

Supporting Information

Comparison of PsbQ and Psb27 in photosystem II provides insight into their roles

Christopher J. Gisriel¹ and Gary W. Brudvig^{1,2,*}

¹Department of Chemistry, Yale University, New Haven, CT 06520, USA.

²Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA.

*To whom correspondence should be addressed: gary.brudvig@yale.edu

Supplementary Fig. 1 Sequence alignment of PsbQ(') and Psb27.

Supplementary Fig. 2 Comparison of the CP43 luminal domain between apo-PSII, Psb27-PSII, and mature PSII.

Supplementary Fig. 3 Structure-based sequence alignment of PsbQ(') and Psb27.

Supplementary Fig. 4 Sequence alignment of PsbQ(') from different organisms.

Supplementary Fig. 5 Phylogenetic relationships of PsbQ(') and Psb27.

Supplementary Fig. 6 Calculate PIPER binding energies for PsbQ(').

Supplementary Fig. 7 Example of PsbQ(') N-terminal extension in plants and algae compared to cyanobacteria.

Supplementary Fig. 8 Sequence alignment of Psb27 subunits from different organisms.

Supplementary Fig. 9 Calculated PIPER binding energies for Psb27.

Supplementary Fig. 10 Lack of electrostatic potential corresponding to a Cl⁻ ion bound to Psb27 in the dimeric Psb27-PSII complex.

Supplementary Table 1 Sequence identity matrix of selected PsbQ(') and Psb27 calculated using Clustal Omega.

Supplementary Table 2 Sequence identity matrix of selected PsbQ(') and Psb27 sequences calculated using structure-based alignments from PROMALS3D.

Supplementary Table 3 Sequence identity matrix for PsbQ(') sequence alignment.

Supplementary Table 4 Number of amino acids missing from molecular structures of PsbQ(').

Supplementary Table 5 Sequence identity matrix for Psb27 homolog alignment.

Supplementary Figures

```

S. 6803 PsbQ ----- 0
T. elongatus PsbQ ----- 0
T. vulcanus PsbQ ----- 0
C. caldarium PsbQ' -----MFVGTSGIFGERLCSRL--MIKRAQGVRSVTRGRARVTTA 38
C. gracilis PsbQ' -----MKSATICFVA--LAASASAFAP---EA-NNAR 25
C. reinhardtii PsbQ -----MALASKVAT-RPAVASRRGA 19
P. sativum PsbQ MAQAMASSMAGCLRGCSSTVLEGLSLQFSGPN---RL--SLLHGNTNNV-NKVTRRSSVT 54
S. oleracea PsbQ MAQAMAS-MAG-LRG-ASQAVLEGLSLQISGSN---RL--SGPTTSRVAV-PKMGLNIRAQ 51
A. thaliana PsbQ MAQAVTS-MAG-LRG-ASQAVLEGLSLQINGSN---RL--NISRV-SVGS-QRTGLVIRAQ 50
S. sp. PCC 6803 Psb27 ----- 0
T. elongatus Psb27 ----- 0
T. vulcanus Psb27 ----- 0
C. reinhardtii Psb27 -----MASITC-----SSGKAADVSRAKTARPVARRMSV-----VA 31
S. oleracea Psb27 -----MASCSLIT-----PTSCLKLLSIKSLIPTTTSAAP-----ST 33
A. thaliana Psb27 -----MASASATA-----TLLKPNLPPHKPTIIA--SSVSP-----PL 31

```

```

S. 6803 PsbQ -----MSRLRSLLSLILVLVTTV-----LVSCSSPQVEI--PT-T-----YS 34
T. elongatus PsbQ ----MLRLNRKSLISILLSVVALI-----LVGCGGPSATTPPPP-T-----YS 38
T. vulcanus PsbQ ----MLRLNRKSLISILLSVVALI-----LVGCGGPSATTPPPP-T-----YS 38
C. caldarium PsbQ' KLADDTPLDRRQLLQVAATSATV-----LIGKALPSSAAGEPKMSFFGADAPSSFFTYN 93
C. gracilis PsbQ' VATSINAEGRRVFGKIAAAGAAF-----LPVAANAAVGESPRFSVFLVGDGTSYSEG 79
C. reinhardtii PsbQ VVVRASGESRRRAVLGGLLASAVAAPKAALAL-TPVDLFD--DRS---VRDRGFDLIYE 73
P. sativum PsbQ VRAQQQESRRRAVIGLVATGLVSSSFVQAVLAEAIPIKVGPPPLS---GGLPGTLNSDE 111
S. oleracea PsbQ QVSAEAEYSRRAMLGFVAAGLASGSFVKAVLAEARPIVVGPPPLS---GGLPGTENSdq 108
A. thaliana PsbQ QNVSVPESSRRSVLGLVAAGLAGGSFVKAVFAEAIPIKVGPPPLS---GGLPGTDNSDQ 107
S. sp. PCC 6803 Psb27 -----MSFLKNQLSR-----LLALIL-----VVAIGLTAC-DSGTGLTGN 34
T. elongatus Psb27 -----MKRFWAM-----VCALFL-----SVSLLTSCANVPTGLTGN 32
T. vulcanus Psb27 -----MKRFWAM-----VCALFL-----SVSLLTSCANVPTGLTGN 32
C. reinhardtii Psb27 SAAPQGNASRRELLGLSVAASLLSS-----RPAHAIFGFADDNTALFDT 77
S. oleracea Psb27 SAPAAASAFTRRQVSVTA--LTAAIL-----SPLLSSAAHAAADDE 72
A. thaliana Psb27 PPPRRNHLRRDFLSLAATSTLLTQSI-----QFLAPAPVSAAEDEE 73

```

:

```

S. 6803 PsbQ -----PEKIAQL-----QVYVNP---I-----AVARDGMEKRL 59
T. elongatus PsbQ -----ELQITRI-----QDYLRD---I-----EKNAERF-ADL 62
T. vulcanus PsbQ -----ELQITRI-----QDYLRD---I-----EKNAERF-ADL 62
C. caldarium PsbQ' -----EREGEVYKGITPELLNEYRASING-----AKKRIEDA 126
C. gracilis PsbQ' AAYGTDQADKLYSPYSVSPGEKSLYKPDNAEYLARKKAVL-----AETKNRLQKI 131
C. reinhardtii PsbQ -----ARDLDLPQN--VREGFTQARASLDETCKRVKSEARIDADD 112
P. sativum PsbQ -----ARDLKLP-L--KERFFIQP-LAPTEAAARTKESAKE-IVAA 147
S. oleracea PsbQ -----ARDGTLPYT--KDRFYLQP-LPTEAAQRAKVSASE-ILNV 145
A. thaliana PsbQ -----ARDFSLA-L--KDRFYIQP-LSPTTEAAARAKDSAKE-IINV 143
S. sp. PCC 6803 Psb27 -----YSQDTLTVIAT--LREAIIDLQDAP-----NRQEVQDFTARGQI 70
T. elongatus Psb27 -----FREDTLALISS--LREAIALPENDP-----NKKAAQAEARKKL 68
T. vulcanus Psb27 -----FREDTLALISS--LREAIALPENDP-----NKKAAQAEARKKL 68
C. reinhardtii Psb27 -----YTAETSAILDK--VKVTLALDKDDP-----AKEDSVKGLRKDI 113
S. oleracea Psb27 -----YVKETKEVISK--VRSTLNKNKSDP-----DVADAVTELRAVS 108
A. thaliana Psb27 -----YIKDTSAVISK--VRSTLSMQKTDp-----NVADAVAELEAS 109

