YHVEANEQPEA--ARAILKKOYEIMEKRLLNIIIT-----  
 Flavobacterium hemJ 22 -----LFYVPRLFM--YOIEASOKPEPD--RSILGDQALMTKRLWKIITWPSMILASLFAFMMH  
 Rhodobacter hemJ 37 -----LFYLPRLFV--YHAEVVKGTGTETDTLF-----QTMERLLRAIMNPAMIATWIFGL-ALV  
 Bombyx nuom 163 NDLNTMTIYFLKFFNMNMY-----LLYISMIFAFVLKMPMYFVHLWLPKAHVEAPVSGSMILAGIMLKLGGYGLLRVMI-FLOE-VNLKLNWIWI-IIS  
 Drosophila nuom 163 NKIGSMNFYLMNMFNFYD-----LLYFCLLCAFLVKMPMFLVHLWLPKAHVEAPVSGSMILAGIMLKLGGYGLLRVIS-FLQL-MNLKYSFVWI-SIS  
 Homo nuom 169 NTLGSLNILLTLTAQELSNSWANNLMWLAYTMAFMVKMPLYGLHLWLPKAHVEAPIAGSMVLA AVLLKLGGYGMRLTL-ILNP-LTKHMAYPFL-VLS  
 Synechocystis nuom 189 ----DVNTFDMSAIAAKDIPVNLQLLLYAGFLIAYGVKLPFPLHTWLPDAHGEATAPAHMLLAGILLKMGYALLRMNVMGLPD-AHAVFAPVLV-ILG  
 Escherichia nuom 197 YNATGVWTFNYEELLNTPMSSGVEYLLMLGFFIAFAVKMPVPLHGWLPDHSAQAPTAGSDVLAGILLKTAAYGLLRFSLPLFPN-ASAEFAPIAM-WLG



