A TDDFT investigation of the Photosystem II reaction center: Insights into the precursors to charge separation

Maeve A. Kavanagh[a,b] [ID], Joshua K. G. Karlsson[c,1] [ID], Jonathan D. Colburn[d,1] [ID], Laura M. C. Barter[a,b,2] [ID], and Ian R. Gould[a,b,2] [ID]

[a]Department of Chemistry, Molecular Sciences Research Hub, Imperial College London, W12 0BZ London, United Kingdom; [b]Institute of Chemical Biology, Molecular Sciences Research Hub, Imperial College London, W12 0BZ London, United Kingdom; [c]Molecular Photonics Laboratory, School of Natural and Environmental Sciences, Newcastle University, Newcastle Upon Tyne NE1 7RU, United Kingdom; and [d]School of Chemistry, University of St. Andrews, St. Andrews KY16 9ST, Scotland

Photosystem II (PS II) captures solar energy and directs charge separation (CS) across the thylakoid membrane during photosynthesis. The highly oxidizing, charge-separated state generated within its reaction center (RC) drives water oxidation. Spectroscopic studies on PS II RCs are difficult to interpret due to large spectral congestion, necessitating modeling to elucidate key spectral features. Herein, we present results from time-dependent density functional theory (TDDFT) calculations on the largest PS II RC model reported to date. This model explicitly includes six RC chromophores and both the chlorin phytol chains and the amino acid residues <6 Å from the pigments' porphyrin ring centers. Comparing our wild-type model results with calculations on mutant D1-His-198-Ala and D2-His-197-Ala RCs, our simulated absorption-difference spectra reproduce experimentally observed shifts in known chlorophyll absorption bands, demonstrating the predictive capabilities of this model. We find that inclusion of both nearby residues and phytol chains is necessary to reproduce this behavior. Our calculations provide a unique opportunity to observe the molecular orbitals that contribute to the excited states that are precursors to CS. Strikingly, we observe two high oscillator strength, low-lying states, in which molecular orbitals are delocalized over Chl_{D1} and Phe_{D1} as well as one weaker oscillator strength state with molecular orbitals delocalized over the P chlorophylls. Both these configurations are a match for previously identified exciton–charge transfer states $(Chl_{D1}{}^{+}Phe_{D1}{}^{-})^{*}$ and $(P_{D2}{}^{+}P_{D1}{}^{-})^{*}$. Our results demonstrate the power of TDDFT as a tool, for studies of natural photosynthesis, or indeed future studies of artificial photosynthetic complexes.

photosynthesis | TDDFT | Photosystem II | charge-separation precursors | structure–function relationship

During oxygenic photosynthesis, a photon of sunlight may be absorbed by any of the >250 chlorophylls or carotenoids in the Photosystem II (PS II) complex (1). This energy is then rapidly transferred to the reaction center (RC), where a remarkable charge separation (CS) process takes place, resulting in a radical pair species with oxidative ability, which is unparalleled in nature (+1.1 eV) (2). The RC pigments are bound by the protein chains D1 and D2, themselves being arranged in two nearly symmetrical branches (3). Each branch hosts a central chlorophyll a (Chl) molecule (P_{D1} and P_{D2}, respectively), a second Chl molecule (Chl_{D1} and Chl_{D2}), and a pheophytin (Phe) molecule (Phe_{D1}, Phe_{D2}), which together comprise the six core pigments of the PSII RC. Beyond these, a plastoquinone (Q_A and Q_B) and a further peripheral Chl ($Chlz_{D1}$ and $Chlz_{D2}$) are found on each chain.

CS proceeds from an initial delocalized excited state (ES) (2), identified as a pair of exciton–charge transfer states $(Chl_{D1}{}^{\delta+}Phe_{D1}{}^{\delta-})^{*}$ and $(P_{D2}{}^{\delta+}P_{D1}{}^{\delta-})^{*}$ (4), down one of two charge-transfer (CT) pathways (5–7) (Schemes 1 and 2), producing radical pair species $P_{D1}{}^{+}Phe_{D1}{}^{-}$.

One of the two CT pathways may be favored due to the specific conformation of the protein (7) or the wavelength of incident light (8). Vibronic coupling between states has been identified (9), which facilitates switching between the two available pathways (10) to further assist efficient CS.

Histidine (His) residues ligate the two P chlorophylls in the wild-type (W-T) crystal structure (3). These ligands have high polarizability and will stabilize the positive charge on the P chlorophylls during the formation of the primary radical pair state $P^{+}Phe_{D1}{}^{-}$ (11). The relative perturbation of the W-T absorption spectrum following mutation near the P_{D1} and P_{D2} sites, respectively, suggests that the cationic charge in the radical pair is in fact shared by the two P chlorophylls, with around 80% localized on P_{D1} and 20% on the P_{D2} (11–14). In particular, replacement of these His residues in site-directed mutagenesis studies (11, 15, 16) has allowed the spectral features of the P chlorophylls' absorption to be identified, with P_{D1} absorption peaks at 672.5 nm (at 80 K) and 433 nm (at 298 K) and P_{D2} at 436 nm (at 298 K) (11).

The powerfully oxidizing cation of the $P^{+}Phe_{D1}{}^{-}$ radical pair proceeds to abstract an electron from substrate water, which is

Significance

Examining the excited states of the Photosystem II reaction center furthers our understanding of available charge separation pathways that lead to successful photosynthesis. Our results comprise the largest complete model of the Photosystem II reaction center to be described using time-dependent density functional theory reported in the literature to date. We reveal the molecular orbitals contributing to the excited states that are precursors to charge separation. We demonstrate that our model can successfully predict the action of specific mutations, a valuable tool for the agricultural industry. These models may also be beneficial in informing the design of artificial photosynthetic complexes as well as enhanced bioengineered photosystems.

BIOPHYSICS AND COMPUTATIONAL BIOLOGY

$$(Chl_{D1}{}^{\delta+} Phe_{D1}{}^{\delta-})^* \rightarrow Chl_{D1}{}^+ Phe_{D1} \rightarrow P_{D1}{}^+ Phe_{D1}{}^-$$

Scheme 1.

bound at the oxygen-evolving complex via a tyrosine residue, Tyr_Z (17). After four successive excitations and electron-transfer cycles, two water molecules are oxidized to form one oxygen molecule as a byproduct (18, 19). Meanwhile, $Phe_{D1}{}^-$ transfers an electron to Q_A, which in turn reduces the loosely bound Q_B. After a second successive excitation and electron-transfer cycle, $Q_B{}^-$ is further reduced to $Q_B{}^{2-}$ and becomes protonated to plastoquinol. Plastoquinol then leaves the binding site to carry its reducing equivalents over to Photosystem I and is replaced by a new plastoquinone (20).

Site-Directed Mutagenesis

Due to the considerable congestion of the PS II RC absorption spectrum, identifying intermediate states in the CT process poses a difficult challenge (1). Site-directed mutagenesis has therefore been employed to pinpoint the cofactors that are implicated in these intermediates (4, 11, 15, 16, 21–24).

A number of investigations have selected the ligands of P_{D1} and P_{D2} (residues D1-198 and D1-197) as the sites for mutation. These sites each host His residues in the W-T and examples can be found in the literature of their mutation to alanine (Ala), asparagine (Asn), leucine, serine (Ser), glutamine (Gln), and glutamate (Glu) (4, 11, 15, 16, 23, 24). In the case of mutation to smaller amino acids Ala and Asn, there is likely to be a water molecule coordinating the P chlorophyll Mg atom (11, 13, 15, 16, 23) in place of the nitrogen on the His. In this work, we have chosen to model two mutants, D1-His-198-Ala (H198A) and D2-His-197-Ala (H197A), each with water molecules inserted to coordinate the Mg atoms.

