Supporting information for:

Rhodopsin Absorption from First Principles:

Bypassing Common Pitfalls

Omar Valsson,[†] Pablo Campomanes,[‡] Ivano Tavernelli,[‡] Ursula Rothlisberger,[*,‡]

and Claudia Filippi[*,†]

MESA+ Institute for Nanotechnology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands, and Laboratory of Computational Chemistry and Biochemistry, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

E-mail: ursula.roethlisberger@epfl.ch; c.filippi@utwente.nl

[*]To whom correspondence should be addressed

[†]MESA+ Institute for Nanotechnology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

[‡]Laboratory of Computational Chemistry and Biochemistry, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

1 Computational Details

1.1 ZINDO, TDDFT, and CC2 calculations

We employ the Gaussian 09 code[S1] to perform the ZINDO calculations, using the default settings, and the TDDFT calculations with the CAM-B3LYP[S2] and LC-ωPBE[S3,S4] functionals, and the 6-31+G* basis set. For the CC2 calculations, we use the TURBOMOLE code[S5] where we employ the resolution-of-the-identity (RI) approximation[S6] and the ANO-L-VDZP[S7] basis set. Since a corresponding RI auxiliary basis set is not available for the ANO-L-VDZP basis set, we use the aug-cc-pVQZ auxiliary basis set.[S8,S9] We have adopted and validated this procedure in our recent work on cyanine dyes.[S10] The TDDFT/MM and CC2/MM calculations are performed by including the partial MM charges as point charges.

1.2 CASPT2 Calculations

For the CASPT2 calculations, we use the MOLCAS 7.2[S11] code. Unless otherwise stated, we employ the default IPEA zero-order (S-IPEA) Hamiltonian[S12] and report the single-state (SS) CASPT2 values. We use a constant imaginary level shift[S13] of 0.1 a.u. and the ANO-L-VDZP[S7] basis set. We use the Cholesky decomposition of the two-electron integrals[S14] with the threshold of 10^{-4}. The default convergence criteria are used for all calculations. We do not correlate as many of the lowest σ-orbitals as there are heavy atoms in the molecule. In the CAS active space, we include all 12 π electrons in the reference, and test the use of 12 to 15 active π orbitals, resulting in a CAS(12,12) to CAS(12,15) expansion. We use a state-average (SA) CASSCF wave function with equal weights over either the two (S_0 and S_1) or three (S_0, S_1, and S_2) lowest-energy states. The CASPT2/MM calculations are performed by including the partial point charges using the ESPF[S15] electrostatic coupling scheme implemented in Molcas.

1.3 NEVPT2 Calculations

The *n*-electron valence state perturbation theory[S16–S19] (NEVPT2) calculations in the strongly contracted (SC) variants are performed using the ORCA 2.8 code.[S20,S21] In the CASSCF calculations with the ORCA code, the RIJCOSX approximation[S22] for the CASSCF steps and the RI approximation[S6] for the integral transformation steps are used. Since a corresponding RI auxiliary basis set is not available for the ANO-L-VDZP basis set, we use the aug-cc-pVTZ auxiliary basis set.[S8,S9] We have adopted and validated this procedure in our recent work on cyanine dyes.[S10] The orbital energies for the doubly occupied and virtual orbitals appearing in the Dyall Hamiltonian[S23] (used for the definition of the zero order Hamiltonian in NEVPT2) are obtained by the diagonalization of a generalization of the Fock operator[S24] (canonical orbital option in ORCA).

In the construction of the third- and fourth-order density matrices, the CASSCF wave function is truncated so that only configurations with a weight larger than a threshold of 10^{-8} are kept. In the computation of the NEVPT2 excitation energies, we use the same CASSCF wave functions as in the CASPT2 calculations described above. The NEVPT2/MM are performed by including the partial MM charges as point charges.

1.4 Quantum Monte Carlo Calculations

The program package CHAMP[S25] is used for the QMC calculations. We employ scalar-relativistic energy-consistent Hartree-Fock pseudopotentials[S26] where the carbon, nitrogen, and oxygen 1*s* electrons are replaced by a non-singular *s*-non-local pseudopotential and the hydrogen potential is softened by removing the Coulomb divergence. Different Jastrow factors are used to describe the correlation with different atom types and, for each atom type, the Jastrow factor consists of an exponential of the sum of two fifth-order polynomials of the electron-nuclear and the electron-electron distances, respectively.[S27] In all DMC calculations we use a CAS(12,12) expansion and a SA-CASSCF wave functions optimized over two states with equal weights. The starting determinantal components are obtained in CASSCF calculations which are performed with the program GAMESS(US).[S28] The final CAS expansions are expressed on the weighted-average CASSCF natural orbitals. The CAS wave functions of the ground and excited states are truncated with 0.04 threshold on the CSF coefficients and the union set of surviving CSF's for the states of interest are retained in the QMC calculations. The Jastrow correlation factor and the CI coefficients are optimized by energy minimization in a state-averaged sense within VMC with equal weights.[S29] The pseudopotentials are treated beyond the locality approximation[S30] and an imaginary time step of 0.08 a.u. is used in the DMC calculations. We use the Gaussian basis sets[S26] constructed for our pseudopotentials and, specifically, the cc-pVDZ basis augmented with s and p diffuse functions[S31] on the heavy atoms.

The QMC/MM calculations are performed using the electrostatic coupling scheme as in the QM/MM approach employed in the CPMD code.[S32] For the starting trial wave function, CASSCF/MM calculations are performed within GAMESS(US), and the divergence of point charges at the origin is removed to resemble the embedding scheme employed within CPMD.

1.5 Excited-state QM/MM simulations

In the QM/MM excited-state calculations, we employ two different QM regions, one including only the chromophore (RPSB) and the other including the chromophore and the Glu113 counter-ion (RPSB+Glu113). For the RPSB chromophore, the QM/MM boundary is along C_γ-C_δ bond of the Lys296 residue and, for Glu113, it is along the C_α-C_β bond. The valency of the QM region is saturated by adding hydrogen capping atoms at the QM/MM boundary. The MM environment is

modeled by employing partial point charges for the MM atoms with the parameters taken from the Amber force field.

To avoid over-polarization of the QM region, we turn off (set to zero) the partial point charges for a small group of MM atoms at the QM/MM boundary. To do this in a balanced way, we turn off whole charge groups which have a total charge close to zero (as defined in the Amber force field). Therefore, the MM charges for the following atoms are set to zero:

- RPSB QM region:
  - For the Lys296 residue, C_γ together with the two bonded hydrogens (CG, HG2, HG3)
- RPSB+Glu113 QM region:
  - For the Lys296 residue, C_γ together with the two bonded hydrogens (CG, HG2, HG3)
  - For the Glu113 residue, C_α and N on the backbone together with the two bonded hydrogens (CA HA N H), which are all in the same charge group

However, we note that over-polarization of the QM region does not seem to be a serious issue for the excited-state calculations: In CAM-B3LYP calculations for Frame 4 where the charges are not turned off, we obtain excitations (for both QM regions) that differ by only 0.01 eV from the values obtained with the charges set to zero. Finally, we stress that setting these charges to zero is only performed for the excited-state QM/MM calculations and that the ground-state QM/MM MD simulations with CP2K (as described in the paper) are performed with the full MM charges.

1.6 Comparison with Previous CASPT2/MM Studies

The CASSCF/MM and CASPT2/MM calculations in the "Comparison with Previous CASPT2/MM and Other Studies" section are performed with a Molcas-Tinker[S35] interface where the MOLCAS 7.2 package is coupled with a modified version of the MM package Tinker 4.2.[S36]

In all geometrical optimizations of the Section "Comparison with Previous CASPT2/MM and Other Studies", only the chromophore, the Lys296 side-chain, and two nearby waters (wat2a and wat2b) are allowed to move while the rest of the protein is kept frozen at the crystallographic coordinates. In the ground-state geometrical optimization of the QM part with Molcas, we employ the 6-31G* basis set and use either the CASSCF method (with a single-state wave function and a CAS(12,12) expansion) or the DFT method (with the BLYP[S37,S38] functional).

We consider two structural models for the complete protein. The first model is given by the Rhodopsin structure from refs S33, S34, which was kindly provided by N. Ferré and is described in the same references. The model was constructed from chain A of the 1U19[S39] crystallographic structure, and only the chromophore included in the QM region with the QM/MM boundary along the C_δ-C_ε bond of the Lys296 residue. The MM environment was described with the Amber

and TIP3P force fields. In the original paper,[S33] the structure of this model was optimized at the CASSCF/Amber level, where only the chromophore, the Lys296 side-chain, and two nearby waters (wat2a and wat2b) were allowed to relax while the rest of the protein was kept frozen at the crystallographic coordinates.

In the construction of the second model, we also start from chain A of the 1U19[S39] crystallographic structure as done in Ref. S33, but add the hydrogens according to the protonation states we employ for all other calculations described in the paper. Then, we follow the recipe of Ref. S33 and relax all hydrogens and waters in an MM energy minimization, keeping the rest of the protein fixed. Only the chromophore is included in the QM region and the QM/MM boundary is along the C_γ-C_δ bond of the Lys296 residue. The MM environment is described using the Amber and TIP3P force fields.

In order to be compatible with the previous CASPT2/MM studies in refs S33, S34, we present always the SS-CASPT2 excitations, and use the same SA-CASSCF wave function as employed there, that is, a CAS(12,12) expansion over three equally weighted states (S_0, S_1, and S_2). To compute the CASPT2 excitations, we employ either the older (0-IPEA) zero-order Hamiltonian without IPEA shift used in refs S33, S34 or the standard S-IPEA zero-order Hamiltonian which we employ in all other sections here. We use a constant imaginary level shift[S13] of 0.1 a.u. and either the 6-31G* or the ANO-L-VDZP[S7] basis set.

2 Vertical Excitation Energies for RPSB in the Gas Phase

Table S1: Vertical excitation energies (eV) of the 11-*cis* retinal protonated Schiff base (RPSB) in the gas phase computed on DFT, MP2, and CASSCF ground-state equilibrium geometries. The CASPT2, NEVPT2, and DMC results are from Ref. S40. We report the results obtained employing single-state (SS) CASPT2 and strongly contracted (SC) NEVPT2. The bond length alternation (BLA), the C_6-N_{16} distance, and the angle of the β-ionone ring ($\phi_{\text{C5-C6-C7-C8}}$) are also listed.

	Geometry				
	DFT/BLYP (0%)	DFT/B3LYP (20%)	DFT/M06-2X (54%)	MP2	CASSCF
BLA (Å)	0.024	0.033	0.051	0.044	0.101
$d_{\text{C6-N16}}$ (Å)	11.73	11.62	11.52	11.53	11.50
$\phi_{\text{C5-C6-C7-C8}}$ (°)	-29.7	-33.5	-38.0	-40.5	-68.8
Vertical excitation energies (eV)					
CASPT2/0-IPEA	1.77	1.85	1.94	1.89	2.27
CASPT2/S-IPEA	2.12	2.20	2.30	2.24	2.61
NEVPT2	2.18	2.26	2.33	2.27	2.60
DMC	2.22(4)	2.37(3)	-	-	-
TDDFT/B3LYP	2.27	2.29	2.25	2.22	1.89
TDDFT/CAM-B3LYP	2.36	2.43	2.49	2.44	2.73
TDDFT/LC-ωPBE	2.39	2.48	2.61	2.54	3.01

For the BLYP, B3LYP, and M06-2X DFT functionals, the percentage of exact exchange is reported in parenthesis. The TDDFT excitations are obtained with the 6-31+G* basis set. The CASPT2 and NEVPT2 excitations are obtained with the ANO-L-VDZP basis set and the DMC excitations with D+ basis set. The bond length alternation (BLA) is defined as the difference between the averages of the single and double carbon-carbon bonds lengths, and computed including the bonds between C_5 and C_{15}.

