Supporting Information: Environment-Induced Exciton Renormalization in the Photosystem II Reaction Center

Tucker Allen,*,† Barry Y. Li,† Nadine C. Bradbury,‡ and Daniel Neuhauser†

†*Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, CA, 90095, USA*

‡*Department of Chemistry, Princeton University, Princeton, NJ, 08544, USA*

E-mail: tuckerallen27@ucla.edu

Date: February 23, 2026

I. Methodology

We summarize here the key ingredients to the TDHF@v_W method that enable a large-scale comparative study of the isolated and protein-embedded PSII-RC. In Section A, we describe the mixed deterministic/sparse-stochastic approach to efficiently evaluate exchange integrals within a plane-wave representation. In Section B, we briefly discuss how the anisotropic response kernel v_W is generated through short-time stochastic time-dependent Hartree (TDH) dynamics. In Section C, we outline how to obtain the linear-response optical absorption cross-section and exciton amplitudes through an iterative Chebyshev polynomial expansion of spectral operators.

A. Mixed Deterministic and Stochastic Evaluation of Exchange Integrals

In this section, we explain the stochastic-fragmented approach, which is a key ingredient in our method. It allows the calculation of accurate exchange interactions with little numerical effort, by replacing the millions of explicit k-vector calculations with around $10,000-15,000$ "fragments", which are each quite easy to act on.

Within a plane-wave grid-based approach, generation of exchange integrals is costly due to many basis elements. Here we discuss a mixed deterministic/sparse-stochastic approach, developed in Refs.[1–3] and applied in Refs.[4–7] First, given a general exchange interaction kernel, $u(k)$, where k denotes a reciprocal lattice vector, we divide the kernel between a few long-wavelength, low-k components that are evaluated deterministically, and the remainder high-k terms are treated through a sparse-stochastic basis that is system-size independent, yet describes the full exchange interaction.

To achieve this fragmentation of $u(k)$, the following resolution-of-the-identity is introduced:

$$I = \sum_{k_{\text{low}}} |k_{\text{low}}\rangle \langle k_{\text{low}}| + \sum_{k_{\text{high}}} |k_{\text{high}}\rangle \langle k_{\text{high}}| . \tag{1}$$

In this basis a general interaction becomes:

$$\begin{aligned} u(k) = &\sum_{k_{\text{low}}} |k_{\text{low}}\rangle u(k_{\text{low}}) \langle k_{\text{low}}| \\ &+ \sum_{k_{\text{high}}} \sqrt{|u(k_{\text{high}})|} |k_{\text{high}}\rangle \langle k_{\text{high}}| \sqrt{|u(k_{\text{high}})|}. \end{aligned} \tag{2}$$

The key ingredient is then that the high-k space is compressed by using short, "fragmented", random vectors:

$$\alpha(k_{\text{high}}) = \pm\sqrt{\frac{N_{k_{\text{high}}}}{L}} A_\alpha(k_{\text{high}}), \tag{3}$$

where $N_{k_{\text{high}}}$ is the number of high-k terms being sampled, L is the length of each fragment,

and $A_\alpha(k_{\rm high}) = \pm 1$ for $k_{\rm high}$ within the length-L fragment, and 0 outside the fragment. The length L is much smaller than the extent of the $k_{\rm high}$-grid, enabling a large number of fragments to sample efficiently the $k_{\rm high}$-space. The high-k contribution to $u(k)$ is now a set of N_α states, $\{|\zeta\rangle\}$, with components

$$\langle k_{\rm high}|\zeta\rangle = \sqrt{|u(k_{\rm high})|}\ \alpha(k_{\rm high}). \tag{4}$$

The full auxiliary basis, $N_\xi = N_{k_{\rm low}} + N_\alpha$, is then

$$|\xi\rangle = \{\sqrt{|u(k_{\rm low})\,|}\ |k_{\rm low}\rangle\} \oplus \{|\zeta\rangle\}, \tag{5}$$

and the interaction is now a sum over separable terms:

$$u = \sum_\xi |\xi\rangle\langle\xi|. \tag{6}$$

This approach to evaluating exchange integrals is implemented both in the ground-state RSH-DFT stage, using the near-gap hybrid DFT method (detailed in[2]), and the linear-response TDHF@v_W spectral simulations, detailed in.[4] In the present study, the sparse-stochastic basis consists of $N_\alpha = 5,000$ vectors and the total number of $|\xi\rangle$ terms is then between $10,000 - 15,000$. For comparison, a fully deterministic calculation would involve $N_k = 17,808,384$ and $38,223,360$ plane waves for the isolated and protein-embedded PSII-RC, respectively. The mixed deterministic/sparse-stochastic basis therefore corresponds to less than 0.1% of the deterministic plane-wave basis-set size.

B. Stochastic Fitting of Screened Coulomb Operator W

To describe dielectric screening in large biological systems at the random-phase approximation (RPA) level, we evaluate the action of the screened Coulomb operator W on stochastic occupied-occupied pair densities using stochastic TDH propagation. Technical details are

given in Refs.;[8,9] here we focus on the stochastic fitting procedure used to construct the screened-exchange kernel v_W. The $v_W(r-r')$ potential masks atomic details but was found by us to be very accurate for large systems as it gives the overall k-dependent polarization.

An optimal fit of v_W to W is derived by minimizing the following objective:[10]

$$J = \sum_{ij} (\phi_i \phi_j | (W - v_W)^2 | \phi_i \phi_j), \tag{7}$$

where $\{\phi_i\}$ are a set of mean-field occupied molecular orbitals (MOs), obtained from a prior RSH-DFT calculation. Next, we replace the deterministic occupied-occupied MO products with a stochastic realization. We define a stochastic process, and two independent stochastic vectors,

$$\begin{aligned} \bar{\beta}(r) &= \sum_l \bar{\beta}_l \phi_l(r), \\ \bar{\bar{\beta}}(r) &= \sum_l \bar{\bar{\beta}}_l \phi_l(r), \end{aligned} \tag{8}$$

where the sums extend over occupied states and $\bar{\beta}_l = \pm 1$, $\bar{\bar{\beta}}_l = \pm 1$. Using the identity $\left\{\bar{\beta}_i \bar{\beta}_j\right\} = \left\{\bar{\bar{\beta}}_i \bar{\bar{\beta}}_j\right\} = \delta_{ij}$, where curly brackets denote an average over many stochastic instances, it follows that

$$\left\{\bar{\beta}_i \bar{\bar{\beta}}_j \beta(r)\right\} = \phi_i(r)\phi_j(r), \tag{9}$$

where $\beta(r) \equiv \bar{\beta}(r)\bar{\bar{\beta}}(r)$. Upon replacement of the deterministic pair-densities with $\beta(r)$ and moving to reciprocal space, it can be shown that:

$$v_W(k) = \left\{ \frac{\beta^*(k) \langle k|W|\beta\rangle}{|\langle k|\beta\rangle|^2} \right\}_\beta, \tag{10}$$

where $\beta(k)$ is the momentum-space representation of a random occupied-occupied pair density. The action $\langle k|W|\beta\rangle$ is first calculated in real-space, with a short-time stochastic TDH propagation. For further technical details on the real-time dynamics, we refer to Refs.[5,11]