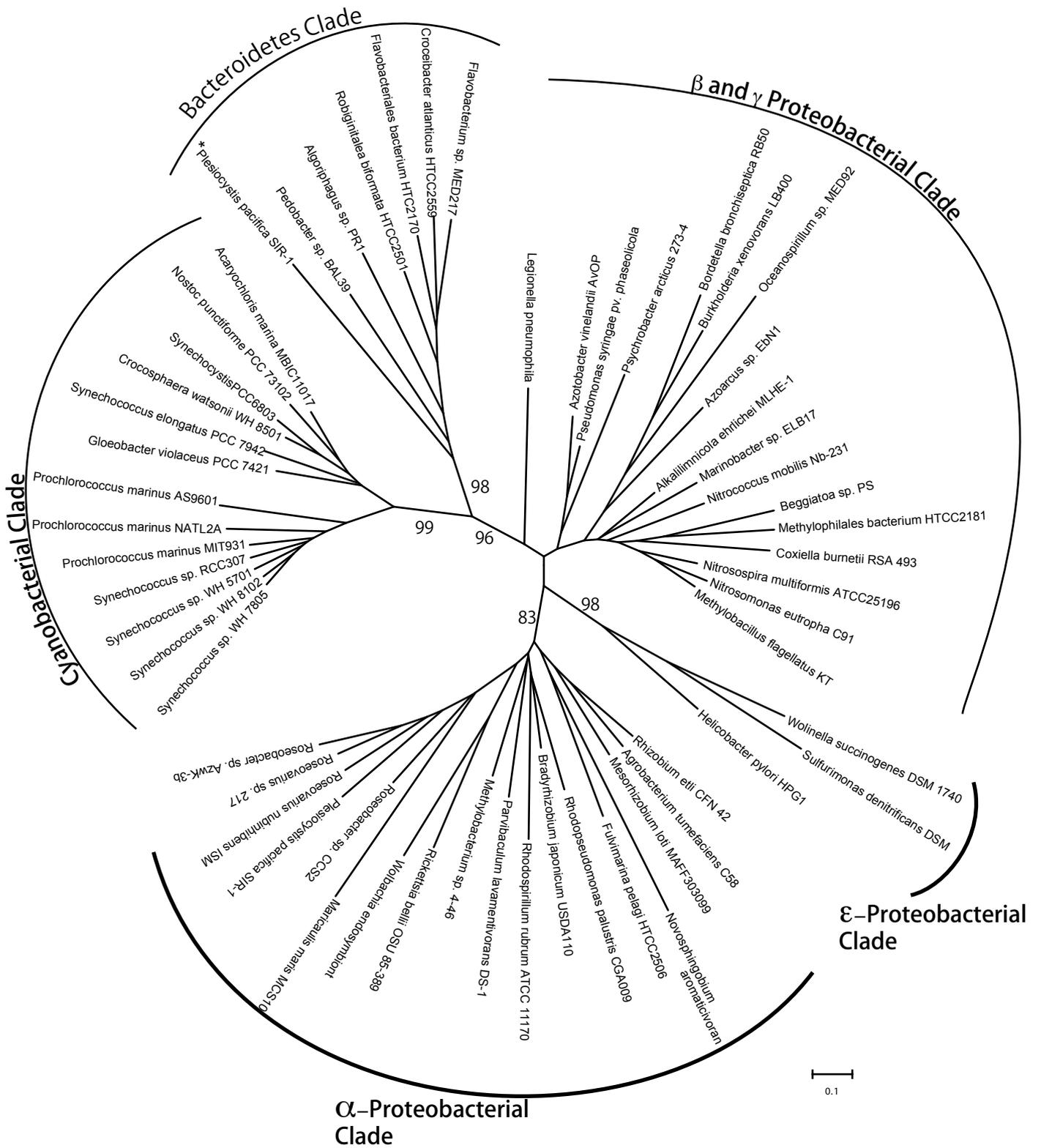
Synechocystis\_hemJ 61 -TPGMVVTVAMAIGLIFTEPEILKSGWLHIKLT FVALLLL--YHFYCGRVMKKLAQGESQWSGOQFRALN-----  
 Gloeobacter hemJ 61 -TPGMVLTVAMAVGMLVVQPDWLKAGWLHIKLT-LVLLM-GYHFYCMRLRTQLAAGTCRWGPKQLRALN-----  
 Flavobacterium hemJ 80 IVPGLLNAPWMIKLG FVLLYAYHFKNH-----QIFROLOS GNFKYTTKFMRIWN-----  
 Rhodobacter hemJ 90 FTPGIVDWSMLWPWT-KAASVIAM-TAFHMWLAARRKDF AAGTNSRAGR TYRMMN-----  
 Bombyx nuom 254 LVGGLFISMKFCFCQV-DMKSLIAYSSVAHMSM VIGGIMVMN-YWGFIGSYILMISHGLCSSGMFCLANIN YERLHSRSLYINSGMMNFMPMSLWFFLL  
 Drosophila nuom 254 LVGGVLVSLVCLRQT-DLKAL IAYSSVAHMGIVLSGLLTMT-YWGLCGSYTLMIAHGLCSSGLFCLANVSYERLGSRSMLINKGLLNFMPSMTLWFFLLS  
 Homo nuom 266 LWGMIMTSSICLRQT-DLKSLIAYSSISHMALVVTAILIQT-PWSFTGAVILMIAHGLTSSLLFCLANSNYERTHSRIMILSQGLQTLPLMAFWLLAS  
 Synechocystis nuom 284 VVNI IYAAFTSFAQR-NLKRKIAYSSISHM GFLIGLASFT-DLGM SGAMLOMISHGLIGASLFFMVGATYDRHTLMLDEMGGIGQMKKGFAMWTACS  
 Escherichia nuom 295 VIGIFYGAWMAFAQT-DIKRLIAYTSVSHM GFLIAIYTG-SQLAYQGA VIOMIAHGLSAAGLFI LCGQLYERIHTRDMRMMGGLWSKMKWLPALSLFFA

