ELECTRONIC SUPPLEMENTARY INFORMATION

Elucidating Substrate Binding in the Light-Dependent Protochlorophyllide Oxidoreductase

Penelope Pesara[1], Katarzyna Szafran[2], Henry C. Nguyen[2], Abhishek Sirohiwal[3], Dimitrios A. Pantazis[1*], Michal Gabruk[2*]

[1] Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470 Mülheim an der Ruhr, Germany.

[2] Department of Plant Physiology and Biochemistry, Uniwersytet Jagielloński, Poland.

[3] Department of Biochemistry and Biophysics, Arrhenius Laboratory, Stockholm University, 10691 Stockholm, Sweden.

* Michal Gabruk (michal.gabruk@uj.edu.pl), Dimitrios A. Pantazis (dimitrios.pantazis@kofo.mpg.de)

Computational Details

System Preparation. The molecular mechanics (MM) model used in this work is based on the cryo-EM structure of LPOR (PDB ID: 7JK9).[1] We extracted a monomer from the large oligomeric structure which included the protein itself and three co-factors (Pchlide, NADPH and MGDG). The Pchlide is present in the mode B in the cryo-EM structure. Mode A was created using cryo-EM refinement with "flipped" orientation of the Pchlide and we further removed the high-energy structural contacts with a systematic MM energy minimization that allowed relaxation and molecular adaptations of the pigment binding pocket. In addition to the WT systems, we also created MM set-up for Y276F and Y177F mutant for the LPOR with Pchlide in mode B, in order to study the role of these point mutations on the ligand binding affinity (see binding energy results and discussion for more details). The protocol described hereafter applies to MM set-ups of LPOR with Pchlide oriented in both mode A and mode B, and Y276F and Y177F systems.

The water molecules are not resolved in the cryo-EM structure, therefore we hydrated the internal cavities using a thorough hydration protocol involving the three-dimensional reference interaction site model (3D-RISM) and integration of the Monte Carlo (MC) method with molecular dynamics (MC/MD).[2-4] Both techniques work in complement to each to other, i.e. 3D-RISM predicts hydration content on static structures which further acts as reference for a dynamic hydration/de-hydration protocol in the MC/MD. This combined strategy has already shown success in predicting the hydration content of membrane-bound Photosystem II.[5] The protonation states of protein residues are assigned from the *reduce* program of AmberTools18.[6] The missing tail of MGDG was completed using the *pymol* suite.[7] We added two Na^+ ions to neutralize the system. After that, the complete system was embedded in a rectangular water box and the minimum distance between the solute and box edge was set at 12 Å. The entire system contained a total of 72 270 atoms.

The electrostatic charges for the Pchlide, NADPH and LMG co-factors were based on the Merz–Kollman Restrained Electrostatic Potential (RESP) methodology.[8] In the first step, we fully optimized the cofactors at the B3LYP/def2-SVP level of theory.[9, 10] After that, single-point HF/6-31G* calculations were performed on the optimized structure and subsequently the RESP fitting of the charges was performed using Multiwfn.[11] The bonded parameters for the NADPH and MGDG co-factors were based on the GAFF2 (general Amber force field) force-field,[3] while we adapted the Chl *a* force-field for Pchlide parameterization.[12] The force-field for the protein was based on the Amber14SB,[13] and the TIP3P model,[14] was chosen to represent the water. Joung–Cheatham parameters for monovalent ions compatible with the TIP3P water model were employed.[15, 16]

Classical Molecular Dynamics Simulations. In the first step, all hydrogens atoms were energy-minimized for a total of 2000 steps and all non-hydrogens atoms were restrained to their positions with a force constant of 50 kcal mol^{-1} $Å^{-2}$. In the next step, all atoms of the system were optimized for a total of 2000 steps except the Pchlide, NADPH and MGDG co-factors and backbone atoms (CA, CB, C, O and N) which were restrained with a force constant of 50 kcal mol^{-1} $Å^{-2}$. This restraining force is maintained throughout the simulation protocol. During the equilibration phase, the system is slowly heated from 10 to 300 K within 50 ps, and further propagated for another 50 ps at 300 K in the *NVT* ensemble. The temperature during this procedure is controlled using Langevin dynamics[12] with a collision frequency of 5 ps^{-1}. In the next step, the system is propagated for another 1 nanosecond in the *NPT* ensemble. Thereafter, we evoked the MC/MD module implemented in AMBER18. The steric grid for the MC moves was chosen to precisely cover the entire protein volume. The number of Monte-Carlo (MC) move attempts in each MC cycle was set at 1,000,000, whereas the number of MD steps in each MC cycle was set at 1,000. After that, we initiated production level simulations for 200 nanoseconds using the *NPT* ensemble. During the *NPT* production simulations, the collision frequency of the Langevin thermostat is lowered to 1 ps^{-1}. The Berendsen barostat[17] was used to regulate the pressure isotropically with a relaxation time of 2 ps and maintained at 1 bar. The Particle Mesh Ewald (PME) approach[18] was used to treat electrostatic interactions with a 12 Å cutoff. In MD simulations a long-range dispersion correction based on an analytical integral assuming an isotropic, uniform bulk particle distribution beyond the cutoff is added to the van der Waals energy. The SHAKE algorithm[19] was used to constrained the bond length involving hydrogen atoms. Integration time-step of 1 fs was used throughout the MD. The energy minimizations were performed using the CPU version while equilibration simulations were performed using the GPU version of the *pmemd*[20-22] engine of the AMBER18 package. The production run for 200 ns was performed for WT system with pchlide in mode A and mode B. In case of Y276F and Y177F, we followed the similar protocol and propagated the system for 50 ns. For the QM/MM computations of WT system with pchlide in mode A and mode B, we extracted the last structural configuration from the first 10 ns of the production MD simulations.

QM/MM Geometry Optimization. All calculations were performed by employing the additive Quantum Mechanics/Molecular Mechanics (QM/MM) multiscale model with electrostatic embedding as implemented in ORCA 5.0.[23, 24] Structural analyses were performed for both binding modes A and B. For geometry optimization, density functional theory (DFT) computations were performed using the PBE functional.[25] Amber14SB and associated parameters were used as described above for the MM part. The size of our system prevents the use of large basis sets, therefore def2-SVP basis set was used for geometry optimizations. The QM region contains Pchlide, truncated NADPH, truncated MGDG, and complete amino acids (except backbone atoms) within 9 Å from the central Mg atom of Pchlide.

