PNAS
www.pnas.org

## Supplementary Information for

**Structural and functional insights into the unique CBS-CP12 fusion protein family in cyanobacteria**

Claudia Hackenberg, Johanna Hakanpää, Fei Cai, Svetlana Antonyuk, Caroline Eigner, Sven Meissner, Mikko Laitaoja, Janne Jänis, Cheryl A. Kerfeld, Elke Dittmann, Victor S. Lamzin

Corresponding authors: Claudia Hackenberg, Victor Lamzin
Email: hackenberg@embl-hamburg.de, victor@embl-hamburg.de

**This PDF file includes:**

## SI Materials and Methods

### DNA manipulation and generation of protein expression strains

Total DNA from *Microcystis aeruginosa* PCC 7806 (hereafter *Microcystis*) was isolated and purified according to ref. 1. All other DNA techniques, such as plasmid isolation, transformation of *Escherichia coli*, ligations and restriction analysis (enzymes were obtained from Thermo Fisher Scientific or New England Biolabs), were standard methods. The coding sequences of selected genes (CP12-CBS-N/C: IPF_2164, encoded on contig C319, NCBI id: AM778949, hereafter CBS-CP12, PRK: IPF_5236) were amplified by PCR using Taq polymerase (Thermo Fisher Scientific), DNA of *Microcystis* as template and gene-specific primers with added *Nco*I and *Not*I cleavages sites (CBS-CP12-fw ccatggTGCTGCAAGCTCAAGAAATTATG, CBS-CP12-rev gcggccgcTTAATCGTCATAAATGCGGCA, CP12del-fw ccatggCCAAGCGGTTGTTCATTGAAGAT, CBSdel-rev gcggccgcTTAGGGCTTTTCGACAAAATC, PRK-fw ccatggTGGCCAATAAGCCAGAGCGCGTG, PRK-rev gcggccgcTTAGACTGACGCTGCCACTTT). For the cloning of CP12del a fragment of 244 nt of the C-terminal part of CBS-CP12, and for CBSdel a fragment of 406 nt of the N-terminal part of CBS-CP12 were generated. Genes encoding TrxA (IPF_2161) and GAPDH (IPF_4508) with added cleavage sites were purchased (Eurofins Genomics and GenScript). The verified coding sequences were transferred into expression vector pETM-22 (CBS-CP12), pETM-20 (PRK), pETM-30 (CP12del), pET-MBP (GAPDH) or pETM-11 (CBS-CP12, CBSdel, TrxA) using *Nco*I/*Not*I.

### Protein expression and purification

All proteins were expressed as fusion proteins with affinity and solubility N-terminal tags containing a protease cleavage site between the target protein and tag, leading to the addition of 1-4 amino acid residues (Table S1). GAPDH-MBP was expressed in *E. coli* C41 (DE3) (Lucigen) (2) and TB-medium supplemented with 100 µg/ml kanamycin. Cells were grown at 37˚C until the optical density at 750 nm (OD750) reached 0.6. Protein expression was induced by addition of 0.5 mM IPTG (isopropylthio-b-galactoside) and performed at 30˚C for 24 hrs. All other proteins were expressed in *E. coli* strain LOBSTR (Kerafast) (3) for 6-8 hrs at 30˚C following up to 65 hrs at 20˚C and 140 rpm in 500 ml auto-induction medium (4), containing 10 g/l trypton/peptone, 5 g/l yeast extract, 25 mM $Na_2HPO_4$, 25 mM $KH_2PO_4$, 50 mM $NH_4Cl$, 5 mM $Na_2SO_4$, 2 mM $MgSO_4$, 0.05 % (w/v) glucose, 0.2 % (w/v) lactose, 0.5 % (w/v) glycerol, 50 mM $FeCl_2$, 20 mM $CaCl_2$, 10 mM $MnCl_2$, 10 mM $ZnCl_2$, 2 mM $CoCl_2$, 2 mM $CuCl_2$, 2 mM $NiCl_2$, 2 mM $NaMoO_4$, 2 mM $Na_2SeO_3$ and 2 mM $H_3BO_3$ supplemented with 100 µg/ml kanamycin or ampicillin. Cells were centrifuged at 12,227 x g, 15 min, at room temperature and stored at -80˚C. The pellet was thawed and resuspended in buffer 1 (100 mM Bicine/KOH pH 7.8, 40 mM KCl, 10 mM beta-mercaptoethanol) supplemented with 1 mM EDTA (Ethylenediaminetetraacetic acid), broken by sonication (2x 3 min, 60%, 3 sec on, 7 sec off, on ice) and centrifuged at 43,667 x g, 20 min, 4˚C.

The supernatant was filtrated through 0.45 µm membrane and subsequently purified by metal ion affinity chromatography with cOmplete™ His-tag purification resin (Roche) equilibrated with buffer 1. Contaminants were eluted 2x with 10 mL buffer 1 containing 5 mM and 10 mM imidazole,

respectively. Target protein was eluted in several steps with 10 mL buffer 1 containing 200 mM imidazole. Eluates were combined and concentrated by ultra-filtration with Amicon-Ultra centrifugal filter units. The fusion proteins were digested with either 3C (Human Rhinovirus 3C) or TEV (Tobacco Etch Virus) protease in a ratio of 1:25 protease to protein to cleave off the tag (Table S1) during dialysis in buffer 1 (approximately 400x the volume of the protein sample) for approximately 16-18 hrs at 4˚C. The resulting protein mixture was applied to cOmplete™ His-tag purification resin equilibrated with buffer 1, to separate the un-tagged protein from the His-tagged protease and the His-tag and then washed once with 10 mL buffer 1 containing 10 mM imidazole. Both flow-through fractions, containing un-tagged target protein, were collected, combined and concentrated by ultra-filtration. The protein was further purified by size-exclusion chromatography using AKTA Explorer or Pure FPLC systems, and HiLoad 16/60 Superdex 200 pg, HiLoad 16/60 Superdex 75 pg or Superdex 200 Increase 10/300 GL columns (GE Healthcare) equilibrated with buffer 1. Target protein containing fractions were pooled, concentrated by ultra-filtration, supplemented with 20 % (v/v) glycerol, flash frozen in liquid $N_2$ and stored at -80˚C. All purification steps were performed on ice or at 4˚C. All protein fractions were checked for purity using SDS-PAGE and staining by Instant Blue (Expedeon). Protein concentrations were determined using NanoDrop ND-1000 (Thermo Fisher Scientific) and the respective extinction coefficient of the target proteins. For all further analyses, protein aliquots were thawed and either dialyzed for 16-18 hrs at 4˚C or subjected to buffer exchange using Econo-Pac 10DG columns (Bio-Rad) in the respective buffer.

**Crystallization**

CBS-CP12 was concentrated by ultra-filtration to 6.1 mg/ml in buffer 2 (50 mM Bicine/KOH pH 7.8, 40 mM KCl) for hexagonal crystals; and to 7.3 mg/ml in buffer 2 supplemented with 10 mM TCEP (tris(2-carboxyethyl)phosphine) for derivative and native orthorhombic crystals. Proteins were incubated for 16-18 hrs at 4˚C in the respective buffer prior to crystallization. Derivative and native orthorhombic crystals grew from 0.3 M sodium acetate, 0.1 M HEPES/NaOH (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid), pH 8.0, and 20% and 25% PEG2000MME (Polyethylene glycol), respectively. Hexagonal crystals grew from 0.15 M KSCN, 0.1 M HEPES/NaOH, pH 7.0 and 18% PEG3350 containing 0.1 mM guanidine hydrochloride and protein incubated with 0.132 mM AMP (Adenosine monophosphate) for about 10-30 min prior crystallization. Crystallization method was hanging drop vapor diffusion in 24-well plates at 19˚C. For derivatization crystals were soaked for 30 min in a solution containing 25% PEG2000MME, 0.2 M sodium acetate, 0.1 M HEPES/NaOH, pH 8.0 and 0.1 M Yb-$HPDO_3A$ (Ytterbium ligated to 10-(2-hydroxyprolyl)-1,4,7,10-tetraazacyclododecan-1,4,7-triacetic acid) and flash-frozen in $N_2$. Native orthorhombic data and derivative data after soaking were collected without additional cryoprotectant. For native hexagonal data the crystals were cryoprotected by soaking in crystallization solution with 25% ethylene glycol.

**Data collection, structure determination and model refinement**

Data were collected at 100 K at P13 beamline (5) (EMBL-Hamburg/DESY). 1800 images were collected with an oscillation range of 0.1° for native data sets and 0.2° for derivative data set, using 40 ms exposure per frame. Wavelength was 0.976 Å (12.7 keV), 0.918 Å (13.5 Å) and 1.385 Å (8.95 keV) for native orthorhombic, native hexagonal and derivative data sets, respectively.