Evidence from experimental studies of the His-198-Ala mutant (replacement of the P_{D1} His ligand with an Ala) show that this mutation increases the stability of the P^+ cation (11, 15, 16) possibly by stabilization of charge onto the new water ligand to the P_{D1} chlorophyll (13). The mutation is observed to increase the quantum yield of the primary radical pair state $P^+Phe_{D1}{}^-$, which is due to a lower free energy of the primary radical pair (15). Also observed for this mutant is a lower fluorescence yield (15) and a greatly increased rate of recombination of the radical pair state relative to the W-T (11). Despite increasing radical pair yield, overall, the photosynthetic activity for this mutant would be expected to be lower than for the W-T, as the driving force for secondary CT step $Y_ZP^+ \rightarrow Y_Z{}^+P$ is diminished. This is evidenced by the oxygen evolution rate for His-198-Ala mutant being 50 to 70% that observed for the W-T (11).

In the case of the analogous D2 branch mutant His-197-Ala, a slightly increased rate of recombination of the radical pair suggests that the P^+ cation has been stabilized by the mutation but to a lesser extent than the D1 branch mutation (11). This is explained by the mutation stabilizing the P_{D2} chlorophyll to cation localization and the positive charge of the resultant P^+ cation being more evenly shared by P_{D1} and P_{D2}.

Explicit quantum mechanical (QM) modeling of these mutant RCs allows us to closely examine the effect of the mutation on the molecular orbitals, something that has not been achievable before now.

Models of the RC

Complementing the vast body of experimental work on the RC are various computational studies of the system (2, 5, 6, 10, 13, 14, 25–35). Early theoretical studies of the PSII RC revealed that there must be some key structural differences between this and the purple bacterial RC. This was proven first by the Multimer model (2), which considered weak excitonic interactions between

the majority of RC cofactors and could produce a good approximation of the experimental absorption spectrum despite there being no crystal structure data available at the time. As improvements in the resolution of crystal structures of PSII were published (3, 36, 37), it enabled new models to be built, based upon these parameters (13, 26–28, 32, 33). These new models incorporated the coupling between pigments (32) and/or the atomic positions of the pigments themselves (13, 26, 33) as determined from the crystal structure. Among the latter, the size of the model became restrictive, and care had to be taken in choosing how much of the protein environment could be included and at what level of theory (26–28). Several authors opted for an idealized model to reduce computational cost (26, 28); however, in one example, it was shown that inclusion of the phytol chains in single residue QM calculations emphasized the preference for the active branch CT pathway over the inactive (33). Additionally, there is some debate in the literature as to whether geometry optimization (13, 28, 33), or relaxation using molecular dynamics (27, 34, 38), of the crystal structure atomic positions is necessary to produce an accurate model of the PS II RC or whether the crystal structure geometries are sufficient (25, 26). A number of alternative mathematical models have also been designed where model parameters are based upon fits to experimental spectroscopic data (5, 10, 29, 35, 39).

We believe the RC lends itself best to treatment as a supermolecule in order to accurately describe its ESs, which are likely to be delocalized across multiple pigments due to the weak coupling between all core RC pigments. So far, this has been a difficult computational task, due to the high computational demand associated with a quantum system of this size. The Multimer model (2, 32) and its relatives (5, 10, 29, 35, 39), for example, use simple dipole approximations for the chromophores (33, 35, 40) and consider only local excitations of the chromophores (2), a single CT state (5), or only active branch chromophores (10, 29). Other QM models have broken up the QM region into constituent chromophores (14, 33, 38, 41, 42) and/or used hybrid methods to mitigate the large computational costs (13, 14, 28, 38). The six RC chromophores (without phytol chains) were included in a supermolecular configuration interaction singles calculation in 2011 (26). Previously, the largest (~300 atoms) time-dependent density functional theory (TDDFT) model of the RC includes six chromophores and two quinones reduced to their rings only (no side chains or phytol chains included in the QM region) (28).
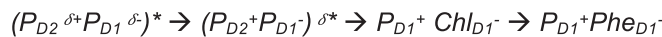
Our most complete QM model, reported here, comprises 6 chromophores (including side chains and full phytol chains) along with 23 whole residues and 2 water molecules, which have been explicitly included (1,299 atoms), in keeping with recent insight (43–45). Therefore, the present work constitutes the largest complete ES QM model of the PSII RC reported to date, using the robust method of TDDFT.

TDDFT is based on the ground state method of density functional theory, which uses charge density as an analog to the electronic wavefunction. TDDFT provides a method to solve the many body time dependent Schrödinger equation as a way to calculate excitation energies and oscillator strengths (46).

Long-range hybrid-corrected functional wB97x-D was employed in this work as this has been shown to be effective for diffuse (delocalized) and CT-like ESs (47, 48).

Results

Four W-T Models. Four W-T models were made, labeled henceforth as models 1 to 4 (details in Table 1), all including the six core chlorins (P_{D1}, P_{D2}, Chl_{D1}, Chl_{D2}, Phe_{D1}, and Phe_{D2}) and four chlorophyll

$$(P_{D2}{}^{\delta+}P_{D1}{}^{\delta-})^* \rightarrow (P_{D2}{}^+P_{D1}{}^-)^{\delta*} \rightarrow P_{D1}{}^+ Chl_{D1}{}^- \rightarrow P_{D1}{}^+Phe_{D1}{}^-$$

Scheme 2.

ligands: His-198; His-197 coordinating P_{D1} and P_{D2}, respectively; and two water molecules coordinating the Chl_{D1} and Chl_{D2} chlorophylls.

Models 1 and 3 had the phytol chains of the six chlorins truncated up to the methyl group on C7, and for model 2 and 4, they were left as full-length chains (see structures of model 1 and 4 in Fig. 1 *A* and *B*). Models 3 and 4 included a further 21 whole, neutral amino acid residues (terminated COOH/NH$_2$), including the nearest residues to the centers of the six porphyrin rings (within 6 Å radius), in addition to Gln-130 (D1 branch) and Gln-129 (D2 branch) (Fig. 1*B*). For a list of residues included please, see *SI Appendix*, Table S0.

ESs are characterized by a weighted combination of transitions, represented by pairs of occupied and virtual molecular orbitals (MOs), where the weights are given by the computed configuration-interaction coefficients. Each ES can be labeled as an excitation of the cofactor, which contributes most highly to the occupied MOs involved in the major transitions (highest configuration-interaction coefficient) of that ES. The contribution of each cofactor to an MO was determined as the sum of

atomic contributions to that MO over all atoms in that cofactor (49).

The six lowest-energy ESs of the four W-T models, as calculated using TDDFT with the wB97x-D (47, 48) functional and 6-31G (d, p) (50–56) basis set, are described here. The major transitions and oscillator strengths for states 1 to 6 are presented in Table 1.