The active MM region (region that is allowed to be energy-minimized during QM/MM optimization) contains complete residues with any atom within 5 Å from the QM region. The rest of the MM region was kept fixed during QM/MM optimizations optimization. The regions were further manually adjusted to be chemically meaningful. The final size of the QM/MM system was 72270 atoms, where the QM subsystem contains 401 atoms including 22 link atoms. The QM and active MM regions are visualized in Fig. S1. Due to the high computational cost of DFT in large systems using large QM and active MM regions, the optimization was approached in a stepwise manner. First, NADPH was optimized at the QM level using a charge of -4 for the QM subsystem while the rest of the system was kept fixed. The resulting PDB was taken for the optimization of MGDG at the QM level. Finally, Pchlide and the side chains of important amino acid residues within ca. 6 Å to the central Mg atom, as well as waters within this radius, were optimized with an active MM region of 3 Å around the QM system including complete residues. This QM region contains full Pchlide and the amino acids: F327, F330, F316, T335, T230, A237, I229, V240, Y177, Y276, K277, P179, E318 and H319 as well as NADPH and MGDG. The resolution of the identity approximation was used with appropriate auxiliary basis sets (def2/J).[26]

Long-range electrostatic interactions are essential for modeling biological processes and in our work we made use of the electrostatic embedding scheme, where the QM electron density interacts explicitly with the MM charges. In the additive QM/MM multiscale approach, link atoms are created and the charges are corrected with a charge shifting scheme. Since the QM/MM calculation of such large systems as the LPOR, the calculation of nonbonded interactions can become a bottleneck, we made used force switching for the Lennard-Jones (LJ) interaction and force shifting for the electrostatic interaction. For the former, a smooth switching function is used with two cutoffs, the first at the distance at which the switching function is turned on (here starting at 10 Å), and the second at the distance where the quantity reaches 0 without introducing any discontinuities (here after 12 Å). For the electrostatics, one cutoff was used for the distance where the electrostatic potential is shifted to zero (here at 12 Å). We note that no cutoffs for electrostatic interactions were used in the case of single-point calculations reported in this work.

Binding Energy Calculations. The binding free energy corresponding to formation of the LPOR complex, involving Pchlide as the ligand and apo-LPOR as the protein, is determined using the MM-PBSA approach.[27-29] This combines Molecular Mechanics, implicit Generalized Born (GB)/Poisson-Boltzmann (PB) solvation schemes[30] and solvent accessibility surface area calculations to estimate the binding free energy according to equation 1:

$$\Delta G_{\text{bind}} = \Delta G_{\text{complex}} - (\Delta G_{\text{protein}} + \Delta G_{\text{ligand}}) \qquad (1)$$

Further theoretical details regarding the MM-PBSA approach can be found elsewhere.[28, 31] Here a single-trajectory approach is employed, where only the complex form is propagated, eliminating the need for separate molecular dynamics (MD) simulations for the ligand and protein. All calculations are performed using the parallelized MMPBSA.py module[32] of the Amber/21 package. Solvation free energies are computed using both the Generalized Born (GB) and Poisson-Boltzmann (PB) implicit solvation schemes. Binding energy calculations for the specific system were carried out on a set of equidistant snapshots (total 500), with a temporal interval of 100 ps, extracted from the initial 50 ns of the NPT simulations. The calculated binding energies were subsequently averaged over this ensemble of snapshots. In order to establish control, we conducted analogous computations over a period of 50-100 ns, 100-150 ns and 150-200 ns (each set consisting of 500 snapshots), yielding comparable outcomes (see Table S4). The solute dielectric constant is set to 2.0 in case of the MM/PBSA computations.[33] The entropic contribution (ΔS) to the binding free energy is excluded due to its high computational expense and slow convergence, as the focus is primarily only on the relative binding affinity trends. The calculations are conducted for both orientations of Pchlide discussed in the introduction (modes A and B) and also for Y276F and Y177F system with Pchlide bound in mode B orientation. Binding energy calculation for all other mutants (F330A, Y177A, V240A, T230A, Y276A, Q331A, T335A, F316A, F327A, C309A, Y276F and Y177F) were performed with the "Alanine scanning" functionality of *MMPBSA.py* module. To gain molecular level insights, binding free-energies were further decomposed into per-residue contributions.[34, 35]

Interaction Energy Calculations and Local Energy Decomposition Analysis.

Local Energy Decomposition (LED)[36-38] analysis at the coupled-cluster DLPNO-CCSD(T)[39-42] was used to analyze the binding in terms of wave function derived interaction energies. This methodology, relying on a local coupled-cluster method with singles, doubles, and perturbatively included triples excitations,[39, 41] is a generally applicable interaction decomposition approach. Within the LED analysis the system is divided into chemically meaningful fragments, where first the nuclei and subsequently the molecular orbitals (MOs) are assigned to the corresponding fragments. The DLPNO-CCSD(T) interaction energies between Pchlide and a list of surrounding amino acids as well as the lipid MGDG and cofactor NADPH were computed and decomposed according to the LED scheme. Doing so, it is possible to quantify key substrate–residue non-covalent interactions. A similar study has been done by Beck et al. on nicotine and imidacloprid binding to the nicotinic acetylcholine receptor.[38]

Within a supramolecular approach, the Enzyme–Substrate (ES) binding energy ΔE_{bind} can be computed at the DLPNO-CCSD(T) level as the difference between the energy of the ES adduct ($E_{\text{tot}}^{\text{SE}}$) and that of the isolated E ($E_{\text{tot}}^{\text{E}}$) and S ($E_{\text{tot}}^{\text{E}}$) fragments frozen in their in-adduct geometry:

$$\Delta E_{bind} = E_{tot}^{SE} - E_{tot}^{S} - E_{tot}^{E} \quad (2)$$

E_{tot}^{SE} can be decomposed into intra-fragment contributions (energy of the substrate E_{tot}^{E} and the energy of the residues E_{tot}^{R} in their adduct geometry) plus a series of inter-fragment contributions, substrate-residue (E_{tot}^{SR}) and residue-residue (E_{tot}^{RR}) interactions. Similarly E_{tot}^{E} can be decomposed into intra residue $E_{tot}^{E(R)}$ and inter residue-residue $E_{tot}^{E(R,R)}$ contributions. After insertion into (2) and rearrangement, the following equation is obtained describing the binding energy as demonstrated by Beck et al.[38]:

$$\Delta E_{bind} = \Delta E_{tot}^{S} + \sum_R \Delta E_{tot}^{R} + \sum_R \Delta E_{tot}^{RR} + \sum_R \Delta E_{tot}^{E(R)} \quad (3)$$

Where ΔE_{tot}^{S} and ΔE_{tot}^{R} show the change in energy of the substrate and the residues upon the possible binding modes, and $\Delta E_{tot}^{(R,R)}$ represents the change in residue-residue interaction in the active site.