Data were processed with XDS (6) and scaled with Aimless (7). The derivative structure was solved as single wavelength anomalous diffraction (SAD) using the Crank2 pipeline (8) resulting in a model with about 180 residues built in three chains of the asymmetric unit (chain A residues 2-182, chain B residues 3-179, chain C residues 1-179). All models were refined with Refmac (9) in the CCP4-suite (10) iterated with solvent adjustment as implemented in ARP/wARP (11). The maps were inspected with Coot (12). The derivative structure could be directly refined into native orthorhombic data (Table S2). Due to structural flexibility, there are differences between the CP12 domains within the asymmetric unit of the native model. In one out of three subunits, the entire visible part (chain A residues 1-178) is well ordered and the majority of the side chains can be unambiguously fitted to the density. Electron density is less refined for the remaining two subunits, but shows the presence of the two helices (chain B residues 4-178, chain C residues 1-176). The model was truncated to regions visible in the electron density (chain break between residues 161 and 162 in chain B and 161-165 in chain C), refined and manually adjusted using the same protocol as for the derivative model. The hexagonal model (1 chain in the asymmetric unit) was solved with molecular replacement using chain A of the native structure as a search model, resulting in a model with residues 3-177 (Table S2).

**Isothermal Titration Calorimetry (ITC)**

CBS-CP12 and CP12del were dialyzed in buffer 2 for 16-18 hrs at 4˚C, CBSdel in 50 mM Bicine/KOH pH 7.8, 200 mM KCl for 16-18 hrs at 21˚C. AMP, ADP, ATP and cAMP were dissolved in the respective dialysis buffer. Calorimetric measurements were carried out using a MicroCal VP-ITC. The cell contained the protein, the syringe the ligand and the reference cell distilled water. Titrations were performed with 20 µM CBS-CP12 and 400 µM AMP, 100 µM CBSdel and 700 µM AMP, 100 µM CP12del and 800 µM AMP, 50 µM CBS-CP12 and 600 µM ADP, 20 µM CBS-CP12 and 500 µM ATP, and 20 µM CBS-CP12 and 400 µM cAMP, respectively. Each experiment was performed at a constant temperature of 25°C and consisted of 45 injections of 6.3 µL aliquots, repeated every 240 s. The interaction of CBS-CP12 with ATP was also performed at 25°C and consisted of 28 injections of 10 µL aliquots, repeated every 220 s. The heat of dilution, measured by control experiments in which the ligand was injected into a buffer-filled cell, was subtracted. Signals recorded in each experiment were integrated using OriginPro® 7 software supplied with the instrument. The standard thermodynamic binding parameters, variations of enthalpy ($\Delta H^0$), Gibbs free energy ($\Delta G^0$), entropy ($\Delta S^0$), dissociation constant ($K_D$) and the number of binding sites ($N$), were obtained by nonlinear regression of the integrated heat plots, according to the 'two sets of sites' or 'one set of sites' model of the OriginPro® 7 software. Values reported represent thermodynamic parameters as obtained by the OriginPro® 7 software for three individual

measurements per protein (Table S3).

**Analytical size exclusion chromatography**

Superdex 200 Increase 10/30 GL, Superdex 75 10/30 GL columns (GE Healthcare), AKTA Explorer or PURE FPLC systems were used. Columns were equilibrated in buffer 2 and calibrated with a gel filtration standard (Bio-Rad). All experiments were performed in buffer 2. The oligomeric state of CBS-CP12 with AMP was tested by incubating 50 µM CBS-CP12 (1.16 mg/ml) and 100 µM AMP for 3 hrs at 4˚C. Formation of ternary complex was tested according to ref. 18, by incubating 30 µM GAPDH (1.1 mg/ml), 30 µM PRK (1.14 mg/ml), 25 mM DTTox (trans-4,5-Dihydroxy-1,2-dithiane) and 0.5 mM $NAD^+$ (β-Nicotinamide adenine dinucleotide) with 30 µM CP12del (0.26 mg/ml) or 30 µM CBS-CP12 (0.7 mg/ml), with and without 0.5 mM AMP for 2 hrs at 21˚C. Interaction of CBS-CP12 and TrxA was tested under native conditions with 50 µM CBS-CP12 (1.16 mg/ml) and 100 µM TrxA (1.23 mg/ml) incubated for 2 hrs at 21˚C. Reduced CBS-CP12 was obtained by incubation with 1 mM DTT (1,4-dithiothreitol) for 3 hrs at 4˚C and the column was equilibrated with buffer 2 containing 1 mM DTT. Interaction of CP12del and CBSdel was tested under native conditions by incubating 75 µM CP12del (0.74 mg/ml) and 50 µM CBSdel (0.67 mg/ml) for 16-18 hrs at 21˚C. Flow rates of 0.9, 1 or 0.25 (complex formation) ml/min were applied. Proteins, ligands, protein-protein and protein-ligand mixtures were loaded individually. Elution profiles were detected by UV absorbance at 280 nm or 254 nm. Absorption patterns were normalized and superimposed. Estimated molecular masses are reported as mean ± SD of 2 up to 5 technical replicates.

**Native mass spectrometry (MS)**

The protein samples were desalted/buffer exchanged to 100 mM ammonium acetate ($NH_4OAc$) buffer, pH 8.0 (pH adjusted using ammonia solution), using a PD Miditrap G-25 (Sephadex G-25M) column (GE Healthcare, Sweden). Protein concentrations of the eluted fractions were determined by UV absorbance at 280 nm using the sequence-derived extinction coefficient. All reagents and solvents were HPLC quality. Mass spectra were measured using a 12-T Bruker solariX XR Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer (Bruker Daltonics, Bremen, Germany), equipped with an Apollo-II electrospray ion source and a dynamically harmonized Paracell ICR cell. The instrument was calibrated externally using perfluoroheptanoic acid (PFHA) clusters (Sigma-Aldrich) and operated using ftmsControl 2.2. The mass spectra were further analyzed and processed using Bruker DataAnalysis 4.4 software. The protein solution at 4.7 µM hexamer concentration with and without 200 µM AMP was directly infused into the ion source at a flow rate of 2 µL/min. Drying and nebulizing gas were set at 4 L/min and 1 bar, respectively. The ion source temperature was kept at 200°C. For native MS measurements, the instrument parameters were optimized for the detection of high-mass ions. The time-of-flight was set to 3.0 ms, and the RF frequencies were set as follows: octopole: 5 MHz, collision cell: 1.4 MHz and transfer optics: 2 MHz. To aid desolvation, the collision cell voltage was set to -5.0 V, and the skimmer1 value was varied between 30-70 V. The ion accumulation time in the collision cell was

set to 0.5 s. For each spectrum, a total of 512 time-domain transients (1 MWord) were summed and automatically processed to a final 2 MWord magnitude mode data. The mass range was set at *m/z* 387-8000.

**Quantitative RT-PCR**

*Microcystis* cultures were grown in BG-11 medium (13) at room temperature and white fluorescent light (Philips Master TL-D 865). For investigation of diurnal gene expression, three cultures were incubated for six weeks before the experiment at 15 µmol photons $m^{-2}$ $s^{-1}$, 60 rpm, 25˚C and a 12 h light-dark regime, maintaining OD750 between 0.4 and 0.8. Sampling was conducted over a 24 h period at time points: 06:30 (30 minutes before light-onset), 08:30, 11:30, 13:00, 16:00, 18:30 (30 minutes before dark), 20:30 and 01:00. High-light stress was induced by transferring two cultures with OD750 of 0.4 from 15 µmol photons $m^{-2}$ $s^{-1}$ (control) to 300 µmol photons $m^{-2}$ $s^{-1}$ for 3 hrs. Iron limitation was induced by cultivating two cultures in iron-free BG-11 medium for 15 days at 15 µmol photons $m^{-2}$ $s^{-1}$. Final OD750 was 1.1.

Samples of 25-30 mL were collected and centrifuged at 4°C, 4.000 x g for 3 min. Total RNA was extracted after hot TRIzol® (Thermo Fisher Scientific) and chloroform treatment using the RNeasy kit (Qiagen) according to the manufacturers protocol. RNA concentration was determined spectrophotometrically (NanoDrop 2000, Thermo Fisher Scientific). Purified RNA was free of DNA as tested by PCR. First strand cDNA synthesis of 1 µg (stress experiments) and 2 µg (diurnal expression) RNA was done with Maxima® reverse transcriptase (Thermo Fisher Scientific) and random hexamer primers according to the manufacturers protocol. Real-Time PCR was performed with SensiMix™ SYBR® Low-ROX Kit (Bioline) and LightCycler 480 (Roche) using gene-specific primer (RT-*cbs-cp12*-fw TTTGGCCACGATCCTAAAAC, RT-*cbs-cp12*-rev GAGCAAATAACCGGGCTACA, RT-*cp12*-fw GCGGTGGAAGAATTACAAGC, RT-*cp12*-rev TCGTCATAAACGCGACACTC, RT-*trxA*-fw CCAAACGCTTTGGGATTAGA, RT-*trxA*-rev CGGCTTCCGTAAATTCTTCA, RT-*rnpB*-fw GGGGTAAGGGTGCAAAGGT, RT-*rnpB*-rev AGACCAACCTTTGTCCCTCC). *RnpB* was used as reference. Data was converted (LC480conversion) and processed in LinRegPCR (14). Values reported are mean ± SD of 3 technical replicates (Table S4).