For all W-T models, the brightest states, the first, second, and fifth, correspond to the Phe excitations and the Chl_{D1} excitation. From Table 1, we note that a higher oscillator strength is calculated for state 5 (Chl_{D1} excitation) for the full phytol chain models 2 and 4 in comparison for models 1 and 3. Whereas adding amino acids to the models results in higher oscillator strengths for states 1 and 2 (Phe excitations): without the phytol chains (models 1 and 3), we see an increase of oscillator strengths for states 1 and 2 from 0.39 and 0.4 to 0.41 and 0.55 on adding the amino acids. Similarly, with the phytol chains (model 2 and 4), states 1 and 2 increase in oscillator strength from 0.31 to 0.37 and from 0.4 to 0.42, respectively (Table 1).

Table 1. Excited state energies, oscillator strengths, and dominant transitions for states 1 to 6 for W-T models 1 to 4

Excited State	1	2	3	4	5	6
Model 1: 2 amino acids, truncated phytol chains						
Transition Energy (nm)	673.74	668.84	667.22	655.88	651.56	646.08
Oscillator Strength	0.39	0.4	0.13	0.08	0.37	0.15
Dominant Transition	Phe_{D2}	Phe_{D1}	P_{D1}	P_{D2}	Chl_{D1}	P_{D2}
Model 2: 2 amino acids, full phytol chains						
Transition Energy (nm)	677.30	670.54	669.01	657.06	654.72	646.40
Oscillator Strength	0.31	0.4	0.1	0.06	0.5	0.15
Dominant Transition	Phe_{D2}	Phe_{D1}	P_{D1}	Chl_{D2}	Chl_{D1}	P_{D2}
Model 3: 23 amino acids, truncated phytol chains						
Transition Energy (nm)	684.91	681.07	673.90	667.67	657.22	654.59
Oscillator Strength	0.41	0.55	0.07	0.12	0.38	0.08
Dominant Transition	Phe_{D2}	Phe_{D1}	P_{D1}	P_{D2}	Chl_{D1}	Chl_{D2}
Model 4: 23 amino acids, full phytol chains						
Transition Energy (nm)	692.46	687.23	675.93	665.76	660.40	655.24
Oscillator Strength	0.37	0.42	0.09	0.18	0.41	0.14
Dominant Transition	Phe_{D1}	Phe_{D2}	P_{D1}	P_{D2}	Chl_{D1}	Chl_{D2}

Results are as calculated using wB97x-D/6-31G (d, p). Bright oscillator strength states are colored green with darkest green indicating the highest oscillator strength state. Dominant transitions are also colored by cofactor as in Figs. 2*A* and 4: Phe_{D1}, blue; Chl_{D1}, aquamarine; P_{D1}, red, P_{D2}, pink; Chl_{D2}, orange; and Phe_{D2}, yellow. States 7 to 22 for all models were also calculated; these results are discussed further in subsection Higher Energy ESs and full results are presented in *SI Appendix*, Tables S1–S4 *A–C*.
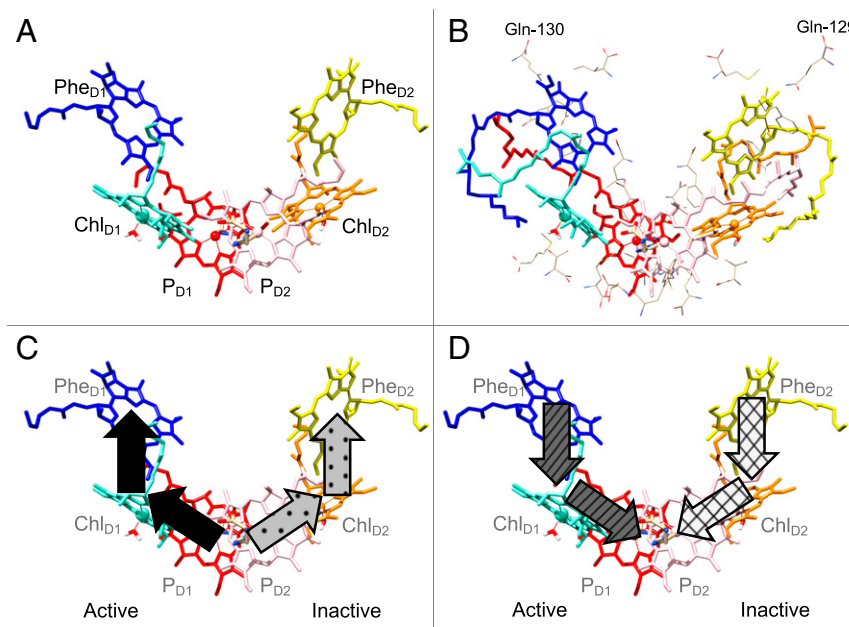
Fig. 1. (*A*) Model 1 structure including the six chlorins with phytol chains truncated up to the methyl group on C7, two waters coordinating the Chl_{D1} and Chl_{D2}, and two His residues coordinating the P chlorophylls. Cofactors are colored as in Table 1 and Figs. 2*A* and 4: Phe_{D1}, blue; Chl_{D1}, aquamarine; P_{D1}, red; P_{D2}, pink; Chl_{D2}, orange; and Phe_{D2}, yellow. (*B*) Model 4 including the six cofactors with full phytol chains, two water molecules, and 23 amino acid residues (wire structure). All residues lie within 6 Å of the porphyrin ring-centers except Gln-130 and Gln-129 (labeled). (*C*) Black arrows show CT up the active branch (from P chlorophylls toward Chl_{D1} or Phe_{D1} or from Chl_{D1} toward Phe_{D1}), dotted arrows show CT up the inactive branch (from P chlorophylls toward Chl_{D2} or Phe_{D2} or from Chl_{D2} toward Phe_{D2}). (*D*) Striped arrows show CT down the active branch (from Phe_{D1} chlorophylls toward Chl_{D1} or P chlorophylls or from Chl_{D1} toward P chlorophylls), and hashed arrows show CT down the inactive branch (from Phe_{D2} chlorophylls toward Chl_{D2} or P chlorophylls or from Chl_{D2} toward P chlorophylls).

When comparing models 3 and 4, it can be seen that the energies of the Phe excitations are shifted by addition of phytol chains when the 23 amino acids are present, including the phytol chains raises the energy of Phe_{D2} excitation above Phe_{D1} (Table 1). When including the surrounding amino acids in models 3 and 4, the Chl_{D2} excitation

appears highest in energy, whereas for the two models without the amino acids, the Chl_{D2} excitation is lower in energy (model 2) or does not dominate any of these states (model 1).