For the respective isolated and protein-embedded PSII-RC systems, the stochastic fit captures 76.8% and 84.3% of the overall screened interaction W. The residual, $\langle (W - v_W)^2 \rangle / \langle W^2 \rangle$, decreases from 23.2% to 15.7% upon protein embedding, indicating that the translationally invariant representation becomes more accurate in the larger system.

C. Iterative Chebyshev Approach to Optical Spectra and Exciton Amplitudes

We compute optical spectra and frequency-resolved exciton amplitudes using an iterative Chebyshev polynomial expansion of the full TDHF@v_W Liouvillian. The full Liouvillian satisfies the following eigenvalue problem:[12]

$$\mathcal{L} \begin{pmatrix} f^+ \\ f^- \end{pmatrix} = \hbar\omega \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} f^+ \\ f^- \end{pmatrix} \tag{11}$$

where the two-particle Liouvillian is expressed as a block matrix:

$$\mathcal{L} = \begin{pmatrix} A & B \\ -B & -A \end{pmatrix}, \tag{12}$$

and f^+, f^- are the positive- and negative-frequency transition eigenvectors, respectively. For spin-singlet excitons, the A and B matrices are:

$$\begin{aligned} A_{ia,jb} &= (\varepsilon_a - \varepsilon_i)\delta_{ij}\delta_{ab} + 2(ia|jb) - (\phi_a\phi_b|v_W|\phi_i\phi_j) \\ B_{ia,bj} &= 2(ia|bj) - (\phi_a\phi_j|v_W|\phi_i\phi_b), \end{aligned} \tag{13}$$

where the electron repulsion integrals are:

$$(ia|jb) = \int \phi_i(r)\phi_a(r)|r - r'|^{-1}\phi_j(r')\phi_b(r')dr'dr, \tag{14}$$

and $(ia|jb)$=$(ia|bj)$. The screened exchange integrals for the diagonal portion of $\mathcal{L}$ are:

$$(\phi_a\phi_b|v_W|\phi_i\phi_j) = \int \phi_a(r)\phi_b(r)v_W(r-r')\phi_i(r')\phi_j(r')dr'dr, \tag{15}$$

with analogous elements for the off-diagonal coupling part. We have introduced valence $\varepsilon_{i,j,...}$ (occupied) and conduction $\varepsilon_{a,b,...}$ (virtual) energies and associated real-valued orbitals from a preceding RSH-DFT calculation. In this work, we use the CAM-LDA0 hybrid functional,[13] with hybrid-exchange parameters: $\alpha = 0.19, \beta = 0.46, \gamma = 0.33$ Bohr^{-1}.

To iteratively solve for the excitation energies and associated amplitudes of $\mathcal{L}$, we define an initial exciton vector (assuming laser polarization in the x-direction for simplicity):

$$|\chi_{ia}\rangle = \begin{pmatrix} f_{ia}^{+} \\ f_{ia}^{-} \end{pmatrix} = \begin{pmatrix} +\langle\phi_a|x|\phi_i\rangle \\ -\langle\phi_a|x|\phi_i\rangle \end{pmatrix}. \tag{16}$$

Then, the absorption cross-section is obtained via

$$\sigma(\omega) \propto \omega\langle\chi_{ia}|\delta(\mathcal{L}-\omega)|\chi_{ia}\rangle, \tag{17}$$

where one evaluates the delta function by finding the Chebyshev residues:

$$R_n^{\pm} = \sum_{ia}\langle f_{ia}^{\pm}|T_n(\tilde{\mathcal{L}})|f_{ia}^{\pm}\rangle, \tag{18}$$

with $T_n(\tilde{\mathcal{L}})$ being the n-th Chebyshev polynomial, and $\tilde{\mathcal{L}}$ the scaled Liouvillian with eigenvalues between the interval $[-1,+1]$. For numerical details on the matrix-vector applications, we refer to Ref.[4] One then can write the correlation function as:

$$\sigma(\omega) = \sum_{n=0}^{N_{\text{Cheb}}} c_n(\omega)(R_n^{+} + R_n^{-}), \tag{19}$$

where $c_n(\omega)$ are smoothly-decaying numerical coefficients[14] and the Chebyshev length, N_{Cheb},

is set to 2000 terms. This value for N_{Cheb} results in a peak broadening of roughly 0.1 eV for absorption spectra.

In addition to the Chebyshev residues used to construct $\sigma(\omega)$, the same recursion yields frequency-resolved exciton amplitudes at essentially no additional cost. During the Chebyshev propagation we generate the vectors $|u_n\rangle \equiv T_n(\tilde{\mathcal{L}})|\chi\rangle$ for $n = 0, \ldots, N_{\mathrm{Cheb}}$. For sampling a specific frequency ω, a filtered exciton vector is then obtained as $|f(\omega)\rangle \equiv \delta(\mathcal{L} - \omega)|\chi\rangle \approx \sum_{n=0}^{N_{\mathrm{Cheb}}} c_n(\omega)|u_n\rangle$, using the same Chebyshev coefficients $c_n(\omega)$ used for the absorption cross-section. The components $f_{ia}^{\pm}(\omega)$ provide the exciton transition amplitudes in the valence-conduction basis and are normalized such that $\sum_{ia} |f_{ia}(\omega)|^2 = 1$.

From the normalized amplitudes we define the real-space transition density $\rho(r;\omega) = \sum_{ia}[f_{ia}^{+}(\omega) + f_{ia}^{-}(\omega)]\phi_i(r)\phi_a(r)$. Exciton delocalization is quantified via the participation ratio, $\mathrm{PR}(\omega) = 1/\sum_{ia} |f_{ia}(\omega)|^4$.