```

```

S. 6803 PsbQ QGLIADQNVDVTQTYIHGFLGQLRRDMLGLASSLL-----PKDQDKAKTLAKEVFGHL 112
T. elongatus PsbQ EVSVAKGDWQEQEARNIMRGLGEMLMDMRALNRNLL-----AKDQPTPTALTRALTDFF 115
T. vulcanus PsbQ EVSVAKGDWQEQEARNIMRGLGEMLMDMRALNRNLL-----AKDQPTPTALTRALTDFF 115
C. caldarium PsbQ' GEAIAKKSWEDVRSALRLAVGTLRSVCSKVNSYAM---AQATKQEKQNIKAYREFLKR 183
C. gracilis PsbQ' PAYVDKKEWFNVKDELTRYMYETRGAVRSLSSSVT-----Q-----KEKAEVFFRAL 178
C. reinhardtii PsbQ DVFIQKSYWTEAREQLRRQVGTLRFDLNTLAS-----TKEKEAKKAALGLRKEFIQAV 165
P. sativum PsbQ KKFIDQKAWPFLQNDLRLRAGYLRDYLKTIIS-----SKPKDQKQSLKELTDKLFQDI 200
S. oleracea PsbQ KQFIDRKAWPSLQNDLRLRASYLRYDLKTVIS-----AKPKDEKKSQELTSLKLFSS 198
A. thaliana PsbQ KSFIDKKAWPYQNDLRLRASYLRYDLNTVIS-----AKPKEEKQSLKDLTAKLFQTI 196
S. sp. PCC 6803 Psb27 NDYISRYRRKGDG--GGLKSFTMTQALNSLAGHYYSYGA-RPIPEK-LKKRLQLEFTQAE 127
T. elongatus Psb27 NDFFALYRRDSSL-RSLSSFMTQALNSLAGHYSSYPN-RPLPEK-LKARLEQEFKQVE 125
T. vulcanus Psb27 NDFFALYRRDSSL-RSLSSFMTQALNSLAGHYSSYPN-RPLPEK-LKARLEQEFKQVE 125
C. reinhardtii Psb27 NNWVAKYRREPKV-SGKPSFGNTYSALNALAGHFNSFGATAPIPKK-RLERLQKLEDDAT 171
S. oleracea Psb27 NFWVAKYRKEKNL-LGRPSFRDMYSALNAVSGHYISFGFTSPIPAK-RRARILEEMDSVE 166
A. thaliana Psb27 NSWVAKYRKEKAL-LGKASFRDIYSALNAVSGHYISFGFTAPIPAK-RKARILEEMETA 167

```

:

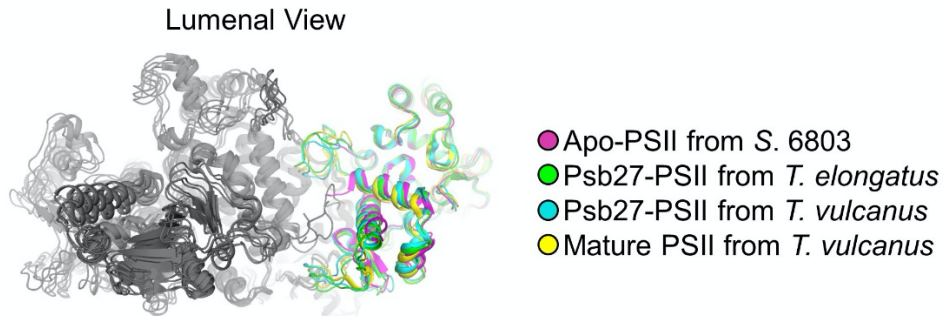
```

S. 6803 PsbQ ERLDAAAKDRNGSQAKIQYQEAALADFDNFLNLLPQAS 149
T. elongatus PsbQ LKIDQGADLDSVTVAQEGFREAEADFKAAYLNSLPELS 152

```

<i>T. vulcanus</i> PsbQ	LKIDQGADLDSVTVAQEGFREAEADFKAYLNSLPELS	152
<i>C. caldarium</i> PsbQ'	EDMDFARMKDQDRAERLRTASISALDNWMSVVGI--	218
<i>C. gracilis</i> PsbQ'	EDTYGAATLKKGDAVKASNDKAIAALDAFTATL----	211
<i>C. reinhardtii</i> PsbQ	EDLDFALREKDAQASAACKLEITKAKLDSVLA AVL---	199
<i>P. sativum</i> PsbQ	SNLDHAAKIKSPSEAEKYYAIAVSTLNDVLSKIA---	234
<i>S. oleracea</i> PsbQ	DNLDHAAKIKSPTEAEKYYGQTVSNINEVLAKLG---	232
<i>A. thaliana</i> PsbQ	DNLDYAARSKSPDAEKYYSETVSSLNVLAKLG---	230
<i>S. sp. PCC 6803</i> Psb27	RSIERGV-----	134
<i>T. elongatus</i> Psb27	LALDREAKS-----	134
<i>T. vulcanus</i> Psb27	LALDREAKS-----	134
<i>C. reinhardtii</i> Psb27	LLLTRNR-----	178
<i>S. oleracea</i> Psb27	KALLRGR-----	173
<i>A. thaliana</i> Psb27	KALTRGR-----	174

Supplementary Fig. 1 Sequence alignment of PsbQ(') and Psb27. PsbQ(') sequences are colored in blue and Psb27 sequences are colored in green. The sequence alignment was performed in Clustal Omega (Sievers et al. 2011), and the Clustal Omega conservation identifiers are listed below each section, although they are mostly blank which means that most positions do not exhibit similarity. The corresponding sequence identity matrix is shown in **Supplementary Table 1**. *Synechocystis* 6803 is abbreviated *S. 6803*.