We note that the PBE ground-state equilibrium geometry has the following geometrical parameters: BLA $= 0.022$ Å, $d_{\text{C6-N16}} = 11.68$ Å, and $\phi_{\text{C5-C6-C7-C8}} = -28.6°$. These geometrical parameters are very similar to the ones for the BLYP geometry, and, therefore, the PBE and BLYP ground-state equilibrium geometries should yield very similar excitation energy in the gas phase.

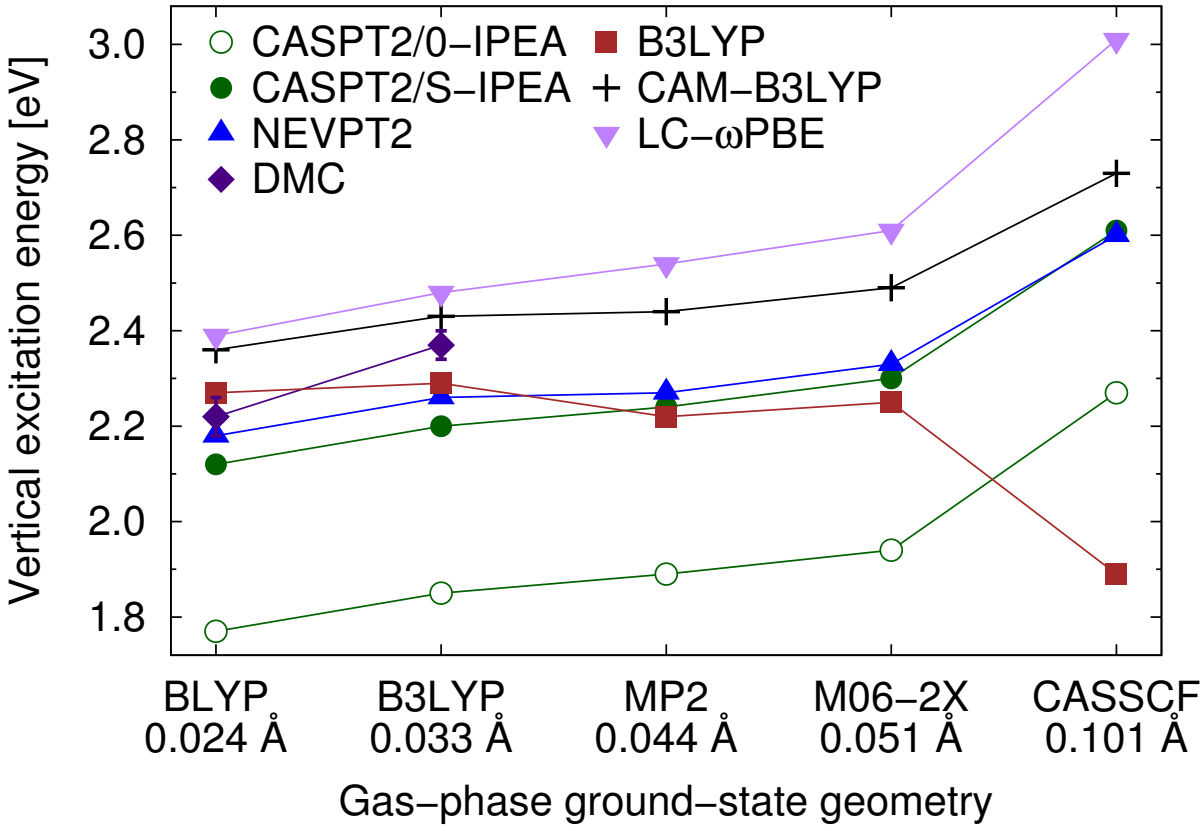

Figure S1: Vertical excitation energies (eV) of the 11-*cis* retinal protonated Schiff base (RPSB) in the gas phase, computed with TDDFT, CASPT2, NEVPT2, and DMC. See the caption of Table S1 for further information. For the ground-state geometries on the x-axis, we also report the bond length alternation (BLA) in Å.

In Figure S1, we observe that TDDFT in combination with the B3LYP functional displays the incorrect dependence on the BLA of the chromophore since the TDDFT/B3LYP excitation energy decreases with increasing BLA at variance with all highly-correlated approaches (this failure has been noted before in the literature[S41]). On the other hand, TDDFT in combination with the long-range corrected CAM-B3LYP and LC-ωPBE functionals leads to the correct dependence on the BLA. Therefore, for retinal systems, the use of the B3LYP functional is not suitable to compute the excitation energies, and a more appropriate choice is to employ the long-range corrected CAM-B3LYP and LC-ωPBE functionals. This warning against the use of B3LYP holds especially for TDDFT calculations in the protein environment, where the the BLA of retinal is enhanced as compared to the gas phase.

3 Cluster Models

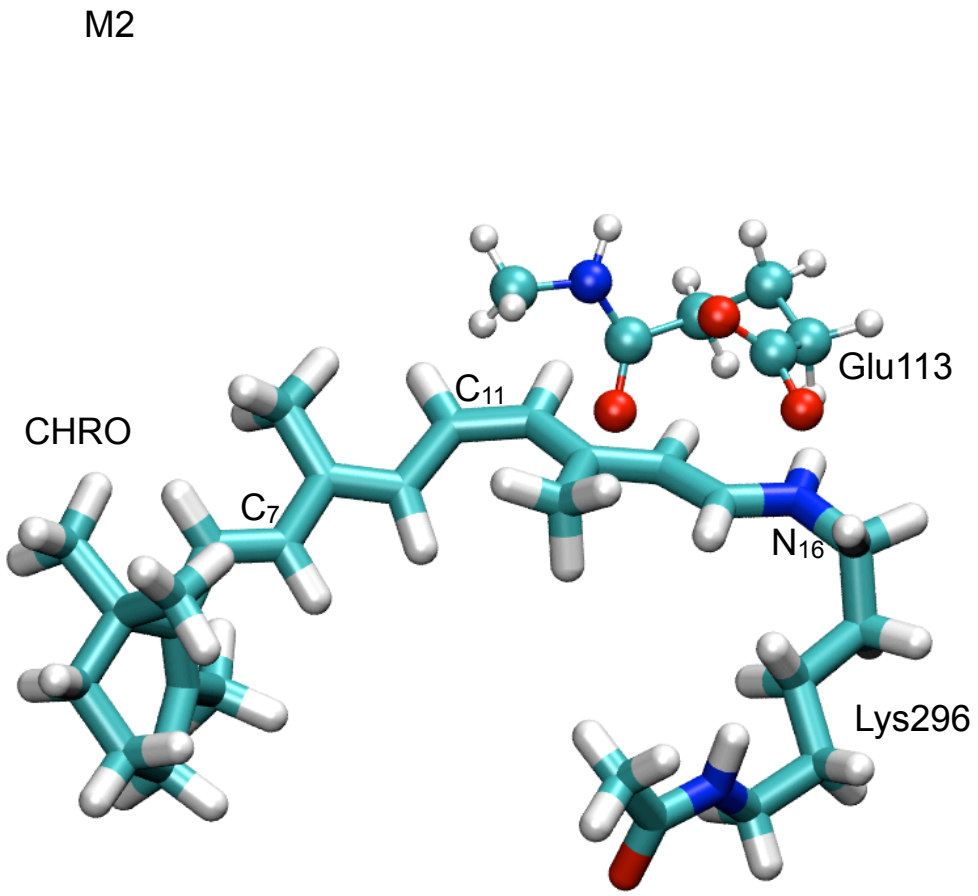

Figure S2: The M2 (CHRO + Glu113) cluster model employed in the Zindo and TDDFT calcuations. The waters are not shown for clarity.

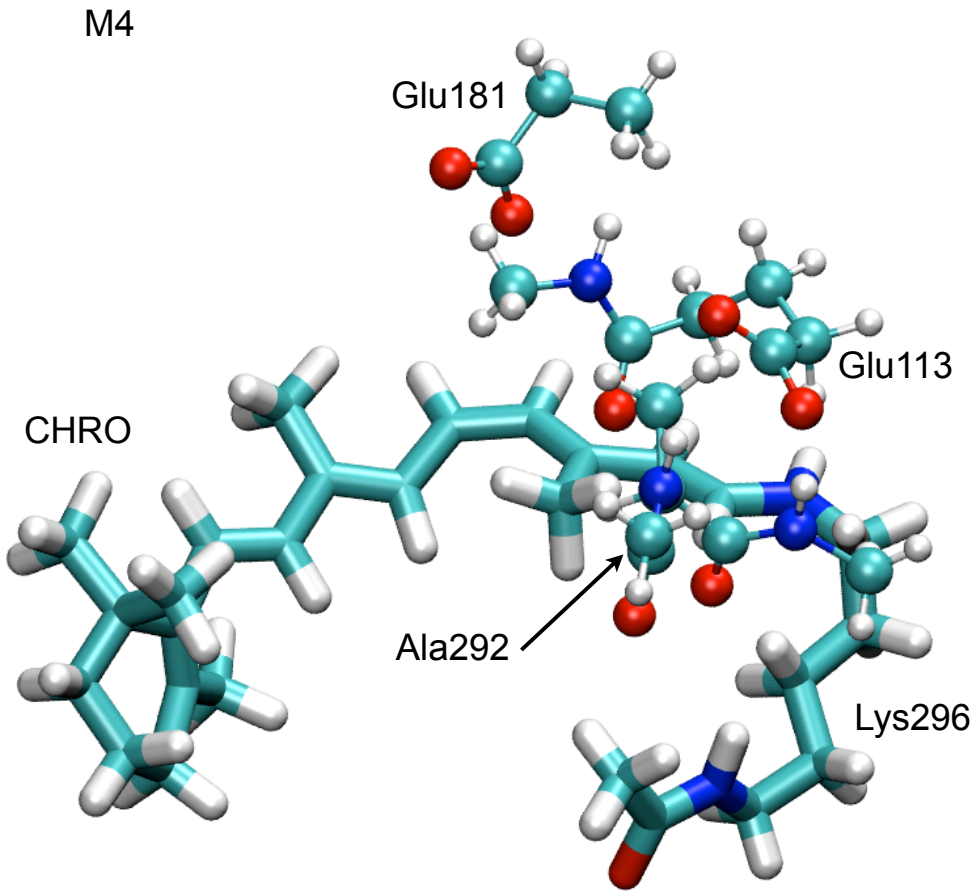

Figure S3: The M4 (M2 + Glu181 + Ala292) cluster model employed in the Zindo and TDDFT calcuations. The waters are not shown for clarity.