II. Workflow of PSII-RC Geometry Extraction

The structural model of the PSII-RC, based on chlorophyll-a (CLA) and its surrounding protein environment, is derived from an experimentally resolved PSII crystal structure reported in Ref.[15] Starting from the full biological PDB, all CLA residues associated with the hexameric pigment site are identified,[15] and a local protein environment is extracted using a radial cutoff that retains only residues in direct spatial proximity to the pigments. This procedure preserves the native relative positions and orientations of the chromophores while reducing the system size to a computationally tractable cluster.

The resulting coordinates are further processed to ensure chemical completeness, including fragment cleanup and hydrogen addition. For the multi-chlorin system, we employ a QM-optimized structure in which the internal geometries of the pigment macrocycles are replaced by high-level QM-optimized structures taken from prior reference calculations by Förster and Visscher,[16] while preserving their positions within the protein scaffold. The re-

maining protein environment is subsequently relaxed using the extended tight-binding force-field approach (xTBFF) as implemented in the ORCA 6.0 package,[17,18] with the pigment "core" held fixed.

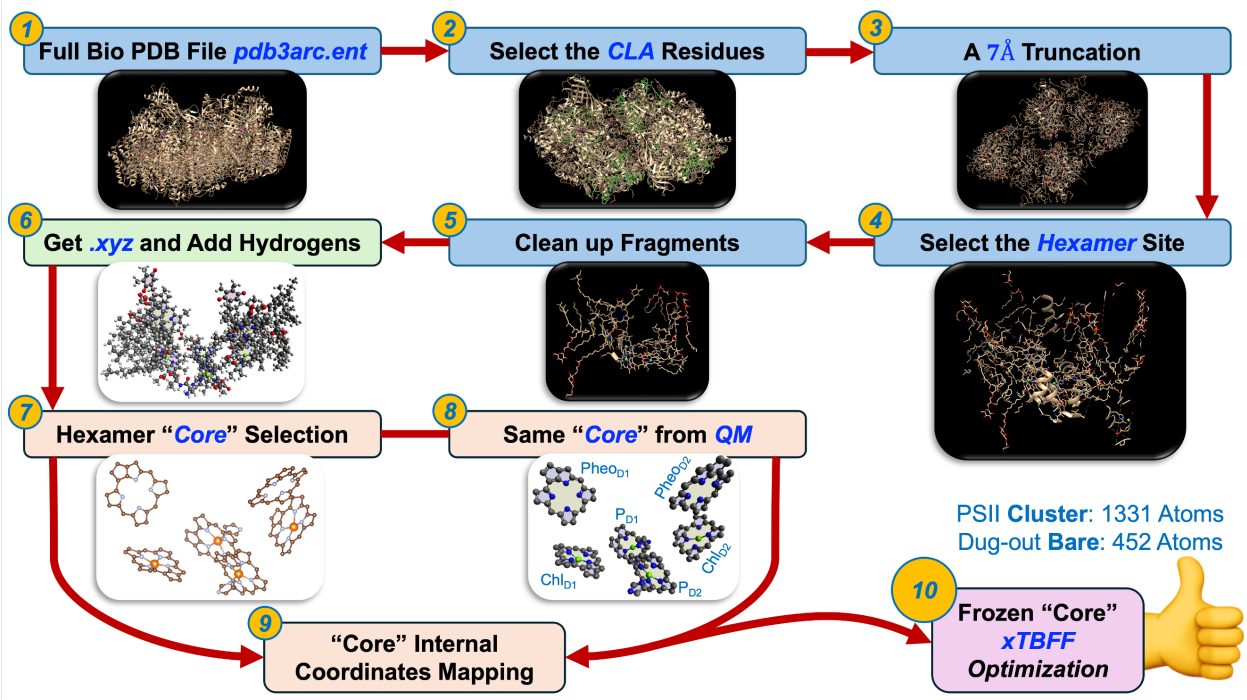

Figure 1: Schematic workflow for constructing the isolated and protein-embedded PSII-RC from the full experimental PDB structure, including pigment selection, protein truncation, core replacement with quantum-optimized geometries, and environment relaxation using xTBFF with a frozen pigment core.

Fig. 1 summarizes the construction and preparation of a protein-embedded Photosystem II reaction center (PSII-RC). Starting from a full experimental PDB structure, the workflow isolates the relevant pigment-protein region, reduces the system size through spatial truncation, enforces chemical consistency, and finally combines high-level quantum-optimized pigment geometries with a low-cost force-field relaxation of the environment using xTBFF,[17] while keeping the chromophore core frozen. The detailed workflow is listed below:

- **Step 1: Biological PDB structure.** The workflow begins from the experimentally resolved PSII structure (`pdb3arc.ent`) from Ref.[15] and its online Data Deposits. This file contains the complete protein complex, cofactors, pigments, and associated

residues, represented in crystallographic coordinates. At this stage, the system is chemically heterogeneous, extremely large ($\sim 55,000$ atoms), and not directly suitable for quantum calculations.

- **Step 2: Selection of CLA (chlorophyll-a) residues.** From the full PDB, all CLA residues corresponding to chlorophyll-a pigments are identified and selected. This step isolates the electronic subsystem of interest while preserving the original pigment positions and orientations relative to the protein scaffold.

- **Step 3: Spatial truncation (7-Å cutoff).** A 7-Å radial truncation is applied around the selected CLA pigments. All protein residues and cofactors with at least one atom within this cutoff are retained, while the remainder of the PSII complex is removed. This step reduces the system size while preserving the local electrostatic and steric environment that directly influences pigment geometry and electronic structure. The truncation is purely geometric and does not yet enforce chemical capping or saturation.

- **Step 4: Selection of the hexamer site.** From the truncated region, the specific chlorophyll hexamer site is selected. This hexamer corresponds to the functionally relevant pigment cluster responsible for excitonic and charge-transfer processes in PSII. The surrounding protein residues retained at this stage provide an explicit, local environment for the pigments, resulting in a mixed pigment-protein cluster.

- **Step 5: Fragment cleanup and chemical repair.** Truncation inevitably breaks covalent bonds within protein residues. In this step, all dangling bonds are identified, fragmented residues are capped or repaired, and obvious steric clashes and crystallographic artifacts are removed.

- **Step 6: XYZ generation and hydrogen addition.** The cleaned protein-embedded PSII-RC is converted from PDB to XYZ format. Hydrogen atoms are added to all other

atoms according to standard valence rules, yielding a chemically complete structure. At this point, the protein-embedded PSII-RC contains 1331 atoms, including pigments, protein residues, and hydrogens.