Some states display a high degree of mixing, which is omitted for clarity in Table 1, and a complete picture for model 4 is presented
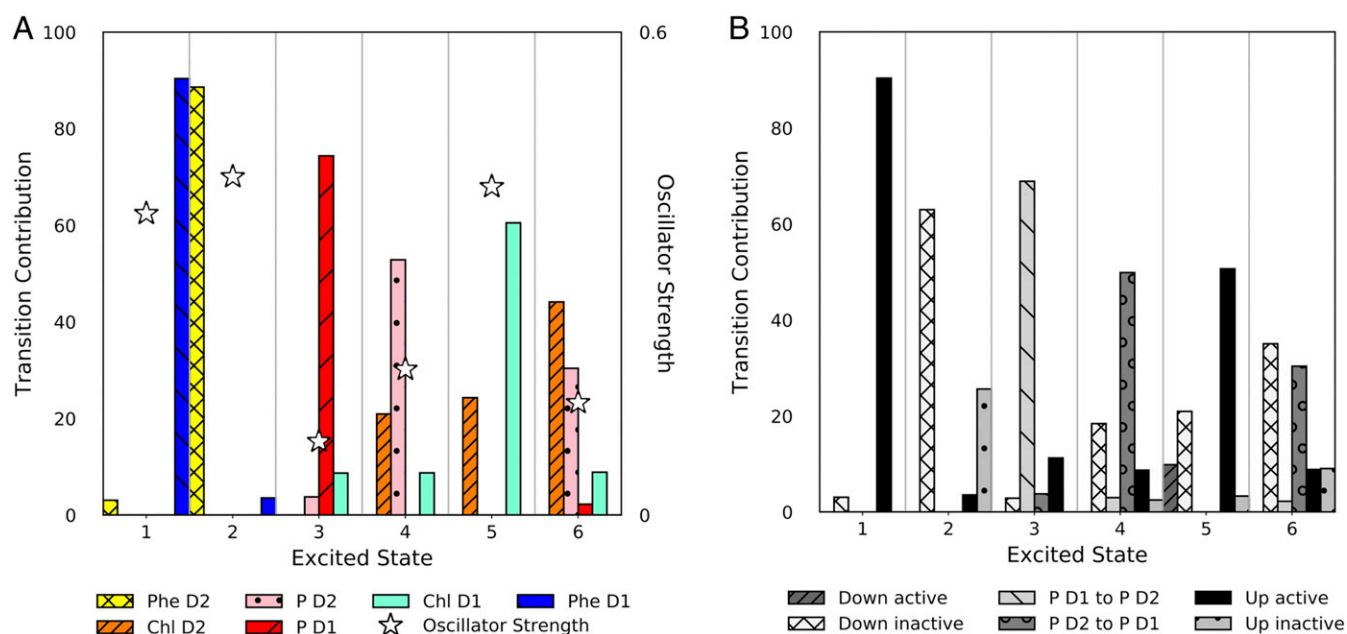
Fig. 2. (*A*) Contribution from occupied orbitals on each cofactor for transitions contributing to each state for W-T model 4—with 23 amino acids and full phytol chains. Oscillator strength for each state is plotted on secondary *x* axis in white stars. (*B*) CT character of transitions contributing to each state for W-T model 4—with 23 amino acids and full phytol chains. Direction of CT is shaded as in Fig. 1 *C* and *D*. (States 7 to 22 for model 4 were also calculated; these results are discussed further under the subheading Higher Energy ESs and full results for all models are presented in *SI Appendix*, Tables S1–S6 *A* and *B*.)

Kavanagh et al.

in Fig. 2. The MOs for the transitions with the highest contribution to states 1 and 5 of W-T model 4 are presented in Fig. 3.

Individual transitions, such as those pictured in Fig. 3, were analyzed for the extent of CT between cofactors using MultiWfn (49) by examining the cofactor contributions to the orbitals. For example, consider a transition where the occupied and virtual MOs are delocalized over Chl_{D1} and Phe_{D1} on the active branch. If the Phe_{D1} contribution to the virtual MO is higher than the Phe_{D1} contribution to the occupied MO, we would identify this transition to show a degree of CT character, where electron density is directed up the active branch from Chl_{D1} to Phe_{D1} during the course of the transition (as in Fig. 1C). If the Chl_{D1} contribution increases from the occupied to the virtual MO, then we would identify that this transition has CT character where electron density is directed down the branch.

In the Qy region (comprising first six ESs), it was found that transitions for which CT travels up the D1 branch (from P chlorophylls to Phe_{D1}) are dominant for two high oscillator-strength states (states 1 and 5) for model 4 (Fig. 2B). For models 1 to 3, transitions with CT up the D1 branch dominate one state only (*SI Appendix*, Figs. S1–S3B).

For whole ESs, which are described by a linear combination of transitions, we analyzed the extent of CT using the method described in (57). We found that the CT character of these vertical excitations was minimal at low energies (*SI Appendix*, Tables S1–S4D).

Higher Energy ESs. For all models, higher-energy ESs were evaluated (22 ESs in total). The ES energies and oscillator strengths obtained from our models can be broadened by Gaussian functions to give a calculated absorption spectrum (Fig. 4).

We observe three distinct bands, labeled from lowest to highest energy as the Qy, Qx, and Soret bands in our calculated absorption spectrum (Fig. 4). The Qy region comprises ESs 1 to 6 and is described in detail in Table 1 and Fig. 2. The states in the

Qx region mirror the Qy states, in that they correspond to an excitation of each of the six chlorins (see all results in *SI Appendix*, Tables S1–S4 A–C and Figs. S1–S4 A and B and MO diagrams for model 4 in *SI Appendix*, Fig. S4C). The remaining ESs, 13 to 22, form the Soret band. Six states out of the 10 in this region (443 to 396 nm) have their highest contributions from P_{D1} chlorophyll excitations (Fig. 4, Soret region; red dashed lines: states 14, 15, 17 to 19, and 22). The remaining four states have their highest contribution from P_{D2} (state 13 and 16) or Phe excitations (states 20 and 21).

Modeling Single Residue Mutations. Mutating the residues ligating to the P chlorophylls from His to alanine (His-198-Ala and His-197-Ala) on either the D1 or the D2 branches of the PSII RC, gave two mutant RC models (Fig. 5).

The high energy Soret band of the spectrum for our models 1 to 4 corresponds to a high density of P chlorophyll excitations (see W-T model 4 results in Fig. 4; also see *SI Appendix*, Tables S1–S4 A–C). When the states corresponding to P_{D1} excitations were isolated and the sum of these Gaussian-broadened states was plotted for the W-T and the two mutants it could be seen that the H198A mutation resulted in an increase in energy of the P_{D1} excitations overall, whereas the H197A mutation decreased the energy of the P_{D1} excitations overall (Fig. 6).

Discussion

We set out to probe the ESs of the PSII RC core pigments, taking a large region (up to 1,299 atoms including H for model 4) of the crystal structure as our starting point; 22 ESs were calculated and the effects of including pigment phytol chains and surrounding protein environment in the form of whole residues was investigated.