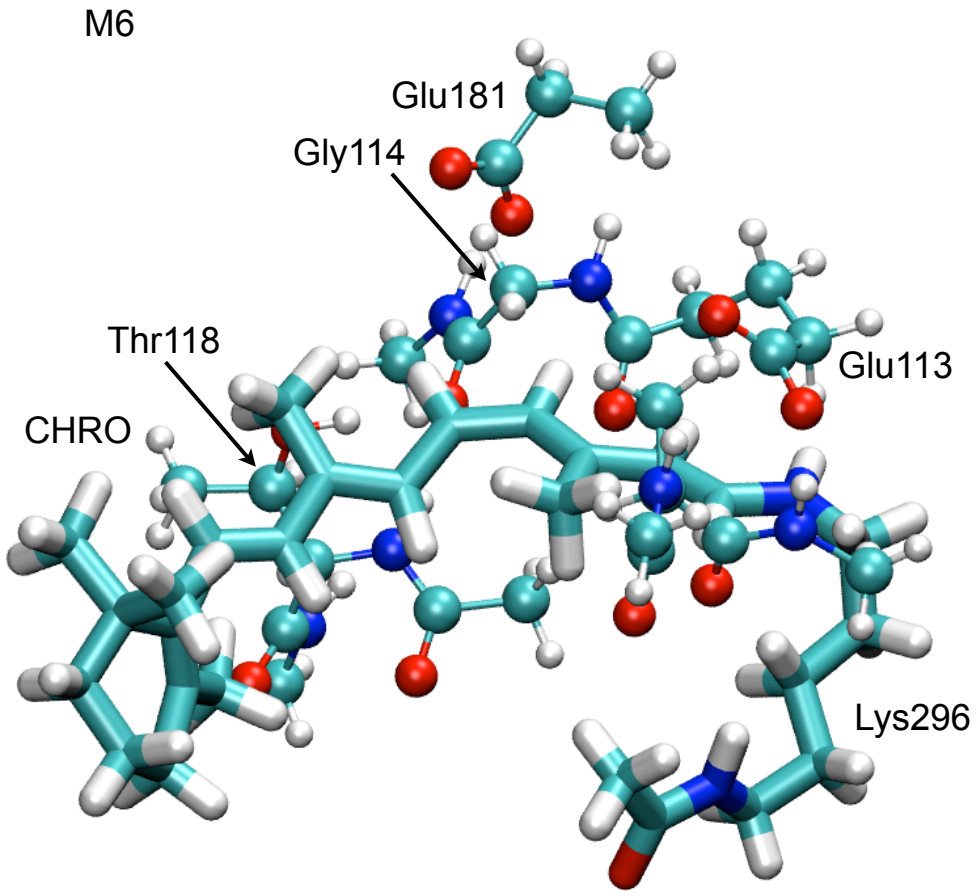

Figure S4: The M6 (M4 + Thr118 + Gly114) cluster model employed in the Zindo and TDDFT calcuations. The waters are not shown for clarity.

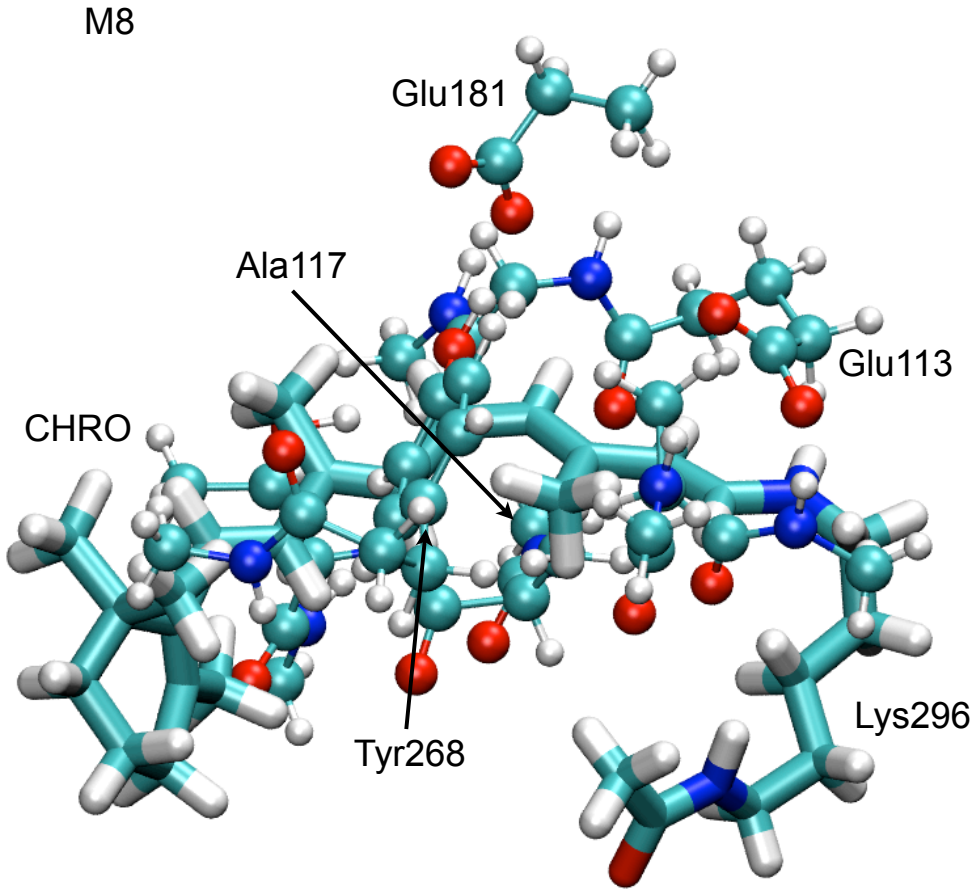

Figure S5: The M8 (M6 + Ala117 + Tyr268) cluster model employed in the Zindo and TDDFT calcuations. The waters are not shown for clarity.

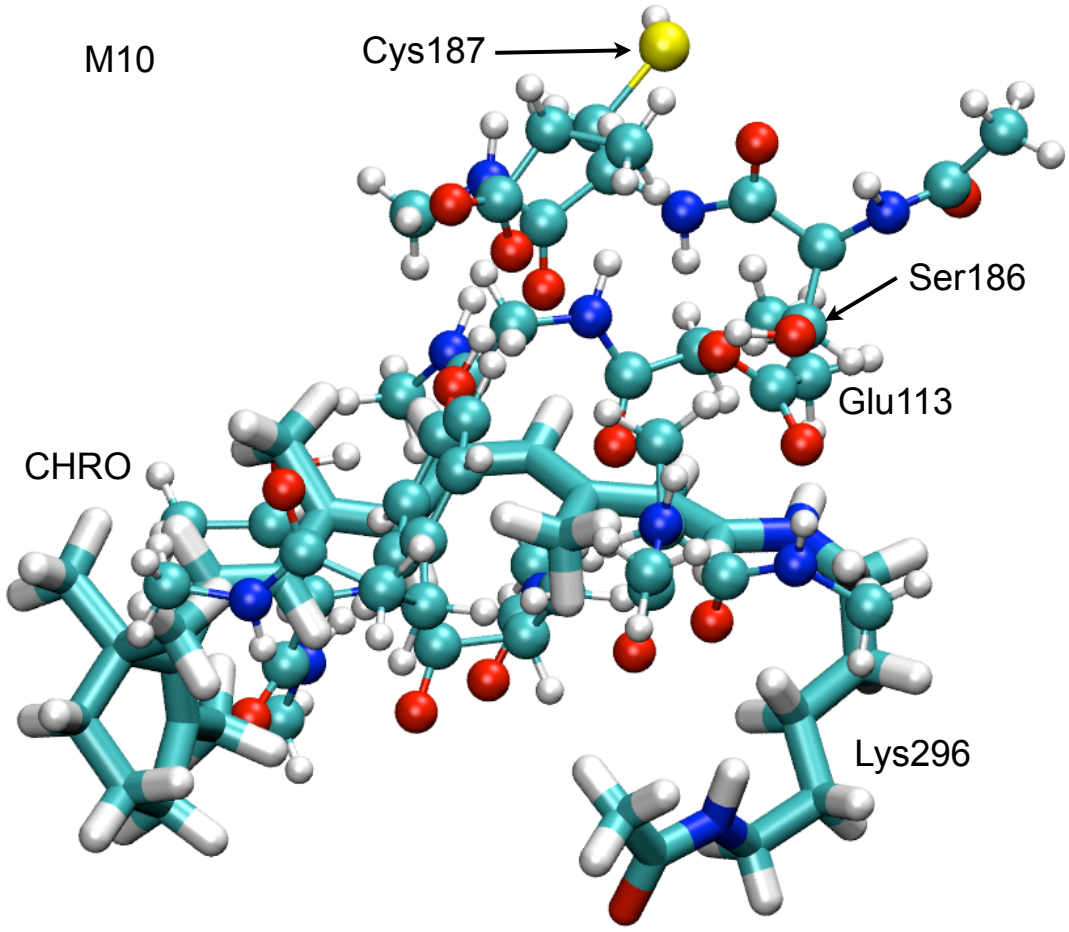

Figure S6: The M10 (M8 + Cys187 + Ser186) cluster model employed in the Zindo and TDDFT calcuations. The waters are not shown for clarity.

4 Bond lengths for the RPSB in the Protein.

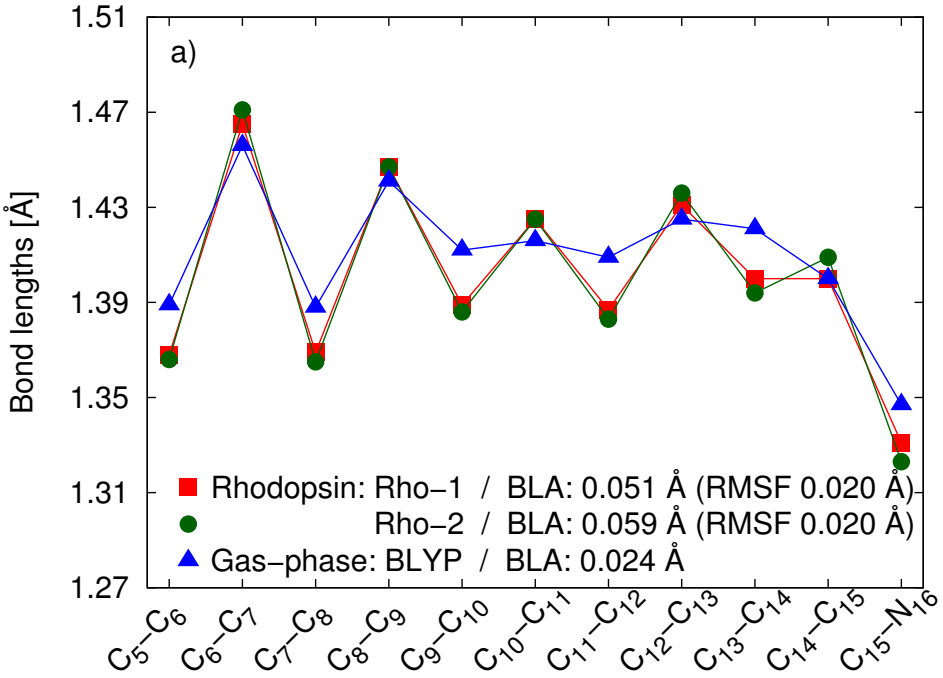

Figure S7: Average bond lengths of the RPSB chromophore from the **Rho-1** and **Rho-2** QM/MM MD simulations. The root mean square fluctuations (RMSF) for the average bond lengths are around 0.03 Å (not shown). For comparison, we also present the bond lengths of RPSB in the gas phase from Ref. S42, which are obtained in a geometrical optimization with DFT/BLYP (i.e. a GGA functional similar to the one employed in the QM/MM MD simulations).