Synechocystis\_hemJ 128 -----EAPTILLVVIVLLAVFKNNLPLDATTWLIVALVIAMAASIQLYAKKRRRDQALLTEQQAASAQN-----  
 Gloeobacter hemJ 128 -----EAPTILLVTIVLLAVFKNDLPTDATAWIVFGLVISFAVTIQLYARKRRLDKEKQLASQGGQ-----  
 Flavobacterium hemJ 130 -----EGPTLILFAVIFLVITKSATNWIWGLAGLICLAVLLMLGIKLYKSIREKNPNA-----  
 Rhodobacter hemJ 142 -----ELPTLLMLVIVFSAVAKWNFWDF-----  
 Bombyx nuom 352 SSNMAAPPSLNLLGEISLINS-LVSWSNISMILLMMISFFSAGYSLYLYSYIQHG---KYFQGLYSFYVGS SREYLLLFLHWFPLNMMILKVDFISIW F-  
 Drosophila nuom 352 SANMAAPPTLNLLGEISLLNS-IVSWSWISMILLSFLSFFSAAYTLYLYSFSQHG---KLFSGVYSFSSGKIREYLLMLLHWLPLNLLILKSESFMLWL-  
 Homo nuom 364 LTNLALPPTINLLGELSVLVT-TFSWSNITLLLTGLNMLVTALYSLHMFTTTQWG---SLTHHINMKPSFTRENTLMFMHLSPI LLLSLNPDIITGFSS  
 Synechocystis nuom 382 LASLALPGMSGFVAELMV FVG FATS DAYNLVFR TIVVVL MGVGVILTP IYLLSMLREMLYGPENEELVNHTNLVDVEPREVFIIGCLLVPIIGIGFYPKL  
 Escherichia nuom 393 VATLGMPGTGNFVGEFMILFG-SFQVVPVITVISTFGLV FASVYSLAMLHRA YFG-KAKSQIASQELPGMSLRDVFMI LLLVLLVLLGFYPPQILDTSH

Synechocystis sp. PCC6803 1 -----MAYVWFKAFHILIGV VVWFAGLFYLVRLFYVHAEADQEPPEAKTILKKQYELMEKRLVNIITTPGMVVTVMAMAIGLIFTEPEILKSGWLEHKL  
 Crocosphaera watsonii WH8501 1 -----MAYVWFKSFHLIGV VVWFAGLFYLVRLFYVHAEASEKKEPEQSILKAQYELMEKRLVNIITTPGMVVTVMAMAIGLVSTPEEVLKSGWLEHKL  
 Acaryochloris marina MBIC11017 1 -----MAYVWFKAFHILIGV VVWFAGLFYLVRLFYVHVEANEPEEPAPAILKQOYELMEKRLVNIITTPGMVVTVMAMAIGLIVTQEDLLKQGWLVKLL  
 Nostoc punctiforme PCC73102 1 -----MAYVWFKAFHIVGIV VVWFAGLFYLVRLFYVHVEANLEPEPAQTILKNOYQIMEKRLVNIITTPGMVVTVMAMAIGLITTEPDVLEKSGWLEHKL  
 Synechococcus elongatus PCC7942 1 MRSLEADRVAAYWFKAFHIVGIV VVWFAGLFYLVRLFYVHVEANDKPEPARTILQEQYELMEKRLVNIITTPGMVVTVMAMAIGLIVTQEDVLEKSGWLEHKL  
 Gloeobacter violaceus PCC 7421 1 -----MAYLWFKAFHIVGIV VVWFAGLFYLVRLFYVHVEANEQPEAARAILKKQYELMEKRLVNIITTPGMVVTVMAMAIGLVVQPDVLEKSGWLEHKL  
 Synechococcus sp. WH7805 1 MTL---PPEAYLWFKTHHIVGIV VVWFAGLFYLVRLFYVHVEAEMDPEL S LAFKSOYELMEKRLVNIITTPGMVVTVMAMAIGLIVVQPTVLEKSGWLEHKL  
 Synechococcus sp. WH5701 1 MTLPALPPEAYLWFKTHHIVGIV VVWFAGLFYLVRLFYVHVEAEALEPLRGAFOOYELMEKRLVNIITTPGMVVTVMAMAIGLVVQPTVLEKSGWLEHKL  
 Synechococcus sp. RCC307 1 MAL-SLPPEAYLWFKTHHIVGIV VVWFAGLFYLVRLFYVHVEAEQEQPIRGAFOOYELMEKRLVNIITTPGMVVTVMAMAIGLVVQPTVLEKSGWLEHKL  
 Synechococcus sp. WH8102 1 MTL---PPEAYLWFKTHHIVGIV VVWFAGLFYLVRLFYVHVEAEALEPLRGAFOOYELMEKRLVNIITTPGMVVTVMAMAIGLVVQPTVLEKSGWLEHKL  
 Prochlorococcus marinus NATL2A 1 ---MNFSPETYLWFKSFHIIIGV VVWFAGLFYLVRLFYVHVEVQTKQBEIREFVENKQYELMEKRLVNIITTPGMVVTVMAMAIGLVVQPTVLEKSGWLEHKL  
 Prochlorococcus marinus MIT9313 1 ----MAEAFLWFKSLHIIIGV VVWFAGLFYLVRLFYVHVEAEALEPLRGAFOOYELMEKRLVNIITTPGMVVTVMAMAIGLVVQPTVLEKSGWLEHKL  
 Prochlorococcus marinus AS9601 1 ----MAEAFLWFKSLHIIIGV VVWFAGLFYLVRLFYVHVEAEALEPLRGAFOOYELMEKRLVNIITTPGMVVTVMAMAIGLVVQPTVLEKSGWLEHKL

