The Structure of CyanoP at 2.8 Å: Implications for the evolution and function of the PsbP subunit of Photosystem II

Franck Michoux[§],

Kenji Takasaka[§a], Marko Boehm[§], Peter J. Nixon[§] and James W. Murray*[¶]

Divisions of Biology[§] and Molecular Biosciences[¶], Wolfson Biochemistry Building, Imperial College London, S. Kensington Campus, London, SW7 2AZ, UK

a: current address, Division of Bioscience, Graduate School of Natural Science and Technology/Faculty of Science; Okayama University, Okayama 700-8530, Japan

**: To whom correspondence should be addressed. Telephone: +44 (0)20 75948895 Email: j.w.murray@imperial.ac.uk*

Supplementary Material

Materials and Methods

Cloning, expression and purification of CyanoP.

The DNA sequence corresponding to the CyanoP homologue in *Thermosynechococcus elongatus* (*tlr2075*) was cloned into a modified version of pRSET-A (kindly provided by Dr Ernesto Cota, Imperial College London, UK) without its predicted signal peptide and the first cysteine. The corresponding PCR fragment was amplified from genomic DNA using Phusion polymerase (Finnzymes, Finland) and primers CyanoP-XhoI-F (5'-TATATACTCGAGTCGGCCACCAGTGGGTTACAGGCCTATGTCGA-3') and CyanoP-EcoRI-R (3'-TATATAGAATTCTTATTAATATACAGTAAACGAGGACACAA-3'), double digested by XhoI/EcoRI and ligated (Quick Ligation Kit, New England Biolabs, UK) into XhoI/EcoRI linearised pRSET-A. The vector was then transformed into KRX cells (Promega, UK).
Expression of CyanoP in transformed cells was induced with 2 g/L of rhamnose and cells were grown at 18 ºC overnight. Cells were then lysed with a sonicator (Sonics and Materials, CT, USA) in lysis buffer (50 mM Tris-HCl pH 7.9, 500 mM NaCl, 1 mM $MgCl_2$) supplemented with a Complete Protease Inhibitor Cocktail Tablet – EDTA (Roche, UK). The supernatant was mixed with a Ni-IDA resin (Generon, UK). Non-specifically bound proteins were removed by washing 3 times with wash buffer (20 mM Tris-HCl pH 7.9, 500 mM NaCl, 60 mM imidazole) and CyanoP was eluted with elution buffer (20 mM Tris-HCl pH 7.9, 500 mM NaCl, 1M imidazole). Purified CyanoP was further dialysed overnight at 4 ºC in 20 mM Tris-HCl pH 7.9, 500 mM NaCl (Figure S2a). The His-tag present in the construct was removed by thrombin (GE healthcare, UK) digestion at a ratio of 1 unit of thrombin per 100 µg of purified CyanoP. Proteolysis was performed overnight at room temperature and the digested sample was reloaded onto a Ni-IDA column. The flow through, containing the His-Tag cleaved CyanoP, was concentrated to 24 mg/ml with centrifugal concentrator devices (Sartorius, Germany) (Figure S2b).
Concentrated samples were placed in crystallisation screens using a Mosquito® robot (TTP LabTech, UK).

Crystallization and structure solution.

The protein at around 24 mg/ml was screened against several commercially available screens using sitting-drop vapour diffusion. Needle-shaped crystals suitable for data collection appeared in 0.2 M zinc acetate, 0.1 M sodium cacodylate pH 6.5, 14.4 % w/v PEG 8K, 20 % v/v glycerol (Fig S2c). Smaller crystals also appeared in other screen conditions containing zinc ions. Crystals only appeared in conditions containing zinc ions. Crystals were cryoprotected for a few minutes in a solution made from 10 µl of the mother liquor with 3 µl glycerol added, then flash frozen into liquid nitrogen. Elliptical litholoops (Molecular Dimensions, Ltd), were used to mount the long thin crystals. Data were collected on beamline I02 at the Diamond synchrotron at the peak energy of Zn K-edge as determined by a fluorescence scan. The data were integrated and processed using MOSFLM (*1*) and programs of the CCP4 suite (*2*). Zinc sites were found and phases calculated using the Phenix AutoSol Wizard (*3*). The structure was built into the density-modified experimental map manually using COOT (*4*), and refined with REFMAC (*5*) against the SAD phase target. MolProbity (*6*) was used for validation. Data collection and refinement statistics are shown in Table 1. Structures have been aligned and compared using alignments based on secondary-structure matching (*7*). Structure figures were prepared using PyMol (*8*), sequences were aligned using the structure based "expresso" mode of T-Coffee (*9*) with the higher plant PsbP structure and the CyanoP structure described here, and presented using TeXshade (*10*).