We observe that the pattern of the average bond lengths from the QM/MM MD simulations display a clear alternation between single and double bonds. Furthermore, as compared to the gas phase, the interaction with the protein results in an enhanced bond length alternation.

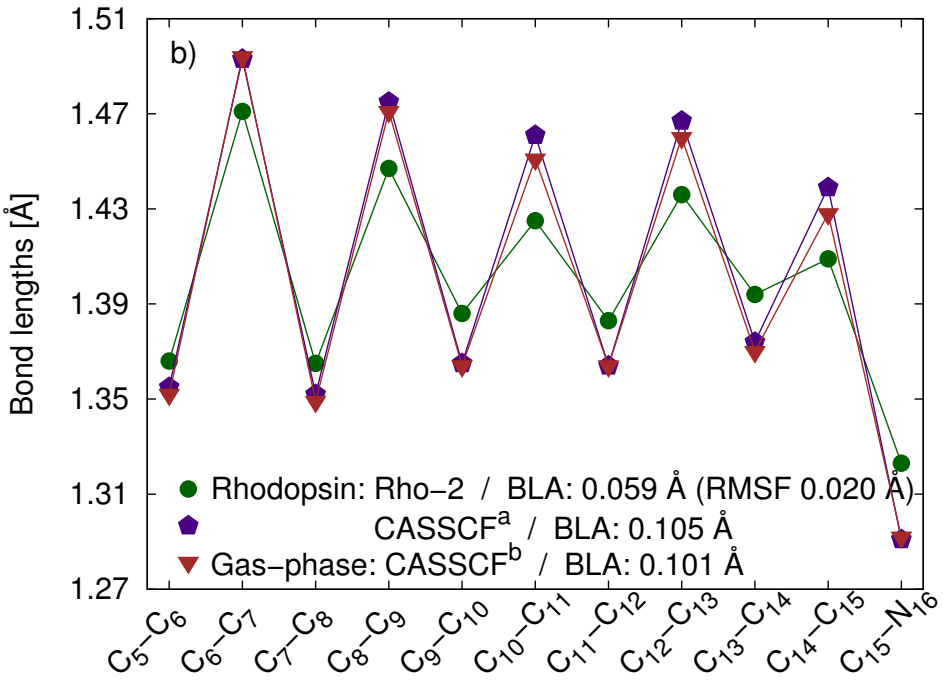

Figure S8: Average bond lengths of the RPSB chromophore from the **Rho-2** QM/MM MD simulations. For comparisons, we show the bond lengths obtained with the commonly employed CASSCF methods, both in Rhodopsin[S33] and the gas phase.[S43]

We observe the well known behavior of CASSCF, which yields an enhanced bond length alternation as compared to DFT. Furthermore, for CASSCF, there is almost no difference between the bond lengths in the gas phase and in the protein environment. We note that in our recent studies,[S40,S44] we have shown that CASSCF gives an inadequate structural description of RPSB in the gas phase.

5 ZINDO Absorption Spectra

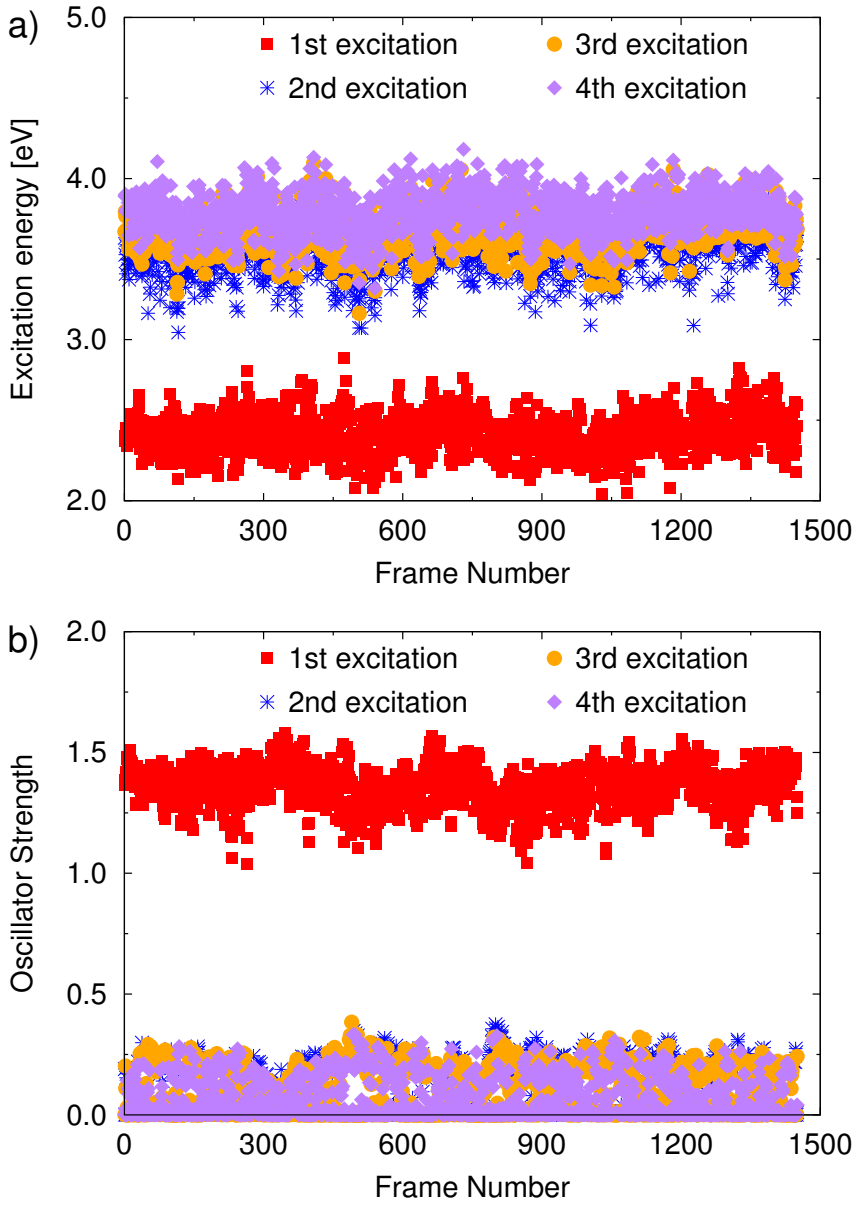

Figure S9: ZINDO excitation energies [panel a)] and corresponding oscillator strengths [panel b)] computed on the frames from the **Rho-2** trajectory. The lowest excitations (red, square symbols) correspond to the relevant, bright $S_0 \rightarrow S_1$ excitation. We observe that the excitation energies and oscillator strengths of the lowest excitations are clearly separated from the higher excitations. Therefore, the inclusion of states higher than S_1 does not appreciably affect the ZINDO absorption spectra.

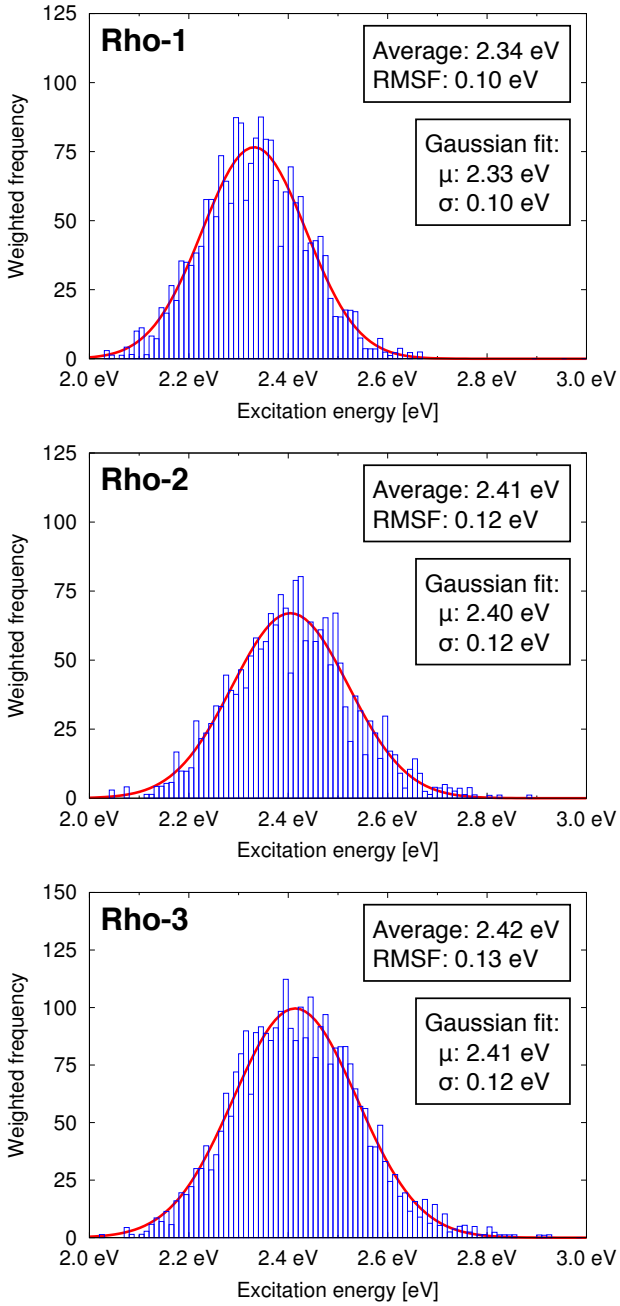

Figure S10: ZINDO absorption spectrum for the **Rho-1**, **Rho-2**, and **Rho-3** trajectories obtained as a histogram of the four lowest excitations weighted with the corresponding oscillator strength (blue boxes). For each trajectory, we report the average and root mean square fluctuation (RMSF) for the bright ($S_0 \rightarrow S_1$) excitation energy. We also show a Gaussian fit of the histogram (red line) and report the mean (μ) and the standard deviation (σ) obtained in the fit. For the histogram, we use bins of 0.01 eV and equidistant frames taken every 5 fs. We note that including only the bright ($S_0 \rightarrow S_1$) excitation energy in the histogram gives the same spectrum as using all four excitations. This is consistent with the fact that in the ZINDO calculations, the higher excitations are clearly separated from the lowest, bright excitation (see Figure S9).