- **Step 7: Definition of the hexamer "core".** A core region of 186 atoms is defined, consisting exclusively of the chlorophyll macrocycles and their central Mg atoms (and 2 directly bonded substituents). This core represents the electronic subsystem to be treated at high quantum-chemical accuracy.

- **Step 8: Replacement with a high-level QM-optimized "core".** The geometries of the 186-atom chlorophyll core are replaced by high-level QM-optimized structures taken from an independent Ref.[19] This step ensures that the internal bond lengths, angles, and macrocycle planarity of each chlorophyll are described at a reliable quantum level; errors introduced by force-field or low-level optimization of the pigments are eliminated. The chromophore labeling in Fig. 1 follows the convention of Ref.[20]

- **Step 9: Internal coordinate mapping.** After replacement, a one-to-one mapping of internal coordinates is carried out between the original and substituted cores. This guarantees consistency of atom ordering, bonding topology, and connectivity between the pigment core and the surrounding environment. The mapping step is essential to avoid artificial strain or discontinuities at the pigment–protein boundary.

- **Step 10: xTBFF optimization with frozen "core".** Finally, the full protein-embedded PSII-RC is optimized using xTBFF (extended tight-binding force field),[17,18] while freezing all atoms in the 186-atom chlorophyll core. Only the protein environment and peripheral atoms are allowed to relax.

The extended tight-binding framework (xTB), interfaced with ORCA 6.0, is a semiempirical quantum method that solves a simplified electronic Schrödinger equation using parameters fitted to DFT reference data. As a result, xTB retains an explicit electronic struc-

ture description, including electron density, self-consistent atomic charges, and orbital energies, while naturally allowing for bond formation and bond breaking without predefined connectivity.[17,18] The xTB force-field mode (xTBFF) builds on this quantum-mechanical foundation while adopting a force-field-like representation for large-scale structural relaxation. Although xTBFF organizes interactions in terms of bond stretches, angle bends, and torsional terms, the associated parameters are not fixed but are dynamically derived from the underlying xTB electronic structure.[17] Electrostatic interactions are governed by self-consistent charges rather than predefined atom types. Consequently, xTBFF preserves chemically meaningful bonding information and environmental responsiveness that are absent in classical force fields, which rely on static atom types, fixed parameters, and manual reparameterization.[17] Both coordinate sets (in Å) are provided in the additional supporting materials as `psiirc_embedded.xyz` and `psiirc_isolated.xyz`.

III. Basis-set Extrapolation for Optical Spectra

The optical spectra provided in the main text use a transition space of $N_v = 200$ valence orbitals and $N_c = 400$ conduction orbitals to construct the two-particle Liouvillian. For both the isolated and embedded PSII-RC models, this exciton basis-set size is sufficient to converge the spectral shapes below 3 eV. We have confirmed that increasing N_c beyond 400 does not affect either the spectral shapes or peak positions for low-energy (< 3 eV) excitations. The consistent exciton basis-set size also allows us to directly compare the transition densities and participation ratios of each system.

We include a rigid shift in the reported spectra by performing a series of calculations for different N_v and N_c. Here, we perform a simple linear extrapolation to converge the absolute peak positions of spectra. Table 1 shows peak positions as a function of N_v, where we vary the number of valence orbitals while doubling the number of conduction orbitals for each selected N_v, i.e., $N_c = 2N_v$. This convention is arbitrary, and we have confirmed for both

systems that spectra are converged within 0.01 eV even when going beyond $N_c = 2N_v$ for all low-energy (< 3 eV) excitations.

Table 1: Onset absorption peak energies calculated using the TDHF@v_W method with different values of N_v and N_c for the isolated and protein-embedded PSII-RC at different polarization directions. Note that $N_c = 2N_v$ holds in all cases.

	N_v	N_c	**Isolated Peak (eV)**	**Embedded Peak (eV)**
x**-pol**	50	100	-	1.84
	100	200	1.79	1.82
	150	300	1.73	1.82
	200	400	1.71	1.79
	300	600	1.68	1.77
y**-pol**	50	100	-	1.90
	100	200	1.89	1.90
	150	300	1.87	1.87
	200	400	1.84	1.87
	300	600	1.82	1.84
	400	800	1.79	-
z**-pol**	50	100	-	1.84
	100	200	1.76	1.82
	150	300	1.73	1.82
	200	400	1.71	1.79
	300	600	1.68	1.79

The linear extrapolation is performed with the target function, $\omega = a(1/N_v) + b$, where ω is the onset absorption energy in eV, a and b are regression parameters. Table 2 shows the linear extrapolation results and the predicted peaks at $N_v \to N_v^{\max} = N_{\text{occ}}$.

Table 2: Results from linear extrapolation for the isolated and protein-embedded PSII-RC at different polarization directions.

System	Polarization	a	b	$N_v = N_v^{\max}$	**Peak at $N_v^{\max}$ (eV)**
Isolated	x-pol	16.3	1.63	638	**1.65**
	y-pol	12.6	1.77	638	**1.79**
	z-pol	11.7	1.65	638	**1.67**
Embedded	x-pol	7.54	1.75	1619	**1.76**
	y-pol	8.23	1.82	1619	**1.82**
	z-pol	5.14	1.77	1619	**1.78**

In addition to converging with respect to N_v, we also converge the number of low-k vectors, $N_{k_{\text{low}}}$, used to evaluate the screened exchange matrix elements entering $\mathcal{L}$. Ex-

change is evaluated in a plane-wave basis; in the fully deterministic limit this corresponds to $(2)^3 N_x N_y N_z$ reciprocal vectors, where the $(2)^3$ comes from doubling the real-space cell in each direction using the Martyna–Tuckerman construct to minimize grid-reflection effects.[21]

To check convergence of peaks (i)-(iv) in the main text, we increased $N_{k_{\mathrm{low}}}$ (with fixed $N_v = 100$, $N_c = 200$) from $4683 \rightarrow 37523$ for the isolated PSII-RC and from $10091 \rightarrow 41299$ for the embedded PSII-RC, and observed no discernible change in the y-polarized spectra; we therefore consider the results converged with respect to $N_{k_{\mathrm{low}}}$ at the values used in the main text. The total auxiliary basis is $N_\xi = N_{k_{\mathrm{low}}} + N_\alpha$, with $N_\alpha = 5000$ sparse vectors, giving a total of $10,000 - 15,000$ auxiliary vectors. A fully deterministic plane-wave calculation would require millions of basis elements; our approach achieves comparable accuracy with only thousands.