Supplementary Fig. 2 Comparison of the CP43 luminal domain between apo-PSII, Psb27-PSII, and mature PSII. Only the core subunits D1, D2, CP47, and CP43 are shown. These are superimposed by the D1 and CP43 subunits only, and the CP43 subunits are colored according to the key. The apo-PSII structure corresponds to structure PDB 6WJ6, the Psb27-PSII structures correspond to PDB 7NHP for *T. elongatus* and PDB 7CZL for *T. vulcanus*, and the mature PSII structure corresponds to PDB 7D1T. *Synechocystis* 6803 is abbreviated *S. 6803*.

```

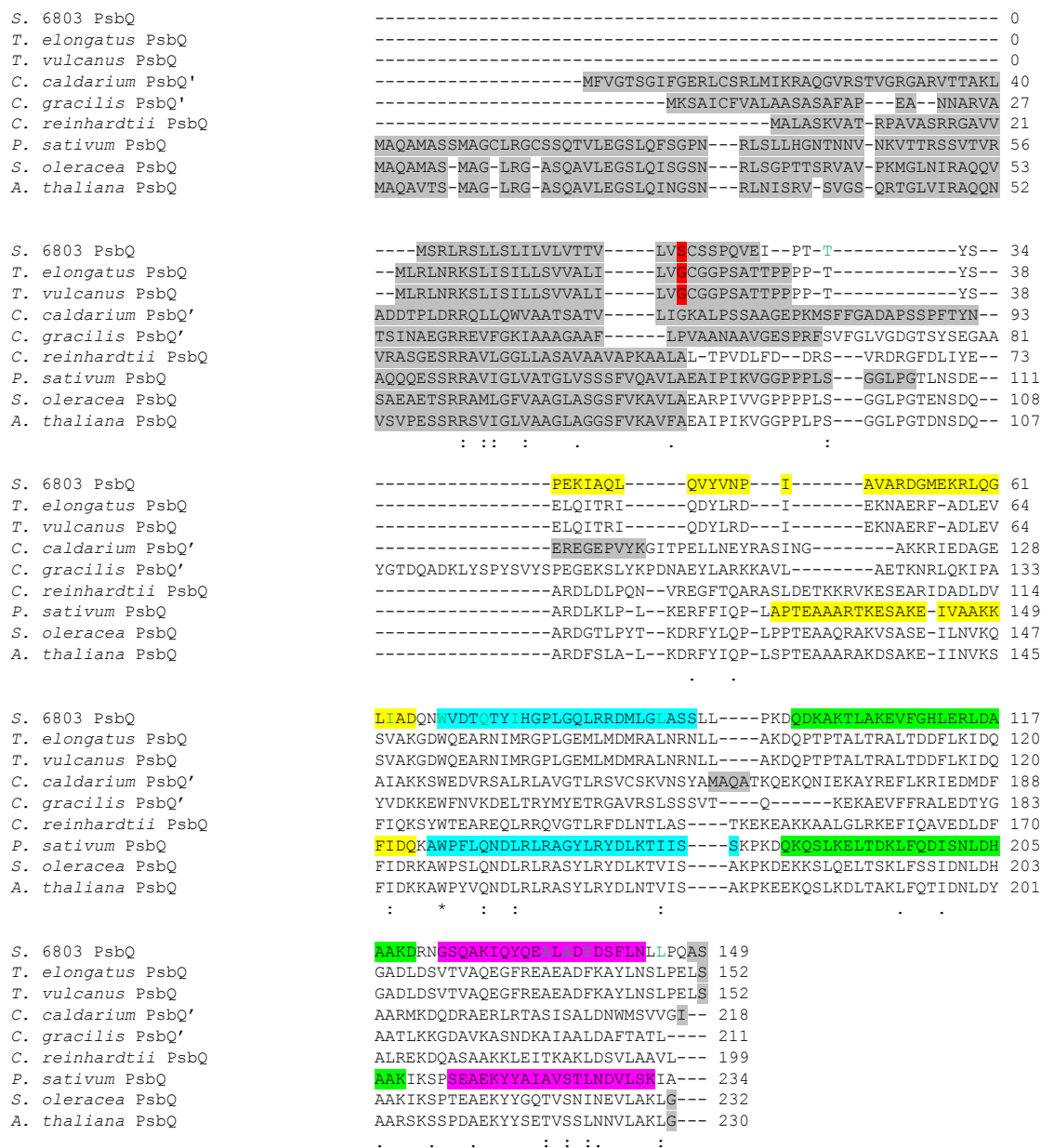
S. 6803 PsbQ -----IPT-----TYSPE-----KIAQLQVY 16
T. elongatus PsbQ -----PPT-----YSE--LQI-----TRIQDY 15
T. vulcanus PsbQ -----PPT-----YSE--LQI-----TRIQDY 15
C. caldarium PsbQ' -----LNEY 4
C. gracilis PsbQ' -SVFGLIGDGTYSYSEGAAAYGTQADKLYSPYSVYSPEGEKSLYKPD-----SAEYLARK 53
C. reinhardtii PsbQ -----LTPVDLFD-----RSVRD-----RQFDLIYE 22
P. sativum PsbQ -----TLNSDEARDLKL-PLKERFFIQPLAPTEA 28
S. oleracea PsbQ EARPIVVG-PPPLPSGGLPG-----TENSQARDGTLPYTKDRFYLQPLPTEA 48
A. thaliana PsbQ (h) EAIPKIVG-GPPLPSGGLPG-----TDNSQARDPFL-ALKDRFYIQLPSPTA 47
S. sp. PCC 6803 Psb27 -----CDSGTG---LTGNY 11
T. elongatus Psb27 -----PTG--LTGNE 8
T. vulcanus Psb27 -----CANVPTG--LTGNF 12
C. reinhardtii Psb27 (h) -----DTY 3
S. oleracea Psb27 (h) -----DEY 3
A. thaliana Psb27 -----EEY 3

S. 6803 PsbQ -----VNPFI-----AVARDGMEKRLQGLIADQNWVDTQTYIHGFLGQLRRDML-SI 61
T. elongatus PsbQ -----LRDI-----EKNAERF-ADLEVSVAKGDWQEARNIMRGPLGEMLMMDMR-AL 59
T. vulcanus PsbQ -----LRDI-----EKNAERF-ADLEVSVAKGDWQEARNIMRGPLGEMLMMDMR-AL 59
C. caldarium PsbQ' -----RASI-----NGAKKRIE-DAGEAIAKKSWEDEVRSALRLAVGTLRSVCS-KV 48
C. gracilis PsbQ' -----KAVL-----AETKNRLQ-KIPGVYDKKWFNVKDELTRYMYETRGAVR-SL 97
C. reinhardtii PsbQ -----ARDLDPQNVREGFTQARASLDETKKRVKSEARIDADLDVFIQKSYWTEAREQLRRQVGTLRFDLN-TL 91
P. sativum PsbQ -----AART-----KESAKEIV-AAKFFIDQKAWPFQNDLRLRAGYLRDYDLK-TI 72
S. oleracea PsbQ -----AQRA-----KVSASEIL-NVKQFIDRKAWSLQNDLRLRASYLRYDLK-TV 92
A. thaliana PsbQ (h) -----AARA-----KDSAKEII-NVKSFDKKAWSYVQNDLRLRASYLRYDLN-TV 91
S. sp. PCC 6803 Psb27 -----SQDT-----LTVIATLR-EAIDLPO--DAPNRQEVQDARGQINDYISRYR 54
T. elongatus Psb27 -----REDT-----LALISSLR-EAIALP--ENDFNKKAQAQAEARKKLNDFFALYR 51
T. vulcanus Psb27 -----REDT-----LALISSLR-EAIALP--ENDFNKKAQAQAEARKKLNDFFALYR 55
C. reinhardtii Psb27 (h) -----TAET-----SAILDKVK-VTLALD--KDDPAKEDSVKGLRDKDINNVAKYR 46
S. oleracea Psb27 (h) -----VKET-----KEVISKVR-STLNKN--KSDPDVADAVTELRAATSNFVWAKYR 46
A. thaliana Psb27 -----IKDT-----SAVISKVR-STLSMQ--KTDPNVADAVAEALREASNSVWAKYR 46

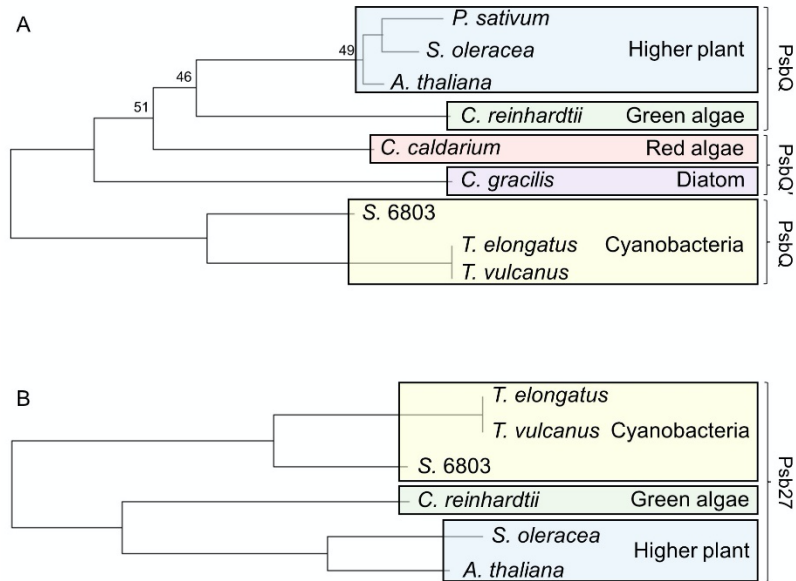
S. 6803 PsbQ -----ASSLLPKDQDKARTLAKEVFGHLERLDAAAK-DRN-SSQAKIQYQEAALADFDSEFLN-LLPQ----- 119
T. elongatus PsbQ -----NRNLLAKDQPTPTALTRALTDFFLKIQQGAD-LDSVTVAQEGFREAEADFKAYLNSL--PEL-- 118
T. vulcanus PsbQ -----NRNLLAKDQPTPTALTRALTDFFLKIQQGAD-LDSVTVAQEGFREAEADFKAYLNSL--PEL-- 118
C. caldarium PsbQ' -----NSYATKQEKQNIKAYREFLKRIEDMDFAAR-MKDQDRAERLRTASISALDNWMSVVG----- 105
C. gracilis PsbQ' -----SSSVTQKEK-----AEVFFRALEDTYGAAT-LKKGDAVKASNDKAIALDAFTATL----- 147
C. reinhardtii PsbQ -----ASTKEKEAKKAALGRKEFIQAVEDLDFALR-EKDQASAAKKLEITKAKLDSVLA AVL----- 148
P. sativum PsbQ -----ISSKPKDQKQSLKELTDKLFQDISNLDHAAK-IKSPSEAEKYAIAVSTLNDVLSKIA----- 129
S. oleracea PsbQ -----ISAKPKDEKKSLELTSKLFSSIDNLDHAAK-IKSPTAEKYYGQTVSNI NEVLAKL----- 148
A. thaliana PsbQ (h) -----ISAKPKEEKQSLKDLTAKLFQTDNLDYAAR-SKSSPDAEKYYSETVSSLNVLAKL----- 147
S. sp. PCC 6803 Psb27 -----R-KGDAGGL-KSFTTMQTALNSLAGYYSY-TSYGAR-PIP-EKLLKRLQLEFTQAERSIERGV--- 110
T. elongatus Psb27 -----RDDSLRSL--SSFMTMQTALNSLAGHYSSY-PNR-PLP-EKLLKARLEQEFKQVELAL--DREAKS 109
T. vulcanus Psb27 -----RDDSLRSL--SSFMTMQTALNSLAGHYSSY-PNR-PLP-EKLLKARLEQEFKQVELAL--DREAKS 113
C. reinhardtii Psb27 (h) -----REPKVS-GK-PSFGNTYSALNALAGHFN SFGATA-PIP-KKRLERLQKELDDATLLL-TRNR-- 103
S. oleracea Psb27 (h) -----KEKNLLGR--PSFRDYSALNAVS GHYISFGPTS-PIP-AKRRARILEEMDSVEKAL-LRGR-- 103
A. thaliana Psb27 -----KEKALL-GK-ASFRDIYSALNAVS GHYVSFGPTA-PIP-AKRRARILEEMETAEBKAL-TRGR-- 103