Supplementary Tables

Table S1. CyanoP - Crystallographic data collection and refinement statistics.

Crystal parameters		
Space group		P 3_2 2 1
Cell dimensions		a=b=64.95 Å c=81.80 Å α=β=90º, γ=120º
Data collection		
Beamline		Diamond Light Source I02
Wavelength (Å)		1.28220
Resolution (Å)		56.3-2.80 (2.95-2.80)
Unique observations		5133 (728)
R_{merge}		0.11 (0.45)
Mean (<*I*>/ σ*I*)		12.5 (4.8)
Completeness (%)		98.9 (98.4)
Multiplicity		6.8 (4.8)
Refinement		
R_{work} / R_{free} (%)		19.2 / 23.8
Number of protein residues		155
RMSD stereochemistry		
	Bond lengths (Å)	0.017
	Bond angles (°)	1.6
Ramachandran analysis (%)		
	Residues in outlier regions	0%
	Residues in favoured regions	97.35%
	Residues in allowed regions	2.65%

Numbers in parentheses refer to the outermost resolution shell.
R_{merge} = Σ |*I* -<*I*>| / Σ *I* where *I* is the integrated intensity of a given reflection and <*I*> is the mean intensity of multiple corresponding symmetry-related reflections.
R_{work} = Σ ||F_o| - |F_c|| / Σ F_o where F_o and F_c are the observed and calculated structure factors respectively.
R_{free} = R_{work} calculated using ~5% random data excluded from the refinement.
RMSD stereochemistry is the deviation from ideal values.

Table S2. Conservation of zinc-binding residues by PsbP-like protein family, as classified by (*11*). No single site is conserved across all families.

Zn site	Residue in CyanoP	Fully conserved families	Partially conserved families
1	Asp31	G, E, D	H1
1	Asp34	E	H1
1	Asp54	E	
1'	Glu164		
2	Glu87	A	B, E
2	Asp91		
2	His58		H2
3	Glu170	E, G	E, G, H1
4	His142	A, B, D, H1, H2	E
4	Glu163	E, F, H1, H2	A
5	Asp59	A, E, H1, H2	C

Supplementary Figures

Figure S1. Differential protein extraction of membranes. Crude thylakoid membranes were isolated from *Thermosynechococcus elongatus* by glass bead breakage and differential centrifugation (12). An amount of crude thylakoid membranes corresponding to 20 µg of Chl *a* were resuspended in 50 µl extraction buffer (EB; 25 mM MES pH=6.5, 1 mM ε-amino caproic acid) and subsequently mixed with further 50 µl extraction buffer containing various additives, such as: 4 M NaCl or 40 mM CAPS pH=12.0. Samples were then subjected to two consecutive freeze (30 min at -80 °C) / thaw (30 min at RT) cycles (F/T) and separated into a soluble (S) and a pellet (P) fraction by ultracentrifugation (TLA 120.1; 100,000 g, 20 min, 4 °C). Samples were analysed on 18 % (w/v) polyacrylamide SDS PAGE gels (13), which were either Coomassie stained (a) or used for immunoblotting with the indicated specific antibodies (b) anti-D1 (c) anti-PsbO (d) anti-CyanoP (Boehm *et al.*, unpublished). The gel and immunoblots show that CyanoP is strongly associated with the thylakoid membrane, like D1, and unlike PsbO, which is removed by high salt or high pH washes.