**Phosphoribulokinase assay**

PRK activity was determined at 30˚C and 340 nm essentially as described in ref. 15. All reactions were performed in 100 µL of 50 mM Bicine/KOH pH 7.8, 40 mM KCl, 5 mM DTT, 0.5 mM EDTA, 1 mM $MgCl_2$, 1 mM NADH (reduced β-Nicotinamide adenine dinucleotide), 2 mM ATP, 2 mM ribose-5-phosphate, 2 mM phospho*enol*pyruvate, 1U phosphoriboisomerase, 3U lactate dehydrogenase and 2U pyruvate kinase (all enzymes from Sigma-Aldrich). Protein concentrations were 300 nM PRK, 600 nM TrxA, 150, 300 and 600 nM CBS-CP12, 300 nM CP12del and 300 nM CBSdel. AMP concentrations were 1, 2 and 4 mM. PRK was pre-incubated for 15 min at 30˚C with DTT, TrxA, CBS-CP12, CP12del, CBSdel and AMP, respectively, to ensure maximum activity (15). A Varioskan Flash micro-plate reader (Thermo Scientific) was used. Values reported are mean ±

SD of 5-15 technical replicates.

**Pull-down assays**

His-tagged and un-tagged proteins were expressed and purified as described. *Microcystis* cells were grown to high cell density according to ref. 16. Cells were centrifuged 5 min at 6.000 x g, washed with 10 mM Bicine/KOH pH 7.6, 1 mM EDTA, 1 mM DTT, 1 mM PMSF (phenylmethylsulfonyl fluoride), and lysed in a cooled cell disruptor (Constant Systems) at 25 kPSI. The extract was centrifuged at 21.000 x g, 4°C for 30 min. Proteins from the supernatant were precipitated with 50 % sat. ammonium sulfate at 4°C.

*Microcystis* protein extract and His-CBS-CP12 were dialyzed in buffer 2 for 16-18 hrs, 4˚C prior to the experiment. Ten mg His-CBS-CP12 (bait control, bait+prey), 31 mg (prey control) and 63 mg *Microcystis* protein extract (bait+prey) and 1 ml equilibrated cOmplete™ His-tag purification resin each were used. Washing and elution were done with 1 ml of 5 and 200 mM imidazole, respectively, in buffer 2. Bait control and bait+prey: Resin was incubated with His-CBS-CP12 for 1.5 hrs, 4˚C under rotation and washed 1-5x. Bait control: His-CBS-CP12 was immediately eluted 4x. Bait+prey and prey control: *Microcystis* protein extract was applied on resin containing His-CBS-CP12 or on fresh resin and incubated for 16-18 hrs, 4˚C under rotation. Resins were washed 4x and protein was eluted 4-5x. Interaction of CBS-CP12 and TrxA was tested by incubating His-CBS-CP12 with un-tagged TrxA, and un-tagged CBS-CP12 with His-TrxA. Bait and prey control were done in buffer 2. Interaction was studied in buffer 2 with 25 mM DTTox (oxidizing) or 25 mM DTT (reducing), with and without 1 mM AMP. cOmplete™ His-tag purification resin, 250 or 500 µL, were equilibrated and proteins were washed with the respective buffer. Proteins were used at final concentrations of 50 µM each. Protein-mixes were incubated at 21˚C, 1,200 rpm for 1 hr without and another 30 min with the resin, washed and then eluted with 250 or 500 µL of the respective buffer containing 5, 10, 15 and 20 mM imidazole (W1-W4, respectively), and 200 mM imidazole (E).

**Bioinformatics**

All CBS-CP12 genes were retrieved from Integrated Microbial Genomes (https://img.jgi.doe.gov) by using the co-presence of the two CBS domain (pfam00571) and the CP12 domain (pfam02672) as a search requirement. There were a total of 232 CBS-CP12 genes identified in 333 cyanobacterial genomes, displayed in a phylogenetic tree adapted from ref. 17 (Fig. S1). Letters referring to the phylogenetic subclades are according to ref. 18. Deduced CBS-CP12 protein sequences were subject to multiple sequence alignment performed by using T-Coffee at EMBL-EBI server (19, 20). Based on the alignment, each CBS-CP12 sequence is assigned to one of five variants: 46 CBS-CP12-N/C (CP12 domain has both cysteine pairs), 40 CBS-CP12-0 (no cysteine pairs in the CP12 domain), 123 CBS-CP12-N (possessing the first cysteine pair in the CP12 domain), 14 CBS-CP12-N-tr ("tr" for "truncation"; protein sequences end after the AWD_VEEL motif and therefore only have the first cysteine pair) and 9 CBS-CP12-tr ("tr" for "truncation"; protein sequences end before the first cysteine pair of the CP12 domain). From 34 genomes, a total of 36

CBS-CP12 genes are co-present with TrxA and/or ferredoxin genes in the same cluster. In all but two cases, there is a gene containing two Bateman modules (4 CBS domains) in the gene neighborhood. Deduced protein sequences of these 36 genes are used to generate a Hidden-Markov-Model (HMM). To avoid bias caused by over-representation of closely related species, redundant sequences were removed by running in-house BLASTCLUST (parameters: the sequence length to be covered was 95% and the score identity threshold was 90%). After running BLASTCLUST, 10 unique sequences are aligned again. The resulting alignment file was fed into HMMBUILD at the Mobyle Portal server (21) with enone turned off. Finally, the HMM is visualized by using HMM logo-Mat at the Skylign server (22).

**Figure preparation**

All images showing molecular models were generated using UCSF Chimera (23).

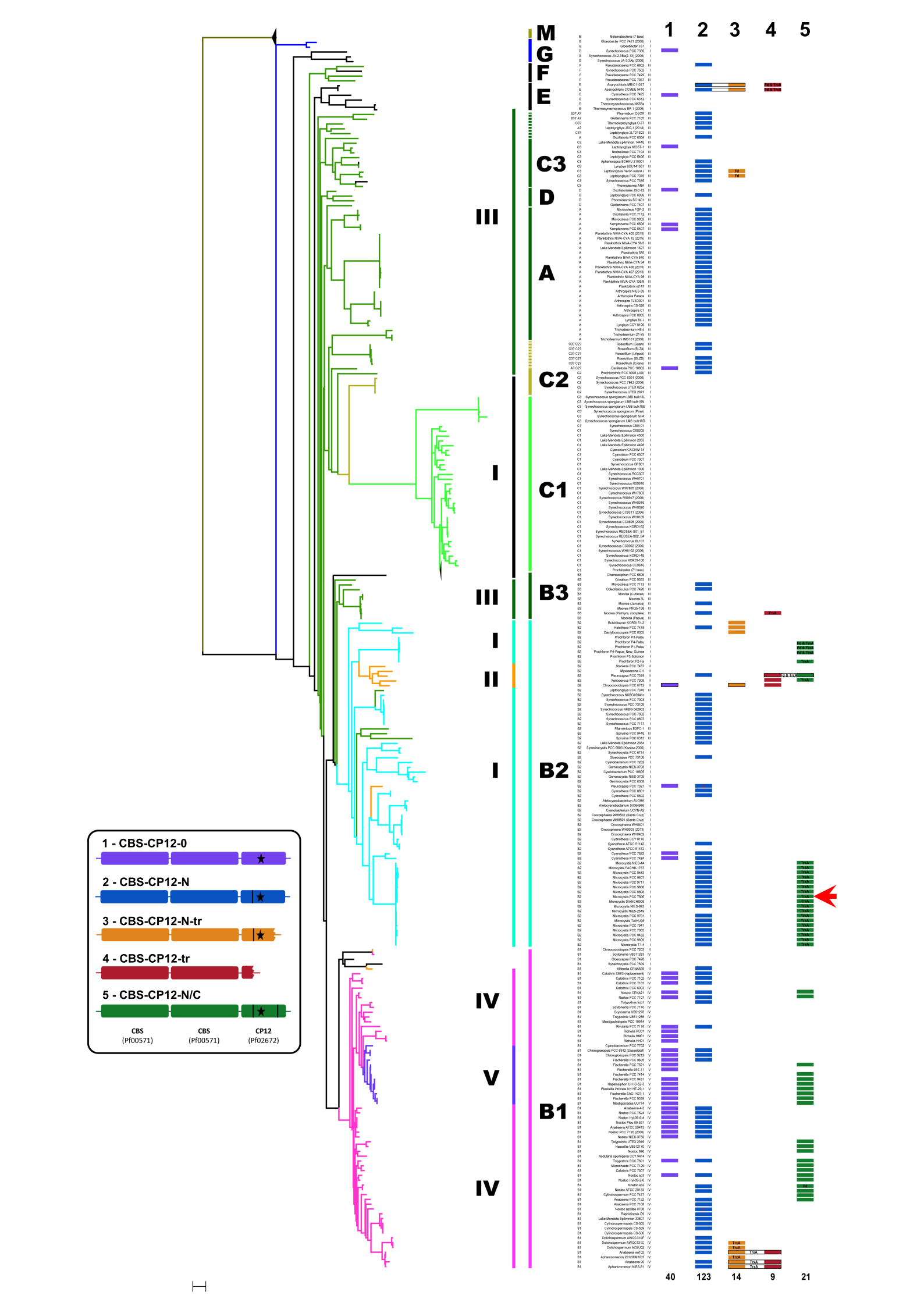
1 - CBS-CP12-0
2 - CBS-CP12-N
3 - CBS-CP12-N-tr
4 - CBS-CP12-tr
5 - CBS-CP12-N/C
CBS (Pf00571)
CBS (Pf00571)
CP12 (Pf02672)
M
G
F
E
C3
D
III
A
C2
I
C1
III
B3
I
II
I
B2
IV
V
B1
IV
1
2
3
4
5
40
123
14
9
21