```

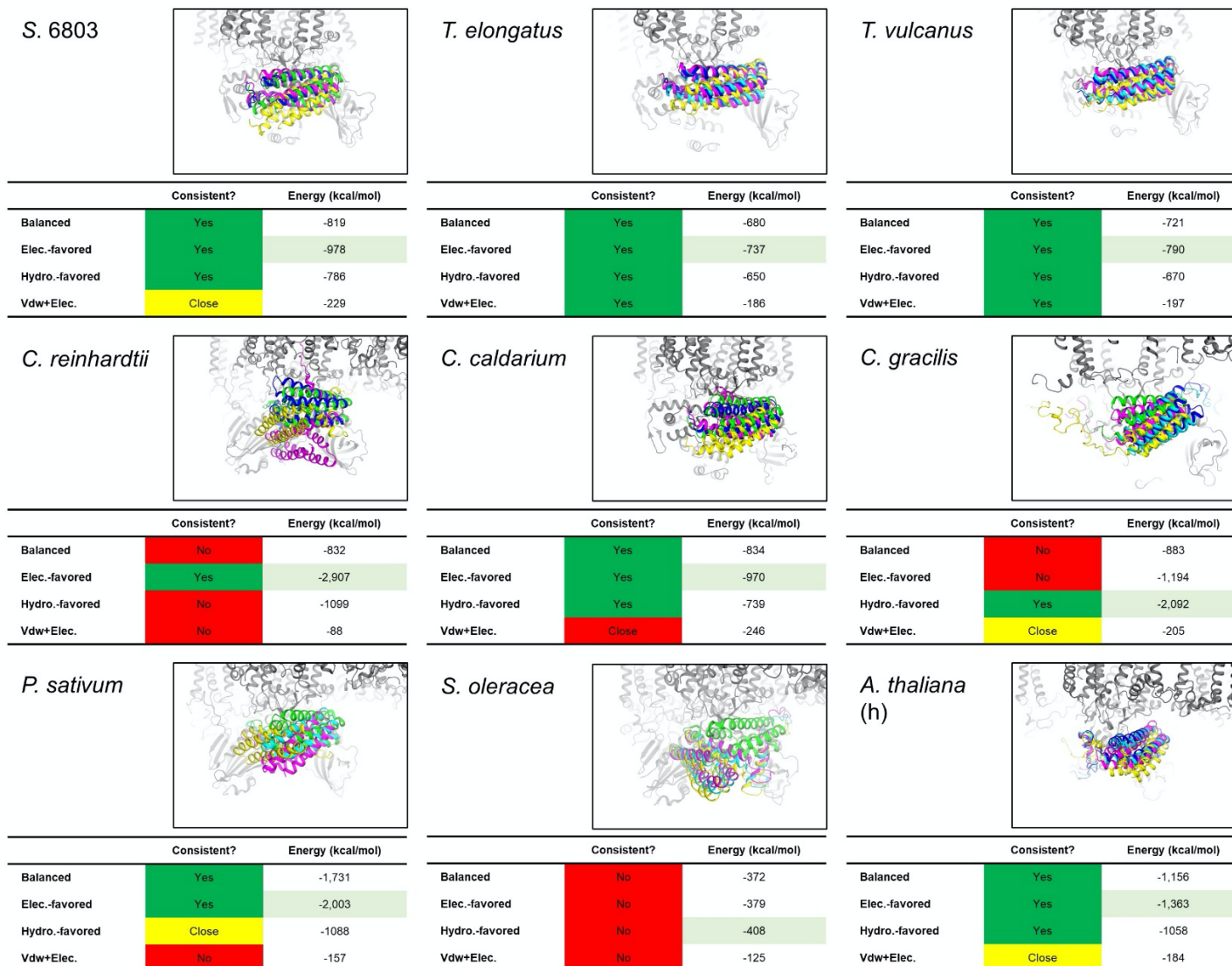
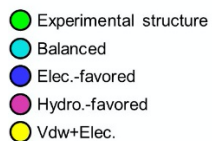
Supplementary Fig. 3 Structure-based sequence alignment of PsbQ(′) and Psb27. PsbQ(′) sequences are colored in blue and Psb27 sequences are colored in green. The amino acids that correspond to α -helices in the four-helix bundle for the structures of PsbQ from *Synechocystis* 6803 (PDB 7N8O) and Psb27 from *T. elongatus* (PDB 7NHP) are highlighted yellow. The alignment was performed in PROMALS3D (Pei et al. 2008). *Synechocystis* 6803 is abbreviated *S. 6803*.