Active Branch Excitations Are Favored When the Local Environment Is Included. We examined the effects of the environment on oscillator strengths of the ESs as well as the atom contributions to the
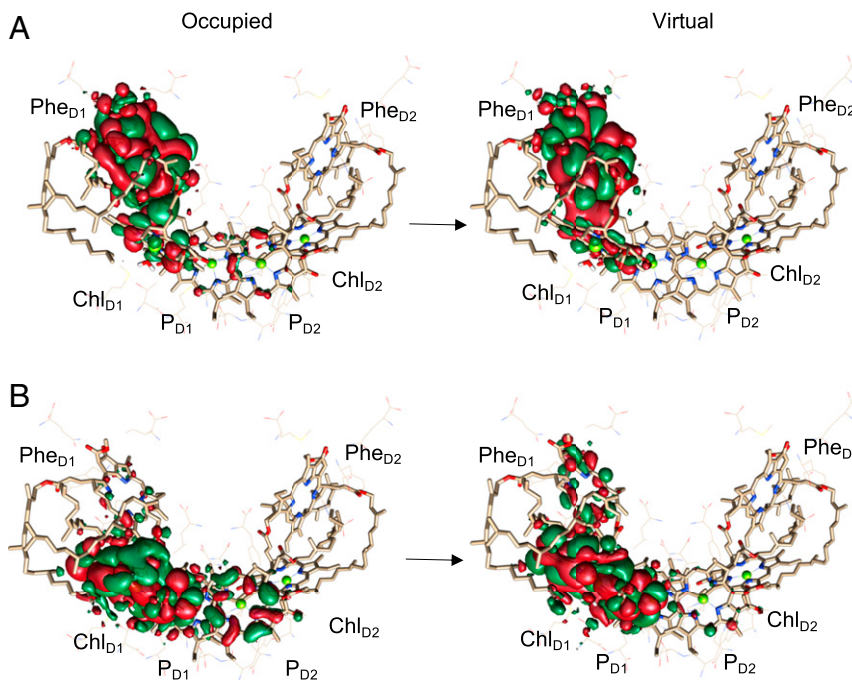
Fig. 3. The molecular orbitals, states obtained from model 4 (23 amino acids and full phytol chains), participating in the most dominant transitions for the two brightest ESs in the Qy region, pictured with an iso surface value of ±0.002. (A) The transition contributing 64.9% to state 1 for model 4, involving Phe_{D1} excitation with delocalization across Chl_{D1} and P chlorophylls. (B) The transition contributing 50.6% to state 5 for model 4 involving Chl_{D1} excitation with delocalization across P chlorophylls and Phe_{D1}.
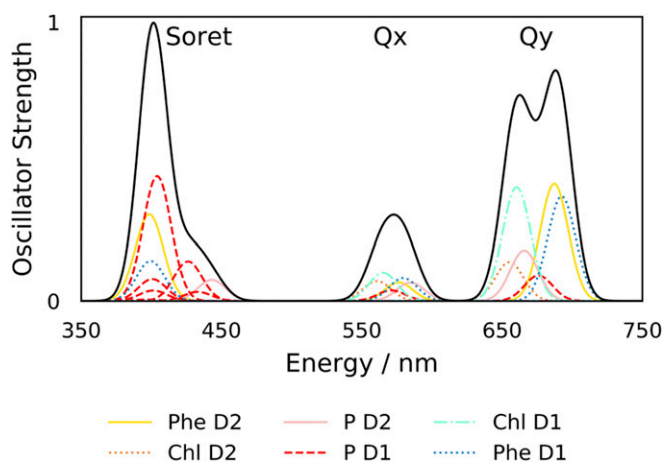
Fig. 4. W-T Model 4 (23 amino acids, full phytol chains) calculated absorption spectrum includes three distinct bands, from lowest to highest energy, the Qy, Qx, and Soret regions. Black solid spectrum, overall calculated absorption spectrum; colored lines:, single ESs are colored according to the cofactor of the dominant transition as in Table 1 and Fig. 1: Phe_{D1}, blue; Chl_{D1}, aquamarine; P_{D1}, red, P_{D2}, pink; Chl_{D2}, orange; and Phe_{D2}, yellow.

occupied and virtual MOs. We find that our largest model (model 4), which includes the phytol chains and 23 amino acids, demonstrates the strongest preference for active branch excitations.

From Table 1, we see that for all our models the high oscillator-strength states are Phe and Chl_{D1} excitations. Upon inclusion of the phytol chains the Chl_{D1} (state 5) oscillator strength is increased. Investigation of each transition's direction of CT (Fig. 2B and *SI Appendix*, Figs. S1–S3B) revealed that inclusion of both the phytol chains and the amino acids increases the contribution of transitions which direct CT up the active branch. Both state 1 and 5 are dominated by transitions of this nature for model 4 (Fig. 2B). Transitions that progress toward successful generation of the radical pair $P_{D1}^{+}Phe_{D1}^{-}$ are therefore favored when the local environment is included.

The lowest-energy ES has been assigned to Chl_{D1} excitation (27); our results obtained using model 4 are indicative of agreement as the lowest energy ES is identified as a Phe_{D1} excitation with delocalization across Chl_{D1} (Table 1 and Fig. 3). In contrast, models 1 to 3 have their lowest-energy ES attributed to Phe_{D2} on the inactive branch. Our calculations reveal the sensitivity of the site energies to cofactor configurations. We note that, while all cofactor excitations within the Qy region lie close in energy, the specific conformation of the pigments in the crystal structure (3) may lead to a higher site energy for localized Chl_{D1} excitations than the Phe_{D1} in our calculations.

A Closer Examination of the Active Branch Excitations. TDDFT allows us to examine the effect of initial excitation and thus the precursor states to the experimentally observed radical pair state intermediates. Two of the brightest states, obtained from model 4, are delocalized across the D1 branch cofactors (Fig. 3). This configuration is consistent with the exciton–CT ES ($Chl_{D1}^{\delta+}Phe_{D1}^{\delta-}$)* (Scheme 1), previously identified by stark spectroscopy on site-directed mutants (4). This state is the precursor to the dominant CS pathway.

Initial CS may proceed via one of two pathways (4): via either Scheme 1 or 2 in which CS originates in the P chlorophyll pair from ($P_{D2}^{\delta+}P_{D1}^{\delta-}$)*. Thus, we are unsurprised to find states involving excitation of the P chlorophylls in the darker states 3, 4, and 6, signifying the existence of another pathway (Fig. 2A; for MO diagrams for model 4, see *SI Appendix*, Fig. S4C), which may be more favorable for a different realization of the disorder of the RC (1, 7).

Simulating the Absorption Spectrum of the RC. The experimental low-temperature absorption spectrum of the smallest charge-separating complex that can be isolated from PSII, the D1/D2 cytochrome (Cyt) b559 PSII RC complex, has three distinct peaks in the region from 700 to 400 nm (58, 59) (Table 2). These are labeled, from low to high energy, the Qy, the Qx, and the Soret region. It should be noted that the D1/D2 Cyt b559 RC contains six Chl and two Phe and thus contains two peripheral chlorophyll molecules, which are not included in our models. These two additional chlorophyll molecules are very loosely coupled to the RC pigments and are thus commonly excluded from models of the RC (2, 6, 28).

Our calculated spectrum (Fig. 4) has three distinct bands in the same region 350 to 750 nm; states 1 to 6 make up the high oscillator-strength Qy band and are discussed in detail above; states 7 to 12 all have relatively low oscillator strength and fall within the Qx region; and states 13 to 22 comprise the Soret band (see *SI Appendix*, Fig. S4 *A and B* and Table S4 *A–C* for the nature of these ESs). As can be seen in Fig. 4, our model is able to reproduce the main features of the D1-D2-Cyt b559 complex low-temperature ultraviolet-visible absorption spectrum (58).

Diner et al. have shown that $P^{+} - P$ absorption-difference spectroscopy on W-T RCs shows a bleach in the Soret region at 433 nm, indicating that the chlorophyll bearing the cation (P_{D1} and, to a lesser extent P_{D2}) is absorbing in this region (11). Our model finds that eight states out of the ten in the region of 443 to 396 nm have their highest contributions from P chlorophyll excitations. Of these, six states are dominated by P_{D1} excitations. Gaussian broadened peaks for states in the Soret region have been summed to give the blue spectral line in Fig. 6.

Mutant Reaction Centre Models. With our chosen method of TDDFT we were able to investigate the effect on the ground state absorption of the cofactors following mutation of the residues ligating to the P chlorophylls from His to alanine (His-198-Ala and His-197-Ala). The energies and oscillator strengths of the states attributed to P_{D1} excitation were broadened using Gaussian functions and their sum plotted for model 4 W-T, His-198-Ala and His-197-Ala. The wavelengths of the center of the absorption peaks of these simulated P_{D1} bands were determined from the spectra obtained using model 4 for the W-T and mutant RCs (Fig. 6). It can be seen that the D1 branch mutation, His-198-Ala, has shifted the band to higher energies whereas the D2 branch mutant has had the opposite effect.