IV. Dominant Transitions in Isolated and Protein-embedded PSII-RC

This section lists the dominant valence-to-conduction orbital transition contributions for the low-energy excitations of the isolated and protein-embedded PSII-RC, along with the corresponding weights and participation ratios.

Table 3: Contributions of individual valence (VMO) and conduction (CMO) molecular orbitals to the first x-polarized excitation for the isolated PSII hexamer (Isolated) and the protein-embedded (Embedded) PSII-RC. The numbers in the VMO and CMO columns indicate the MO index relative to the HOMO (0) or LUMO (0). $|A|^2$ denotes the transition amplitude of each configuration (only dominant transitions with normalized $|A|^2 \geq 0.01$ are shown). PR denotes the participation ratio of the exciton, with larger values indicating contributions from a greater number of VMO to CMO transitions and increased exciton delocalization in the transition basis.

Isolated Peak 1 x-pol (1.66 eV)			**Embedded Peak 1 x-pol (1.75 eV)**		
VMO	CMO	$\|A\|^2$	VMO	CMO	$\|A\|^2$
0 (HOMO)	+5	0.029	0 (HOMO)	+5	0.019
−1	+6	0.046	0 (HOMO)	+7	0.015
−1	+11	0.333	0 (HOMO)	+12	0.027
−2	+11	0.061	−1	+5	0.064
−6	+9	0.016	−2	+2	0.122
−6	+11	0.377	−2	+10	0.016
−6	+16	0.035	−3	+5	0.021
−7	+17	0.017	−5	+5	0.023
−8	+11	0.018	−5	+12	0.044
			−7	0 (LUMO)	0.124
			−7	+6	0.118
			−9	+6	0.013
			−10	+2	0.015
			−10	+11	0.076
			−11	0 (LUMO)	0.144
			−11	+6	0.076
PR = 3.8			**PR = 11.7**		

Table 4: Molecular orbital transition analysis for the first y-polarized excited-state.

Isolated Peak 1 y-pol (1.79 eV)			**Embedded Peak 1 y-pol (1.82 eV)**		
VMO	CMO	$\|A\|^2$	VMO	CMO	$\|A\|^2$
0 (HOMO)	+4	0.018	0 (HOMO)	+5	0.028
0 (HOMO)	+5	0.091	0 (HOMO)	+7	0.174
−1	+5	0.143	−1	+7	0.018
−1	+11	0.106	−1	+13	0.112
−2	+5	0.012	−3	+7	0.041
−2	+11	0.061	−3	+13	0.094
−4	+5	0.011	−4	+3	0.201
−5	+5	0.063	−5	+3	0.010
−5	+11	0.034	−5	+7	0.024
−6	+3	0.036	−6	+4	0.016
−6	+9	0.043	−8	+9	0.014
−7	0 (LUMO)	0.068	−9	+3	0.019
−7	+6	0.021	−9	+11	0.141
−8	+3	0.061			
−8	+9	0.029			
−11	0 (LUMO)	0.028			
−11	+6	0.046			
PR = 15.3			**PR = 8.6**		

Table 5: Molecular orbital transition analysis for the first z-polarized excited-state.

Isolated Peak 1 z-pol (1.67 eV)			**Embedded Peak 1 z-pol (1.77 eV)**		
VMO	CMO	$\|A\|^2$	VMO	CMO	$\|A\|^2$
−3	+5	0.289	−1	+5	0.021
−3	+12	0.024	−2	+2	0.110
−4	+6	0.021	−2	+10	0.014
−4	+11	0.156	−5	+12	0.014
−5	+11	0.029	−7	0 (LUMO)	0.172
−8	+12	0.010	−7	+6	0.165
−9	+5	0.177	−9	+6	0.018
−9	+11	0.017	−10	+2	0.014
−10	+5	0.023	−10	+11	0.068
−10	+12	0.174	−11	0 (LUMO)	0.200
			−11	+6	0.106
PR = 5.8			**PR = 7.9**		

Table 6: Molecular orbital transition analysis for the x-polarized "hump" transition.

Isolated Hump x-pol (1.79 eV)			**Embedded Hump x-pol (1.83 eV)**		
VMO	CMO	$\|A\|^2$	VMO	CMO	$\|A\|^2$
0 (HOMO)	+5	0.054	0 (HOMO)	+7	0.035
−1	+5	0.079	−1	+13	0.021
−1	+11	0.056	−3	+3	0.012
−2	+11	0.034	−3	+13	0.018
−5	+3	0.011	−4	+3	0.433
−5	+5	0.033	−4	+11	0.023
−5	+9	0.013	−5	+3	0.020
−5	+11	0.019	−9	+3	0.042
−6	+3	0.158	−9	+11	0.296
−6	+9	0.151			
−8	+3	0.179			
−8	+9	0.105			
PR = 9.4			**PR = 3.6**		

Table 7: Molecular orbital transition analysis for the right shoulder ("hump") z-polarized excited state.

Isolated Hump z-pol (1.78 eV)			**Embedded Hump 1 z-pol (1.84 eV)**		
VMO	CMO	$\|A\|^2$	VMO	CMO	$\|A\|^2$
−4	0 (LUMO)	0.011	−3	+3	0.013
−5	0 (LUMO)	0.025	−4	+3	0.477
−7	0 (LUMO)	0.357	−4	+11	0.027
−7	+6	0.129	−5	+3	0.022
−9	0 (LUMO)	0.014	−9	+3	0.046
−9	+6	0.010	−9	+11	0.323
−11	0 (LUMO)	0.138			
−11	+2	0.010			
−11	+6	0.213			
PR = 4.8			**PR = 3.0**		

Table 8: Molecular orbital transition analysis for the second x-polarized excited state.

Isolated Peak 2 x-pol (2.01 eV)			**Embedded Peak 2 x-pol (2.02 eV)**		
VMO	CMO	$\|A\|^2$	VMO	CMO	$\|A\|^2$
-3	$+8$	0.048	-2	$+1$	0.077
-4	0 (LUMO)	0.017	-2	$+8$	0.044
-4	$+2$	0.338	-2	$+10$	0.329
-4	$+7$	0.038	-10	$+1$	0.416
-5	$+2$	0.063	-10	$+4$	0.031
-6	$+3$	0.012	-10	$+10$	0.024
-8	$+3$	0.011			
-9	$+2$	0.039			
-9	$+7$	0.251			
-10	$+1$	0.066			
-11	$+7$	0.011			
PR = 5.2			**PR = 3.4**		

Table 9: Molecular orbital transition analysis for the second y-polarized excited state.