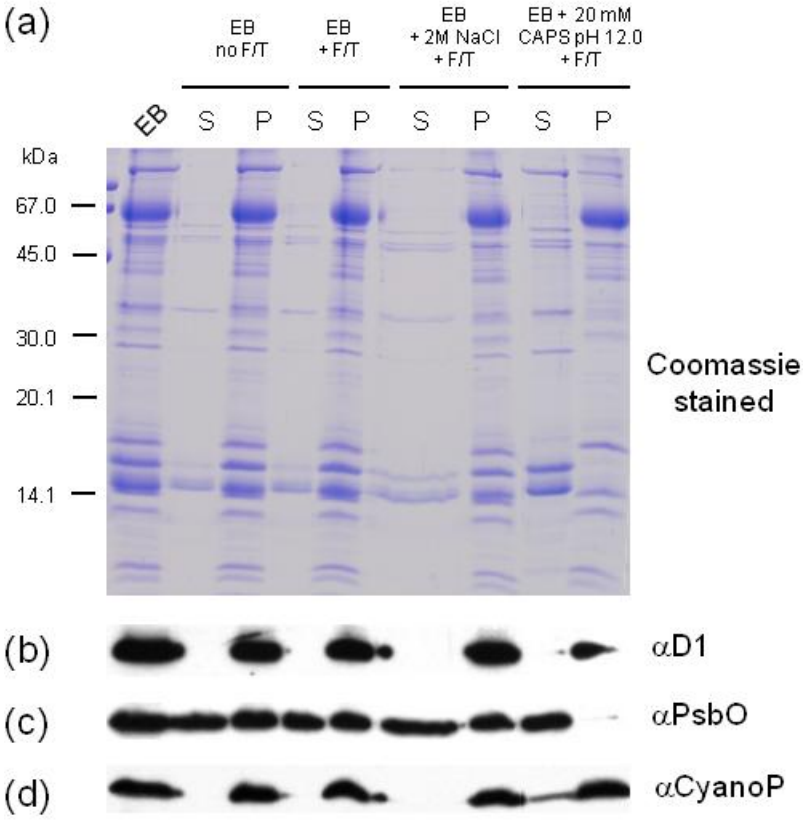
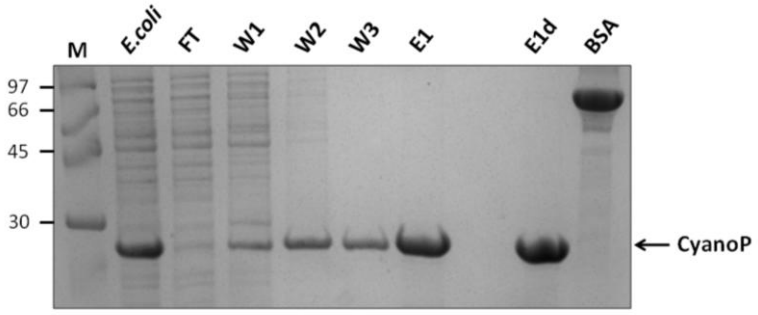
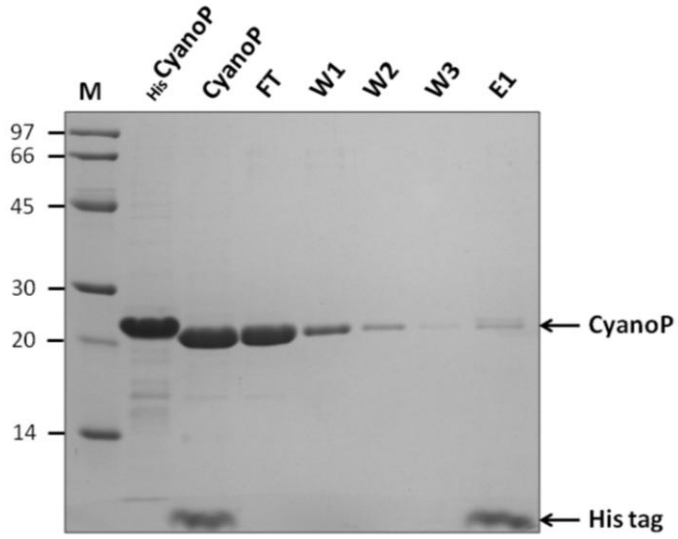


Figure S2. Purification and Crystallization of CyanoP.

(a) Purification of *T. elongatus* His-tagged CyanoP overexpressed in *E. coli*. Soluble protein samples taken at different time points during the affinity purification of CyanoP were loaded onto a 16 % (w/v) SDS-PAGE gel and visualised by Coomassie blue staining. Samples included KRX *E. coli* cell soluble proteins (*E .coli*), flow through (FT), the three washes (W1, W2 and W3), the elution before (E1) and after (E1d) dialysis along with 10 µg of bovine serum albumin (BSA).



(b) Removal of the His-tag from CyanoP. The N-terminal His-tag of CyanoP ($_{His}$CyanoP) was removed by thrombin cleavage (CyanoP) and the His-tag was separated from the sample using a Ni-NDA column in a similar fashion as for the purification of CyanoP. Cleaved CyanoP was found in the flow through (FT) and was used for further crystallisation studies. Each removal step was analysed on a 16 % SDS-PAGE gel and visualised by Coomassie blue staining.