**Fig. S1. Distribution of CBS-CP12 variants and CBS-CP12/TrxA gene cluster types in the cyanobacterial phylum.** CBS-CP12 variants found in 333 cyanobacterial genomes displayed phylogenetically. The tree was adapted from ref. 17. Domain compositions of the five identified CBS-CP12 variants are shown in symbol legends: each CBS-CP12 contains two CBS domains (PF00571) and one CP12 domain (PF02672). Black vertical bars represent the first and second cysteine pair of the CP12 domain; a black star represents the AWD_VEEL core sequence of the CP12 domain. 1 - full-length CBS-CP12-0 lacking both cysteine pairs in the CP12 domain; 2 - full-length CBS-CP12-N containing the N-terminal cysteine pair in the CP12 domain; 3 - shorter CBS-CP12 variant CBS-CP12-N-tr with truncated C-terminal CP12 domain containing the N-terminal cysteine pair; 4 - CBS-CP12-tr with a very short CP12 domain; 5 - full-length CBS-CP12 variant CBS-CP12-N/C (as the presented CBS-CP12 of *M. aeruginosa* PCC 7806) containing both cysteine pairs in the CP12 domain. Numbers at the top of the tree refer to the respective CBS-CP12 variant; numbers at the bottom of the tree represent the number of genomes in which the respective CBS-CP12 variants is present. A red arrow points to the CBS-CP12 of *M. aeruginosa* PCC 7806 presented in this manuscript. Roman numerals refer to the morphological section: I - unicellular species; II – unicellular species, multiple fission; III – filamentous cyanobacteria without cell differentiation; IV – multicellular and differentiated cyanobacteria, e.g. heterocysts; V – multicellular and differentiated cyanobacteria capable of branching. Letters refer to the phylogenetic subclades according to ref. 17. Each colored bar in the phylogenetic tree represents one copy of a certain gene, unless marked: tow gene copies of CBS-CP12-N are identified in *Nostoc azollae* 0708 and *Pleurocapsa* sp. PCC 7327. In some cases, two CBS-CP12 variants are adjacent to each other in the genome; then a black box is added to include both variants. The presence of a thioredoxin (TrxA) and/or ferrodoxin (Fd) gene in the gene neighborhood of a CBS-CP12 gene is marked on the colored bar. CBS-CP12-0 and CBS-CP12-N have not been observed to be co-present with neither TrxA nor Fd genes. CBS-CP12/TrxA gene cluster types are presented in Fig. 1C. CBS-CP12-N-tr, CBS-CP12-tr and CBS-CP12-N/C mostly occur in addition to CBS-CP12-0 or CBS-CP12-N, except in *Dactylococcopsis* PCC 8305, *Rubidibacter* KORDI 51-2, *Tolypothrix* UTEX 2349, *Hassallia* VB512170, *Nostoc* 996, *Microchaete* PCC 7126, *Calothrix* PCC 7507, *Nostoc* Hyl-09-2-6, *Cylindrospermu*m PCC 7417 and *Aphanizomenon* 2012/KM1/D3. Whenever a TrxA and/or a ferredoxin gene is co-present in the same gene cluster with CBS-CP12-N-tr, CBS-CP12-tr or CBS-CP12-N/C, a gene containing four CBS domains (or two Bateman modules) is immediately upstream but on the opposite strand (see Fig. 1B), with only two exceptions: *Prochloron didemni* P2-Fiji and *Nostoc* sp. sp2.

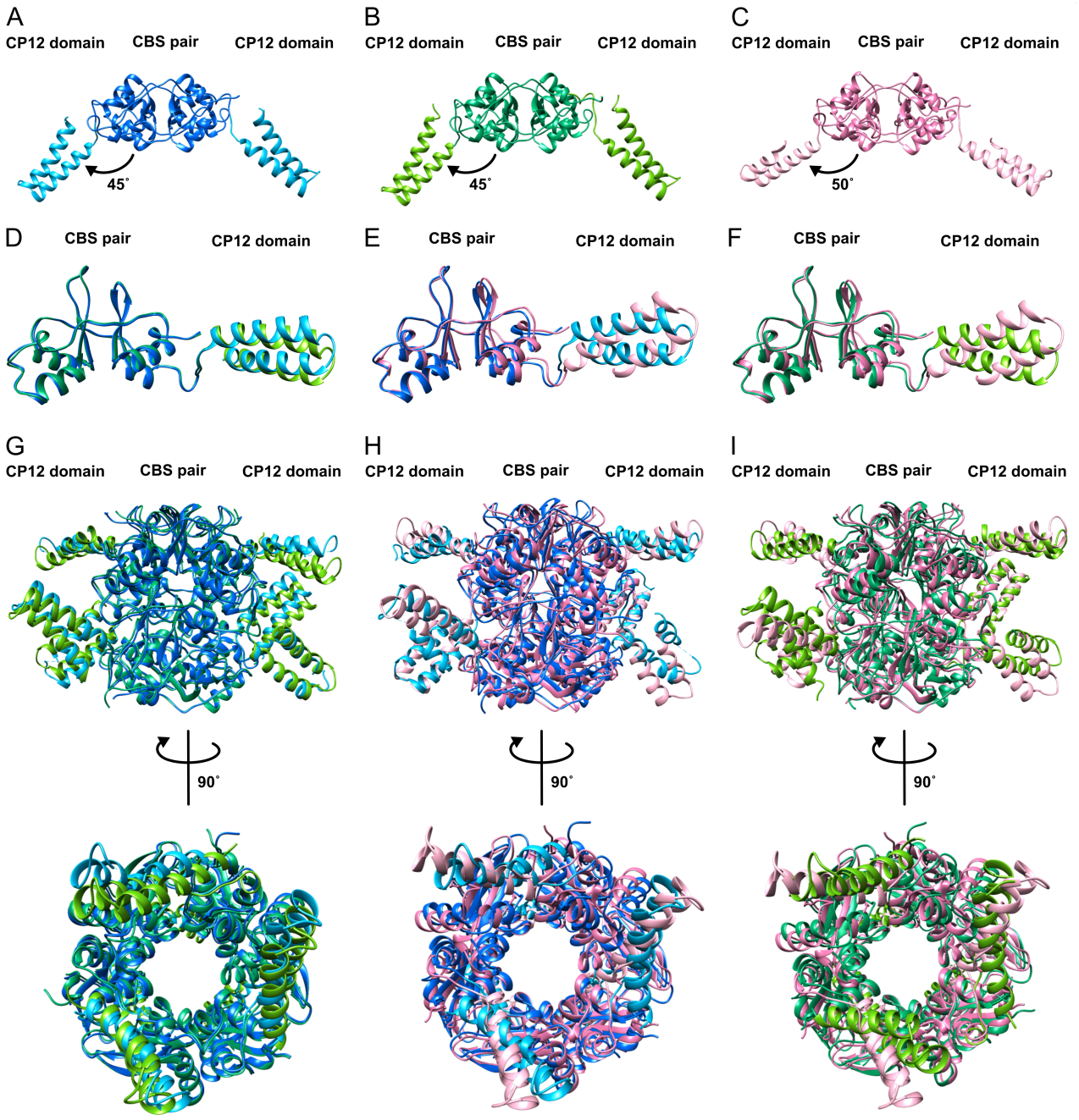


**Fig. S2. Comparison of the native orthorhombic, derivative orthorhombic and native hexagonal structure of CBS-CP12.** (A) Native orthorhombic structure. (B) Derivative orthorhombic structure. (C) Native hexagonal structure. Superimposition of one CBS-CP12 subunit of the (D) native orthorhombic (blue) and derivative orthorhombic (green) structure, (E) native orthorhombic (blue) and native hexagonal (pink) structure, and (F) native hexagonal (pink) and derivative orthorhombic (green) structure. Superimposition of the CBS-CP12 hexamer of the (G) native orthorhombic (blue) and derivative orthorhombic (green) structure, (H) native orthorhombic (blue) and native hexagonal (pink) structure, and (I) native hexagonal (pink) structure and derivative orthorhombic (green) structure, in two different orientations related by a 90° rotation around a vertical axis.