Isolated Peak 2 y-pol (1.98 eV)			**Embedded Peak 2 y-pol (2.01 eV)**		
VMO	CMO	$\|A\|^2$	VMO	CMO	$\|A\|^2$
-3	$+1$	0.044	-2	$+2$	0.077
-3	$+8$	0.276	-2	$+9$	0.044
-4	$+2$	0.015	-2	$+11$	0.328
-6	$+3$	0.012	-10	$+2$	0.412
-6	$+9$	0.012	-10	$+5$	0.034
-7	0 (LUMO)	0.020	-10	$+11$	0.024
-7	$+6$	0.044			
-8	$+1$	0.018			
-8	$+3$	0.015			
-10	$+1$	0.378			
-10	$+8$	0.012			
-11	0 (LUMO)	0.046			
-11	$+6$	0.014			
PR = 4.4			**PR = 3.5**		

Table 10: Molecular orbital transition analysis for z-polarized excited state.

Isolated Peak 2 z-pol (1.99 eV)			**Embedded Peak 2 z-pol (2.04 eV)**		
VMO	CMO	$\|A\|^2$	VMO	CMO	$\|A\|^2$
-3	$+1$	0.039	-2	-2	0.078
-3	$+8$	0.247	-2	$+5$	0.045
-6	$+3$	0.048	-2	$+7$	0.334
-6	$+9$	0.029	-10	-2	0.421
-7	0 (LUMO)	0.016	-10	$+1$	0.035
-7	$+6$	0.033	-10	$+7$	0.025
-8	$+1$	0.030			
-8	$+3$	0.030			
-8	$+9$	0.041			
-10	$+1$	0.326			
-10	$+3$	0.013			
-10	$+8$	0.012			
-11	0 (LUMO)	0.037			
-11	$+6$	0.012			
PR = 5.6			**PR = 3.4**		

Table 11: Molecular orbital transition analysis for the third peak. Only the y-polarized data is shown as the spectra is isotropic for the isolated and protein-embedded PSII-RC individually.

Isolated Peak 3 y-pol (2.29 eV)			**Embedded Peak 3 y-pol (2.23 eV)**		
VMO	CMO	$\|A\|^2$	VMO	CMO	$\|A\|^2$
0 (HOMO)	0 (LUMO)	0.014	0 (HOMO)	$+5$	0.400
0 (HOMO)	$+5$	0.565	0 (HOMO)	$+7$	0.026
0 (HOMO)	$+10$	0.052	-1	$+5$	0.312
0 (HOMO)	$+11$	0.039	-1	$+12$	0.026
0 (HOMO)	$+16$	0.026	-1	$+13$	0.012
-1	$+4$	0.022	-3	$+5$	0.018
-1	$+5$	0.103	-5	$+5$	0.108
-1	$+11$	0.011	-6	$+5$	0.011
-2	$+4$	0.012			
-2	$+5$	0.052			
-5	$+5$	0.031			
PR = 3.0			**PR = 3.7**		

V. Frontier Molecular Orbitals

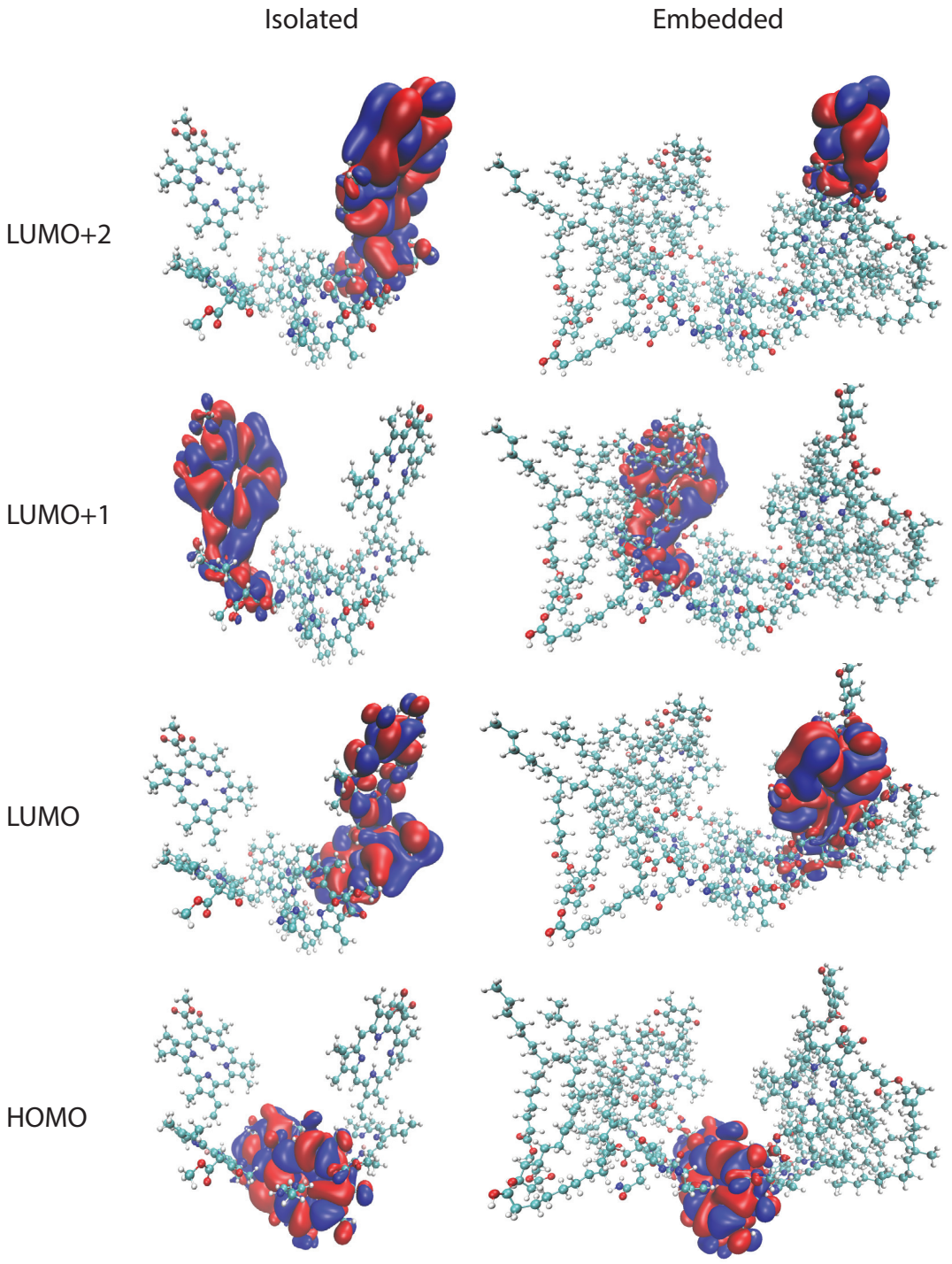

Figure 2: Molecular orbital plots for the isolated PSII-RC (left) and protein-embedded PSII-RC (right) using the CAM-LDA0 RSH functional. Isovalues of ± 0.002 Bohr$^{-3/2}$ are used for all orbitals with blue used for positive values and red for negative values.

