Supporting Information

# Crystal structures of cyanobacterial light-dependent protochlorophyllide oxidoreductase

Chen-Song Dong, Wei-Lun Zhang, Qiao Wang, Yu-Shuai Li, Xiao Wang, Min Zhang, and Lin Liu

This Supporting Information includes:





www.pnas.org/cgi/doi/10.1073/pnas.1920244117

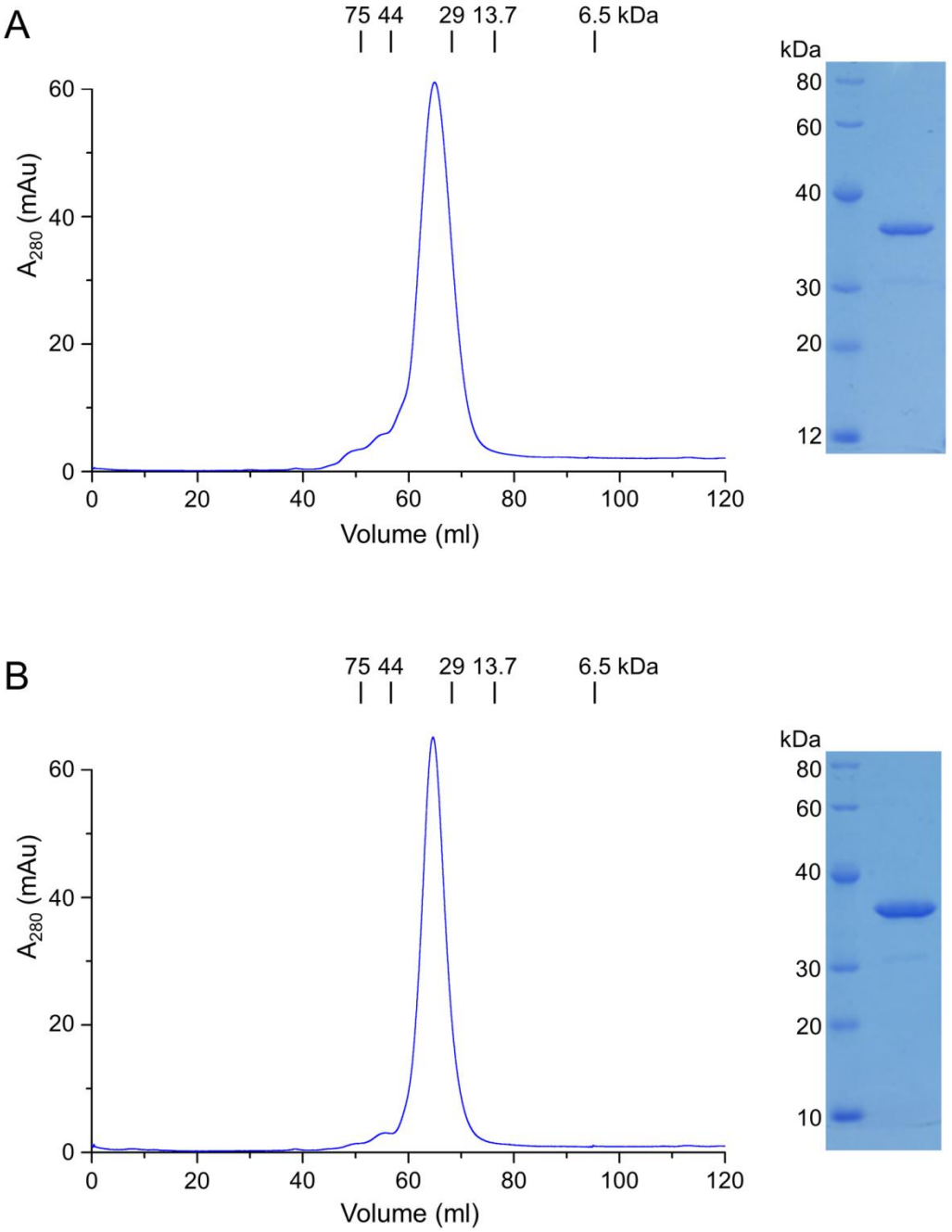


**Fig. S1.** Purification of *Sy*LPOR (A) and *Te*LPOR (B). The SEC profiles on a HiLoad Superdex 75 pg column (GE Healthcare) were recorded as absorbance at 280 nm. The flow rate was 1 ml $min^{-1}$. The molecular weight of protein standards (GE Healthcare) was labeled. The purity of the purified protein (peak fraction) was shown by the SDS-PAGE.