Supplementary Fig. 4 Sequence alignment of PsbQ(′) from different organisms. Residues with green font in the *Synechocystis* 6803 sequence exhibit sequence similarity $\geq 50\%$ between sequences which also corresponds to the labeled residues in Fig. 2. For reference, residues within each of the four α -helices in the four-helix bundle of the PsbQ subunits in the *Synechocystis* 6803 PSII structure (PDB 7N8O) and the *P. sativum* PSII structure (PDB 5XNL) are highlight in yellow, cyan, green, and magenta for α_1 , α_2 , α_3 , and α_4 , respectively. Residues highlighted in red represent cleavage sites of signaling sequences identified by the SigP server (Almagro Armenteros et al. 2019). Residues not modeled in the corresponding molecular structures (or homology model) are highlighted in grey. The alignment was created using Clustal Omega (Sievers et al. 2011) and the sequence identifiers are shown below each section of the alignment. *Synechocystis* 6803 is abbreviated *S. 6803*. Note that a multiple sequence alignment was also performed without the signal sequences, but the outcome was similar.

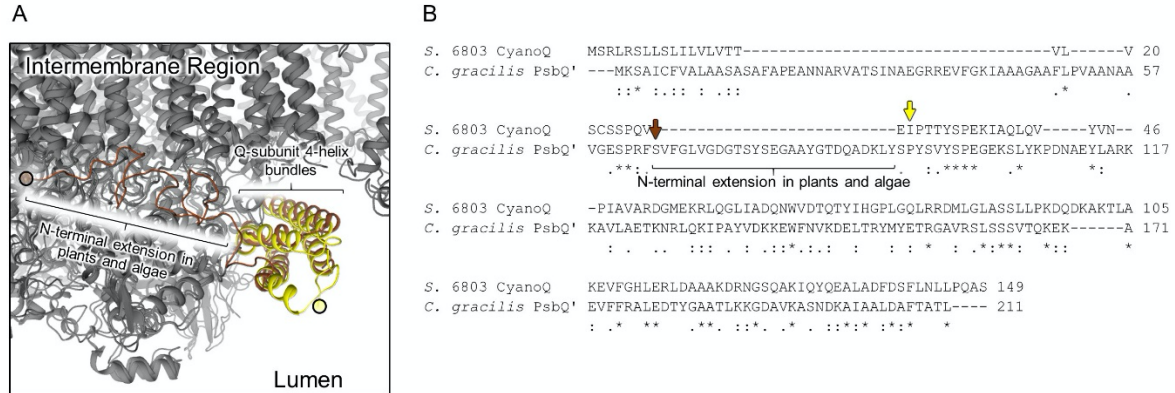


Supplementary Fig. 5 Phylogenetic relationships of PsbQ(′) and Psb27. The evolutionary history for PsbQ(′) (**A**) and Psb27 (**B**) were inferred as described in **Methods** using MEGA7 (Kumar et al. 2016). The percentage of replicate trees in which the associated taxa are clustered together in the bootstrap test (100 replicates) are shown next to the branches except for those where that value is 100%. The tree with the highest log likelihood is shown (-3,533.66 and -1,715.90 for PsbQ(′) and Psb27, respectively). The Gamma distribution +G parameter was 2.2160 and 2.6728 for PsbQ(′) and Psb27, respectively. The trees are drawn to scale, with branch lengths measured in the number of substitutions per site. The analyses involved nine amino acid sequences for PsbQ(′) and six amino acid sequences for Psb27. There was a total of 272 positions in the final data set for PsbQ(′) and 182 positions in the final dataset for Psb27. *Synechocystis* 6803 is abbreviated *S. 6803*.

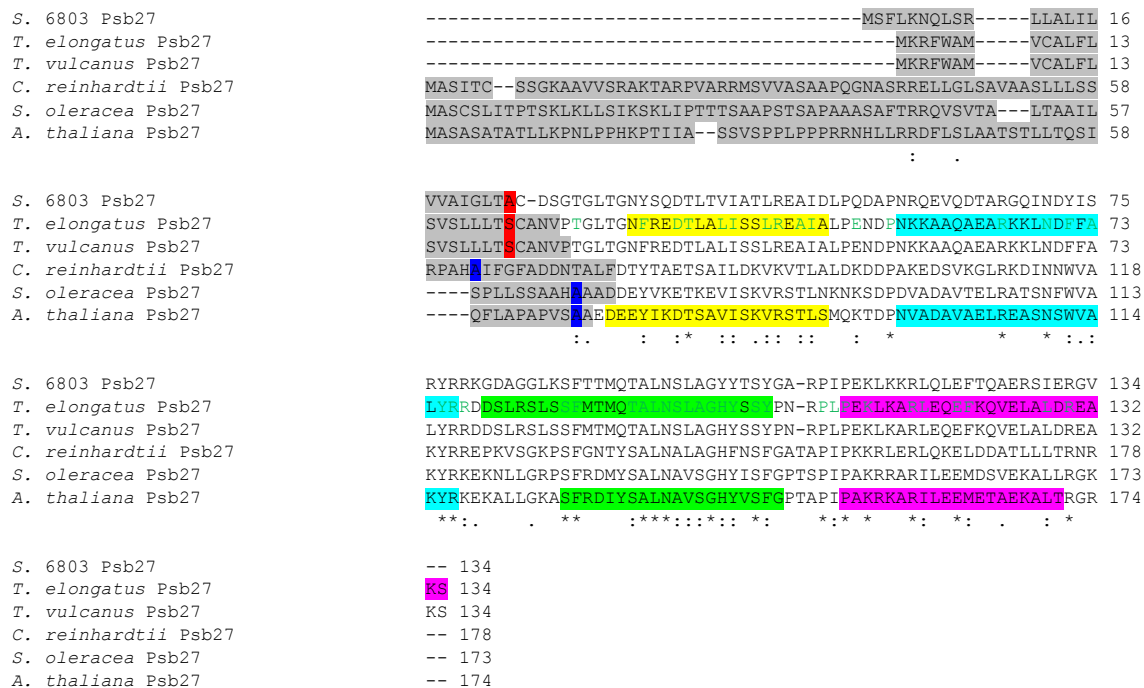


Supplementary Fig. 6 Calculate PIPER binding energies for PsbQ('). Each panel shows an image of the docking simulation of PsbQ(') (colored) onto the PSII core (grey). The green ribbons correspond to the experimentally determined structures. The cyan, blue, magenta, and yellow ribbons correspond to the top model