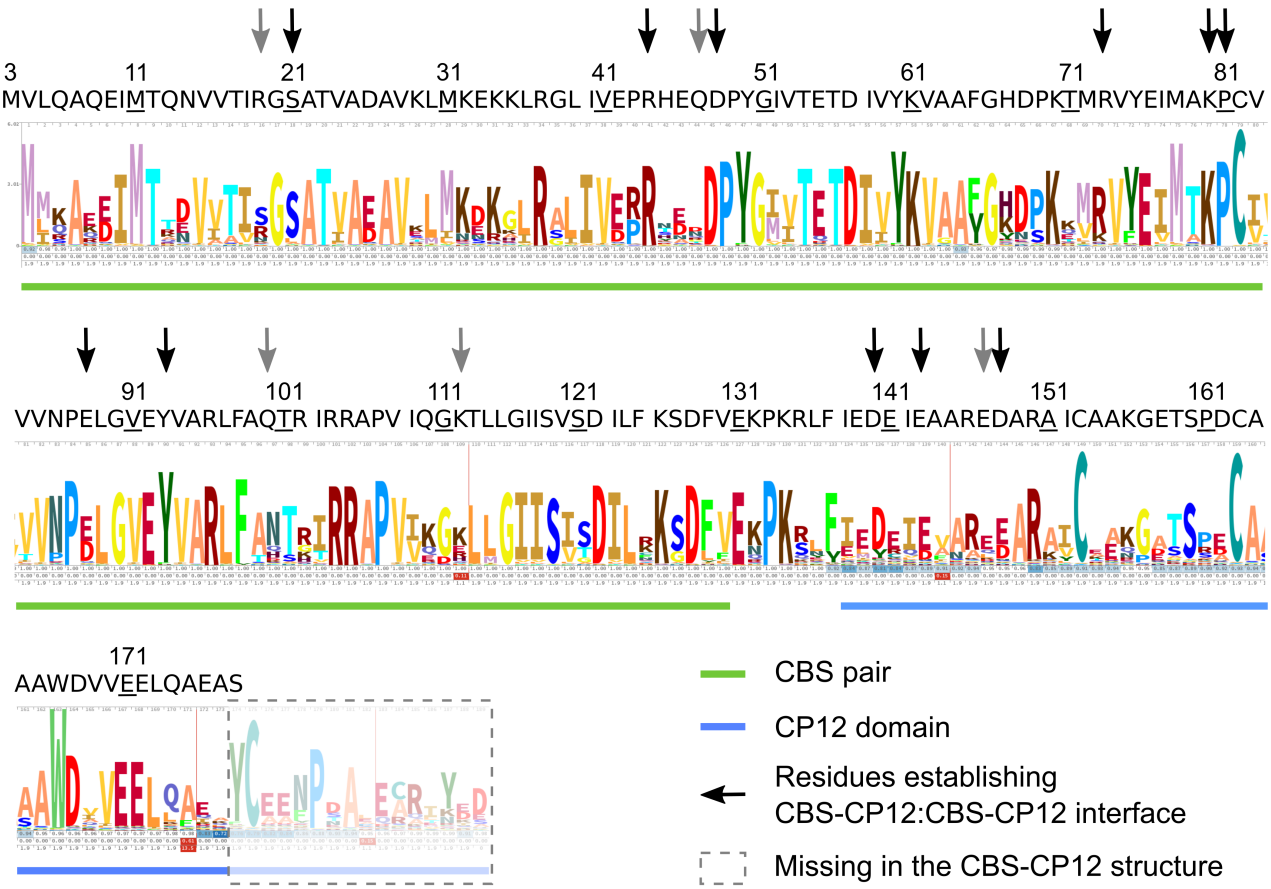


**Fig. S3. Conservation of residues establishing the CBS-CP12:CBS-CP12 interface.** Of the 33 genomes containing a gene cluster with CBS-CP12 and TrxA (Fig. 1C), ten non-redundant sequences, with 95% length coverage and 90% identity cut-off, were used to build the Hidden-Markov-Model. Shown is the amino acid sequence of the heterologously expressed CBS-CP12 protein of *M. aeruginosa* PCC 7806 (IPF_2164). Arrows highlight the residues establishing the CBS-CP12:CBS-CP12 interface (Figs. 2H-J). Black arrows represent conserved residues, gray arrows not conserved residues.

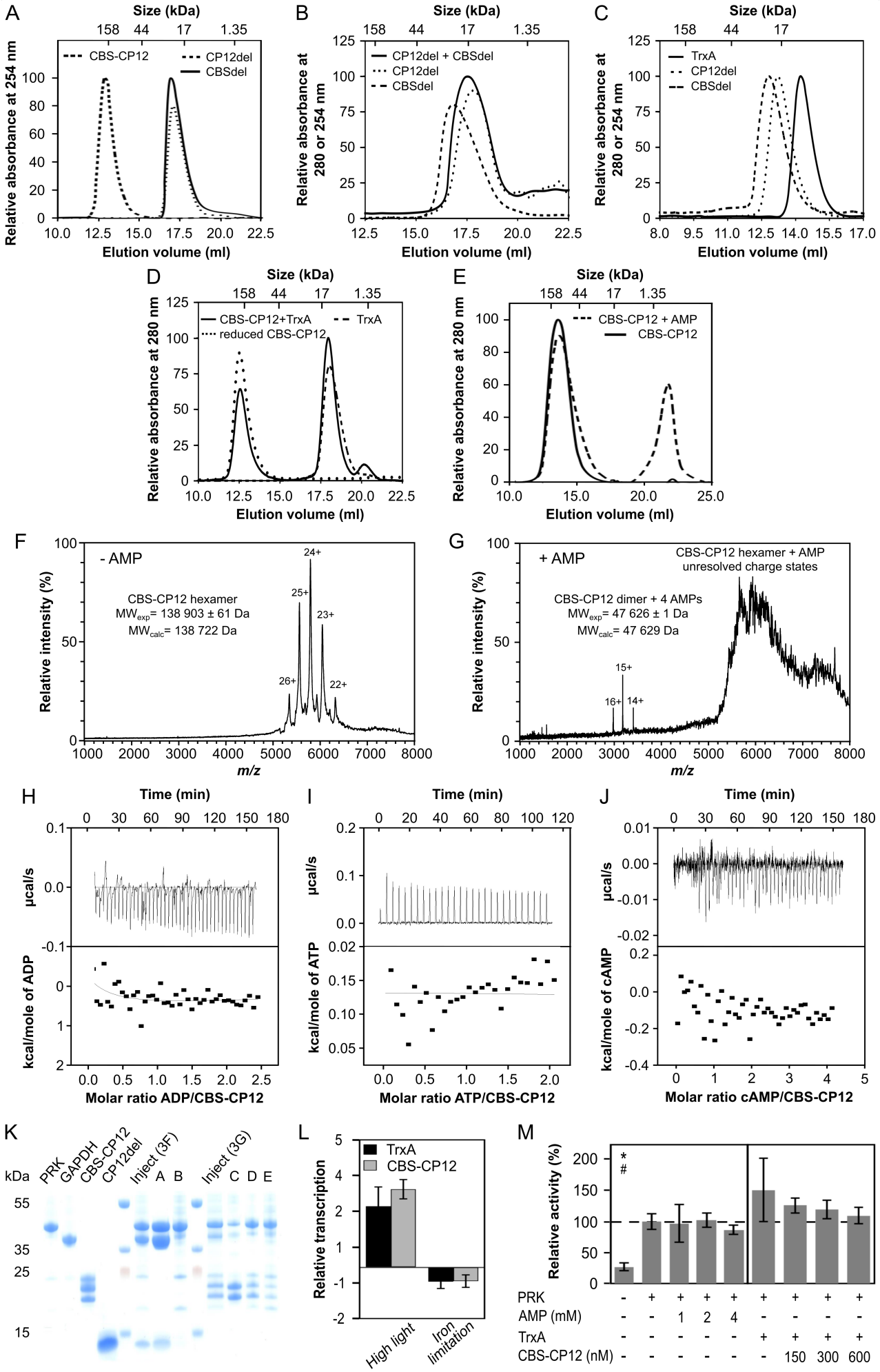
A
Size (kDa)
158 44 17 1.35
CBS-CP12
CP12del
CBSdel
Relative absorbance at 254 nm
Elution volume (ml)
B
Size (kDa)
158 44 17 1.35
CP12del + CBSdel
CP12del
CBSdel
Relative absorbance at 280 or 254 nm
Elution volume (ml)
C
Size (kDa)
158 44 17
TrxA
CP12del
CBSdel
Relative absorbance at 280 or 254 nm
Elution volume (ml)
D
Size (kDa)
158 44 17 1.35
CBS-CP12+TrxA
TrxA
reduced CBS-CP12
Relative absorbance at 280 nm
Elution volume (ml)
E
Size (kDa)
158 44 17 1.35
CBS-CP12 + AMP
CBS-CP12
Relative absorbance at 280 nm
Elution volume (ml)
F
- AMP
CBS-CP12 hexamer
MWexp= 138 903 ± 61 Da
MWcalc= 138 722 Da
24+
25+
23+
26+
22+
Relative intensity (%)
m/z
G
+ AMP
CBS-CP12 hexamer + AMP
unresolved charge states
CBS-CP12 dimer + 4 AMPs
MWexp= 47 626 ± 1 Da
MWcalc= 47 629 Da
15+
16+
14+
Relative intensity (%)
m/z
H
Time (min)
µcal/s
kcal/mole of ADP
Molar ratio ADP/CBS-CP12
I
Time (min)
µcal/s
kcal/mole of ATP
Molar ratio ATP/CBS-CP12
J
Time (min)
µcal/s
kcal/mole of cAMP
Molar ratio cAMP/CBS-CP12
K
kDa
PRK
GAPDH
CBS-CP12
CP12del
Inject (3F)
A B
Inject (3G)
C D E
55
35
25
15
L
TrxA
CBS-CP12
Relative transcription
High light
Iron limitation
M
Relative activity (%)
PRK
AMP (mM)
TrxA
CBS-CP12 (nM)