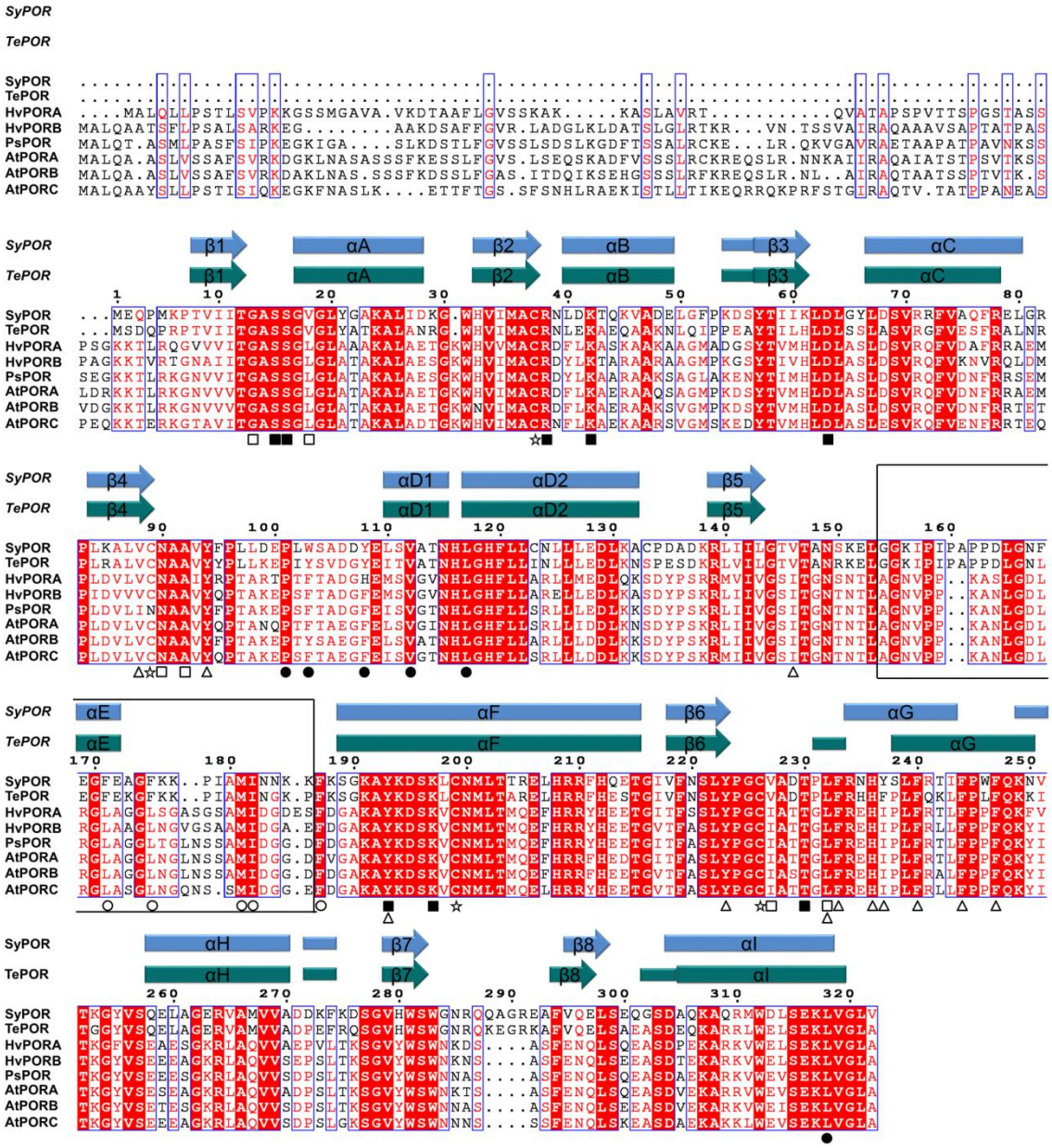


**Fig. S2.** Multiple LPOR sequences alignment. Identical residues are in white on a *red* background; similar residues are in *red* and boxed. Dots indicate gaps introduced during alignment. The secondary structure is calculated by DSSP, and labeled as cylinders for α-helices and arrows for β-strands. The stars denote the four conserved cysteine residues; the rectangles denote residues that interact with NADPH via the side-chain atoms (*black*) or the backbone atoms (*white*); the triangles denote residues that possibly participate in Pchlide binding; the circles denote residues that stabilize the LPOR-specific insertion by hydrophobic interactions, with the insertion