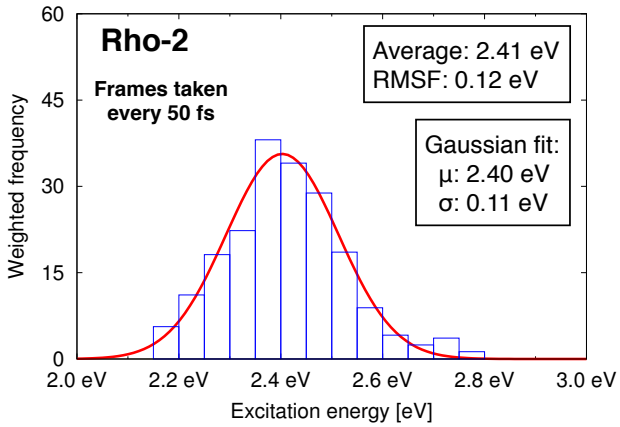

Figure S11: The ZINDO absorption spectrum for the **Rho-2** trajectory obtained using equidistant frames taken every 50 fs (instead of every 5 fs as in Figure S10). For this histogram, we employ bins of 0.05 eV. As discussed in note 84 in the paper, we estimate the excitation energy to be uncorrelated after about 50 fs (10 frames). We note that employing frames taken every 50 fs, as presented in this figure, gives equivalent results (i.e. average, root mean square fluctuation, and Gaussian fit) to the ones we obtain if we employ frames taken every 5 fs.

6 Basis Set Dependence of the Excitations Energies

Table S2: Basis-set dependence of the CAM-B3LYP excitations energies (eV) for Frame 4 from the **Rho-1** run. The results obtained with the default 6-31+G* basis set used in the TDDFT excitations are shown in boldface.

Basis set	QM:	RPSB		RPSB+Glu113		M10
	MM:	None	Full	None	Full	None
6-31G*		2.48	2.74	2.73	2.72	
6-31+G*		**2.45**	**2.70**	**2.69**	**2.71**	**2.56**
6-31++G**		2.43	2.70	2.68	2.70	
cc-pVDZ		2.47	2.73	2.72	2.75	
cc-pVTZ		2.45	2.70	2.69	2.70	
ANO-L-VDZP			2.69			

Table S3: Basis-set dependence of the CASPT2 excitations energies (eV) for Frame 4 from the **Rho-1** run. The QM/MM environment is RPSB/Full and the CAS expansion is a CAS(12,12) over 2 states. The results obtained with the default ANO-L-VDZP basis set used in the CASPT2 excitations are shown in boldface.

Basis set	CASSCF	f	SS-CASPT2	MS-CASPT2
ANO-L-VDZP	**3.27**	**1.46**	**2.72**	**2.79**
aug-ANO-L-VDZP[a,b]	3.28	1.46	2.71	2.78
ANO-L-[4s3p1d]/[3s1p]	3.28	1.46	2.71	2.79
6-31G*	3.31	1.45	2.78	2.85
6-31+G*	3.29	1.45	2.74	2.80
6-31++G**	3.29	1.45	2.73	2.79
cc-pVDZ	3.33	1.47	2.77	2.84

[a] Diffuse functions only on heavy atoms;

[b] Diffuse functions taken from aug-cc-pVDZ;

7 TDDFT Excitation Energies

Table S4: CAM-B3LYP/6-31+G* and LC-ωPBE/6-31+G* excitations energies (eV) and oscillator strengths for Frames 4, 7, and 11 obtained in different QM/MM environments. For the CAM-B3LYP excitation energies, we presents the difference in the dipole moments of the ground and excited states, $|\Delta\mu|$ (D), and the CI coefficient for the HOMO→LUMO excitation (which is a $\pi \rightarrow \pi^*$ excitation localized on the chromophore).

		CAM-B3LYP				LC-ωPBE	
QM/MM	Frame	Exc.	f	$\|\Delta\mu\|$	HOMO→LUMO	Exc.	f
RPSB/None	4	2.45	1.32	6.8	0.69	2.58	1.51
	7	2.42	1.23	6.7	0.68	2.55	1.41
	11	2.36	1.23	6.8	0.69	2.51	1.42
RPSB/Full	4	2.70	1.51	5.0	0.70	2.82	1.60
	7	2.63	1.41	4.7	0.70	2.74	1.47
	11	2.65	1.41	6.2	0.69	2.82	1.53
RPSB+Glu113/None	4	2.69	1.52	5.4	0.70	2.86	0.96
	7	2.72	1.35	3.8	0.70	2.84	1.42
	11	2.72	1.22	2.6	0.69	2.88	1.57
RPSB+Glu113/Full	4	2.71	1.53	5.0	0.69	2.83	1.60
	7	2.62	1.42	5.0	0.70	2.74	1.48
	11	2.65	1.46	6.4	0.69	2.83	1.58
M10 cluster	4	2.56	1.05		0.68	2.66	1.10
	7	2.53	1.00		0.69	2.64	1.03
	11	2.61	1.03		0.69	2.77	1.08

Table S5: CAM-B3LYP/6-31+G* and LC-ωPBE/6-31+G* excitations energies (eV) for Frames in the region of the ZINDO absorption maximum of the **Rho-1**, **Rho-2**, and **Rho-3** runs. We also show the excitations obtained when only the MM charges of the Glu113 counter-ion are included (RPSB/Only Glu113).

Frame	QM:	RPSB			RPSB+Glu113		M10
	MM:	None	Only Glu113	Full	None	Full	None
CAM-B3LYP/6-31+G*:							
Rho-1	1	2.49	2.70	2.68	2.71	2.69	2.58
	2	2.31	2.60	2.58	2.61	2.59	2.51
	3	2.33	2.63	2.67	2.64	2.65	2.53
	4	2.45	2.68	2.70	2.69	2.71	2.56
	5	2.38	2.66	2.69	2.67	2.70	2.57
Rho-2	6	2.36	2.70	2.64	2.71	2.66	2.60
	7	2.42	2.68	2.63	2.72	2.62	2.53
	8	2.38	2.70	2.65	2.71	2.62	2.58
	9	2.38	2.73	2.63	2.78	2.66	2.59
	10	2.38	2.73	2.68	2.78	2.70	2.62
Rho-3	11	2.36	2.67	2.65	2.72	2.65	2.61
LC-ωPBE/6-31+G*:							
Rho-1	1	2.60	2.85	2.83	2.87	2.85	
	2	2.43	2.77	2.74	2.79	2.76	
	3	2.51	2.83	2.84	2.85	2.83	
	4	2.58	2.81	2.82	2.85	2.83	2.66
	5	2.55	2.88	2.89	2.90	2.91	
Rho-2	6	2.53	2.92	2.82	2.97	2.86	
	7	2.55	2.81	2.74	2.84	2.74	2.64
	8	2.51	2.91	2.83	2.98	2.81	
	9	2.54	2.96	2.83	3.02	2.88	
	10	2.54	2.90	2.86	2.97	2.90	
Rho-3	11	2.51	2.84	2.82	2.88	2.83	2.77

8 TDDFT Excitation Energies for Different Chromophore Models

In the M1 to M12 cluster models used in the paper (see for example Figure 4 and 8 in the paper), the retinal chromophore model, denoted as CHRO, is extended to include most of the backbone of Lys296 and, therefore, is slightly larger than the RPSB chromophore used in the conventional QM/MM calculations throughout the paper. The extended CHRO model is cut at the C-C_α bond of Lys296 and the C_α-N bond of Ala295 while the RPSB model is cut at C_γ-C_δ bond of Lys296.

In both cases, the cut is performed some distance away from the π network on the chromophore and, therefore, we do not expect a large difference in the excitations of the two chromophores models when considered in isolation or in the presence of the MM charges. To verify this, we perform CAM-B3LYP/MM calculations using the CHRO chromophore model and, in Table S6, compare these excitation energies to the results obtained with RPSB chromophore model. We observe that without the MM charges (MM None), the difference in the excitations energies between the two chromophore models is 0.03-0.04 eV. Once we include the MM charges (MM Full), the difference is even less or no more than 0.02 eV. Therefore, the two chromophore models yield almost identical excitation energies.

Table S6: CAM-B3LYP/6-31+G* excitations energies (eV) obtained with the two different chromophore models employed in the paper. The RSPB model is used in the conventional QM/MM calculations while the CHRO model is used in the M1 to M12 cluster models. Compared to the RPSB model, the CHRO model is slightly larger since it is extended to include the backbone of Lys296. Note, that for CHRO+Glu113 calculations, we do not include in QM region the 7 waters that are instead included in the M2 cluster presented in Figures 4 and 8.

CHRO chromophore model

QM:	CHRO		CHRO+Glu113	
MM:	None	Full	None	Full
Frame 4	2.42	2.69	2.65	2.70
Frame 7	2.38	2.62	2.69	2.60

RPSB chromophore model

QM:	RPSB		RPSB+Glu113	
MM:	None	Full	None	Full
Frame 4	2.45	2.71	2.69	2.71
Frame 7	2.42	2.63	2.72	2.62

9 CASPT2 Excitation energies

Selection of the Active Space

In Table S7, we present single-state (SS) and multi-state (MS) CASPT2 excitations for Frames 4, 7, and 11 obtained using different SA-CASSCF zero-order wave functions. In Table S7, we underline the CASPT2 excitations which are presented in Table 3 in the paper. In the state average calculations, we employ equal weights over either the two (S_0 and S_1) or three (S_0, S_1, and S_2) lowest-energy states. In general, it has been shown that a minimal CAS space might not always be sufficient[S10,S45–S47] and, therefore, the convergence of the excitation energies must be assessed by including additional virtual orbitals in the active space. To achieve this, we perform calculations with the minimal CAS(12,12) active space (all 12 π electrons in the reference and the same number of active π orbitals) and also with active spaces that include additional π orbitals up to a CAS(12,15) expansion.

For the two-state calculations, the CASSCF excitations decrease by about 0.05 eV when going from the minimal CAS(12,12) active space to the CAS(12,13) active space, which additionally includes the π orbital shown in Figure S13(h). Enlarging the active space further up to a CAS(12,15) leaves the CASSCF excitation energies unchanged. The CASPT2 excitations are less sensitive that the CASSCF ones and there is at most a 0.02 eV difference between values obtained with the CAS(12,12) and CAS(12,13) expansions.

The three-state calculations are more sensitive to the choice of active space especially in the RPSB/Full case: The CASSCF and CASSCF excitation energies decrease by as much as 0.2 and 0.1 eV, respectively, when going from a CAS(12,12) to a CAS(12,13) active space. As in the two-state calculations, further enlarging the active space has almost no effect. The reason for the increased sensitivity of the three-state calculations is that the two excited states S_1 and S_2 are rather close in energy and need to be simultaneously well described by the same CAS expansion. The use of the CAS(12,13) active space lowers the CASSCF excitation energy of the relevant bright state (as compared to the CAS(12,12) active space), improving its description, and yields better CASPT2 excitations. We also observe that for the CAS(12,13) active space, the difference between the two- and three-state CASPT2 excitations is less than for the CAS(12,12) active space.

In summary, we find that the CAS(12,13) is the smallest active space to yield converged CASSCF and CASPT2 excitations and that the two-state and three-state calculations are rather comparable when this active space is employed. We observe a similar behavior for the NEVPT2 excitation presented in Table S8. In the two-state calculations, the NEVPT2 excitations increase by about 0.05 eV when going from a CAS(12,12) to a CAS(12,13) expansion. Consequently, in the paper, we present the single-state CASPT2 and NEVPT2 excitations obtained with the CAS(12,13) and use two states in the state-average calculation.