**Fig. S4.** (A) Analytical S200 SEC elution profile of CBS-CP12, and its individual domains CP12del and CBSdel. CBS-CP12 elutes as hexamer of 143 ± 8.7 kDa (theoretical hexamer size 139 kDa), CP12del elutes at 11.5 ± 1.3 kDa (theoretical monomer size 8.6 kDa) and CBSdel at 15.5 ± 2.0 kDa (theoretical monomer size 14.7 kDa), (mean ± SD, *n* = 3). (B) Analytical S200 SEC elution profile of CP12del, CBSdel and CP12del+CBSdel, incubated in a molar ratio of 1.5:1 on subunit base (75 µM CP12del and 50 µM CBSdel) under native conditions. CP12del and CBSdel do not form a protein complex (*n* = 1). (C) Analytical S75 SEC elution profile of TrxA, CP12del and CBSdel under native conditions. TrxA elutes at 11 ± 0.2 kDa (theoretical size 12.3 kDa), CP12del elutes at 19.6 ± 0.4 kDa and CBSdel elutes at 22 ± 0.5 kDa, (mean ± SD, *n* ≥ 2). (D) Analytical S200 SEC elution profiles of reduced CBS-CP12 (elutes at 180.7 ± 6.6 kDa), TrxA (elutes at 8.4 ± 0.2 kDa) and a CBS-CP12+TrxA mixture (three peaks eluting at 173.1 ± 15.4 kDa and two peaks smaller 10 kDa). CBS-CP12 still elutes as a hexamer and does not form a complex with TrxA (mean ± SD, *n* = 3). (E) Analytical S200 SEC of 50 µM CBS-CP12 ± 100 µM AMP, eluting as hexamer of 135.9 ± 8.5 kDa (- AMP) and 143 ± 8.7 kDa (+ AMP, second peak represents AMP) (mean ± SD, n = 3). (F) Native MS of 4.7 µM native CBS-CP12, appearing at m/z ~5300-6300 (charge states 22+ to 26+) with an average mass of 138 903 ± 61 Da, consistent with a hexameric assembly. The experimentally observed mass is ~183 Da higher than calculated from the six monomers (138 720 Da), mostly likely due to incomplete ion desolvation. Large protein complexes often possess some excess mass accompanied with considerable peak broadening, resulting from incomplete protein ion desolvation and/or ion-pairing (e.g. from buffer salts) during droplet evaporation. (G) Native MS of 4.7 µM CBS-CP12 with 200 µM AMP. CBS-CP12 appears at m/z ~ 5300-8000. The unresolved mass spectral peaks are likely due to the excess amount of sodium ions present in the ligand solution, and prevented exact binding stoichiometry to be determined. Approximately 85% of CBS-CP12 appeared as hexamer and 15% as dimer (charge states 14+ to 16+), with a mass indicating a saturated binding of four AMP per CBS-CP12 dimer. (H-J) ITC titration of (H) 600 µM ADP into 50 µM CBS-CP12, (I) 500 µM ATP into 50 µM CBS-CP12, and (J) 400 µM cAMP into 20 µM CBS-CP12. No interaction of CBS-CP12 with either of the three molecules was observed (*n* = 1). (K) Denaturing SDS-PAGE of the individual proteins, the two peaks obtained in Fig. 3F and the three fractions of the single peak obtained without AMP shown in Fig. 3G. (L) Relative transcription of *cbs-cp12* and *trxA* under high-light stress and under iron limiting conditions *in M. aeruginosa* (mean ± SD, *n* = 6). (M) Activity of 300 nM PRK upon addition of different AMP concentrations, and upon addition of 600 nM TrxA and different CBS-CP12 concentrations. (mean ± SD, *n* ≥ 7, $^{*}P \leq 0.01$ compared to 300 nM PRK, $^{\#}P \leq 0.01$ compared to 300 nM PRK + 600 nM TrxA, Student's t-test).

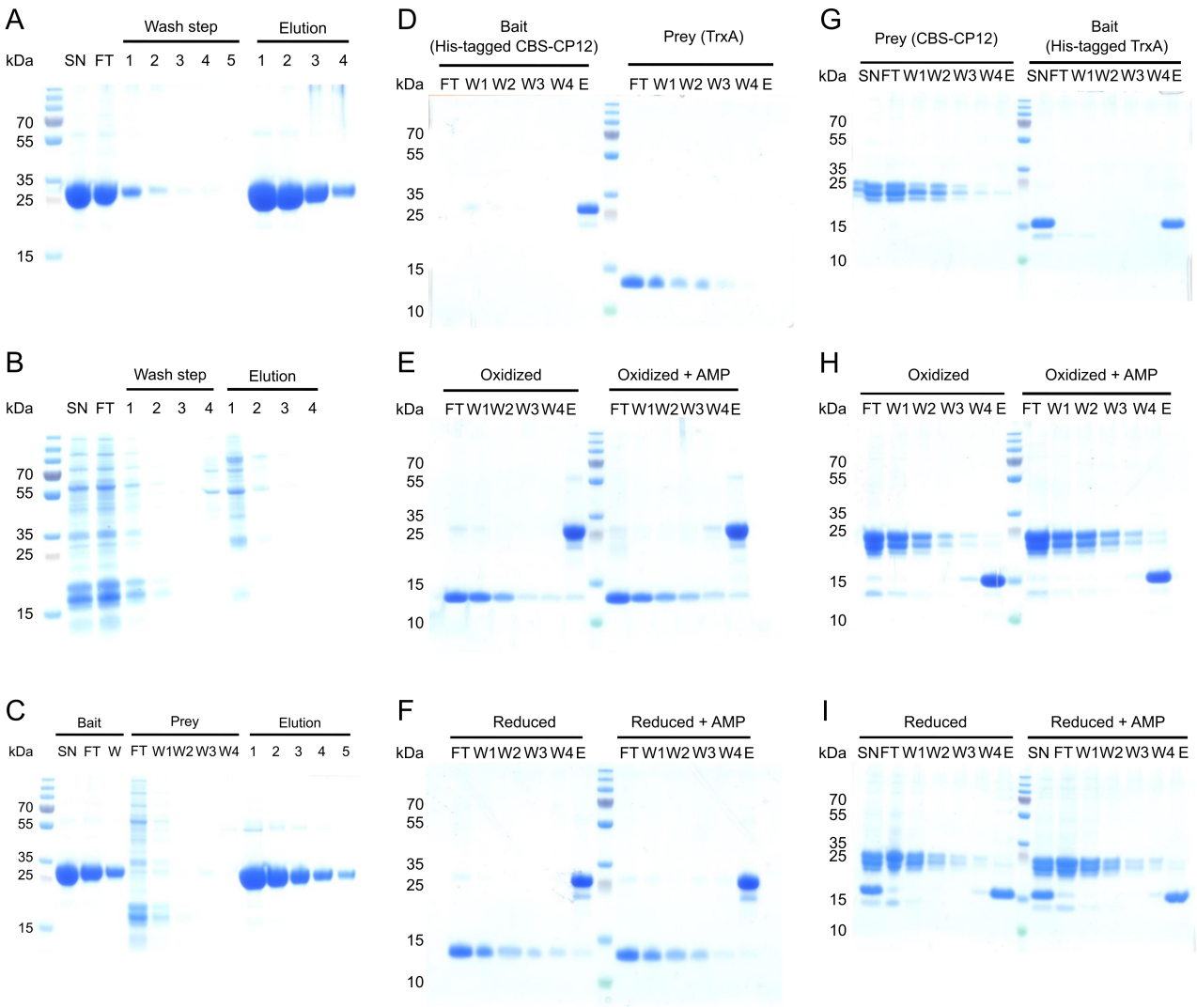


**Fig. S5. Denaturing SDS-PAGE of pull-down assays.** (A-C) Pull-down assay using His-tagged CBS-CP12 (bait) and protein cell extract of *M. aeruginosa* (prey). (A) Bait control using 10 mg His-tagged CBS-CP12. (B) Prey control using 31 mg protein cell extract of *M. aeruginosa.* (C) Pull-down assay using 10 mg His-tagged CBS-CP12 and 63 mg protein cell extract of *M. aeruginosa.* No interaction of His-tagged CBS-CP12 with proteins from *M. aeruginosa* was observed. (D-I) Pull-down assay using (D-F) His-tagged CBS-CP12 (bait) and un-tagged TrxA (prey), and (G-I) His-tagged TrxA (bait) and un-tagged CBS-CP12 (prey). (D) and (G) Respective bait and prey control. (E) and (H) Oxidizing conditions using 25 mM DTTox with and without 1 mM AMP. (F) and (I) Reducing conditions using 25 mM DTT with and without 1 mM AMP. No complex formation of CBS-CP12 and TrxA was observed. SN – supernatant, FT – flow-through, W – wash step, E – elution.