residues in *white* and the core residues in *black*. The 33-residue LPOR-specific insertion is boxed in *black* rectangle. The abbreviations are: *Sy*, *Synechocystis* sp. PCC 6803; *Te*, *Thermosynechococcus elongates*; *At*, *Arabidopsis thaliana*; *Ps*, *Pisum sativum* L.; *Hv*, *Hordeum vulgare*. The alignment is generated with Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) and colored with ESPrispt 3.0 (http://espript.ibcp.fr/ESPript/ESPript/index.php).

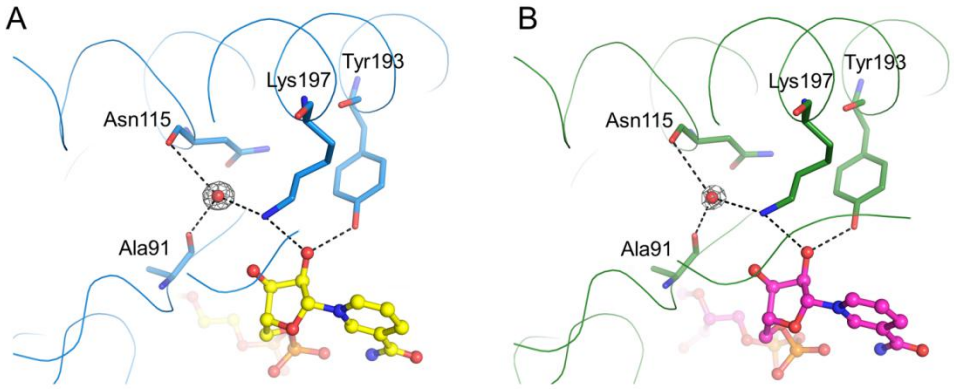


**Fig. S3.** The proton path within *Sy*LPOR (A) and *Te*LPOR (B). The $2F_o-F_c$ electron density map of the water molecule (*red* sphere) bound to Ala91, Asn115, and Lys197 was contoured at 1.0 σ level.