Table S7: CASSCF and CASPT2 excitation energies (eV) on Frames 4, 7, and 11 obtained with the RPSB QM region and different SA-CASSCF wave functions. We underline the CASPT2 excitation energies presented in Table 3 in the paper. The excitations are obtained either without (None) or with (Full) MM charges. We indicate the number of states (St.) and the active space (CAS) used in the SA-CASSCF wave function. We also present the oscillator strengths f and the difference between the ground-state and excited-state dipole moments $|\Delta\mu|$ (D). The S-IPEA Hamiltonian and the ANO-L-VDZP basis set are employed. We use a constant imaginary level shift of 0.1 a.u. In all CASPT2 calculations, we find that the reference weights of the states under consideration are similar and that there are no intruder states.

QM / MM	Fr.	St. / CAS	CASSCF	f^a	$\|\Delta\mu\|^a$	SS-CASPT2	MS-CASPT2
RPSB	4	2 / (12,12)	2.65	1.50	12.9	2.30	2.36
/ None		(12,13)	2.62	1.51	12.6	**2.29**	2.35
		3 / (12,12)	2.77 / 3.64	1.46 / 0.06	14.6 / 0.9	2.23 / 3.36	2.42 / 3.47
		(12,13)	2.66 / 3.63	1.50 / 0.07	13.7 / 1.5	2.26 / 3.34	2.38 / 3.41
	7	2 / (12,12)	2.54	1.45	12.0	2.23	2.30
		(12,13)	2.51	1.47	11.9	**2.22**	2.30
		3 / (12,12)	2.67 / 3.56	1.41 / 0.02	14.1 / 1.0	2.15 / 3.25	2.40 / 3.38
		(12,13)	2.56 / 3.54	1.45 / 0.03	13.3 / 1.1	2.19 / 3.23	2.35 / 3.31
	11	2 / (12,12)	2.54	1.44	13.5	2.19	2.26
		(12,13)	2.51	1.45	13.3	**2.17**	2.25
		3 / (12,12)	2.65 / 3.59	1.39 / 0.08	15.1 / 1.4	2.13 / 3.26	2.32 / 3.36
		(12,13)	2.54 / 3.57	1.43 / 0.09	14.4 / 2.1	2.16 / 3.24	2.29 / 3.31
RPSB	4	2 / (12,12)	3.27	1.46	12.9	2.72	2.79
/ Full		(12,13)	3.22	1.47	12.8	**2.71**	2.78
		(12,14)	3.21	1.46	12.7	2.71	2.78
		(12,15)	3.19	1.47			
		(12,13)	3.22	1.47	12.8	2.38[b]	2.50[b]
		3 / (12,12)	3.52 / 3.59	1.37 / 0.10	13.6 / 1.0	2.62 / 3.38	2.70 / 3.60
		(12,13)	3.35 / 3.59	1.45 / 0.05	13.4 / 1.1	2.67 / 3.38	2.76 / 3.49
		(12,14)	3.33 / 3.59	1.45 / 0.05	13.2 / 1.1	2.68 / 3.37	2.76 / 3.46
		(12,15)	3.32 / 3.60	1.45 / 0.05	13.1 / 1.1		
	7	2 / (12,12)	3.14	1.40	12.5	2.62	2.74
		(12,13)	3.08	1.43	12.5	**2.61**	2.72
		(12,13)	3.08	1.43	12.5	2.31[b]	2.49[b]
		3 / (12,12)	3.40 / 3.50	1.06 / 0.29	10.1 / 3.4	2.68 / 2.98	2.66 / 3.65
		(12,13)	3.21 / 3.48	1.39 / 0.01	13.5 / 1.2	2.57 / 3.23	2.79 / 3.36
		(12,14)	3.19 / 3.49	1.39 / 0.01	13.3 / 1.2	2.56 / 3.22	2.76 / 3.34
	11	2 / (12,12)	3.38	1.37	13.7	2.72	2.80
		(12,13)	3.32	1.40	13.7	**2.72**	2.81
		(12,13)	3.32	1.40	13.7	2.40[b]	2.55[b]
		3 / (12,12)	3.53 / 3.66	0.02 / 1.38	0.6 / 13.8	3.23 / 2.72	2.93 / 3.49
		(12,13)	3.45 / 3.54	1.12 / 0.32	12.1 / 1.4	2.74 / 3.25	2.61 / 3.60
		(12,14)	3.43 / 3.54	1.17 / 0.26	12.5 / 1.0	2.72 / 3.26	2.63 / 3.55

[a] Obtained in a CASSCF calculation;

[b] Obtained with the 0-IPEA zero-order Hamiltonian

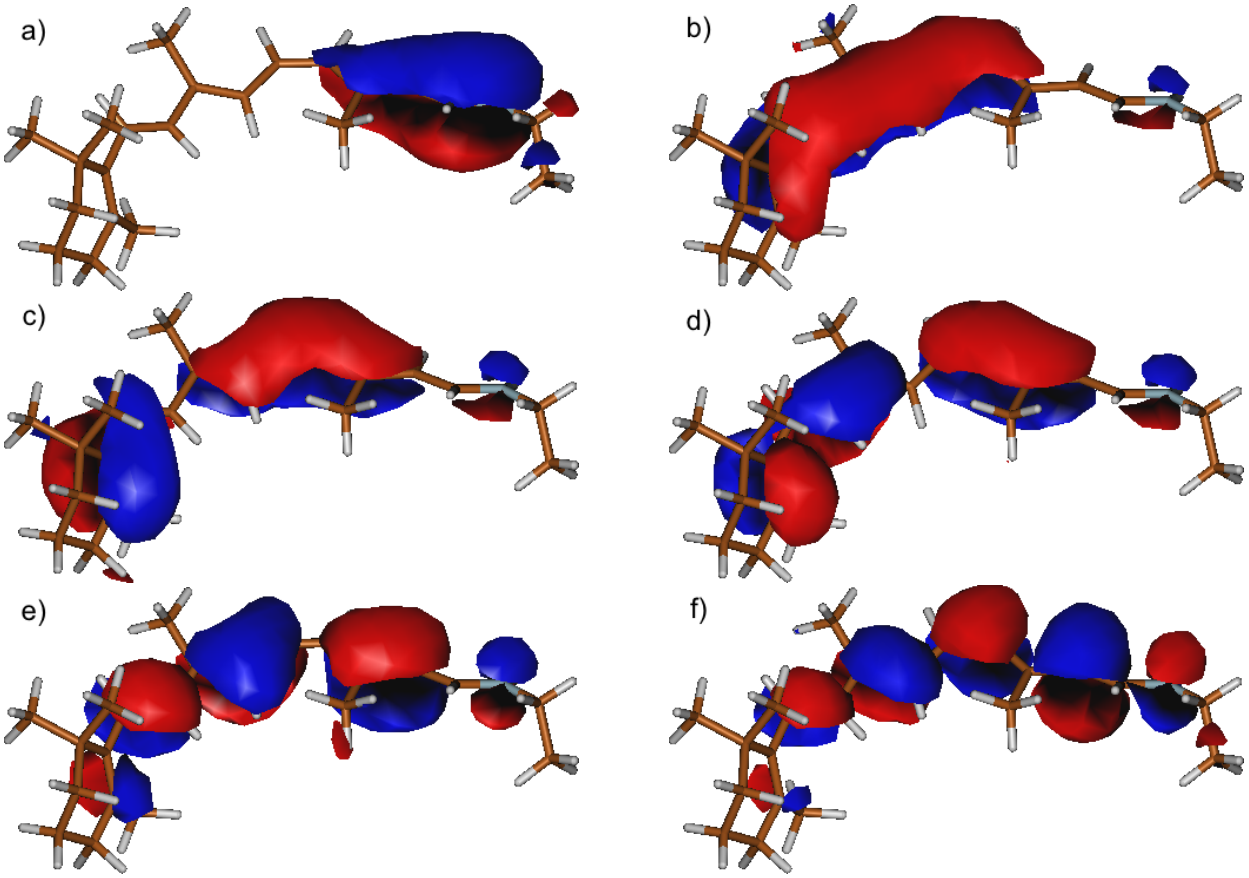

Figure S12: Orbitals in the CAS(12,13) active space. The first six orbitals doubly occupied in the Hartree-Fock determinant are shown.

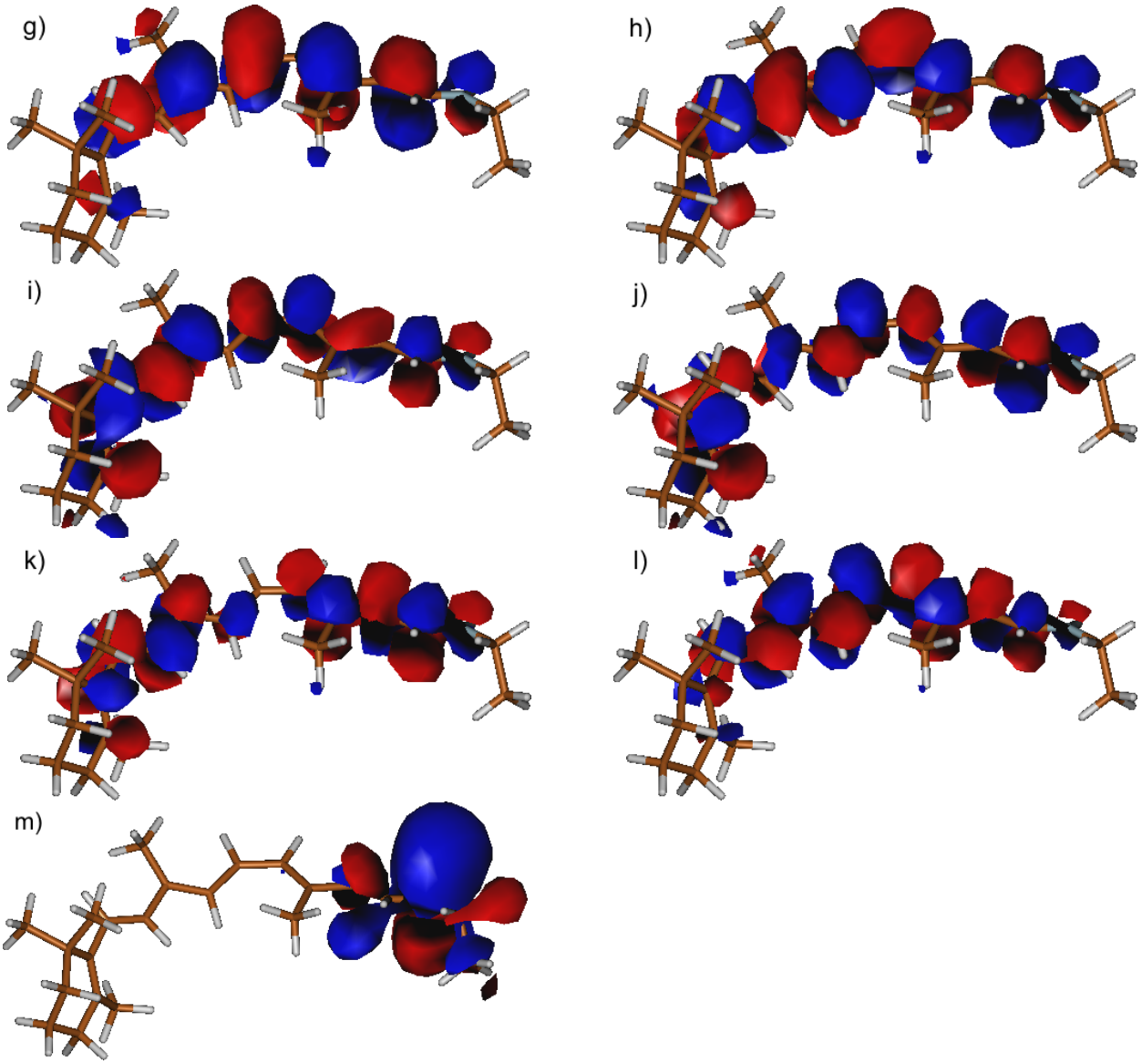

Figure S13: Orbitals in the CAS(12,13) active space. The seven virtual orbitals not occupied in the Hartree-Fock determinant are shown.