Synechocystis sp. PCC6803 93 TFVALLLLYHFFYCGRV M KKLAESESOWSGOQFRALNEAP TLLLVIVLLAVFKNNLP L D A T F W L I V A L V I A M A A S I Q L Y A K K R R R D Q A L L T E Q O K A A S A Q  
 Crocosphaera watsonii WH8501 93 TFVALLLLYHFFYCKKI M KQLASGECOWTGOQFRALNEAP TLLLVIVLLAVFKNNLP L D L T F W L I V A L V I A M A A S I Q L Y A K K R R -----L A Q E K L S S S I  
 Acaryochloris marina MBIC11017 93 AFVALLLLYHFFYCRRL M KKLLEEGTCGWSGOQFRALNEAP TLLLVIVLLAVFKNNLP L D T T V W V I F A M V I A M V V T I Q L Y A K K R K -----R D K E R L A E L T  
 Nostoc punctiforme PCC73102 93 LFVAILLICYHHYCARL M KKLAI GECGWSGOQFRALNEAP TLLLVIVLLAVFKNNLP L D I A A W A I F A M I L M A V T I Q L Y A K K R R -----Q D K E K L T A Q I  
 Synechococcus elongatus PCC7942 101 AFVAVLLICYHHYCARL M KKLQACTCOWTGOQFRALNEAP TLLLVIVLLAVFKNNLP L D T T V W V I F A M V I A M V V T I Q L Y A K K R R -----L D R E R Q E A A A  
 Gloeobacter violaceus PCC7421 93 TIVVLLMAYHFFYCMRL R T Q L A A C T C R W G P K Q L R A L N E A P T L L L V I V L L A V F K N D L P T D A T A W I V F G L V I S E A V T I Q L Y A K R R R -----L D K E K Q L A S Q  
 Synechococcus sp. WH7805 98 AFVALLLYHWFYCYRL M G Q L Q A C H C O W S G K Q L R A L N E P T L L L V I V V M L V V F K S O F P T S A A T W F I V G L V V F M A A S I Q E Y A R W R R -----L K A E S I A A E S  
 Synechococcus sp. WH5701 93 AVVLALLAYHWFYCYRL M G Q L R E G S C R W S G R Q L R A L N E P T L L V L V V M L V V F K G O F P T G A A T W F L V A L V V A M A C S I Q E Y A R W R R -----L R V E Q E A L A G  
 Synechococcus sp. RCC307 92 LFVALLLYHWFYCYRL M G Q L Q R D C N W S G R Q L R A L N E P T L L L V I V V L L V V F K O F P T G A A T W L T V G L V V F M A A S I Q E Y A R W R R -----L R A E R E A A H A  
 Synechococcus sp. WH8102 98 GFVAGLLAYHWFYCYRL M G Q L Q A C T C R T S G K Q L R A L N E P T L L L V I V V M L V V F K S O F P T G A A T W F I V A L V V F M A A S I Q E Y A R W R R -----L R A E A Q A V T G  
 Prochlorococcus marinus NATL2A 98 LFVLELLLYHFFYCYRL M N Q L T N N Q F N E S G O Q L R A L N E P T L L L V V V M L V V F K N O F P T S A A S W L I F G L I L E M A A S I Q E Y A K W R R -----N K K Q L T ---  
 Prochlorococcus marinus MIT9313 96 SFVLGLVIYHYSYCYKI M Y S L Q N C T S T T S A K N L R L L N E P T L L L F I I V L L V I F K N N F P T S I A T W S V V G L I I F M L A S I Q L Y A K I R K K -----N E N S L S N E -  
 Prochlorococcus marinus AS9601 96 SFVLGLVIYHYSYCYKI M Y S L H N C T S S I S A K N L R L L N E P T L L L F I I V L L V I F K N N F P T S V A T W S V V G L I I F M L A S I Q L Y A K I R K K -----N E N S L S N E -