The frontier molecular orbitals (MOs) for the isolated and protein-embedded PSII-RC systems are provided in Fig. 2. The CAM-LDA0 DFT bandgap for the isolated PSII-RC is 3.51 eV, and it decreases to 3.41 eV when the complex is placed in the protein environment. Overall, the environment has a stabilizing effect that pushes the single-particle spectrum to lower energies relative to the vacuum hexamer complex. The frontier orbitals remain predominantly chromophoric in character with partial hybridization to environmental states. Notably, in the protein-embedded PSII-RC the LUMO+2 orbital is localized on the Q_B cofactor. This implies that the environment does not simply shift energies, but induces reordering of electronic states.

VI. Parameterized Results for $\epsilon^{-1}(k)$

The values reported in Table 12 are obtained from an isotropic fit to the k-space inverse dielectric function, $\epsilon^{-1}(k)$, derived from the screened exchange kernel $v_W(k)$. A graphical illustration of the fit is shown in Fig. 3. Definitions of the fitting parameters are given in Ref.[6]

Table 12: The resulting sets of the 7 fitted parameters of $\epsilon^{-1}(k)$ for isolated and protein-embedded PSII-RC systems.

	Isolated PSII-RC	**Embedded PSII-RC**
c_0	0.639	1.087
c_1	−2.875	−4.241
c_2	4.372	5.532
c_3	−2.678	−2.877
c_4	0.010	0.003
k_{mt}	0.820	0.740
γ	0.260	0.260

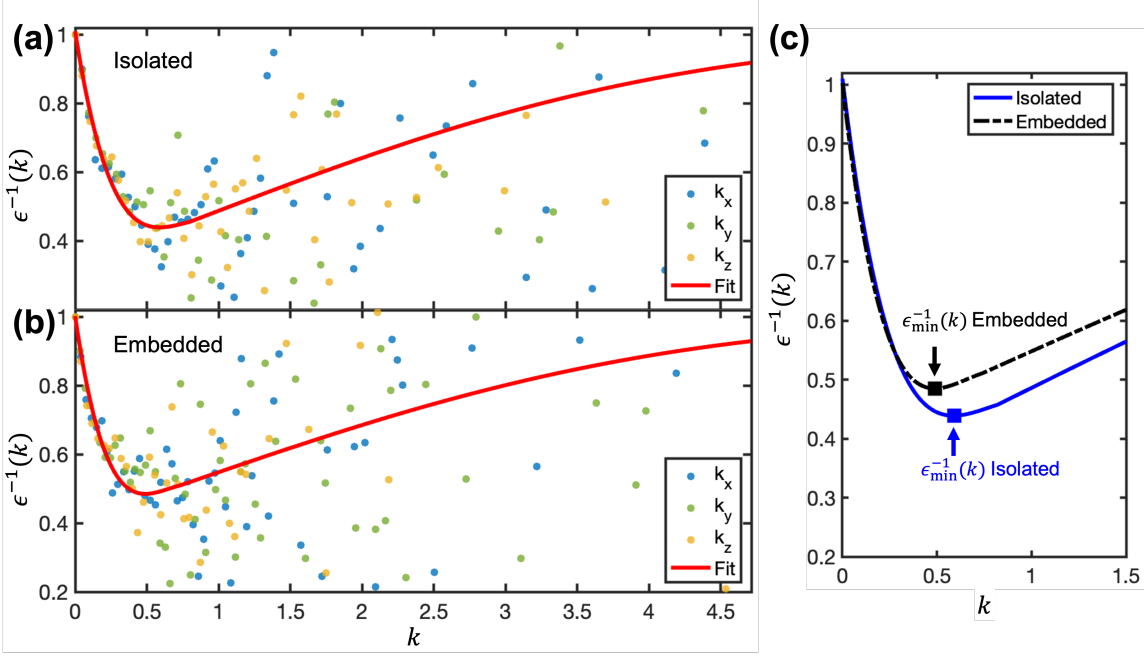

Figure 3: $\epsilon^{-1}(k)$ along $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$ directions (dots) for the isolated **(a)** and protein-embedded PSII-RC **(b)**, k is in units of Bohr^{-1}, and the parameterization is in red lines. **(c)** shows the overlaid parameterizations, with the minimum k-space inverse dielectric value, $\epsilon^{-1}_{\min}(k) \equiv \min[\epsilon^{-1}(k)]$, for the isolated and protein-embedded PSII-RC labeled. These minima correspond to the real-space distances of 5.64 Å for the isolated system and 6.79 Å for the embedded system.

VII. Comparison of Full and TDA TDHF@v_W Spectra

To analyze the role of resonant-antiresonant coupling in large photosynthetic complexes, we provide TDA (Tamm-Dancoff approximation) spectra (Fig. 4) for both the isolated and embedded PSII-RC models. Within the TDA, the coupling matrix B in the Liouvillian operator is neglected. This halves the spectral range and requires half the number of required Chebyshev terms, with 1000 polynomials used in all TDA simulations.