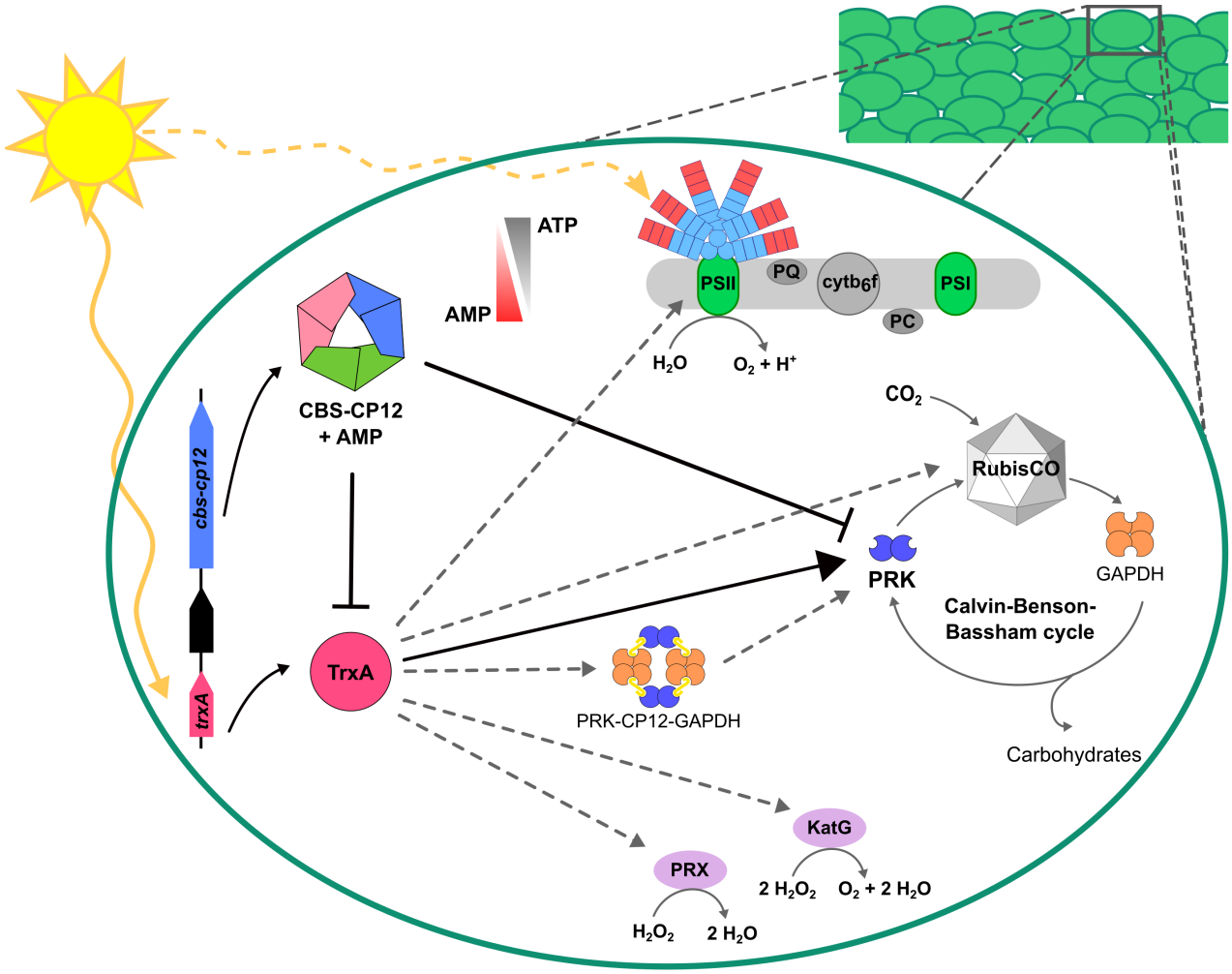


**Fig. S6. Proposed model of the function of CBS-CP12 (IPF_2164) in *M. aeruginosa*.** Our data indicate that CBS-CP12 acts as a redox-dependent, regulatory sensor for intracellular AMP levels and antagonist of TrxA in the regulation of light-dependent processes when the energy levels of the cell are low. The gene cluster containing CBS-CP12 and TrxA is present in scum- and mat-forming, as well as symbiotic cyanobacteria (Fig. 1C), that form dense colonies experiencing pronounced light-gradients. The expression of both proteins is co-regulated and induced by light (Fig. 3H, SI Appendix, Fig. S4L). TrxA then activates photosynthesis and oxidative stress response, amongst others by reducing PRK (Fig. 3I). CBS-CP12 forms functional hexamers (Fig. 2) that bind AMP but not ATP (Fig. 3A, SI Appendix, Fig. S4I), possibly sensing and conveying the cellular energy status through AMP. Contrary to canonical CP12 (24-27) and CP12del, CBS-CP12 does not form a ternary complex with GAPDH or PRK (Fig. 3 D-G), inhibiting both enzymes during the night. Instead, the CBS-CP12/AMP complex inhibits PRK directly (Fig. 3J), but also counteracts the TrxA-dependent stimulation of PRK activity (Fig. 3I), and possibly even of other thioredoxin regulated proteins. Solid lines represent results obtained in this study, dashed lines represent interactions of thioredoxins with various target proteins published elsewhere (26-30). Cytb6f, cytochrome b6f complex; GAPDH; glyceraldehyde-3-phosphate dehydrogenase; KatG, catalase-peroxidase; PC, plastocyanin; PQ, plastoquinone; PRK, phosphoribulokinase; PRX, peroxiredoxin 1-Cys-PRX or Type II PRX; PSI, photosystem I; PSII, photosystem II.

**Table S1. Sizes and protease sites of the investigated proteins**

| Protein | CBS-CP12 amino acid residues | N-terminal tag | Protease site | Size fusion protein (kDa) | Size un-tagged protein (kDa) | Additional amino acids N-terminal (bold) |
|---|---|---|---|---|---|---|
| CBS-CP12 | 1-205 | Trx-His | 3C | 37.3 | 23.12 | **GP**M**V** |
| CP12del | 134-208 | His-GST | TEV | 37.5 | 8.6 | **AMA** |
| CBSdel | 1-133 | His | TEV | 17.79 | 14.74 | **AMV** |
| TrxA | / | His | TEV | 15.33 | 12.28 | **A** |
| Prk | / | Trx-His | TEV | 52.25 | 37.95 | **A**M**V** |
| GAPDH | / | His-MBP | TEV | 79.59 | 36.6 | **A**M**V** |
| CBS-CP12-His | 1-205 | His | / | 26.1 | / | His-tag |

His: Histidine; Trx: Thioredoxin; GST: Glutathione-S-transferase; MBP: Maltose Binding Protein

**Table S2. Data collection, phasing and refinement statistics**

| | **Native (orthorhombic)** | **Native (hexagonal)** | **Derivative (orthorhombic, Yb-ligated)** |
|---|---|---|---|
| **Data collection** | | | |
| Space group | $C222_1$ | $P6_322$ | $C222_1$ |
| Cell dimensions | | | |
| *a*, *b*, *c* (Å) | 81.61,124.02,118.45 | 98.23, 98.23, 100.61 | 79.89,120.73,117.55 |
| $\alpha, \beta, \gamma$ (°) | 90, 90, 90 | 90, 90, 120 | 90, 90, 90 |
| Resolution (Å) | 68.18-2.15 (2.28-2.15) | 100-2.5 (2.6-2.5) | 66.63-2.79 (2.87-2.79) |
| $R_{meas}$ | 0.023 (0.59) | 0.084 (1.35) | 0.055(0.73) |
| $I / \sigma I$ | 14.8 (2.5) | 23.4 (2.6) | 14.54 (1.77) |
| Completeness (%) | 97.6 (96.3) | 99.9 (99.2) | 97.9 (71.7) |
| Redundancy | 3.4 (3.5) | 18.5 (17.6) | 6.5 (2.7) |
| Wilson plot B factor ($Å^2$) | 39.7 | 59.7 | 51.1 |
| **Refinement** | | | |
| Resolution (Å) | 38.6-2.15 (2.21-2.15) | 85.1-2.5 (2.57-2.5) | 10.0-2.79 (2.86-2.80) |
| No. reflections | 30,043 | 9,924 | 13,151 |
| $R_{work}$ / $R_{free}$ (%) | 17.8/23.4 (24.3 /27.8) | 19.9/23.4 (35.3/43.7) | 18.6/21.9 (33.3/38.9) |
| No. atoms | | | |
| Protein | 4001 | 1351 | 4120 |
| Ligand/ion | 4 | / | 73 |
| Water | 191 | 31 | 51 |
| Atomic displacement parameters ($Å^2$) | 53.78 | 74.6 | 53.4 |
| Protein A/B/C | 45.4/55.8/61.0 | 74.75 | 51.7/53.6/55.1 |
| Ligand/ion | | | 60.0 |
| Water | 50.5 | 68.16 | 43.5 |
| R.m.s deviation from target stereochemistry | | | |
| Bond lengths (Å) | 0.015 | 0.014 | 0.013 |
| Bond angles (°) | 1.7 | 1.7 | 1.77 |

*Values in parentheses are for highest-resolution shell.