LED calculations were carried out on the QM/MM models such that all defined fragments used within the LED study were QM-optimized at the PBE/def2-SVP level of theory. Single-point DLPNO-CCSD(T) calculations were performed under the electrostatic influence of the protein matrix using def2-TZVP main basis sets with the RIJCOSX appoximation and with matching def2/J and def2-TZVP/C auxiliary basis sets. Due to a relatively large QM region of 348 atoms (link atoms included), VeryTightSCF criterion was applied in combination with LoosePNO settings, which are sufficient for LED applications. The Pipek–Mezey method was used for orbital localization. The DLPNO-CCSD(T) energy was decomposed into a series of additive contributions corresponding to the interaction of pairs of the defined fragments.[36, 37] The fragmentation includes Pchlide as one fragment, as well as NADPH and the lipid MGDG, and a series of amino acid side chains within a distance of 6 Å to the central Mg atom representing other fragments.

Excited State Calculations and Electrostatic Potential Maps. Excited state calculations were computed in the framework of time-dependent density functional theory (TD-DFT). Full TD-DFT (without the Tamm–Dancoff approximation) calculations were performed using the hybrid functional B3LYP including 20% Hartree-Fock exchange energy, and range-separated functionals CAM-B3LYP[43] and ωB97X-D3BJ[44, 45] which are known to be often superior for spin states and spectroscopic properties. Basis set def2-TZVP and matching auxiliary basis def2/J for employing the RI-J approximation were used for all excited state calculations. Excited state geometries were optimized at the same level of theory. 10 roots were computed and Natural Transition Orbitals (NTOs) were created and visualized for transitions showing significant contributions for the first three excited states. Difference electron density maps were computed for the first three excited states associated with the $S_0 \rightarrow S_1$, $S_0 \rightarrow S_2$, and $S_0 \rightarrow S_3$ transition for all three functionals. The redistribution of electron density upon excitation was studied in gas phase and under the influence of the electrostatic environment of both structural models.

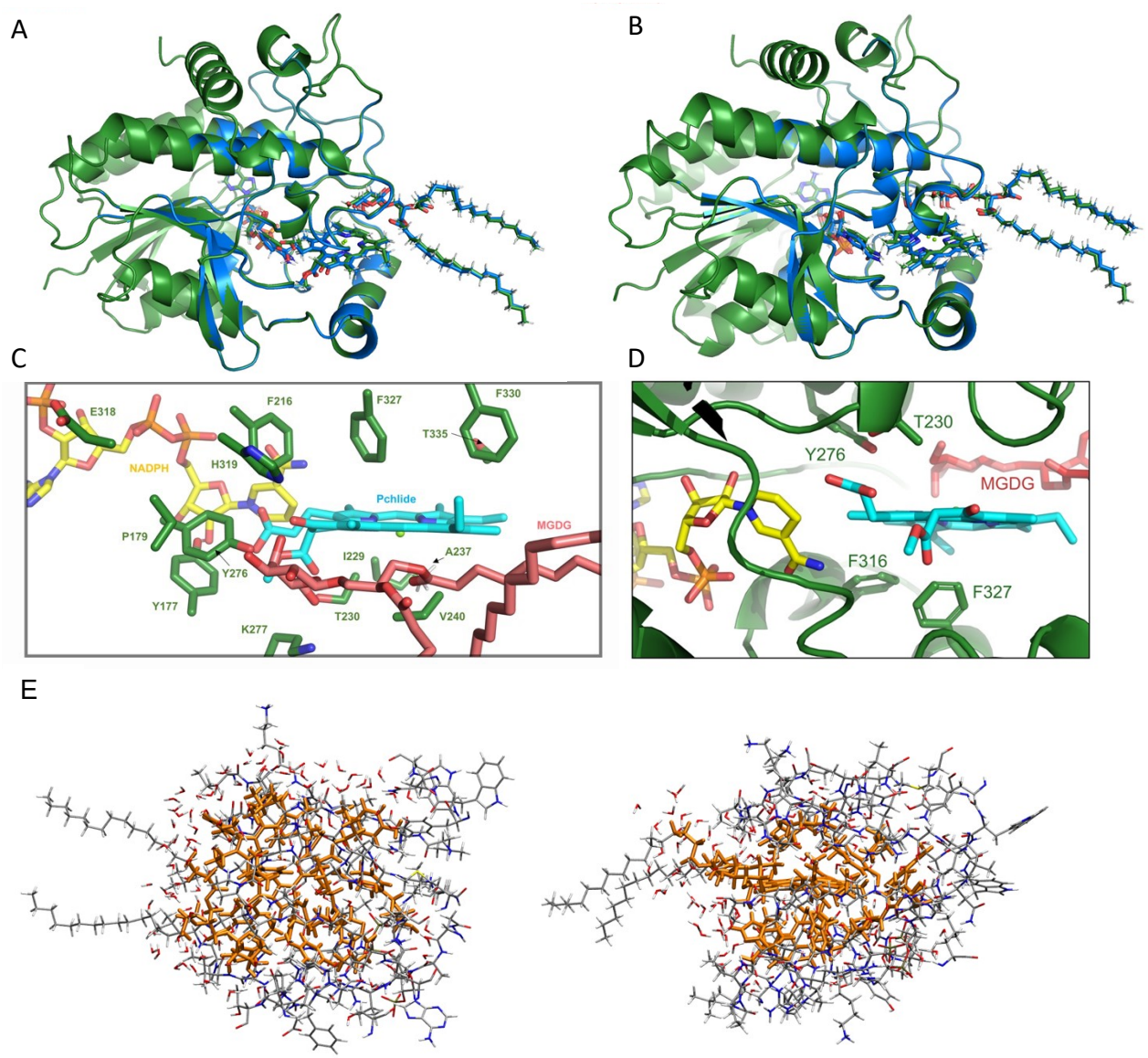

Figure S1. Structural models of LPOR created and used in this work. Green represents the structural configuration derived from the last MD snapshot and blue the QM/MM optimized structure for both Pchlide bonding modes A and B. **C.** Selected QM region in the QM/MM computations shown as sticks as in case for mode B. The same QM region configuration is used for the QM/MM calculations of the mode A. Shown are the QM/MM optimized geometries. D. Pchlide binding in mode A. Detailed view of the pocket highlighting residues with the highest predicted interaction energies. **E.** Active region of the QM/MM setup: the QM subsystem is shown in orange; it was constructed including residues within 9 Å to the central magnesium atom of Pchlide. The active MM region, colored by element, contains residues within a radius of 5 Å from any atoms of the QM region.