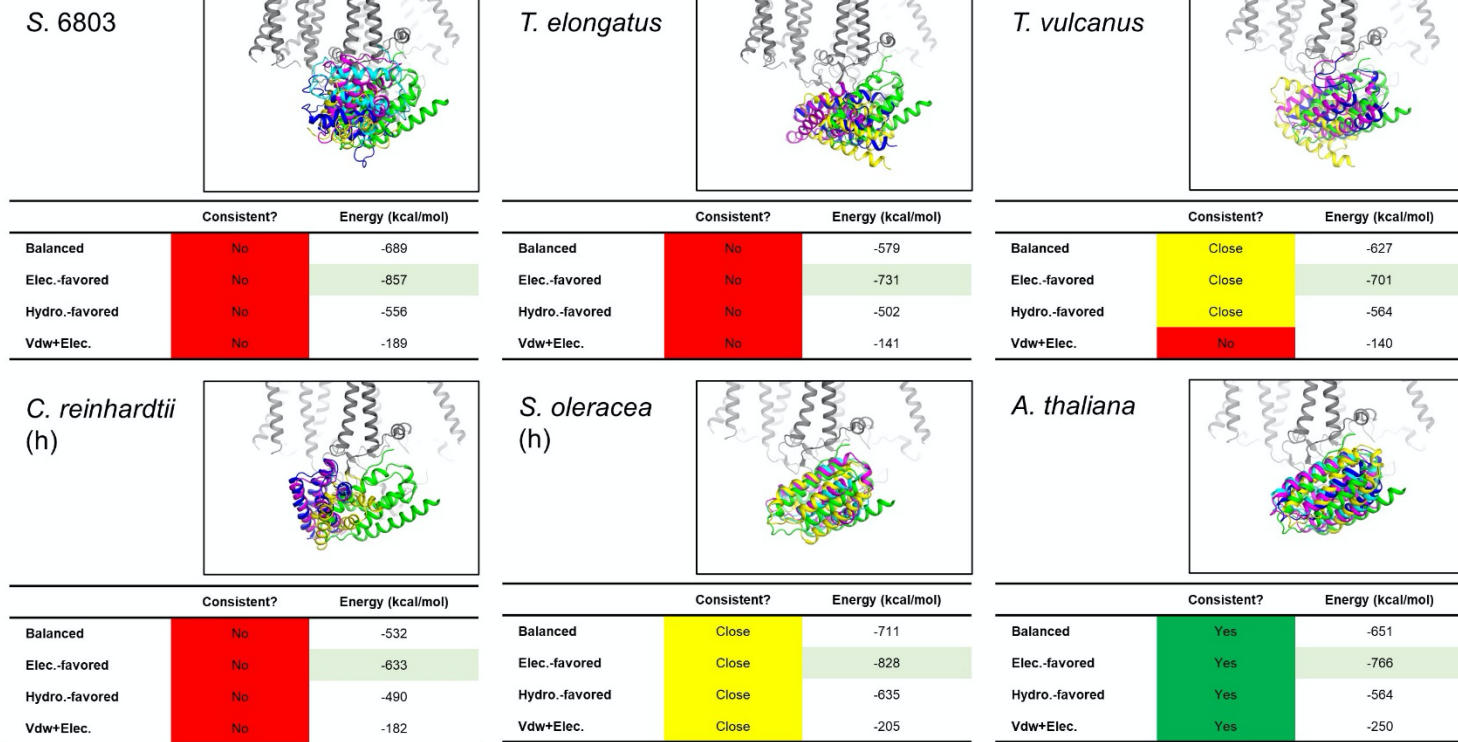
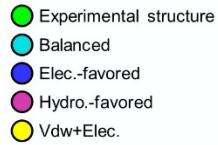
for balanced, electrostatic-favored, hydrophobic-favored, and van der Waals + electrostatic forces (Kozakov et al. 2017). The table below each image shows whether the top model is consistent with the experimental model and the corresponding calculated PIPER binding energy for each coefficient set, with the lowest highlighted in light green. For *A. thaliana*, the docking simulation was performed with a homology model, and is therefore labeled with an “(h)”, and the “Consistent?” column refers to whether the PsbQ(′) orientation is consistent with the closely related PsbQ-containing PSII structures from *P. sativum* and *S. oleracea*. The energy reported in this table is PIPER energy calculated by ClusPro and do not correspond to valid binding free energy, but can be used for general comparison of favorable binding interactions in the docking algorithm (Kozakov et al. 2017). *Synechocystis* 6803 is abbreviated *S.* 6803.



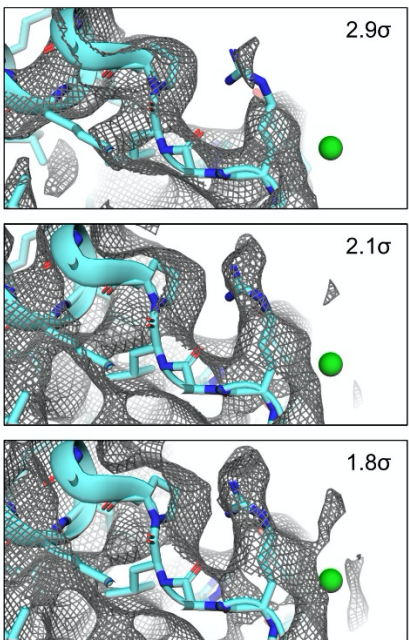
Supplementary Fig. 7 Example of PsbQ(′) N-terminal extension in plants and algae compared to cyanobacteria. **A** The structures of PSII from *Synechocystis* 6803 (7N8O) and *C. gracilis* (PDB 6JLU) are shown superimposed and colored grey except for PsbQ(′) subunit which are colored yellow for *S. 6803* and brown for *C. gracilis*. The closest modeled residue to the N-terminus is shown with a spot colored the same as its corresponding structure. **B** Sequence alignment of PsbQ(′) from *Synechocystis* 6803 and *C. gracilis*. The closest modeled residue to the N-terminus is denoted with an arrow corresponding to the spots in panel A. *Synechocystis* 6803 is abbreviated *S. 6803*.



Supplementary Fig. 8 Sequence alignment of Psb27 subunits. Residues with green font in the *T. elongatus* sequence exhibit sequence similarity $\geq 50\%$ between sequences which also corresponds to the labeled residues in **Fig. 2**. For reference, residues within each of the four α -helices in the four-helix bundle of Psb27 from the *T. elongatus* PSII structure (PDB 7NHP) and the *A. thaliana* Psb27 structure (PDB 5X56) are highlight in yellow, cyan, green, and magenta for $\alpha 1$, $\alpha 2$, $\alpha 3$, and $\alpha 4$. Residues highlighted in red represent cleavage sites of signaling sequences identified by the SigP server (Almagro Armenteros et al. 2019). Residues highlighted in blue may be bacterial peptidase cleavage sites. Residues not modeled in the corresponding molecular structures (or homology model) are highlighted in grey. The alignment was created using Clustal Omega (Sievers et al. 2011) and the sequence identifiers are shown below each section of the alignment. *Synechocystis* 6803 is abbreviated *S. 6803*. Note that a multiple sequence alignment was also performed without the the signal sequences, but the outcome was identical.