(c) The crystals of CyanoP used to solve the structure. The needles are ~200 μm long. The condition was 0.2 M zinc acetate, 0.1 M sodium cacodylate pH 6.5, 14.4 % w/v PEG 8K, 20 % v/v glycerol.

Figure S3. Sequence alignment showing *T. elongatus* CyanoP, and other selected sequences. 1. *T. elongatus* Tlr2057, 2. *Synechocystis* PCC 6803 Sll1418 3. *Gloeobacter violaceus* Gll1440 4. *Cyanidioschyzon merolae* 5. *Chlamydomonas reinhardtii* 6. Spinach *Spinacia oleracea* 7. Tobacco *Nicotiana tabacum.* The secondary-structures of CyanoP are marked above the alignment with the same colour-scheme as Fig. 1, the missing loop (residues 133-137) is marked with a brace. Below the alignment the zinc-ligating residues are shown with a circle and the number of the zinc ligated. Zinc atoms ligated by symmetry-related molecules are shown with a prime. Presequences not present in the mature protein are shaded orange for 'lipobox' proteins and magenta for Tat transported PsbP. The predicted lipidated cysteines are shown in yellow.

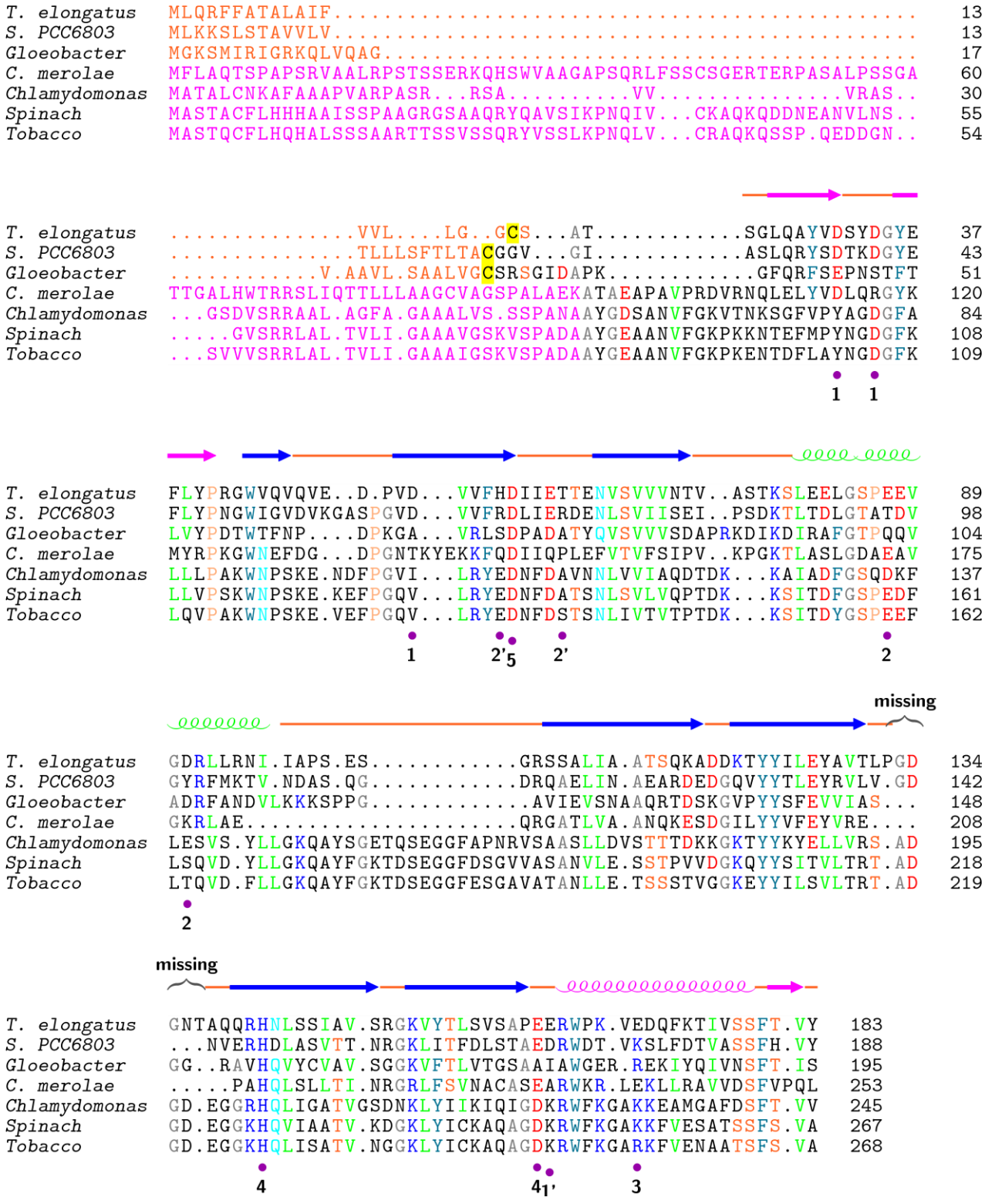


Figure S4. Cartoon view of PsbP coloured by sequence conservation (calculated from the alignment in (11), using TeXshade.) Red residues are unconserved, green are highly conserved. Zinc atoms are marked with ligating residues shown in stick representation.