Table S1. Distances (Å) of selected residues from the central Mg^{2+} ion of Pchlide for both binding modes, obtained from the optimized QM/MM model.

	mode A	**mode B**
Y177	9.0	8.8
S228	8.9	8.7
T230	4.7	4.5
A237	6.0	4.6
V240	4.9	4.4
Y276	5.7	11.3
K277	6.2	8.2
Y306	11.2	8.1
C309	7.1	6.3
L315	9.8	10.6
F316	6.3	7.7
F323	9.1	7.3
F330	6.8	6.4
Q331	5.6	4.3
T335	7.9	6.3

Table S2. Distances (Å) of important residues to C17 and C18 of Pchlide after the QM/MM geometry optimization ("optimized") compared to the starting structure.

	Mode A Optimized		**Mode A Unoptimized**		**Mode B Optimized**		**Mode B Unoptimized**	
	C17	**C18**	**C17**	**C18**	**C17**	**C18**	**C17**	**C18**
NADPH	5.5	5.5	4.3	4.4	5.1	4.8	4.1	3.8
Y276	5.0	4.0	5.5	4.4	7.8	8.6	6.7	7.9
Y306	7.2	8.3	6.6	7.8	6.4	5.3	6.4	5.2
Y177	7.5	6.4	7.1	6.0	6.9	8.3	7.0	8.4
K277	7.9	7.3	8.3	7.3	7.5	8.3	7.1	8.3
C309	4.7	5.2	3.8	4.8	4.3	3.7	5.2	4.2
T230	4.3	4.0	4.8	5.0	3.7	3.6	3.6	4.0
T335	8.2	9.4	8.8	10.1	9.0	8.1	10.2	9.1
F316	5.4	6.3	5.5	4.8	5.2	6.2	5.0	5.6
F323	12.4	11.8	11.1	11.7	7.5	8.4	7.7	8.3
F327	5.2	5.1	4.9	5.1	4.7	5.0	5.2	5.1
F330	8.9	9.5	9.1	9.8	9.7	9.5	10.0	9.6
A237	5.6	6.6	6.5	7.3	6.4	5.5	6.8	6.3
V240	8.2	8.3	8.9	8.6	7.8	8.2	8.5	9.1
G331	5.5	6.3	5.6	6.6	5.8	5.3	6.3	5.6
L315	7.3	6.2	7.1	6.1	6.8	7.8	6.4	7.0
S228	5.4	5.2	6.0	5.8	5.3	5.3	4.4	4.8
MGDG	8.3	7.4	8.0	6.8	7.8	8.9	6.9	8.3

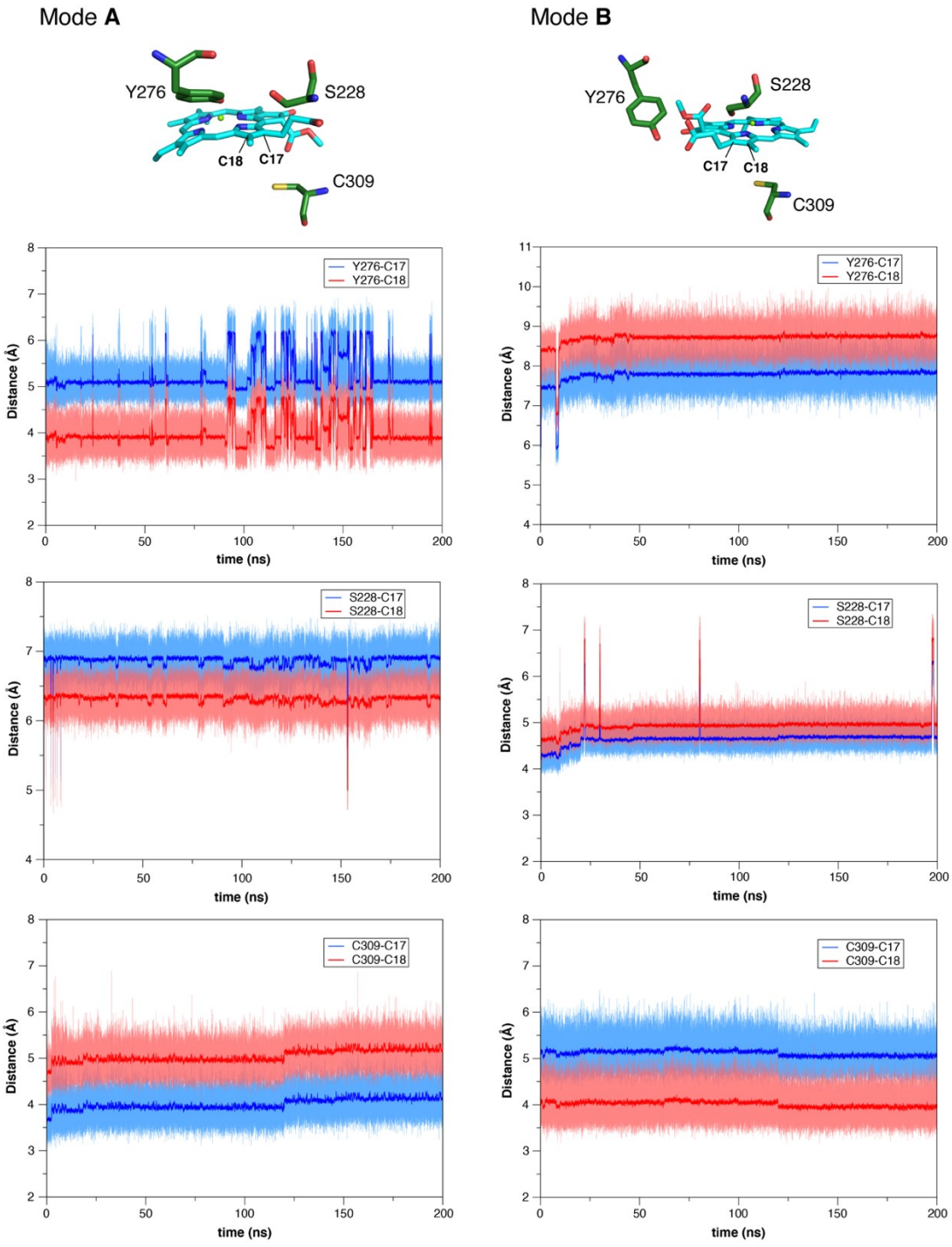

Figure S2. The distances between C17 and C18 atoms of Pchlide and selected residues for the two modes of Pchlide binding determined in the MM-PBSA analysis.