**Table S1.** Constructs information.

| Source organism | *Synechocystis* sp. PCC 6803 | *Thermosynechococcus elongatus* |
|---|---|---|
| DNA source | Synthesized DNA | Synthesized DNA |
| Forward primer* | GGAATTC*CATATG*GAACAGCC GATGAAACC | GGAATTC*CATATG*AGTGACCA GCCGCGCCCGA |
| Reverse primer | GAT*CTCGAG*TTAAACCAGACC AACCAGTTTTTCGC | GAT*CTCGAG*CGCCAGGCCAA CCAGTTTTTC |
| Cloning vector | pUC57 | pUC57 |
| Expression vector | pET-28a(+) | pET-22b(+) |
| Expression host | *Escherichia coli* BL21(DE3) | *Escherichia coli* BL21(DE3) |
| Complete amino acid sequence of the construct produced† | MGSSHHHHHHSSGLVPRGSH MEQPMKPTVIITGASSGVGLY GAKALIDKGWHVIMACRNLD KTQKVADELGFPKDSYTIIKLD LGYLDSVRRFVAQFRELGRPL KALVCNAAVYFPLLDEPLWSA DDYELSVATNHLGHFLLCNLL LEDLKACPDADKRLIILGTVTA NSKELGGKIPIPAPPDLGNFEGF EAGFKKPIAMINNKKFKSGKA YKDSKLCNMLTTRELHRRFHQ ETGIVFNSLYPGCVADTPLFRN HYSLFRTIFPWFQKNVTKGYV SQELAGERVAMVVADDKFKD SGVHWSWGNRQQAGREAFVQ ELSEQGSDAQKAQRMWDLSE KLVGLV | MSDQPRPTVIITGASSGVGLYA TKALANRGWHVIMACRNLEK AEQAAKNLQIPPEAYTILHLDL SSLASVRGFVESFRALNRPLRA LVCNAAVYYPLLKEPIYSVDG YEITVATNHLGHFLLINLLLED LKNSPESDKRLVILGTVTANRK ELGGKIPIPAPPDLGNLEGFEK GFKKPIAMINGKPFKSGKAYK DSKLCNMLTARELHRRFHEST GIVFNSLYPGCVADTPLFRHHF PLFQKLFPLFQKKITGGYVSQE LAGERVAMVVADPEFRQSGV HWSWGNRQKEGRKAFVQELS AEASDEQKARRLWELSEKLVG LALEHHHHHH |

*The restriction site is italicized and underlined.

†Extra amino acids generated during construction are underlined.

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

| | *Sy*LPOR−NADPH | *Te*LPOR−NADPH |
|---|---|---|
| Diffraction Data | | |
| Diffraction source | BL18U1, SSRF | BL17U1, SSRF |
| Detector | Pilatus3 6M | Eiger X 16M |
| Wavelength (Å) | 0.979 | 0.979 |
| Unit-cell parameters | | |
| *a*, *b*, *c* (Å) | 156.0, 57.1, 72.9 | 52.9, 96.2, 141.7 |
| $\alpha$, $\beta$, $\gamma$ (°) | 90, 90, 90 | 90, 90, 90 |
| Space group | P 21 21 2 | P 21 21 21 |
| Resolution (Å) | 50-2.20 (2.28-2.20)* | 50-2.40 (2.49-2.40) |
| Total No. of reflections | 163488 (16060) | 305028 (22810) |
| No. of unique reflections | 33980 (3338) | 27727 (2401) |
| Average redundancy | 4.8 (4.9) | 11.0 (9.5) |
| Mean $I$ / σ$I$ | 15.2 (2.2) | 18.1 (4.5) |
| Completeness (%) | 99.9 (99.9) | 96.9 (85.9) |
| Wilson B-factor (Å$^2$) | 37.26 | 43.01 |
| $R_{merge}$ | 0.107 (0.874) | 0.112 (0.547) |
| $R_{pim}$ | 0.052 (0.428) | 0.036 (0.183) |
| $CC_{1/2}$ | 0.990 (0.686) | 0.988 (0.913) |
| Refinement | | |
| Resolution range (Å) | 46.09-2.20 (2.26-2.20) | 49.57-2.40 (2.51-2.40) |
| No. of reflections used in refinement | 33794 (3279) | 27674 (2275) |
| No. of reflections used for $R_{free}$ | 1710 (173) | 1329 (126) |
| $R_{work}$ | 0.179 (0.256) | 0.181 (0.244) |
| $R_{free}$ | 0.207 (0.289) | 0.221 (0.292) |
| No. of non-hydrogen atoms | 4947 | 5012 |
| protein | 4620 | 4730 |
| ligand | 116 | 96 |
| water | 211 | 186 |
| Average *B* factor (Å$^2$) | 44.00 | 39.69 |
| protein | 43.81 | 39.78 |
| ligand | 39.68 | 33.04 |
| water | 50.58 | 40.86 |
| Model quality | | |
| RMS bond lengths (Å) | 0.004 | 0.004 |
| RMS bond angles (°) | 0.73 | 0.79 |
| Ramachandran plot | | |
| favored (%) | 98.10 | 97.99 |
| allowed (%) | 1.90 | 2.01 |
| PDB code | 6L1G | 6L1H |