Supplementary Fig. 9 Calculated PIPER binding energies for Psb27. Each panel shows an image of the docking simulation of Psb27 (colored) onto the PSII core (grey). The green ribbons correspond to the experimentally determined Psb27-PSII structure from *T. vulcanus* (PDB 7NHP). The cyan, blue, magenta, and yellow ribbons correspond to the top model for balanced, electrostatic-favored, hydrophobic-favored, and van der Waals + electrostatic forces (Kozakov et al. 2017). The table below each image shows whether the top model is consistent with the experimental model and the corresponding calculated PIPER binding energy for each coefficient set, with the lowest highlighted in light green. For *C. reinhardtii* and *S. oleracea*, the docking simulations were performed with homology models, and are therefore labeled with an “(h)”. For all tables in this figure, the “Consistent?” column refers to whether the simulated Psb27 orientation is consistent with the experimental structures of PSII with Psb27 bound (Huang et al. 2021; Zabret et al. 2021). The energy reported in this table is PIPER energy calculated by ClusPro and do not correspond to valid binding free energy, but can be used for general comparison of favorable binding interactions in the docking algorithm (Kozakov et al. 2017). *Synechocystis* 6803 is abbreviated *S. 6803*.



Supplementary Fig. 10 Lack of electrostatic potential corresponding to a Cl⁻ ion bound to Psb27 in the dimeric Psb27-PSII complex. The Psb27-PSII model (PDB 7CZL) in the corresponding map (EMD-30511) is shown at three contour levels. Even at very low contour, no electrostatic potential corresponding to the modeled Cl⁻ ion modeled with Psb27 is present, suggesting that it is incorrectly modeled.

Supplementary Tables

Supplementary Table 1 Sequence identity matrix of selected PsbQ(′) and Psb27 sequences calculated using Clustal Omega (Sievers et al. 2011). Red font shows the calculated sequence identities from the alignment of Psb27 sequences compared to PsbQ(′) sequences from different organisms. Black font shows the calculated sequence identities from the alignment of Psb27 sequences from different organisms (top left) or PsbQ(′) sequences (bottom right). *Synechocystis* 6803 is abbreviated *S.* 6803.

	S. 6803 Psb27	<i>T. elongatus</i> Psb27	<i>T. vulcanus</i> Psb27	<i>C. reinhardtii</i> Psb27	<i>S. oleracea</i> Psb27	<i>A. thaliana</i> Psb27	<i>C. gracilis</i> PsbQ′	S. 6803 PsbQ	<i>T. elongatus</i> PsbQ	<i>T. vulcanus</i> PsbQ	<i>C. caldarium</i> PsbQ′	<i>C. reinhardtii</i> PsbQ	<i>P. sativum</i> PsbQ	<i>S. oleracea</i> PsbQ	<i>A. thaliana</i> PsbQ
S. 6803 Psb27	100	50.38	50.38	29.85	28.46	29.23	11.97	12.87	10.89	10.89	9.92	17.5	12.61	11.67	11.76
<i>T. elongatus</i> Psb27	50.38	100	100	28.03	29.69	32.03	9.4	15	11	11	10.74	19.17	16.81	15	15.97
<i>T. vulcanus</i> Psb27	50.38	100	100	28.03	29.69	32.03	9.4	15	11	11	10.74	19.17	16.81	15	15.97
<i>C. reinhardtii</i> Psb27	29.85	28.03	28.03	100	35.67	38.95	20.67	17.92	15.74	15.74	17.2	18.75	17.72	15.82	16.67
<i>S. oleracea</i> Psb27	28.46	29.69	29.69	35.67	100	56.14	12.59	8.91	9.8	9.8	13.33	13.67	11.11	14.47	12.67
<i>A. thaliana</i> Psb27	29.23	32.03	32.03	38.95	56.14	100	13.19	9.62	10.48	10.48	13.91	16.43	11.04	10.46	13.91
<i>C. gracilis</i> PsbQ′	11.97	9.4	9.4	20.67	12.59	13.19	100	18.25	12.14	12.14	19.47	20.47	17.42	18.44	19.77
S. 6803 PsbQ	12.87	15	15	17.92	8.91	9.62	18.25	100	31.08	31.08	17.36	19.86	19.01	18.31	19.01
<i>T. elongatus</i> PsbQ	10.89	11	11	15.74	9.8	10.48	12.14	31.08	100	100	17.01	16.78	15.17	17.24	16.55
<i>T. vulcanus</i> PsbQ	10.89	11	11	15.74	9.8	10.48	12.14	31.08	100	100	17.01	16.78	15.17	17.24	16.55
<i>C. caldarium</i> PsbQ′	9.92	10.74	10.74	17.2	13.33	13.91	19.47	17.36	17.01	17.01	100	24.73	21.89	22.77	27.5
<i>C. reinhardtii</i> PsbQ	17.5	19.17	19.17	18.75	13.67	16.43	20.47	19.86	16.78	16.78	24.73	100	28.06	27.92	29.23
<i>P. sativum</i> PsbQ	12.61	16.81	16.81	17.72	11.11	11.04	17.42	19.01	15.17	15.17	21.89	28.06	100	66.67	64.35
<i>S. oleracea</i> PsbQ	11.67	15	15	15.82	14.47	10.46	18.44	18.31	17.24	17.24	22.77	27.92	66.67	100	71.3
<i>A. thaliana</i> PsbQ	11.76	15.97	15.97	16.67	12.67	13.91	19.77	19.01	16.55	16.55	27.5	29.23	64.35	71.3	100

Supplementary Table 2 Sequence identity matrix of selected PsbQ(′) and Psb27 sequences calculated using structure-based alignments from PROMALS3D (Pei et al. 2008). Red font shows the calculated sequence identities from the alignment of Psb27 sequences compared to PsbQ(′) sequences from different organisms. Black font shows the calculated sequence identities from the alignment of Psb27 sequences from different organisms (top left) or PsbQ(′) sequences (bottom right). *Synechocystis* 6803 is abbreviated *S.* 6803.