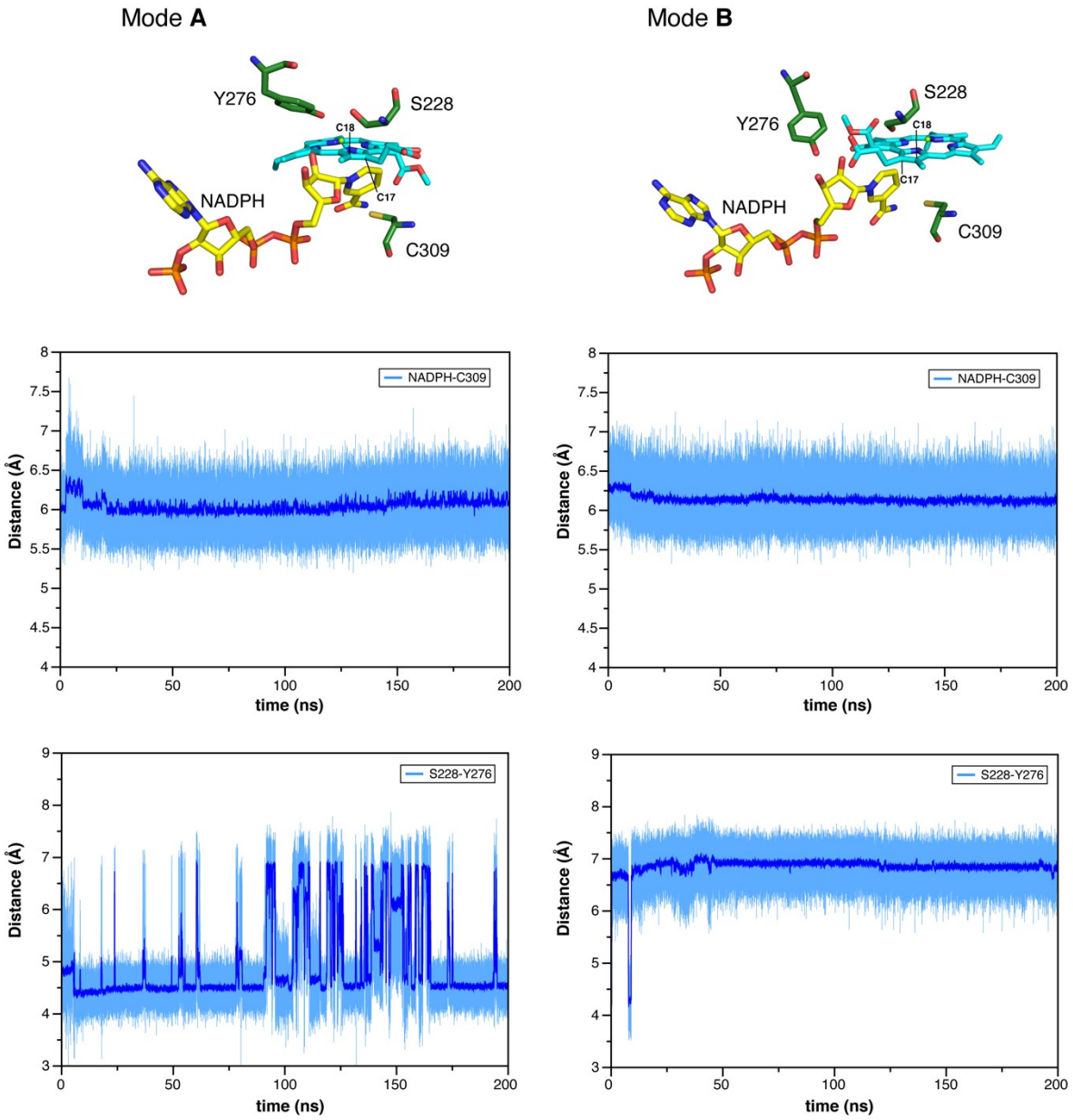

Figure S3. The distances between NADPH–C309 and S228–Y276 for the two modes of Pchlide binding determined in the MM-PBSA analysis.

Table S3. The contribution of selected residues to the Pchlide binding energy (in kcal/mol) as calculated with MM-PBSA and LED QM/MM approaches for binding modes A and B.

	MM-PBSA		LED-QM/MM	
Residue	**mode A**	**mode B**	**mode A**	**mode B**
Y177	-0.2	-4.3	-10.3	-76.6
P179	-0.1	-0.5	-0.6	-0.8
S228	-0.8	0.4	Nd	Nd
I229	-0.7	-0.5	-15.1	-16.2
T230	-1.7	-3.1	-46	-77.9
L236	-0.6	-0.7	Nd	Nd
A237	-2.0	-2.0	-37.1	-30.8
V240	-1.2	-1.5	-32.7	-50.8
A273	-0.2	-1.1	Nd	Nd
Y276	0.1	-0.1	-120	-28.5
K277	3.7	1.7	-13.4	-67.9
K280	0.3	1.4	Nd	Nd
Y306	-1.1	-0.1	Nd	Nd
C309	-1.6	-1.1	Nd	Nd
L315	-0.6	-0.9	Nd	Nd
F316	-1.9	-1.1	-57	-37.4
E318	-0.1	0.3	-15.8	-7.1
H319	0.1	-0.6	-1.8	2.5
F323	-0.5	-1.0	Nd	Nd
F327	-2.3	-2.1	-38.8	-81.7
F330	-0.9	-1.2	-26	-23.1
Q331	-0.5	-1.5	Nd	Nd
T335	-0.8	-1.6	-22.2	-36.1
G337	-0.5	-0.6	Nd	Nd
NADPH	-0.1	1.7	-46.4	-54.3
MGDG	-0.9	0.5	-212	-162.7