10 NEVPT2 Excitation Energies

It has been noted that the NEVPT2 approach can be rather sensitive to the quality of the CASSCF zero-order wave function and give unsatisfactory performance if the correction due to dynamical correlation is large.[S47] To investigate this aspect, we present in Table S8 the NEVPT2 excitations for Frames 4, 7, and 11 obtained with different SA-CASSCF zero-order wave functions. We observe that, for the excitation energies obtained using two states, the correction due to dynamical correlation is always less than 0.7 eV. Furthermore, the difference between the excitation energies obtained with a CAS(12,12) and a CAS(12,13) is less than 0.05 eV, which indicates that the NEVPT2 excitation energies are robust and can be trusted. However, the performance when employing three states is less satisfactory, with larger corrections and considerable differences between the results obtained with the two active spaces. Therefore, as for CASPT2 calculations, we present in Table 3 in the paper, the NEVPT2 excitations obtained using CAS(12,13) over two states which, as explained, are reliable.

Table S8: NEVPT2/SC excitation energies (eV) on Frames 4, 7, and 11 obtained with the RPSB QM region and different SA-CASSCF wave functions. We underline the NEVPT2 excitation presented in Table 3 in the paper. The excitation energies are obtained either without (None) or with (Full) MM charges. We indicate the number of states (St.) and the active space (CAS) used in the SA-CASSCF wave function. We also present the oscillator strengths f. The ANO-L-VDZP basis set is employed.

QM / MM	Fr.	St. / CAS	CASSCF	f^a	NEVPT2/SC
RPSB / None	4	2 / (12,12)	2.65	1.50	2.37
		(12,13)	2.62	1.51	**2.39**
		3 / (12,12)	2.77 / 3.65	1.46 / 0.06	2.23 / 3.34
		(12,13)	2.66 / 3.63	1.50 / 0.07	2.33 / 3.34
	7	2 / (12,12)	2.54	1.45	2.33
		(12,13)	2.51	1.47	**2.35**
		3 / (12,12)	2.66 / 3.56	1.41 / 0.02	2.17 / 3.25
		(12,13)	2.56 / 3.54	1.45 / 0.03	2.27 / 3.25
	11	2 / (12,12)	2.54	1.44	2.26
		(12,13)	2.51	1.45	**2.28**
		3 / (12,12)	2.64 / 3.59	1.40 / 0.08	2.14 / 3.22
		(12,13)	2.54 / 3.57	1.43 / 0.09	2.22 / 3.22
RPSB / Full	4	2 / (12,12)	3.25	1.39	2.73
		(12,13)	3.21	1.48	**2.76**
		3 / (12,12)	3.51 / 3.58	1.39 / 0.08	2.51 / 3.41
		(12,13)	3.34 / 3.58	1.46 / 0.05	2.66 / 3.41
	7	2 / (12,12)	3.13	1.40	2.60
		(12,13)	3.07	1.43	**2.65**
		3 / (12,12)	3.39 / 3.48	0.94 / 0.41	2.69 / 2.92
		(12,13)	3.20 / 3.46	1.39 / 0.01	2.55 / 3.28
	11	2 / (12,12)	3.37	1.40	2.70
		(12,13)	3.32	1.41	**2.74**
		3 / (12,12)	3.51 / 3.66	0.01 / 1.39	3.30 / 2.51
		(12,13)	3.44 / 3.52	1.10 / 0.34	2.69 / 3.30

[a] Obtained in a CASSCF calculation

11 Calculations for the "Comparison with Previous CASPT2/MM and Other Studies" section

Table S9: CASPT2/MM excitation energies (eV) computed with the rhodopsin (1U19) CASSCF/Amber model used in Refs. S33 and S34. The relevant bright state is underlined. We note that Glu181 is protonated in this model. The excitations are obtained using different zero-order Hamiltonians (S-IPEA or 0-IPEA) and basis sets (ANO-L-VDZP or 6-31G*). The calculations are performed using a CAS(12,12) expansion, a SA-CASSCF wave functions optimized over 3 states with equal weights, and a constant imaginary level shift of 0.1 a.u. In all CASPT2 calculations, we find that the reference weights of the states under consideration are similar and that there are no intruder states.

IPEA	Basis set	CASSCF	f	SS-CASPT2
0-IPEA	6-31G*	3.70 / 4.50	0.94 / 0.41	<u>2.38</u> / 3.62
	ANO-L-VDZP	3.66 / 4.49	0.97 / 0.39	<u>2.22</u> / 3.51
S-IPEA	6-31G*	3.70 / 4.50	0.94 / 0.41	<u>2.75</u> / 4.02
	ANO-L-VDZP	3.66 / 4.49	0.97 / 0.39	<u>2.63</u> / 3.96

Table S10: CASPT2/MM excitations (eV) obtained from different rhodopsin (1U19) models. The relevant bright state is underlined. Only the chromophore is optimized with either CASSCF or BLYP, using the protonation state of Glu181 as indicated. We employ the ANO-L-VDZP basis set and the S-IPEA Hamiltonian. The calculations are performed using a CAS(12,12) expansion, a SA-CASSCF wave functions optimized over 3 states with equal weights, and a constant imaginary level shift of 0.1 a.u. In all CASPT2 calculations, we find that the reference weights of the states under consideration are similar and that there are no intruder states.

Residue protonations	Chro. Opt.	Glu181	CASSCF	f	SS-CASPT2
From Refs. S33 and S34	CASSCF	Prot.	3.66 / 4.49	0.97 / 0.39	<u>2.63</u> / 3.96
	CASSCF	Deprot.	3.89 / 4.56	0.93 / 0.46	<u>2.96</u> / 4.02
	BLYP	Prot.	2.95 / 3.34	1.56 / 0.02	<u>2.17</u> / 3.18
	BLYP	Deprot.	3.31 / 3.47	1.27 / 0.28	<u>2.55</u> / 3.20
As in this study	CASSCF	Deprot.	4.04 / 4.55	0.65 / 0.79	<u>3.14</u>[a] / 3.76
		CAS(12,13)	3.88 / 4.46	0.92 / 0.54	3.07 / 4.21
		CAS(12,14)	3.85 / 4.46	0.96 / 0.50	3.06 / 3.94
	BLYP	Deprot.	3.34 / 3.56	0.00 / 1.55	3.16 / <u>2.57</u>

[a] For this case, the second and third states have rather similar oscillator strengths, so the assignment of the bright state is ambiguous. Therefore, to clarify this issue, we perform calculations with larger CAS(12,13) and CAS(12,14) expansions, where additional virtual π orbitals are included. As the CAS increases, the oscillator strength of the second state increases and this state becomes unambiguously the bright state, with an excitation energy of 3.07 eV, that is, rather close to the value of 3.14 eV obtained for the CAS(12,12) expansion. Thus, we assign the second state to be the relevant bright state for all the CAS expansions.

Table S11: CASPT2/MM excitations (eV) obtained using two of the models presented in Table S10. For each model, the geometry of the chromophore is the same as the one relaxed in the presence of the full protein environment (as in Table S10), while the excitations are computed using different MM charges. We employ the ANO-L-VDZP basis set and the S-IPEA Hamiltonian. The calculations are performed using a CAS(12,12) expansion, a SA-CASSCF wave functions optimized over 3 states with equal weights, and a constant imaginary level shift of 0.1 a.u. In all CASPT2 calculations, we find that the reference weights of the states under consideration are similar and that there are no intruder states. We underline the relevant bright state.

Rhodopsin Model / Glu181 / Geo. Opt.	MM charges	CASSCF	f	SS-CASPT2
Refs. S33, S34 / Prot. / CASSCF	None	3.16 / 4.37	1.12 / 0.30	<u>2.46</u> / 3.86
	Glu113	4.31 / 4.97	0.13 / 1.25	3.70 / <u>3.56</u>
	Full	3.66 / 4.49	0.97 / 0.39	<u>2.63</u> / 3.96
From present work / Deprot. / BLYP	None	2.72 / 3.37	1.58 / 0.04	<u>2.11</u> / 3.15
	Glu113	3.28 / 3.57	0.00 / 1.53	3.12 / <u>2.65</u>
	Full	3.34 / 3.56	0.00 / 1.55	3.16 / <u>2.57</u>

12 Geometries for the RPSB+Glu113 QM region

Table S12: Coordiantes (Å) for Frame 1 from the **Rho-1** run. The coordinates are for the RPSB+Glu113 QM region employed in the excited-state calculations.

C	11.453	4.665	4.296
H	11.336	3.666	3.792
H	11.318	5.335	3.492
C	12.945	4.767	4.651
H	13.212	4.167	5.496
H	13.576	4.579	3.758
C	13.294	6.233	5.048
O	13.031	7.135	4.252
O	14.069	6.468	5.955
C	17.743	8.722	8.794
H	18.863	8.651	8.983
H	17.401	7.819	9.343
C	17.430	8.731	7.311
H	18.043	9.565	6.810
H	17.777	7.825	6.700
N	16.033	8.913	6.976
H	15.536	8.017	6.750
C	8.562	18.066	8.504
C	8.291	19.215	9.548
C	9.228	19.054	10.796
C	8.688	17.793	11.487
C	8.614	16.609	10.561
C	8.580	16.704	9.229
C	8.475	15.588	8.302
C	9.244	14.450	8.262
C	9.371	13.448	7.236
C	10.545	12.670	7.189
C	10.702	11.623	6.298
C	11.732	10.765	6.120
C	13.094	10.936	6.485
C	14.011	9.924	6.462
C	15.379	10.022	6.693
C	7.497	18.211	7.350
C	9.905	18.366	7.695
C	8.410	15.319	11.244
C	8.320	13.312	6.175
C	13.458	12.321	6.866
H	7.673	19.080	6.766
H	6.538	18.083	7.710
H	7.532	17.429	6.550
H	10.795	18.542	8.344
H	9.832	19.242	7.020
H	10.220	17.505	7.162
H	7.306	19.274	10.006
H	8.518	20.147	9.093
H	9.211	19.869	11.497
H	10.297	18.795	10.489
H	7.644	17.892	11.895
H	9.362	17.604	12.384
H	8.264	14.455	10.556

Continued on next page

Table S12 – Continued from previous page

H	7.424	15.385	11.815
H	9.337	15.063	11.871
H	7.727	15.766	7.556
H	9.978	14.337	9.084
H	7.863	14.252	6.017
H	7.530	12.662	6.545
H	8.698	13.009	5.188
H	11.276	12.956	7.937
H	9.901	11.265	5.766
H	11.527	9.836	5.662
H	13.443	12.287	7.995
H	14.436	12.628	6.641
H	12.744	13.090	6.567
H	13.727	8.864	6.187
H	16.037	10.913	6.746
H[a]	10.684	4.613	5.066
H[a]	17.370	9.579	9.355

[a] Hydrogen capping atoms at the QM/MM boundry.