	S. 6803 Psb27	<i>T. elongatus</i> Psb27	<i>T. vulcanus</i> Psb27	<i>C. reinhardtii</i> Psb27	<i>S. oleracea</i> Psb27	<i>A. thaliana</i> Psb27	<i>C. gracilis</i> PsbQ′	S. 6803 PsbQ	<i>T. elongatus</i> PsbQ	<i>T. vulcanus</i> PsbQ	<i>C. caldarium</i> PsbQ′	<i>C. reinhardtii</i> PsbQ	<i>P. sativum</i> PsbQ	<i>S. oleracea</i> PsbQ	<i>A. thaliana</i> PsbQ
S. 6803 Psb27	100	44.44	45.3	25.44	23.68	26.32	2.44	7.75	5.43	5.43	6.84	3.77	6.72	5.19	3.27
<i>T. elongatus</i> Psb27	44.44	100	96.46	29.73	32.73	35.14	3.70	6.25	5.51	5.51	6.84	3.16	7.3	6.41	4.52
<i>T. vulcanus</i> Psb27	45.3	96.46	100	28.70	31.58	33.91	3.61	6.06	5.34	5.32	6.61	3.09	7.3	6.41	4.52
<i>C. reinhardtii</i> Psb27	25.44	29.73	28.70	100	46.15	50.49	7.59	7.26	4.1	4.1	6.31	5.84	8.15	5.19	5.88
<i>S. oleracea</i> Psb27	6.45	32.73	31.58	46.15	100	73.08	6.33	6.45	2.46	2.46	9.91	7.79	8.15	5.84	5.23
<i>A. thaliana</i> Psb27	26.32	35.14	33.91	50.49	73.08	100	5.7	3.23	3.28	3.28	8.11	5.19	6.67	5.19	5.23
<i>C. gracilis</i> PsbQ′	2.44	3.70	3.61	7.59	6.33	5.7	100	16.03	8.75	8.75	16.23	12.85	10.06	11.18	13.61
S. 6803 PsbQ	7.75	6.25	6.06	7.26	6.45	3.23	16.03	100	24.22	24.22	16.81	16.11	16.30	13.91	14
<i>T. elongatus</i> PsbQ	5.43	5.51	5.34	4.1	2.46	3.28	8.75	24.22	100	100	12.5	14.19	12.5	11.84	12.58
<i>T. vulcanus</i> PsbQ	5.43	5.51	5.32	4.1	2.46	3.28	8.75	24.22	100	100	12.5	14.19	12.5	11.84	12.58
<i>C. caldarium</i> PsbQ′	6.84	6.84	6.61	6.31	9.91	8.11	16.23	16.81	12.5	12.5	100	22.30	21.71	15.44	20.95
<i>C. reinhardtii</i> PsbQ	3.77	3.16	3.09	5.84	7.79	5.19	12.85	16.11	14.19	14.19	22.30	100	20.12	17.82	21.39
<i>P. sativum</i> PsbQ	6.72	7.3	7.3	8.15	8.15	6.67	10.06	16.30	12.5	12.5	21.71	20.12	100	58.76	58.39
<i>S. oleracea</i> PsbQ	5.19	6.41	6.41	5.19	5.84	5.19	11.18	13.91	11.84	11.84	15.44	17.82	58.76	100	72.97
<i>A. thaliana</i> PsbQ	3.27	4.52	4.52	5.88	5.23	5.23	13.61	14	12.58	12.58	20.95	21.39	58.39	72.97	100

Supplementary Table 3 Sequence identity matrix for the PsbQ(′) sequence alignment. *Synechocystis* 6803 is abbreviated S. 6803.

	S. 6803 PsbQ	<i>T. elongatus</i> PsbQ	<i>T. vulcanus</i> PsbQ	<i>C. caldarium</i> PsbQ′	<i>C. gracilis</i> PsbQ′	<i>C. reinhardtii</i> PsbQ	<i>P. sativum</i> PsbQ	<i>S. oleracea</i> PsbQ	<i>A. thaliana</i> PsbQ
S. 6803 PsbQ	100.00	31.08	31.08	17.36	18.25	19.86	19.01	18.31	19.01
<i>T. elongatus</i> PsbQ	31.08	100.00	100.00	17.01	12.14	16.78	15.17	17.24	16.55
<i>T. vulcanus</i> PsbQ	31.08	100.00	100.00	17.01	12.14	16.78	15.17	17.24	16.55
<i>C. caldarium</i> PsbQ′	17.36	17.01	17.01	100.00	19.47	24.73	21.89	22.77	27.50
<i>C. gracilis</i> PsbQ′	18.25	12.14	12.14	19.47	100.00	20.47	17.42	18.44	19.77
<i>C. reinhardtii</i> PsbQ	19.86	16.78	16.78	24.73	20.47	100.00	28.06	27.92	29.23
<i>P. sativum</i> PsbQ	19.01	15.17	15.17	21.89	17.42	28.06	100.00	66.67	64.35
<i>S. oleracea</i> PsbQ	18.31	17.24	17.24	22.77	18.44	27.92	66.67	100.00	71.30
<i>A. thaliana</i> PsbQ	19.01	16.55	16.55	27.50	19.77	29.23	64.35	71.30	100.00

Supplementary Table 4 Number of amino acids missing from molecular structures of PsbQ(′). *Synechocystis* 6803 is abbreviated S. 6803.

Type	Organism	Protein Data Bank Code	N-term.	C-term.
Cyanobacteria	S. 6803	7N8O	28	2
	<i>T. elongatus/vulcanus</i> *	3ZSU	11	1
Algae	<i>C. reinhardtii</i>	6KAC	51	0
	<i>C. caldarium</i>	4YUU	102	1
	<i>C. gracilis</i>	6JLU	64	0
Plants	<i>P. sativum</i>	5XNL	19	0
	<i>S. oleracea</i>	3JCU	83	1
	<i>A. thaliana</i>	5MDX (homology)	82	1

*The sequences for PsbQ from *T. elongatus* and *T. vulcanus* are identical and are therefore considered to have the same molecular structure.

Supplementary Table 5 Sequence identity matrix for Psb27 homolog alignment. *Synechocystis* 6803 is abbreviated S. 6803.

	S. 6803 Psb27	<i>T. elongatus</i> Psb27	<i>T. vulcanus</i> Psb27	<i>C. reinhardtii</i> Psb27	<i>S. oleracea</i> Psb27	<i>A. thaliana</i> Psb27
S. 6803 Psb27	100.00	50.38	50.38	29.85	28.46	29.23
<i>T. elongatus</i> Psb27	50.38	100.00	100.00	28.03	29.69	32.03
<i>T. vulcanus</i> Psb27	50.38	100.00	100.00	28.03	29.69	32.03
<i>C. reinhardtii</i> Psb27	29.85	28.03	28.03	100.00	35.67	38.95
<i>S. oleracea</i> Psb27	28.46	29.69	29.69	35.67	100.00	56.14
<i>A. thaliana</i> Psb27	29.23	32.03	32.03	38.95	56.14	100.00

Supplementary Bibliography

- Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, et al (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* 37:420–423. doi: 10.1038/s41587-019-0036-z
- Huang G, Xiao Y, Pi X, et al (2021) Structural insights into a dimeric Psb27-photosystem II complex from a cyanobacterium *Thermosynechococcus vulcanus*. *Proc Natl Acad Sci*. doi: 10.1073/pnas.2018053118
- Kozakov D, Hall DR, Xia B, et al (2017) The ClusPro web server for protein–protein docking. *Nat Protoc* 12:255–278. doi: 10.1038/nprot.2016.169
- Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol* msw054. doi: 10.1093/molbev/msw054
- Pei J, Kim B-HH, Grishin N V. (2008) PROMALS3D: A tool for multiple protein sequence and structure alignments. *Nucleic Acids Res* 36:2295–2300. doi: 10.1093/nar/gkn072
- Sievers F, Wilm A, Dineen D, et al (2011) Fast, scalable generation of high quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:1–6.
- Zabret J, Bohn S, Schuller SK, et al (2021) Structural insights into photosystem II assembly. *Nat Plants* 7:524–538. doi: 10.1038/s41477-021-00895-0