Table S4. Binding energy (in kcal/mol) of Pchlide computed using the MM-PBSA and MM-GBSA approaches for both binding orientations of the pigment. The binding energy computations were performed using the first 50 ns of the production trajectory. The results represent an ensemble over 500 equidistant snapshots. Energy decomposition of the binding energy (i.e., Complex - Receptor - Ligand) is also provided.

		vdW	**EEL**	**ΔG_{gas}**	**ΔG_{solv}**	**ΔG_{bind}**
PB	**Mode B (Cryo-EM)**	-67.0309 ±0.1876	-7.7728 ±0.0470	-74.8037 ±0.1978	65.0175 ±0.0742	-9.7862 ±0.2186
	Mode A	-55.0416 ±0.2271	-9.1269 ±0.0502	-64.1685 ±0.2339	62.4552 ±0.0816	-1.7134 ±0.2468
GB	**Mode B (Cryo-EM)**	-67.0309 ±0.1876	-15.5455 ±0.0939	-82.5764 ±0.2179	30.9793 ±0.0595	-51.5971 ±0.2095
	Mode A	-55.0416 ±0.2271	-18.2538 ±0.1004	-73.2954 ±0.2509	37.4506 ±0.0711	-35.8448 ±0.2642

Table S5. Binding energy (in kcal/mol) of Pchlide computed using the MM-PBSA and MM-GBSA approaches for both binding orientations of the pigment. The calculations were performed using snapshots from 0–5 ns, 5–10 ns, 0–50 ns, 50-100 ns, 100–150 ns and 150–200 ns part of the MD trajectories. In each case, the binding energy results represent an ensemble over 500 equidistant snapshots. A solute dielectric constant of 2 is used in MM-PBSA.

Approach	Trajectory	Mode B	Mode A
MM-PBSA	0–5 ns	-2.4644 ±0.1662	6.2906 ±0.1958
	5–10 ns	-2.8591 ±0.1651	6.3334 ±0.1768
	0-50 ns	-9.7862 ±0.2186	-1.7134 ±0.2468
	50-100 ns	-12.2016 ±0.1276	-4.9180 ±0.1396
	100-150 ns	-13.7939 ±0.1317	-6.0024 ±0.1397
	150-200 ns	-15.0446 ±0.1263	-5.9547 ±0.1304
MM-GBSA	0–5 ns	-45.3281 ±0.1757	-27.1794 ±0.2073
	5–10 ns	-44.8850 ±0.1734	-27.3387 ±0.1939
	0-50 ns	-51.5971 ±0.2095	-35.8448 ±0.2642
	50-100 ns	-53.7696 ±0.1321	-38.3830 ±0.1473
	100-150 ns	-54.9484 ±0.1317	-39.6731 ±0.1580
	150-200 ns	-55.6544 ±0.1285	-41.3698 ±0.1365

Table S6. Dependence of the solute dielectric constant on the MM-PBSA binding energy (i.e., Complex–Receptor–Ligand, in kcal/mol). All computations were performed for the snapshots derived from the 0–5 ns part of the full trajectory.

Dielectric constant	Trajectory	Mode B	Mode A
1	0-5 ns	21.5171 ±0.2009	27.0882 ±0.2343
2	0-5 ns	-2.4644 ±0.1662	6.2906 ±0.1958
4	0-5 ns	-14.3125 ±0.1578	-4.0543 ±0.1872

Table S7. Total free-energy decomposition into per-residue contributions for mode A and mode B. Only residues with a significant interaction to the substrate Pchlide in one of the two binding modes over/under a defined threshold (x>0.1; -0.1>x) are shown. All energies in kcal/mol.

Residue	Mode A	Mode B
Tyr 177	-0.55	-3.55
Pro 179	-0.45	-0.31
Ser 228	0.52	-0.21
Ile 229	0.49	-0.83
Thr 230	-2.20	-2.96
Gly 231	0.16	0.01
Asn 232	0.12	0.05
Thr 235	-0.09	-0.12
Leu 236	-0.25	-1.18
Ala 237	-1.20	-2.01
Val 240	-1.51	-1.53
Pro 242	-0.21	-0.27
Lys 243	0.18	0.20
Asp 271	-0.17	-0.27
Ala 273	-0.05	-0.60
Lys 274	0.12	0.06
Tyr 276	-0.94	-1.17
Lys 277	0.62	0.46
Asp 278	-0.11	-0.16
Lys 280	0.95	0.95
Tyr 306	-1.17	-0.08
Gly 308	-0.36	-0.30
Cys 309	-0.82	-1.27
Ile 310	-0.12	-0.10
Leu 315	-0.35	-0.91
Phe 316	-1.58	-0.91
His 319	-0.12	-0.09
Phe 323	-0.62	-0.82
Phe 326	-0.12	-0.12
Phe 327	-2.34	-2.14

Phe 330	-0.89	-1.11
Gln 331	-1.28	-1.83
Thr 335	-0.43	-1.60
Lys 336	0.16	0.16
Gly 337	-0.21	-0.44
Lys 338	-0.11	-0.05
Arg 347	0.12	0.11
Tyr 363	-0.20	-0.01
Asn 375	0.10	0.04
NDP 403	2.45	1.59
LMG 404	0.35	0.19

Table S8. List of primers used for site directed mutagenesis.