**Table S3. Thermodynamic parameters of interaction with AMP at 25˚C, as determined by ITC**. Parameters are calculated per protein monomer for each technical replicate. The best fit of the ITC titration of CBS-CP12 corresponds to a two-site model and for CBSdel to a one-site model (Fig. 3 A and C). *N*, stoichiometry of binding; $K_D$, affinity; $\Delta H^0$, standard enthalpy; $T\Delta S^0$, standard entropy; $\Delta G^0$, standard Gibbs free energy. Values reported represent thermodynamic parameters as obtained by the OriginPro® 7 software for three individual measurements (technical replicates) per protein.

| **Protein** | **Binding Site** | **Replicate** | ***N*** | **$K_D$ (μM)** | **$\Delta H^0$ (kcal mol$^{-1}$)** | **$T\Delta S^0$ (kcal mol$^{-1}$)** | **$\Delta G^0$ (kcal mol$^{-1}$)** |
|---|---|---|---|---|---|---|---|
| CBS-CP12 | 1 | 1 | 1.22 | 0.55 | 8.89 | 17.41 | -8.54 |
| | | 2 | 1.04 | 0.24 | 5.08 | 14.10 | -9.02 |
| | | 3 | 1.17 | 0.44 | 6.99 | 15.65 | -8.67 |
| | 2 | 1 | 1.05 | 1.78 | -20.10 | -12.34 | -7.84 |
| | | 2 | 1.33 | 1.68 | -11.68 | -3.79 | -7.88 |
| | | 3 | 1.12 | 1.89 | -16.56 | -8.77 | -7.81 |
| CBSdel | 1 | 1 | 1.01 | 8.26 | -2.83 | 4.11 | -6.93 |
| | | 2 | 1.19 | 11.81 | -2.95 | 3.79 | -6.72 |
| | | 3 | 1.02 | 18.62 | -2.52 | 3.94 | -6.45 |

**Table S4. Quantitative RT-PCR.** Diurnal rhythm of the transcription of *cbs-cp12, trxA* and *cp12* in three individual cultures of *M. aeruginosa* (mean ± SD, *n* = 3). Values of culture 1 are displayed in Fig. 3H.

| Culture | Time of day | *cbs-cp12* | *trxA* | *cp12* |
|---|---|---|---|---|
| 1 | 1:00 | 0.21 ± 0.14 | 0.20 ± 0.04 | 1.16 ± 0.16 |
| | 6:30 | 1.00 ± 0.00 | 1.00 ± 0.00 | 1.00 ± 0.00 |
| | 8:30 | 8.09 ± 3.29 | 8.70 ± 2.35 | 4.01 ± 0.65 |
| | 10:30 | 8.32 ± 4.19 | 7.55 ± 0.94 | 2.79 ± 1.22 |
| | 13:00 | 6.91 ± 1.70 | 5.69 ± 2.36 | 4.47 ± 1.50 |
| | 16:00 | 1.71 ± 0.81 | 1.24 ± 0.32 | 2.79 ± 0.68 |
| | 18:30 | 1.43 ± 0.48 | 0.62 ± 0.11 | 2.85 ± 0.43 |
| | 20:30 | 1.28 ± 0.53 | 0.91 ± 0.50 | 2.16 ± 0.31 |
| 2 | 1:00 | 2.07 ± 1.20 | 3.71 ± 2.94 | 2.15 ± 2.43 |
| | 6:30 | 1.00 ± 0.00 | 1.00 ± 0.00 | 1.00 ± 0.00 |
| | 8:30 | 34.78 ± 18.58 | 30.41 ± 16.14 | 3.91 ± 2.97 |
| | 10:30 | 33.81 ± 24.23 | 56.70 ± 45.04 | 6.05 ± 7.08 |
| | 13:00 | 28.66 ± 12.22 | 33.14 ± 19.94 | 6.27 ± 4.30 |
| | 16:00 | 3.64 ± 2.12 | 3.45 ± 1.48 | 1.47 ± 1.09 |
| | 18:30 | 4.32 ± 1.93 | 5.00 ± 2.51 | 2.53 ± 1.79 |
| | 20:30 | 2.74 ± 1.61 | 2.96 ± 1.56 | 1.92 ± 1.45 |
| 3 | 1:00 | 0.48 ± 0.14 | 0.57 ± 0.44 | 1.11 ± 0.39 |
| | 6:30 | 1.00 ± 0.00 | 1.00 ± 0.00 | 1.00 ± 0.00 |
| | 8:30 | 5.57 ± 2.32 | 6.61 ± 3.65 | 3.89 ± 1.48 |
| | 10:30 | 1.80 ± 0.50 | 2.57 ± 1.50 | 2.02 ± 0.95 |
| | 13:00 | 4.69 ± 1.01 | 6.94 ± 1.32 | 4.44 ± 0.63 |
| | 16:00 | 1.31 ± 0.46 | 1.29 ± 0.49 | 2.90 ± 0.30 |
| | 18:30 | 0.60 ± 0.28 | 0.77 ± 0.42 | 2.64 ± 1.06 |
| | 20:30 | 0.30 ± 0.07 | 0.41 ± 0.05 | 2.67 ± 0.35 |

## SI References


1. Hisbergues M, Christiansen G, Rouhiainen L, Sivonen K, Börner T (2003) PCR-based identification of microcystin-producing genotypes of different cyanobacterial genera. *Arch Microbiol* 180:402–410.
2. Miroux B, Walker JE (1996) Over-production of proteins in *Escherichia coli*: mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels. *J Mol Biol* 260:289–298.
3. Andersen KR, Leksa NC, Schwartz TU (2013) Optimized *E. coli* expression strain LOBSTR eliminates common contaminants from His-tag purification. *Proteins* 81:1857–1861.
4. Studier FW (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expr Purif* 41:207–234.
5. Cianci M, et al. (2017) P13, the EMBL macromolecular crystallography beamline at the low-emittance PETRA III ring for high- and low-energy phasing with variable beam focusing. *J Synchrotron Radiat* 24:323–332.
6. Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr Sect D Biol Crystallogr* 66:133–144.
7. Evans PR, Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr Sect D Biol Crystallogr* 69:1204–1214.
8. Skubák P, Pannu NS (2013) Automatic protein structure solution from weak X-ray data. *Nat Commun* 4:2777.
9. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr Sect D Biol Crystallogr* 53:240–255.
10. Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr Sect D Biol Crystallogr* 67:235–242.
11. Langer GG, Cohen SX, Lamzin VS, Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* 3:1171–1179.
12. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr Sect D Biol Crystallogr* 66:486–501.
13. Rippka R, Deruelles J, Waterbury JB, Herdman M, Stanier RY (1979) Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. *J Gen Microbiol* 111:1–61.
14. Ruijter JM, Ramakers C, Hoogaars WM, Karlen Y, Bakker O, van den Hoff MJ, Moorman AF (2009) Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res* 37: e45.
15. Rault M, Gontero B, Ricard J (1991) Thioredoxin activation of phosphoribulokinase in a chloroplast multi-enzyme complex. *Eur J Biochem* 197:791–797.
16. Pörs Y, Wüstenberg A, Ehwald R (2010) A batch culture method for microalgae and cyanobacteria with $CO_2$ supply through polyethylene membranes. *J Phycol* 46:825–830.
17. Bao H, et al. (2017) Additional families of orange carotenoid proteins in the photoprotective system of cyanobacteria. *Nat Plants* 3:17089.

18. Shih PM, et al. (2013) Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proc Natl Acad Sci USA* 110:1053–1058.
19. Notredame C, Higgins D, Heringa J (2000) T-Coffee: a novel Method for Fast and Accurate Multiple Sequence Alignment. *J Mol Biol* 302:205–217.
20. Li W, et al. (2015) The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res* 43:W580–W584.
21. Néron B, et al. (2009) Mobyle: a new full web bioinformatics framework. *Bioinformatics* 25:3005–3011.
22. Wheeler TJ, Clements J, Finn RD (2014) Skylign: a tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models. *BMC Bioinformatics* 15:7.
23. Pettersen EF, et al. (2004) UCSF Chimera — A visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612.
24. Wedel N, Soll J, Paap BK (1997) CP12 provides a new mode of light regulation of Calvin cycle activity in higher plants. *Proc Natl Acad Sci USA* 94:10479–10484.
25. Wedel N, Soll J (1998) Evolutionary conserved light regulation of Calvin cycle activity by NADPH-mediated reversible phosphoribulokinase/CP12/glyceraldehyde-3-phosphate dehydrogenase complex dissociation. *Proc Natl Acad Sci USA* 95:9699–9704.
26. Marri L, Trost P, Pupillo P, Sparla F (2005) Reconstitution and Properties of the Recombinant Glyceraldehyde-3-Phosphate Dehydrogenase/CP12/Phosphoribulokinase Supramolecular Complex of Arabidopsis. *Plant Physiol* 139:1433–1443.
27. Marri L, et al. (2009) Prompt and Easy Activation by Specific Thioredoxins of Calvin Cycle Enzymes of *Arabidopsis thaliana* Associated in the GAPDH/CP12/PRK Supramolecular Complex. *Mol Plant* 2:259–269.
28. Lindahl M, Florencio FJ (2003) Thioredoxin-linked processes in cyanobacteria are as numerous as in chloroplasts, but targets are different. *Proc Natl Acad Sci USA* 100:16107–16112.
29. Pérez-Pérez ME, Florencio FJ, Lindahl M (2006) Selecting thioredoxins for disulphide proteomics: Target proteomes of three thioredoxins from the cyanobacterium *Synechocystis* sp. PCC 6803. *Proteomics* 6:S186–S195.
30. Pérez-Pérez ME, Mata-Cabana A, Sánchez-Riego AM, Lindahl M, Florencio FJ (2009) A Comprehensive Analysis of the Peroxiredoxin Reduction System in the Cyanobacterium *Synechocystis* sp. Strain PCC 6803 Reveals that All Five Peroxiredoxins are Thioredoxin Dependent. *J Bacteriol* 191:7477–7489.