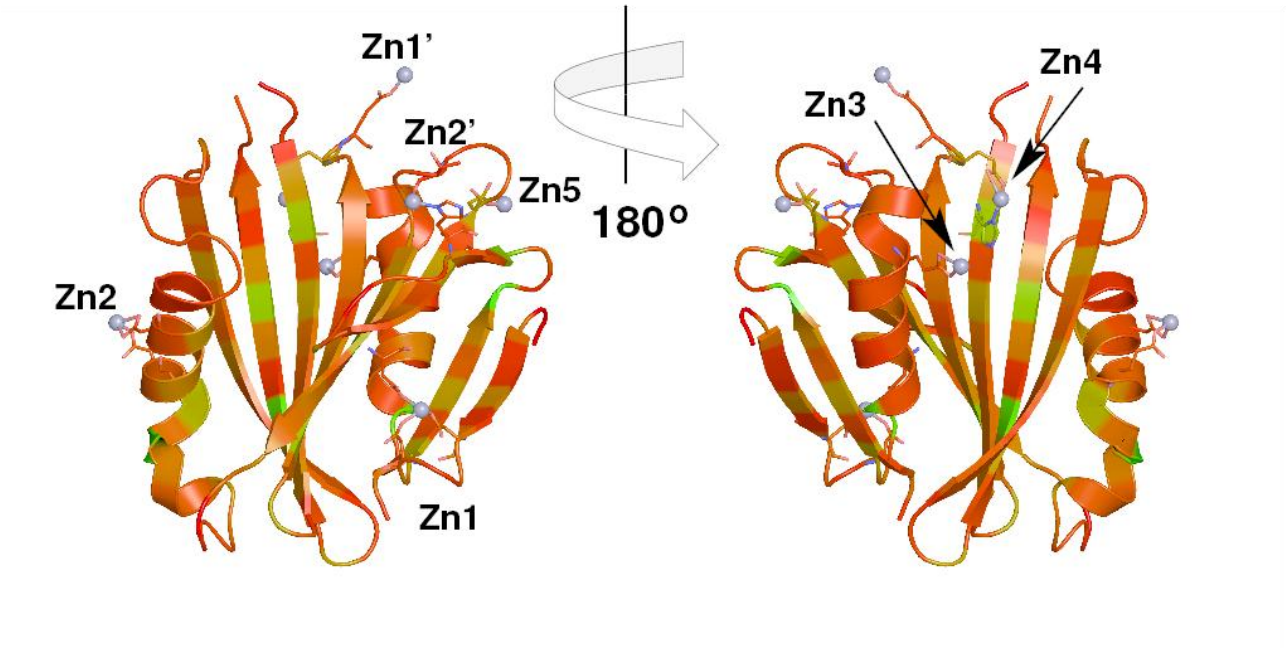


Figure S5. Electron density for the five zinc sites. In blue is 2mFo-DFc density from REFMAC (contoured at 1.8 σ), and green is the phased anomalous difference map (contoured at 10 σ and calculated with the FFT program from the CCP4 suite).