Experimental studies on these mutants have been previously reported by Diner et al. and have shown that the prominent
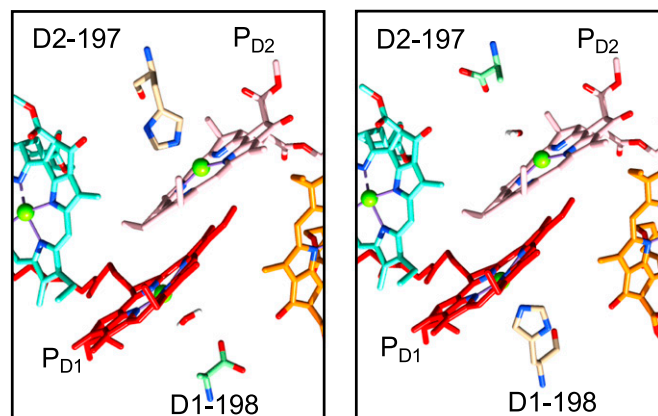
Fig. 5. View of the P chlorophylls with either His at D1-198 (*Left*) or at D2-197 (*Right*) mutated to Ala with a water molecule inserted to ligate to the chlorophyll center for model 4 (23 amino acids, full phytol chains). Colored cofactors are as in Fig. 1; P_{D1} (red) and P_{D2} (pink) and mutated residue His to Ala (green).
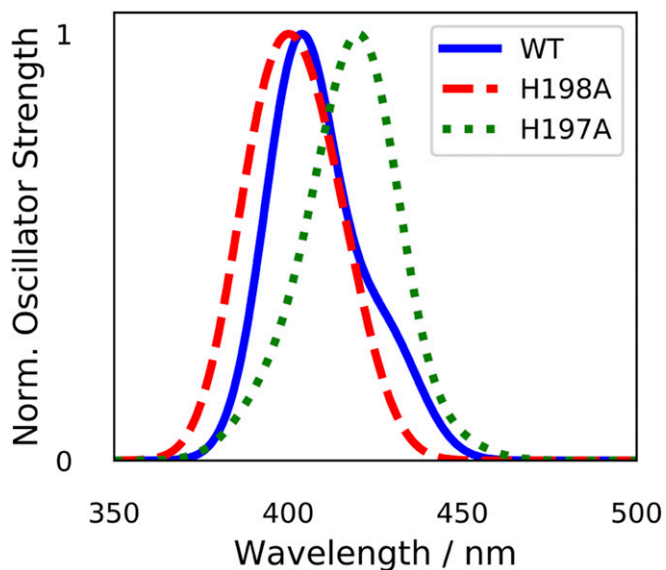
Fig. 6. Simulated P_{D1} absorption bands in the Soret region obtained using model 4 for the W-T RC and two mutants His-198-Ala and His-197-Ala (all with 23 amino acids and full phytol chains): sum of the Gaussian-broadened ESs that are dominated by P_{D1} excitation for W-T (blue solid line), His-198-Ala (red dashed line), and His-197-Ala (green dotted line), where peak heights have been normalized to one for ease of comparison in the energy shifts. See *SI Appendix*, Figs. S1–S3D for model 1 to 3 results.

bleaching band in the $P^+ - P$ absorption-difference spectrum for the W-T RC is blue-shifted from 433 nm to ~431 nm upon mutation of D1 branch mutant His-198 to Ala, whereas for the D2 branch His-197-Ala mutant this band is red-shifted to ~434 nm (11).

The results from our calculations, using Model 4, therefore show agreement with the observed shifts in the experimental $P^+ - P$ absorption band bleaching frequency energies for both of these mutants (Fig. 6), thus demonstrating the predictive capabilities of our model. Interestingly we did not see such an agreement for smaller models 1 to 3 (*SI Appendix*, Fig. S1–S3D), which emphasizes further the importance of including the phytol chains and amino acids in future models of the PSII RC.

In conclusion, we present the results of TDDFT calculations on the largest QM model of the PSII RC reported in the literature to date. Four models of different sizes were built and their ESs examined with TDDFT. The molecular orbitals contributing to two out of the three highest oscillator strength states for our largest model (model 4—full phytol chains and 23 amino acids) are delocalized over the D1 branch cofactors and with transitions having CT character directed up the active branch (Fig. 2B), which is in keeping with known precursor states to CS (4, 7). We observe two different states with this character, one localized on Chl_{D1} and the other on Phe_{D1}.

Table 2. Experimental RC absorption spectrum at 10 K (from ref. 58)

Band	D1-D2-Cyt b559 complex absorption spectrum peak wavelengths (10 K)/nm	Approximate absorbance
Qy	679	0.9
	671	0.8
Qx	542	0.1
Soret	~432	0.8
	~416	1.0

Strikingly, the main features of the experimentally observed low-temperature absorption spectrum of isolated D1 D2 Cyt b559 (containing six Chl and two Phe) are reproduced by all of our models (Fig. 4 and *SI Appendix*, Figs. S1–S3C). Furthermore, models of the two mutant RCs, with phytol chains and amino acids included, are able to produce the expected shifts in the P_{D1} bands in the room temperature absorption spectrum (Fig. 6) compared with W-T, further validating our largest model and demonstrating deeply promising predictive capabilities of using TDDFT.

Materials and Methods

Methods.

Model preparation. In total, 12 different model structures have been prepared, based upon the 1.9-Å resolution crystal structure of PS II from *Thermosynechococcus vulcanus* (3). The RC residues of interest were identified using Chimera (60) and hydrogens were added in GaussView (61). Prior to calculation, the hydrogen positions for each model were optimized using Austin model 1 (AM1) semiempirical theory in Gaussian (62), with heavy atom positions frozen.

For all model structures, the six-core RC chlorins (P_{D1}, P_{D2}, Chl_{D1}, Chl_{D2}, Phe_{D1}, and Phe_{D2}) were included, as were two water molecules coordinating the Chl_{D1} and Chl_{D2} chlorophylls. For models 1 and 3, the phytol chains of the chlorins were truncated up to the methyl group on C7, and for models 2 and 4, they were left as full-length chains (Fig. 1 A and B). For all models, the His ligands (or mutated residues in their place) coordinating the two P chlorophylls, residues D1-198 and D2-197, were included. For half of all models, a further 21 amino acids were included to simulate the surrounding protein environment; 19 of these were chosen as the nearest whole residues to the centers of the 6 chlorin porphyrin rings that lay within 6 Å of the chlorophyll Mg atoms or Phe nitrogen atoms in the crystal structure. In addition, two residues Gln-130 (8 Å away from Phe_{D1} ring) and Gln-129 (7 Å away from Phe_{D2} ring) were included in the larger models as these are known to affect radical pair formation having been targets for site-directed mutations (22). For a list of residues included please, see *SI Appendix*, Table S1. All amino acid residues were neutrally charged for ease of computation with TDDFT, residues were terminated -NH_2/-COOH, and the COOH dihedral angles were set to 0° prior to hydrogen optimization.

Details of the W-T model structures used are presented in Table 1, and the structure of two models (models 1 and 4) can be seen in Fig. 1.