Synechocystis sp. PCC6803 193 N-----  
 Crocosphaera watsonii WH8501 187 D-----  
 Acaryochloris marina MBIC11017 187 ASE-----  
 Nostoc punctiforme PCC73102 187 GQPQEQS--  
 Synechococcus elongatus PCC7942 195 SLEAAPTQS  
 Gloeobacter violaceus PCC7421 187 GGQQ----  
 Synechococcus sp. WH7805 192 GHAS-----  
 Synechococcus sp. WH5701 195 GA-----  
 Synechococcus sp. RCC307 194 G-----  
 Synechococcus sp. WH8102 192 S-----  
 Prochlorococcus marinus NATL2A -----  
 Prochlorococcus marinus MIT9313 -----  
 Prochlorococcus marinus AS9601 -----



**Table S1. Distribution of *slr1790* (*HemJ*), *HemY* and *HemG* homologues among cyanobacteria**

Cyanobacteria	Type of Protox		
	<i>slr1790</i> ( <i>HemJ</i> )	<i>HemY</i>	<i>HemG</i>
<i>Acaryochloris marina</i> MBIC11017	+	-	-
<i>Anabaena variabilis</i> ATCC 29413	+	-	-
<i>Gloeobacter violaceus</i> PCC 7421	+	+	-
<i>Microcystis aeruginosa</i> NIES-843	+	-	-
<i>Nostoc</i> sp. PCC 7120	+	-	-
<i>Prochlorococcus marinus</i> str. AS9601	+	-	-
<i>Prochlorococcus marinus</i> str. MIT 9215	-	-	+
<i>Prochlorococcus marinus</i> str. MIT 9515	-	-	+
<i>Synechococcus elongatus</i> PCC 6301	+	-	-
<i>Synechococcus</i> sp. JA-2-3B'a(2-13)	-	+	-
<i>Synechococcus</i> sp. JA-3-3Ab	-	+	-
<i>Synechocystis</i> sp. PCC 6803	+	-	-
<i>Thermosynechococcus elongatus</i> BP-1	-	+	-
<i>Trichodesmium erythraeum</i> IMS101	-	+	-

Table S2 Distribution of slr1790, HemG and HemY homologues within Bacteroidetes/Chlorobi.

Organism	Group	slr1790(hemJ)	hemG	hemY
<i>Bacteroides fragilis</i> NCTC 9343	Bacteroidetes	-	-	+
<i>Bacteroides fragilis</i> YCH46	Bacteroidetes	-	-	+
<i>Bacteroides thetaiotaomicron</i> VPI-5482	Bacteroidetes	-	-	-
<i>Bacteroides vulgatus</i> ATCC 8482	Bacteroidetes	-	-	-
<i>Candidatus Sulcia muelleri</i> GWSS	Bacteroidetes	-	-	-
<i>Chlorobium chlorochromatii</i> CaD3	Chlorobi	-	-	-
<i>Chlorobium phaeobacteroides</i> DSM 266	Chlorobi	-	+	+
<i>Chlorobium tepidum</i> TLS	Chlorobi	-	-	-
<i>Cytophaga hutchinsonii</i> ATCC 33406	Bacteroidetes	-	-	+
<i>Flavobacterium johnsoniae</i> UW101	Bacteroidetes	+	-	-
<i>Flavobacterium psychrophilum</i> JIP02/86	Bacteroidetes	+	-	-
<i>Gramella forsetii</i> KT0803	Bacteroidetes	+	-	-
<i>Parabacteroides distasonis</i> ATCC 8503	Bacteroidetes	-	-	+
<i>Pelodictyon luteolum</i> DSM 273	Chlorobi	-	-	-
<i>Porphyromonas gingivalis</i> W83	Bacteroidetes	-	-	+
<i>Prosthecochloris vibrioformis</i> DSM 265	Chlorobi	-	-	-
<i>Salinibacter ruber</i> DSM 13855	Bacteroidetes	-	-	+