Primer	Sequence	Annealing temp. (°C)
F323A_F327A_For_t	GCCCGTGCCCTCGCCCCTCCCTTTCAGAAGTACATC	63
F323A_F327A_Rev_t	GGCGAGGGCACGGGCGAGAGGAATGTGCTCTCGG	
Q331 For	ATCACTAAAGGATATGTCTCCGAAAC	65
Q331 Rev	GAAGAGGGCACGGAAGAGAG	
Q331A_For_t	CCTCCCTTTGCGAAGTACATCACTAAAGGATATGTCTCCG AAAC	
Q331A_Rev_t	GTACTTCGCAAAGGGAGGGAAGAGGGCACGGAAGAGAG	
Q331E_For_t	CCTCCCTTTGAGAAGTACATCACTAAAGGATATGTCTCCG AAAC	
Q331E_Rev_t	GTACTTCTCAAAGGGAGGGAAGAGGGCACGGAAGAGAG	

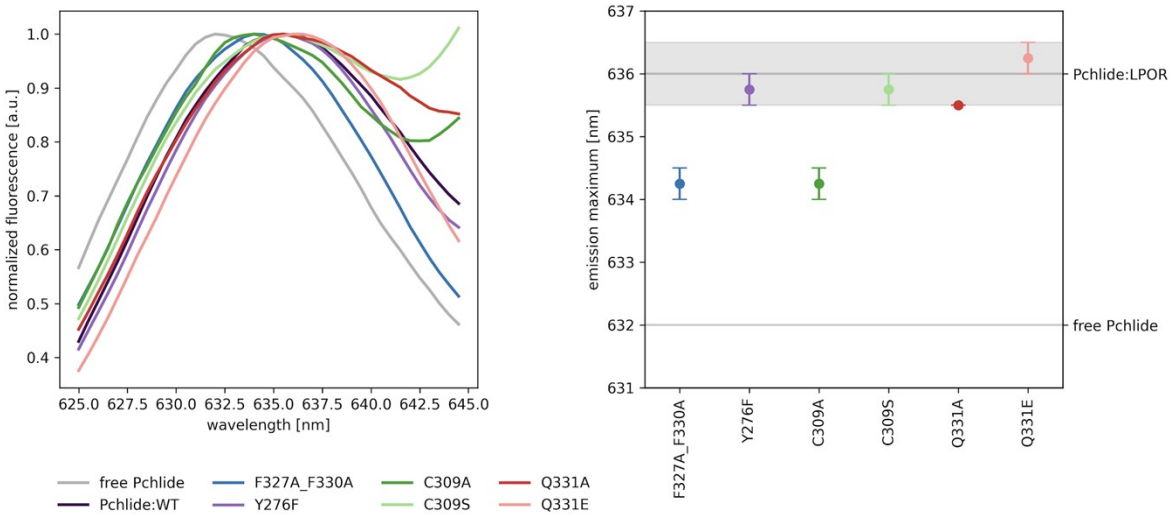

Figure S4. Non-photoactive Pchlide in the analyzed mutants and AtPORB WT. A. The part of the spectra showing non-photoactive Pchlide, normalized at maxima. **B.** The position of the emission maximum of non-photoactive Pchlide in analyzed mutants.

Table S9. Binding energy (in kcal/mol) of Pchlide computed using the MM-PBSA approach WT and several mutations with Pchlide oriented in the mode B. The binding energy computations were performed using the first 50 ns of the production trajectory. The results represent an ensemble over 500 equidistant snapshots. Energy decomposition of the binding energy (i.e., Complex - Receptor - Ligand) is also provided.

System	Binding Energy
WT	-9.7862 ±0.2186
F330A	-8.0077 ±0.2134
Y177A	-4.5880 ±0.1945
V240A	-9.5946 ±0.2181
T230A	-5.3178 ±0.2112
Y276A	-10.1220 ±0.2178
Q331A	-7.4814 ±0.2135
T335A	-8.2947 ±0.2143
F316A	-8.5271 ±0.2180
F327A	-8.0741 ±0.2144
C309A	-8.6033 ±0.2194
Y276F	-7.4859 ±0.1879
Y177F	-3.9382 ±0.1689