In addition to the W-T models detailed above, each of the four model structures in Table 1 were reproduced as two mutants, H198A and H197A. A mutation from D1-198/D2-197-His to Ala was created by replacing the imidazole side chain with a hydrogen. A water molecule was inserted in place of the His nitrogen to ligate the P chlorophyll (see Fig. 5 for the position of these ligands in mutant models 4: H198A and H197A; similar figures are included in *SI Appendix*, Figs. S1–S3E for other models). The atomic positions of the new methyl side chain and water molecule were optimized using AM1 in Gaussian (62) while holding fixed all other atoms in the system.

Computational methods. TDDFT was used to map the 22 lowest energy ESs of each model. ES calculations were carried out with wB97X-D (47, 48)/6-31G (d, p) (50) using Gaussian (62). Model 1 calculations included a diffuse function using 6-31+G (d, p) (50–56) (for results, see *SI Appendix*, Table S7 A–C). Models 1 and 4 calculations were repeated with CAM-b3LYP (63) for comparative purposes (for results, see *SI Appendix*, Tables S8 and S9 A–C).

For each molecular orbital appearing in the TDDFT output, the percentage contribution of each atom in the model to the orbital was analyzed with MultiWfn (49) using Mulliken population analysis (64). The atom contributions were tabulated and the sum of all atom contributions over all of the atoms in each cofactor was calculated for the six cofactors. As such, each orbital can be described by its percentage contributions from each cofactor (*SI Appendix*, Tables S1–S4C). To produce the molecular orbital images in this work, the Cubegen function in Gaussian (62) was used to generate molecular orbital surfaces; these were then viewed using Chimera (60), with the chosen isovalue of +0.002/−0.002 (*SI Appendix*, Fig. S4C).

The extent of CT (57) was evaluated for each state calculated for all models in Gaussian (62) using keyword "pop = DCT." The identity (element, cofactor) of the nearest atom to the center of density depletion and increment from ground state to excited state are presented in *SI Appendix*, Tables S1–S4D. A state was considered CT if the nearest atoms to the tail and head of the vector representing CT extent were located on different cofactors.

1. L. M. C. Barter, D. R. Klug, A unified picture of energy and electron transfer in primary photosynthesis. *Chem. Phys.* **319**, 308–315 (2005).

2. J. R. Durrant *et al.*, A multimer model for P680, the primary electron donor of photosystem II. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4798–4802 (1995).

3. Y. Umena, K. Kawakami, J.-R. Shen, N. Kamiya, Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* **473**, 55–60 (2011).

4. E. Romero *et al.*, Mixed exciton-charge-transfer states in photosystem II: Stark spectroscopy on site-directed mutants. *Biophys. J.* **103**, 185–194 (2012).

5. V. I. Novoderezhkin, J. P. Dekker, R. van Grondelle, Mixing of exciton and charge-transfer states in photosystem II reaction centers: Modeling of stark spectra with modified redfield theory. *Biophys. J.* **93**, 1293–1311 (2007).

6. V. I. Novoderezhkin, E. Romero, J. P. Dekker, R. van Grondelle, Multiple charge-separation pathways in photosystem II: Modeling of transient absorption kinetics. *ChemPhysChem* **12**, 681–688 (2011).

7. E. Romero, I. H. M. van Stokkum, V. I. Novoderezhkin, J. P. Dekker, R. van Grondelle, Two different charge separation pathways in photosystem II. *Biochemistry* **49**, 4300–4307 (2010).

8. A. Pavlou, J. Jacques, N. Ahmadova, F. Mamedov, S. Styring, The wavelength of the incident light determines the primary charge separation pathway in photosystem II. *Sci. Rep.* **8**, 2837 (2018).

9. E. Romero *et al.*, Quantum coherence in photosynthesis for efficient solar-energy conversion. *Nat. Phys.* **10**, 676–682 (2014).

10. V. I. Novoderezhkin, E. Romero, R. van Grondelle, How exciton-vibrational coherences control charge separation in the photosystem II reaction center. *Phys. Chem. Chem. Phys.* **17**, 30828–30841 (2015).

11. B. A. Diner *et al.*, Site-directed mutations at D1-His198 and D2-His197 of photosystem II in Synechocystis PCC 6803: Sites of primary charge separation and cation and triplet stabilization. *Biochemistry* **40**, 9265–9281 (2001).

12. R. Nagao, M. Yamaguchi, S. Nakamura, H. Ueoka-Nakanishi, T. Noguchi, Genetically introduced hydrogen bond interactions reveal an asymmetric charge distribution on the radical cation of the special-pair chlorophyll P680. *J. Biol. Chem.* **292**, 7474–7486 (2017).

13. K. Saito, J. R. Shen, H. Ishikita, Influence of the axial ligand on the cationic properties of the chlorophyll pair in photosystem II from Thermosynechococcus vulcanus. *Biophys. J.* **102**, 2634–2640 (2012).

14. K. Saito *et al.*, Distribution of the cationic state over the chlorophyll pair of the photosystem II reaction center. *J. Am. Chem. Soc.* **133**, 14379–14388 (2011).

15. S. A. P. Merry *et al.*, Modulation of quantum yield of primary radical pair formation in photosystem II by site-directed mutagenesis affecting radical cations and anions. *Biochemistry* **37**, 17439–17447 (1998).

16. M. Sugiura, Y. Ozaki, F. Rappaport, A. Boussac, Corrigendum to "Influence of Histidine-198 of the D1 subunit on the properties of the primary electron donor, P$_{680}$, of photosystem II in Thermosynechococcus elongatus". *Biochim. Biophys. Acta* **1857**, 1943–1948 (2016).

17. B. A. Barry, G. T. Babcock, Tyrosine radicals are involved in the photosynthetic oxygen-evolving system. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7099–7103 (1987).

18. B. Kok, B. Forbush, M. McGloin, Cooperation of charges in photosynthetic O2 evolution-I. A linear four step mechanism. *Photochem. Photobiol.* **11**, 457–475 (1970).

19. P. Joliot, G. Barbieri, R. Chabaud, A new model of photochemical centers in system-2. *Photochem. Photobiol.* **10**, 309–329 (1969).

20. J. Barber, Photosystem II: Its function, structure, and implications for artificial photosynthesis. *Biochemistry (Mosc.)* **79**, 185–196 (2014).

21. J. G. K. Williams, [85] Construction of specific mutations in photosystem II photosynthetic reaction center by genetic engineering methods in Synechocystis 6803. *Methods Enzymol.* **167**, 766–778 (1988).

22. L. B. Giorgi *et al.*, Comparison of primary charge separation in the photosystem II reaction center complex isolated from wild-type and D1-130 mutants of the cyanobacterium Synechocystis PCC 6803. *J. Biol. Chem.* **271**, 2093–2101 (1996).

23. M. Di Donato *et al.*, Primary charge separation in the photosystem II core from Synechocystis: A comparison of femtosecond visible/midinfrared pump-probe spectra of wild-type and two P680 mutants. *Biophys. J.* **94**, 4783–4795 (2008).

24. E. Schlodder *et al.*, Site-directed mutations at D1-His198 and D1-Thr179 of photosystem II in Synechocystis sp. PCC 6803: Deciphering the spectral properties of the PSII reaction centre. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **363**, 1197–1202, discussion 1202 (2008).

25. T. Renger, E. Schlodder, Primary photophysical processes in photosystem II: Bridging the gap between crystal structure and optical spectra. *ChemPhysChem* **11**, 1141–1153 (2010).