*Values in parentheses are for highest resolution shell.

**Table S3.** NADPH environment in *Sy*LPOR and *Te*LPOR.

| Bonding atom | NADPH atom | | Distance (Å) in chain A/B | |
|---|---|---|---|---|
| | PDB V3 (IUPAC) | PDB V2 | *Sy*LPOR | *Te*LPOR |
| Water1 O | O7N | NO7 | 2.63/2.61 | 3.01/2.41 |
| Water2 O | O2N | NO2 | 2.44/2.44 | 2.61/2.50 |
| | O5B | AO5* | 2.85/3.01 | 2.94/2.78 |
| Water3 O | O1A | AO1 | 2.84/3.30 | 2.83/2.52 |
| Water4 O | O1A | AO1 | 2.92/3.58 | – |
| | O3X | AOP3 | 3.11/3.02 | – |
| Water5 O | O3X | AOP3 | 3.06/3.99 | 2.72/2.73 |
| Water6 O | O2X | AOP2 | 2.61/2.54 | 2.69/2.72 |
| Water7 O | O2X | AOP2 | 2.78/2.81 | 2.47/2.69 |
| Water8 O | O2B | AO2* | 2.99/2.95 | 3.19/– |
| | N3A | AN3 | 2.92/2.82 | 2.76/– |
| Water9 O | N6A | AN6 | 2.95/2.99 | 2.79/3.15 |
| Gly13 backbone C=O | O3B | AO3* | 3.21/3.23 | 3.05/2.96 |
| Ser15 γO | O3B | AO3* | 3.08/3.05 | 3.08/2.82 |
| Ser15 γO | O1X | AOP1 | 2.52/2.92 | – |
| Ser16 γO | O2A | AO2 | 2.89/2.96 | 2.85/2.75 |
| Val18 backbone N | O2N | NO2 | 2.93/2.81 | 2.89/2.84 |
| Arg38 ηN | O3X | AOP3 | 3.00/2.62 | 2.71/2.75 |
| Arg38 εN | O2X | AOP2 | 2.80/4.07 | 2.90/4.12 |
| Lys42 ζN | O1X | AOP1 | 2.96/2.77 | 2.79/2.77 |
| Asp63 δO | N6A | AN6 | 3.43/3.29 | 3.22/3.10 |
| Asn90 backbone C=O | O3D | NO3* | 2.68/2.72 | 2.56/2.56 |
| Ala92 backbone N | O4B | AO4* | 3.43/3.29 | 3.15/3.07 |
| Tyr193 ηO | O2D | NO2* | 2.68/2.72 | 2.70/2.64 |
| Lys197 ζN | O2D | NO2* | 3.05/3.05 | 3.18/3.03 |
| Lys197 ζN | O3D | NO3* | 2.95/3.01 | 3.12/3.03 |
| Val227 backbone N | O7N | NO7 | 2.99/2.93 | 2.75/3.08 |
| Thr230 γO | O1N | NO1 | 2.64/2.52 | 2.74/2.77 |
| Leu232 backbone N | O3 | O3 | – | 3.13/3.15 |
| Arg234 $\eta N_1$ | O1X | AOP1 | – | 2.86/2.96 |
| Arg234 $\eta N_2$ | O1A | AO1 | – | 2.87/2.93 |