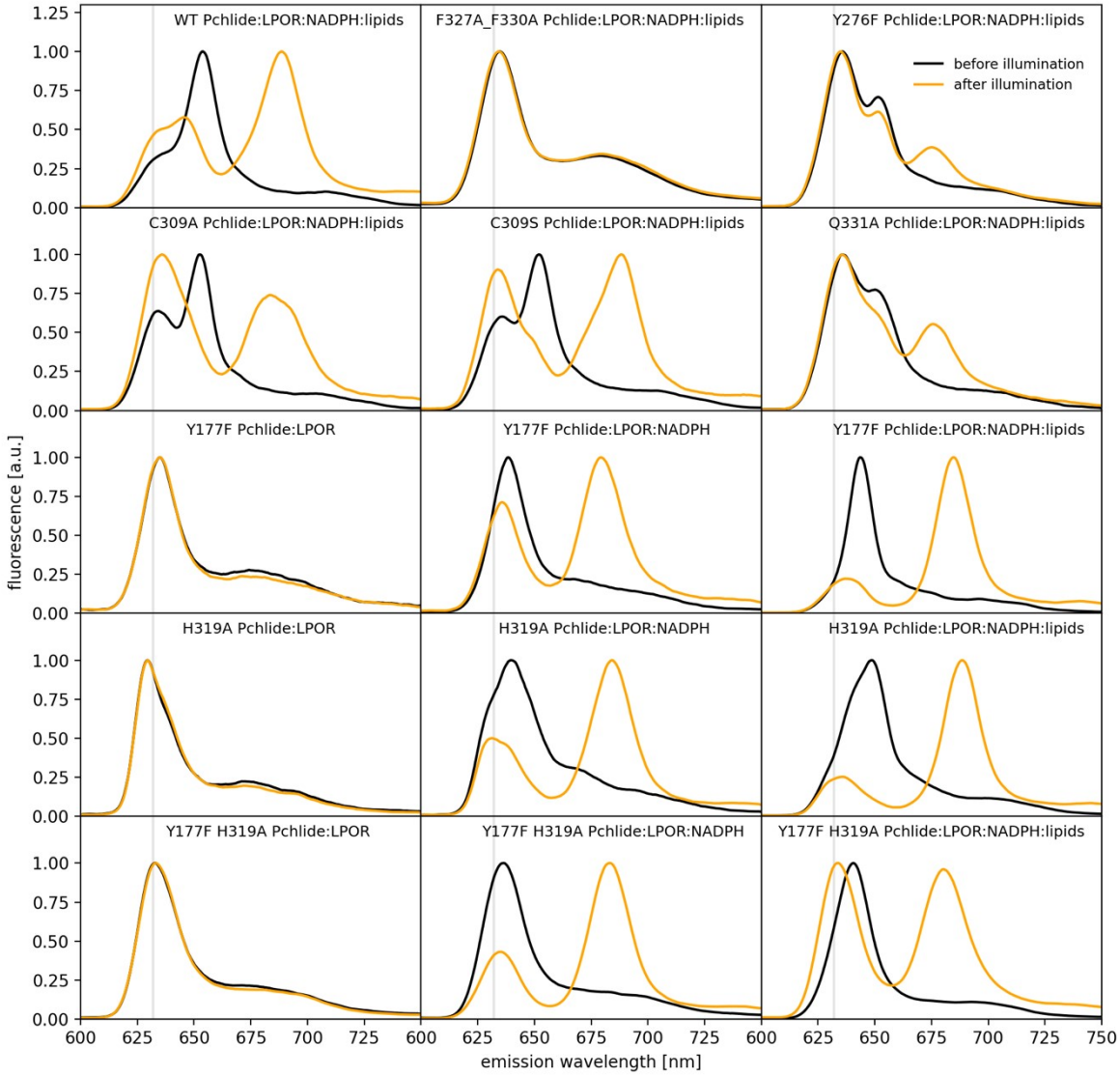

Figure S5. The effect of illumination of the fluorescence spectra of Pchlide in different reaction mixtures. For rows 1 and 2, the reaction was composed of: 15 µM LPOR, 5 µM Pchlide, 200 µM NADPH and 100 µM lipid mix (50mol% MGDG, 35mol% DGDG, 15mol% PG). For rows 3, 4, and 5, specific compositions of the reaction mixtures are provided for each panel; element concentrations mirror those in rows 1 and 2. The vertical grey line denotes the emission maximum of free Pchlide in the buffer.

Movie S1. The artistic visualization of the process of Pchlide binding in mode B based on pdb:9JK7.

References

1. H. C. Nguyen, A. A. Melo, J. Kruk, A. Frost and M. Gabruk, *Nat. Plants*, 2021, **7**, 437-444.
2. I. Y. Ben-Shalom, C. Lin, T. Kurtzman, R. C. Walker and M. K. Gilson, *J. Chem. Theory Comput.*, 2019, **15**, 2684-2691.
3. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.
4. V. Guallar and F. Wallrapp, *J. R. Soc. Interface*, 2008, **5**, 233-239.
5. A. Sirohiwal and D. A. Pantazis, *J. Am. Chem. Soc.*, 2022, **144**, 22035-22050.
6. D. A. Case, T. E. Cheatham III, T. Darden, H. Gohlke, R. Luo, K. M. Merz Jr., A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, *J. Comput. Chem.*, 2005, **26**, 1668-1688.
7. The PyMOL Molecular Graphics System, Version 2.0.
8. C. I. Bayly, P. Cieplak, W. Cornell and P. A. Kollman, *J. Phys. Chem.*, 2002, **97**, 10269-10280.
9. A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648-5652.
10. F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297-3305.
11. T. Lu and F. Chen, *J. Comput. Chem.*, 2012, **33**, 580-592.
12. R. J. Loncharich, B. R. Brooks and R. W. Pastor, *Biopolymers*, 1992, **32**, 523-535.
13. J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696-3713.
14. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926-935.
15. I. S. Joung and T. E. Cheatham III, *J. Phys. Chem. B*, 2009, **113**, 13279-13290.
16. I. S. Joung and T. E. Cheatham III, *J. Phys. Chem. B*, 2008, **112**, 9020-9041.
17. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3690.
18. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577-8593.
19. J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327-341.
20. A. W. Götz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2012, **8**, 1542-1555.
21. R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2013, **9**, 3878-3888.
22. S. Le Grand, A. W. Götz and R. C. Walker, *Comput. Phys. Commun.*, 2013, **184**, 374-380.
23. F. Neese, F. Wennmohs, U. Becker and C. Riplinger, *J. Chem. Phys.*, 2020, **152**, 224108.
24. F. Neese, *WIREs Comput. Mol. Sci.*, 2022, **12**, e1606.
25. J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865-3868.
26. F. Weigend, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1057-1065.
27. R. A. Bryce, I. H. Hillier and J. H. Naismith, *Biophys. J.*, 2001, **81**, 1373-1388.
28. E. Wang, H. Sun, J. Wang, Z. Wang, H. Liu, J. Z. H. Zhang and T. Hou, *Chem. Rev.*, 2019, **119**, 9478-9508.
29. P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case and T. E. Cheatham, *Acc. Chem. Res.*, 2000, **33**, 889-897.
30. N. Homeyer and H. Gohlke, *Molecular Informatics*, 2012, **31**, 114-122.
31. S. Genheden and U. Ryde, *Expert Opinion on Drug Discovery*, 2015, **10**, 449-461.
32. B. R. Miller, III, T. D. McGee, Jr., J. M. Swails, N. Homeyer, H. Gohlke and A. E. Roitberg, *J. Chem. Theory Comput.*, 2012, **8**, 3314-3321.
33. T. Hou, J. Wang, Y. Li and W. Wang, *J. Chem. Inf. Model.*, 2011, **51**, 69-82.
34. A. Metz, C. Pfleger, H. Kopitz, S. Pfeiffer-Marek, K.-H. Baringhaus and H. Gohlke, *J. Chem. Inf. Model.*, 2012, **52**, 120-133.
35. H. Gohlke, C. Kiel and D. A. Case, *J. Mol. Biol.*, 2003, **330**, 891-913.
36. W. B. Schneider, G. Bistoni, M. Sparta, M. Saitow, C. Riplinger, A. A. Auer and F. Neese, *J. Chem. Theory Comput.*, 2016, **12**, 4778-4792.
37. A. Altun, M. Saitow, F. Neese and G. Bistoni, *J. Chem. Theory Comput.*, 2019, **15**, 1616-1632.
38. M. E. Beck, C. Riplinger, F. Neese and G. Bistoni, *J. Comput. Chem.*, 2021, **42**, 293-302.
39. C. Riplinger and F. Neese, *J. Chem. Phys.*, 2013, **138**, 034106.

40. F. Neese, *Coord. Chem. Rev.*, 2009, **253**, 526-563.
41. C. Riplinger, P. Pinski, U. Becker, E. F. Valeev and F. Neese, *J. Chem. Phys.*, 2016, **144**, 024109.
42. C. Riplinger, B. Sandhoefer, A. Hansen and F. Neese, *J. Chem. Phys.*, 2013, **139**, 134101.
43. T. Yanai, D. P. Tew and N. C. Handy, *Chem. Phys. Lett.*, 2004, **393**, 51-57.
44. J.-D. Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6615-6620.
45. Y.-S. Lin, G.-D. Li, S.-P. Mao and J.-D. Chai, *J. Chem. Theory Comput.*, 2013, **9**, 263-272.