26. Y. Kitagawa, K. Matsuda, J. Y. Hasegawa, Theoretical study of the excited states of the photosynthetic reaction center in photosystem II: Electronic structure, interactions, and their origin. *Biophys. Chem.* **159**, 227–236 (2011).

27. L. Zhang *et al.*, Dynamic protein conformations preferentially drive energy transfer along the active chain of the photosystem II reaction centre. *Nat. Commun.* **5**, 4170 (2014).

28. T. J. Frankcombe, Explicit calculation of the excited electronic states of the photosystem II reaction centre. *Phys. Chem. Chem. Phys.* **17**, 3295–3302 (2015).

29. V. I. Novoderezhkin, E. Romero, J. Prior, R. van Grondelle, Exciton-vibrational resonance and dynamics of charge separation in the photosystem II reaction center. *Phys. Chem. Chem. Phys.* **19**, 5195–5208 (2017).

30. E. Romero *et al.*, Quantum–coherent dynamics in photosynthetic charge separation revealed by wavelet analysis. *Sci. Rep.* **7**, 2890 (2017).

31. F. Müh, M. Plöckinger, T. Renger, Electrostatic asymmetry in the reaction center of photosystem II. *J. Phys. Chem. Lett.* **8**, 850–858 (2017).

32. L. M. C. Barter, J. R. Durrant, D. R. Klug, A quantitative structure-function relationship for the photosystem II reaction center: Supermolecular behavior in natural photosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 946–951 (2003).

33. N. Ivashin, S. Larsson, Excitonic states in photosystem II reaction center. *J. Phys. Chem. B* **109**, 23051–23060 (2005).

34. S. Vasil'ev, D. Bruce, A protein dynamics study of photosystem II: The effects of protein conformation on reaction center function. *Biophys. J.* **90**, 3062–3073 (2006).

35. G. Raszewski, B. A. Diner, E. Schlodder, T. Renger, Spectroscopic properties of reaction center pigments in photosystem II core complexes: Revision of the multimer model. *Biophys. J.* **95**, 105–119 (2008).

36. A. Zouni *et al.*, Crystal structure of photosystem II from Synechococcus elongatus at 3.8 A resolution. *Nature* **409**, 739–743 (2001).

37. K. N. Ferreira, T. M. Iverson, K. Maghlaoui, J. Barber, Architecture of the photosynthetic oxygen-evolving centre. *Science* **43**, 1831–1839 (2004).

38. N. Kulik, M. Kutý, D. Řeha, The study of conformational changes in photosystem II during a charge separation. *J. Mol. Model.* **26**, 75 (2020).

39. A. Gelzinis, D. Abramavicius, J. P. Ogilvie, L. Valkunas, Spectroscopic properties of photosystem II reaction center revisited. *J. Chem. Phys.* **147**, 115102 (2017).

40. G. Raszewski, W. Saenger, T. Renger, Theory of optical spectra of photosystem II reaction centers: Location of the triplet state and the identity of the primary electron donor. *Biophys. J.* **88**, 986–998 (2005).

41. H. Ishikita, W. Saenger, J. Biesiadka, B. Loll, E. W. Knapp, How photosynthetic reaction centers control oxidation power in chlorophyll pairs P680, P700, and P870. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 9855–9860 (2006).

42. K. Kawashima, H. Ishikita, Energetic insights into two electron transfer pathways in light-driven energy-converting enzymes. *Chem. Sci. (Camb.)* **9**, 4083–4092 (2018).

43. T. J. Zuehlsdorff, C. M. Isborn, Modeling absorption spectra of molecules in solution. *Int. J. Quantum Chem.* **119**, e25719 (2019).

44. D. R. Widmer, B. J. Schwartz, Solvents can control solute molecular identity. *Nat. Chem.* **10**, 910–916 (2018).

45. K. Saito, T. Suzuki, H. Ishikita, Absorption-energy calculations of chlorophyll a and b with an explicit solvent model. *J. Photochem. Photobiol. Chem.* **358**, 422–431 (2018).

46. E. Runge, E. K. U. Gross, Density-functional theory for time-dependent systems. *Phys. Rev. Lett.* **52**, 997–1000 (1984).

47. J. Da Chai, M. Head-Gorden, Long-range corrected hybrid density functionals with improved dispersion corrections. *J. Chem. Theory Comput.* **10**, 6615–6620 (2008).

48. J. Da Chai, M. Head-Gorden, Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* **128**, 084106 (2008).

49. T. Lu, F. Chen, Multiwfn: A multifunctional wavefunction analyzer. *J. Comput. Chem.* **33**, 580–592 (2012).

50. R. Ditchfield, W. J. Hehre, J. A. Pople, Self-consistent molecular-orbital methods. IX. An extended gaussian-type basis for molecular-orbital studies of organic molecules. *J. Chem. Phys.* **54**, 724–728 (1971).

51. W. J. Hehre, R. Ditchfield, J. a. Pople, Self-consistent molecular orbital methods. XII. Further extensions of gaussian-type basis sets for use in molecular orbital studies of organic molecules. *J. Chem. Phys.* **56**, 2257–2261 (1972).

52. M. M. Francl *et al.*, Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys.* **77**, 3654–3665 (1982).

53. R. C. Binning Jr., L. A. Curtiss, Compact contracted basis sets for third-row atoms: Ga–Kr. *J. Comput. Chem.* **11**, 1206–1216 (1990).

54. J.-P. Blaudeau, M. P. McGrath, L. A. Curtiss, L. Radom, Extension of Gaussian-2 (G2) theory to molecules containing third-row atoms K and Ca. *J. Chem. Phys.* **107**, 5016–5021 (1997).

55. V. A. Rassolov, J. A. Pople, M. A. Ratner, T. L. Windus, 6-31G* basis set for atoms K through Zn. *J. Chem. Phys.* **109**, 1223–1229 (1998).

56. V. A. Rassolov, M. A. Ratner, J. A. Pople, P. C. Redfern, L. A. Curtiss, 6-31G* basis set for third-row atoms. *J. Comput. Chem.* **22**, 976–984 (2001).

57. T. Le Bahers, C. Adamo, I. Ciofini, A qualitative index of spatial extent in charge-transfer excitations. *J. Chem. Theory Comput.* **7**, 2498–2506 (2011).

58. P. J. M. van Kan *et al.*, Time-resolved spectroscopy at 10 K of the Photosystem II reaction center; deconvolution of the red absorption band. *Biochim. Biophys. Acta Bioenerg.* **1020**, 146–152 (1990).

59. M. Germano *et al.*, Selective replacement of the active and inactive pheophytin in reaction centres of Photosystem II by 13(1)-deoxo-13(1)-hydroxy-pheophytin a and comparison of their 6 K absorption spectra. *Photosynth. Res.* **64**, 189–198 (2000).

60. E. F. Pettersen *et al.*, UCSF Chimera–A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

61. R. Dennington, T. A. Keith, J. M. Millam, *GaussView*, (Semichem Inc., Shawnee Mission, KS, 2016).

62. D. J. Frisch *et al.*, *Gaussian 16, Revision B.01*, (Gaussian Inc., Wallingford, CT, 2016).

63. T. Yanai, D. P. Tew, N. C. Handy, A new hybrid exchange-correlation functional using the coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **393**, 51–57 (2004).

64. R. S. Mulliken, Electronic population analysis on LCAO-MO molecular wave functions. I. *J. Chem. Phys.* **23**, 1833–1840 (1955).