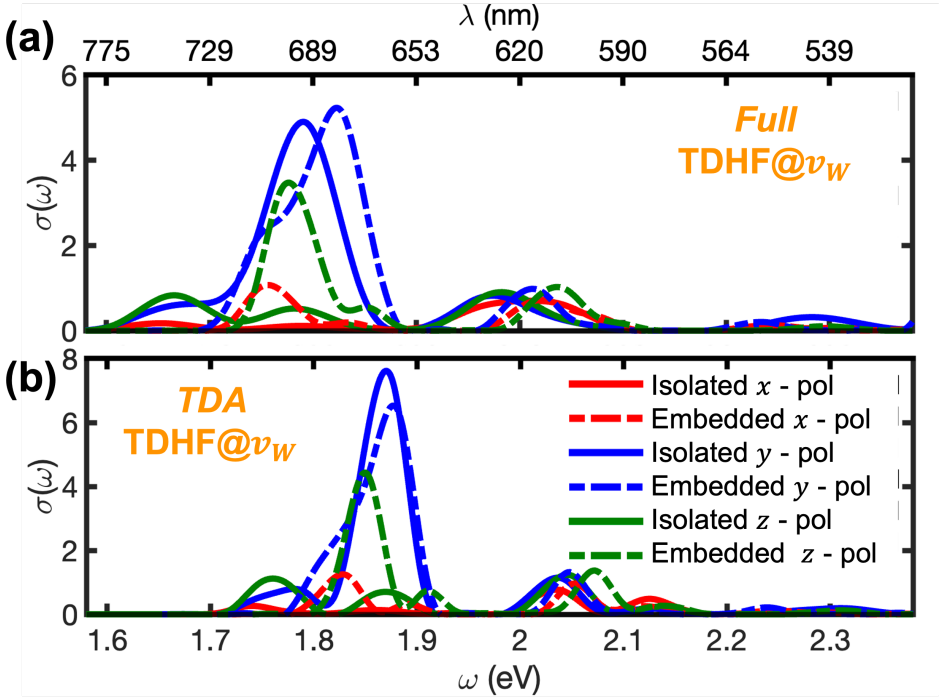

Figure 4: Full TDHF@v_W **(a)** and TDA **(b)** spectra for the isolated and protein-embedded PSII-RC. The same basis-set extrapolation parameters are applied to the TDA spectra.

References

1. Vlček, V.; Li, W.; Baer, R.; Rabani, E.; Neuhauser, D. Swift GW beyond 10,000 electrons using sparse stochastic compression. *Physical Review B* **2018**, *98*, 075107.

2. Bradbury, N. C.; Allen, T.; Nguyen, M.; Neuhauser, D. Deterministic/Fragmented-Stochastic Exchange for Large-Scale Hybrid DFT Calculations. *Journal of Chemical Theory and Computation* **2023**, *19*, 9239–9247.

3. Allen, T.; Li, B. Y.; Duong, T.; Williams, K.; Neuhauser, D. Efficient plane-wave approach to generalized Kohn-Sham density functional theory of solids with mixed deterministic and stochastic exchange. *Phys. Rev. B* **2025**, *112*, 155104.

4. Sereda, M.; Allen, T.; Bradbury, N. C.; Ibrahim, K. Z.; Neuhauser, D. Sparse-Stochastic

Fragmented Exchange for Large-Scale Hybrid Time-Dependent Density Functional Theory Calculations. *Journal of Chemical Theory and Computation* **2024**, *20*, 4196–4204.

5. Bradbury, N. C.; Li, B. Y.; Allen, T.; Caram, J. R.; Neuhauser, D. No more gap-shifting: Stochastic many-body-theory based TDHF for accurate theory of polymethine cyanine dyes. *The Journal of Chemical Physics* **2024**, *161*, 141101.

6. Li, B. Y.; Duong, T.; Allen, T.; Bradbury, N. C.; Caram, J. R.; Neuhauser, D. Parameterized attenuated exchange for generalized TDHF@v_W applications. *The Journal of Chemical Physics* **2025**, *163*, 034102.

7. Chen, K.; Li, B. Y.; Allen, T.; Neuhauser, D. Mixed Planewave and Localized Orbital Basis for Sparse-Stochastic Hybrid Time-Dependent Density Functional Theory. *Journal of Chemical Theory and Computation* **2025**, *21*, 8509–8517.

8. Neuhauser, D.; Gao, Y.; Arntsen, C.; Karshenas, C.; Rabani, E.; Baer, R. Breaking the Theoretical Scaling Limit for Predicting Quasiparticle Energies: The Stochastic GW Approach. *Physical Review Letters* **2014**, *113*, 076402.

9. Bradbury, N. C.; Nguyen, M.; Caram, J. R.; Neuhauser, D. Bethe–Salpeter equation spectra for very large systems. *The Journal of Chemical Physics* **2022**, *157*, 031104.

10. Bradbury, N. C.; Allen, T.; Nguyen, M.; Ibrahim, K. Z.; Neuhauser, D. Optimized attenuated interaction: Enabling stochastic Bethe-Salpeter spectra for large systems. *The Journal of Chemical Physics* **2023**, *158*, 154104.

11. Allen, T.; Nguyen, M.; Neuhauser, D. GW with hybrid functionals for large molecular systems. *The Journal of Chemical Physics* **2024**, *161*, 114116.

12. Negele, J. W. The mean-field theory of nuclear structure and dynamics. *Reviews of Modern Physics* **1982**, *54*, 913–1015.

13. Yanai, T.; Tew, D. P.; Handy, N. C. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chemical Physics Letters* **2004**, *393*, 51–57.

14. Weiße, A.; Wellein, G.; Alvermann, A.; Fehske, H. The kernel polynomial method. *Rev. Mod. Phys.* **2006**, *78*, 275–306.

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

16. Förster, A.; Visscher, L. Quasiparticle Self-Consistent GW-Bethe–Salpeter Equation Calculations for Large Chromophoric Systems. *Journal of Chemical Theory and Computation* **2022**, *18*, 6779–6793.

17. Bannwarth, C.; Caldeweyher, E.; Ehlert, S.; Hansen, A.; Pracht, P.; Seibert, J.; Spicher, S.; Grimme, S. Extended tight-binding quantum chemistry methods. *WIREs Computational Molecular Science* **2021**, *11*, e1493.

18. Neese, F. The ORCA program system. *WIRES Comput. Molec. Sci.* **2012**, *2*, 73–78.

19. Förster, A.; Visscher, L. Quasiparticle Self-Consistent GW-Bethe-Salpeter Equation Calculations for Large Chromophoric Systems. *Journal of Chemical Theory and Computation* **2022**,

20. Sirohiwal, A.; Neese, F.; Pantazis, D. A. Protein Matrix Control of Reaction Center Excitation in Photosystem II. *Journal of the American Chemical Society* **2020**, *142*, 18174–18190.

21. Martyna, G. J.; Tuckerman, M. E. A reciprocal space based method for treating long range interactions in ab initio and force-field-based calculations in clusters. *The Journal of Chemical Physics* **1999**, *110*, 2810–2821.