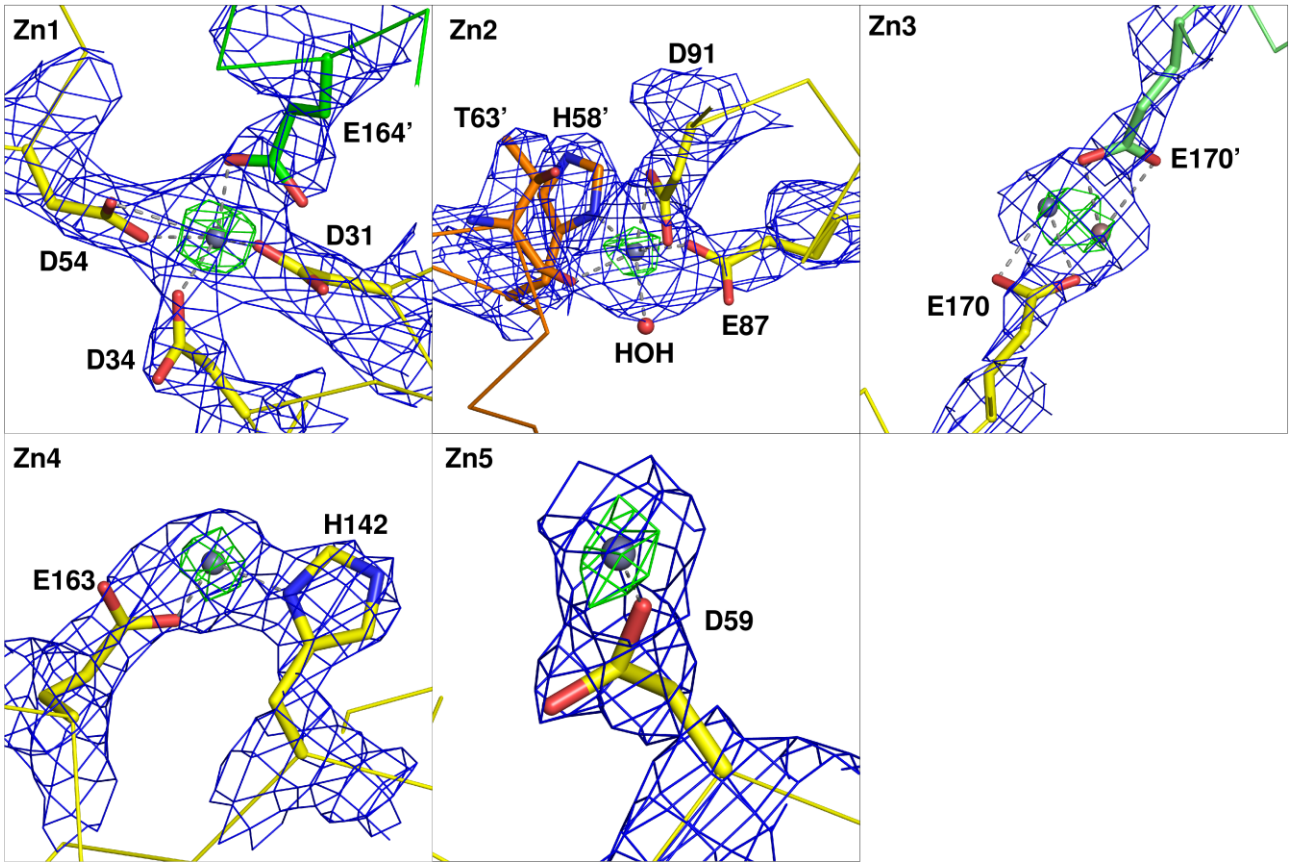


Figure S6. Comparison of CyanoP with other structures. CyanoP is shown in blue and the compared structure in yellow. Each superposition is shown with the C_α RMSD and over how many aligned residues this was calculated.

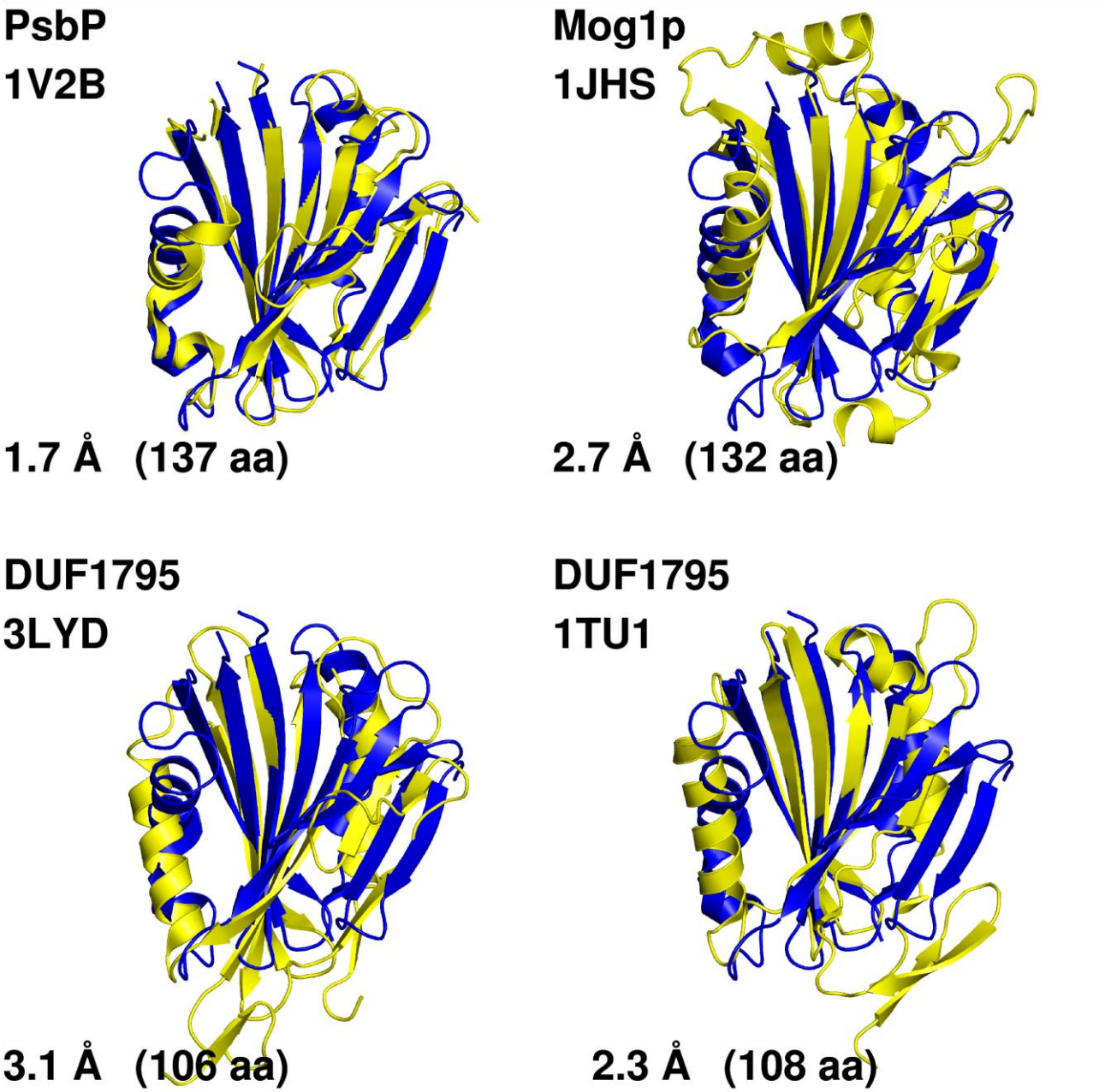
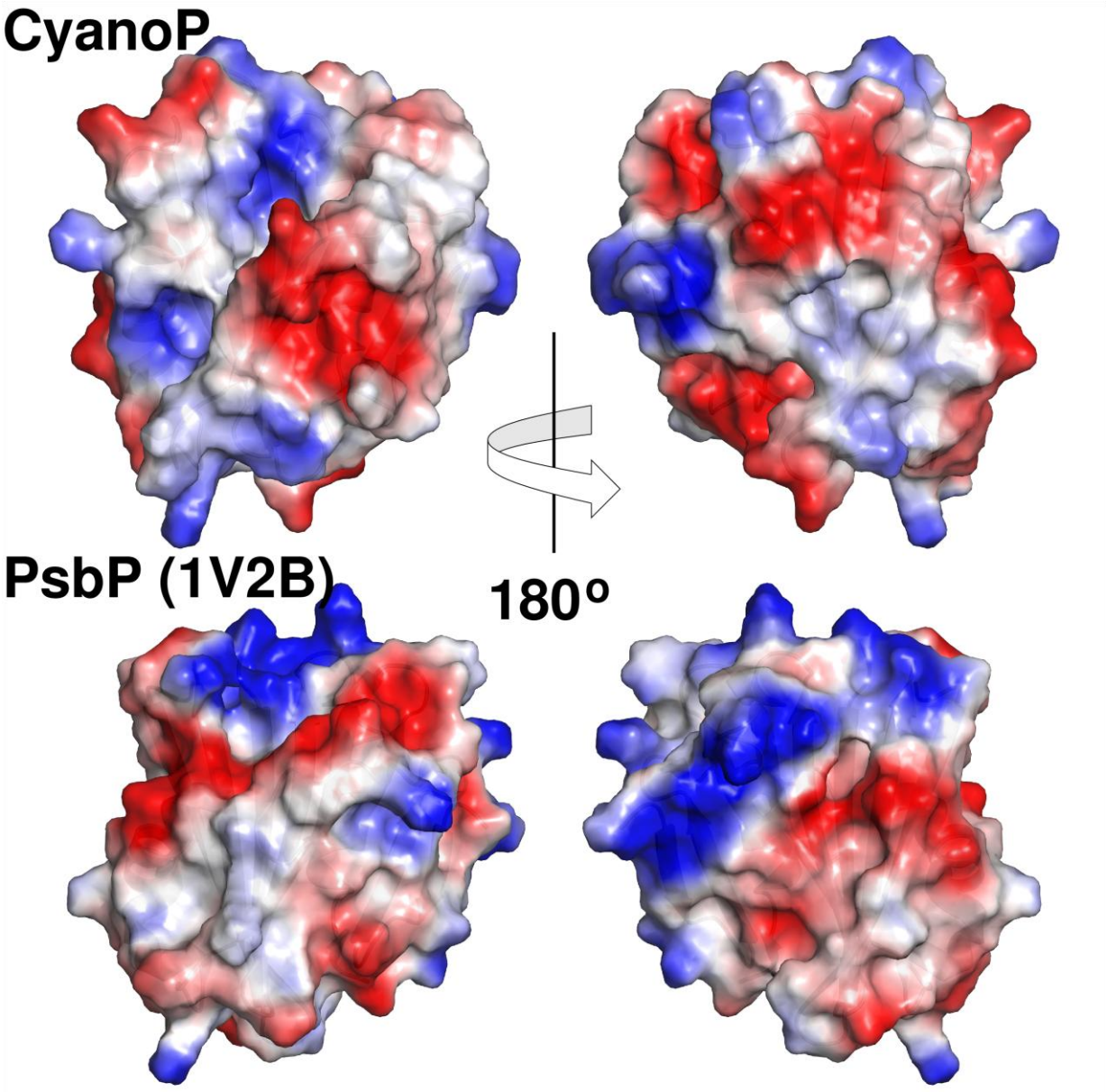


Figure S7. Electrostatic potentials of CyanoP and plant PsbP as calculated in PyMol, incomplete residues were completed with rotameric amino acid positions. The initial protein orientations are as in Fig. 1.



Supplemental References

1. Leslie, A. (1992) Recent changes to the MOSFLM package for processing film and image plate data, *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography*.
2. CCP4 consortium. (1994) The CCP4 Suite: Programs for Protein Crystallography, *Acta Cryst. D50*, 760-763.
3. Adams, P., Afonine, P., Bunkoczi, G., Chen, V., Davis, I., Echols, N., Headd, J., Hung, L., Kapral, G., Grosse-Kunstleve, R., McCoy, A., Moriarty, N., Oeffner, R., Read, R., Richardson, D., Richardson, J., Terwilliger, T., and Zwart, P. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution., *Acta Cryst. D 66*, 213-221.
4. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Cryst. D 66*, 486-501.
5. Murshudov, G., Vagin, A., and Dodson, E. (1997) Refinement of macromolecular structures by the maximum-likelihood method., *Acta Crystallogr D Biol Crystallogr 53*, 240-255.
6. Davis, I., Leaver-Fay, A., Chen, V., Block, J., Kapral, G., Wang, X., Murray, L., Arendall, B., Snoeyink, J., Richardson, J., and Richardson, D. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids, *Nucl. Acids Res. 32*, W375-W383.
7. Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions., *Acta Cryst. D 60*, 2256-2268.
8. Delano, W. (2002) The PyMOL Molecular Graphics System.
9. Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., Keduas, V., and Notredame, C. (2006) Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee, *Nucl. Acids Res. 34*, W604-W608.
10. Beitz, E. (2000) TeXshade: shading and labeling of multiple sequence alignments using LaTeX2e, *Bioinformatics 16*, 135-139.
11. Sato, N. (2010) Phylogenomic and structural modeling analyses of the PsbP superfamily reveal multiple small segment additions in the evolution of photosystem II-associated PsbP protein in green plants., *Mol. Phylogenomics. Evol. 56*, 176-186.
12. Komenda, J., Lupínková, L., and Kopecký, J. (2002) Absence of the psbH gene product destabilizes photosystem II complex and bicarbonate binding on its acceptor side in *Synechocystis* PCC 6803., *Eur. J. Biochem. 269*, 610-619.
13. Boehm, M., Nield, J., Zhang, P., Aro, E., Komenda, J., and Nixon, P. (2009) Structural and mutational analysis of band 7 proteins in the cyanobacterium *Synechocystis* sp. strain PCC 6803., *J. Bacteriol. 191*, 6425-6435.