Table S13: Coordiantes (Å) for Frame 2 from the **Rho-1** run. The coordinates are for the RPSB+Glu113 QM region employed in the excited-state calculations.

C	12.007	4.633	4.487
H	12.133	3.679	3.990
H	11.744	5.351	3.682
C	13.441	4.828	5.068
H	13.583	4.160	5.866
H	14.040	4.622	4.179
C	13.724	6.297	5.487
O	13.516	7.302	4.852
O	14.357	6.361	6.558
C	17.901	8.861	8.947
H	18.999	8.828	9.155
H	17.526	7.965	9.394
C	17.601	8.912	7.494
H	17.925	9.950	7.078
H	18.195	8.166	6.916
N	16.167	8.820	7.142
H	15.709	7.886	7.005
C	8.488	18.087	8.352
C	8.125	19.206	9.400
C	9.108	19.071	10.610
C	8.764	17.830	11.405
C	8.479	16.641	10.507
C	8.592	16.716	9.155
C	8.707	15.643	8.189
C	9.336	14.379	8.308
C	9.498	13.431	7.183
C	10.642	12.614	7.230
C	10.832	11.465	6.421
C	11.953	10.648	6.371
C	13.279	10.988	6.735
C	14.107	9.882	6.768
C	15.421	9.927	7.106
C	7.368	18.033	7.303
C	9.793	18.434	7.726
C	8.308	15.423	11.298
C	8.370	13.241	6.216
C	13.732	12.353	7.103
H	7.226	19.075	6.824
H	6.401	17.693	7.828
H	7.627	17.344	6.481
H	10.633	18.324	8.389
H	9.821	19.437	7.341
H	9.938	17.832	6.823
H	7.028	19.299	9.697
H	8.455	20.133	8.873
H	9.302	20.021	11.122
H	10.109	18.955	10.176
H	7.950	17.970	12.180
H	9.553	17.540	12.045
H	7.860	14.607	10.759
H	7.636	15.673	12.194
H	9.252	15.150	11.761
H	8.545	15.879	7.167

Continued on next page

Table S13 – Continued from previous page

H	9.787	14.057	9.245
H	7.946	14.215	5.954
H	7.528	12.640	6.555
H	8.654	12.765	5.270
H	11.432	12.804	8.002
H	9.945	10.982	6.036
H	11.794	9.759	5.790
H	13.199	13.061	6.453
H	13.607	12.510	8.141
H	14.734	12.598	6.679
H	13.724	8.922	6.463
H	15.802	10.845	7.472
H[a]	11.207	4.585	5.225
H[a]	17.477	9.709	9.485

[a] Hydrogen capping atoms at the QM/MM boundry.

Table S14: Coordiantes (Å) for Frame 3 from the **Rho-1** run. The coordinates are for the RPSB+Glu113 QM region employed in the excited-state calculations.

C	11.938	4.706	4.257
H	11.997	3.704	3.789
H	11.708	5.403	3.467
C	13.331	4.913	4.897
H	13.470	4.223	5.697
H	14.085	4.638	4.194
C	13.734	6.304	5.347
O	13.263	7.235	4.739
O	14.515	6.344	6.286
C	18.117	8.972	9.359
H	19.169	9.243	9.523
H	17.884	7.988	9.816
C	17.827	9.023	7.859
H	18.223	9.836	7.274
H	18.376	8.205	7.429
N	16.389	8.924	7.484
H	15.981	8.005	7.288
C	8.235	18.012	8.728
C	8.020	19.096	9.885
C	9.121	18.938	10.875
C	8.784	17.647	11.727
C	8.462	16.509	10.831
C	8.354	16.662	9.437
C	8.468	15.528	8.573
C	9.322	14.467	8.555
C	9.419	13.408	7.574
C	10.558	12.596	7.716
C	10.929	11.536	6.843
C	12.053	10.807	6.902
C	13.352	10.998	7.461
C	14.212	9.933	7.207
C	15.583	9.996	7.466
C	7.154	18.039	7.716
C	9.561	18.194	7.982
C	8.365	15.145	11.494
C	8.255	13.219	6.588
C	13.800	12.240	8.120
H	6.732	19.019	7.428
H	6.232	17.484	8.053
H	7.391	17.514	6.799
H	10.402	18.054	8.615
H	9.564	19.199	7.460
H	9.626	17.428	7.253
H	7.044	19.021	10.403
H	8.015	20.109	9.459
H	9.020	19.688	11.734
H	10.113	18.942	10.398
H	7.866	17.904	12.260
H	9.636	17.436	12.409
H	7.873	14.438	10.827
H	7.731	15.228	12.404
H	9.415	14.816	11.811
H	7.820	15.489	7.678

Continued on next page

Table S14 – Continued from previous page

H	9.953	14.386	9.463
H	8.232	14.151	6.061
H	7.320	13.106	7.149
H	8.373	12.358	5.935
H	11.300	12.948	8.455
H	10.237	11.269	6.025
H	11.966	9.937	6.218
H	14.897	12.543	8.085
H	13.276	13.129	7.723
H	13.649	12.112	9.181
H	13.846	8.971	6.774
H	16.138	10.917	7.711
H[a]	11.224	4.707	5.081
H[a]	17.603	9.681	10.008

[a] Hydrogen capping atoms at the QM/MM boundry.

Table S15: Coordiantes (Å) for Frame 4 from the **Rho-1** run. The coordinates are for the RPSB+Glu113 QM region employed in the excited-state calculations.

C	11.848	4.731	4.412
H	11.887	3.641	4.205
H	11.740	5.298	3.437
C	13.217	5.072	5.108
H	13.191	4.607	6.095
H	14.018	4.547	4.527
C	13.505	6.541	5.238
O	12.961	7.334	4.430
O	14.363	6.872	6.112
C	18.059	9.191	9.062
H	19.210	9.190	9.204
H	17.754	8.239	9.488
C	17.848	9.312	7.544
H	18.216	10.357	7.229
H	18.478	8.602	6.964
N	16.474	9.226	7.078
H	15.974	8.328	6.994
C	8.358	18.065	8.578
C	8.336	19.238	9.519
C	9.156	19.154	10.799
C	8.691	17.976	11.643
C	8.421	16.781	10.763
C	8.278	16.759	9.404
C	8.404	15.572	8.604
C	9.399	14.626	8.621
C	9.605	13.620	7.602
C	10.768	12.886	7.567
C	10.968	11.849	6.608
C	12.122	11.064	6.400
C	13.442	11.233	6.801
C	14.292	10.135	6.867
C	15.659	10.296	6.993
C	7.090	18.102	7.671
C	9.621	18.141	7.788
C	8.292	15.509	11.690
C	8.440	13.354	6.669
C	14.006	12.662	6.916
H	7.057	19.041	7.207
H	6.191	17.989	8.255
H	7.102	17.256	6.948
H	10.526	18.019	8.449
H	9.536	19.119	7.294
H	9.620	17.291	7.037
H	7.284	19.409	9.796
H	8.575	20.171	9.003
H	9.017	20.146	11.262
H	10.184	19.098	10.551
H	7.972	18.065	12.418
H	9.545	17.696	12.221
H	8.135	14.603	11.171
H	7.527	15.765	12.439
H	9.288	15.311	12.196
H	7.631	15.492	7.786

Continued on next page

Table S15 – Continued from previous page

H	10.139	14.788	9.410
H	8.095	14.370	6.310
H	7.607	13.040	7.277
H	8.577	12.540	5.951
H	11.569	13.060	8.293
H	10.148	11.584	5.994
H	11.866	10.264	5.719
H	14.628	12.775	7.871
H	14.614	12.820	6.009
H	13.209	13.377	6.904
H	13.906	9.088	6.793
H	16.140	11.261	6.936
H[a]	11.012	4.750	5.111
H[a]	17.628	9.971	9.690

[a] Hydrogen capping atoms at the QM/MM boundry.

Table S16: Coordiantes (Å) for Frame 5 from the **Rho-1** run. The coordinates are for the RPSB+Glu113 QM region employed in the excited-state calculations.

C	12.201	4.604	4.357
H	12.231	3.556	4.162
H	12.096	5.061	3.432
C	13.561	5.009	4.930
H	13.680	4.516	5.880
H	14.335	4.658	4.247
C	13.634	6.547	5.200
O	13.335	7.283	4.213
O	14.295	6.951	6.150
C	17.959	9.072	8.859
H	18.997	8.987	8.999
H	17.447	8.107	9.010
C	17.802	9.350	7.325
H	18.244	10.332	7.022
H	18.303	8.639	6.609
N	16.410	9.345	6.901
H	16.022	8.454	6.544
C	8.378	18.171	8.838
C	8.203	19.285	9.944
C	9.170	19.112	11.014
C	9.044	17.794	11.749
C	8.780	16.589	10.767
C	8.655	16.764	9.431
C	8.644	15.696	8.487
C	9.418	14.612	8.624
C	9.528	13.620	7.613
C	10.593	12.744	7.542
C	10.854	11.789	6.491
C	12.002	10.980	6.277
C	13.335	11.245	6.840
C	14.157	10.151	6.556
C	15.536	10.331	6.923
C	7.002	18.036	8.061
C	9.459	18.622	7.924
C	8.636	15.274	11.583
C	8.439	13.505	6.584
C	13.811	12.642	7.258
H	6.450	19.025	8.153
H	6.404	17.289	8.639
H	7.034	17.780	6.999
H	10.374	18.620	8.454
H	9.093	19.564	7.485
H	9.697	17.744	7.298
H	7.193	19.300	10.380
H	8.416	20.257	9.360
H	9.262	19.889	11.787
H	10.226	19.059	10.589
H	8.172	17.747	12.443
H	9.858	17.468	12.422
H	8.062	14.569	11.062
H	7.930	15.611	12.315
H	9.607	14.994	11.935
H	8.114	15.833	7.552

Continued on next page

Table S16 – Continued from previous page

H	10.070	14.560	9.454
H	8.463	14.272	5.781
H	7.427	13.748	7.098
H	8.389	12.471	6.266
H	11.332	12.823	8.340
H	10.007	11.655	5.849
H	11.893	10.226	5.500
H	14.557	12.592	8.041
H	14.391	13.043	6.385
H	12.944	13.294	7.561
H	13.823	9.232	6.149
H	15.893	11.342	7.209
H[a]	11.336	4.687	5.016
H[a]	17.635	9.911	9.474

[a] Hydrogen capping atoms at the QM/MM boundry.

Table S17: Coordiantes (Å) for Frame 6 from the **Rho-2** run. The coordinates are for the RPSB+Glu113 QM region employed in the excited-state calculations.

C	12.300	4.958	3.845
H	12.240	3.999	3.292
H	12.033	5.814	3.084
C	13.676	5.203	4.464
H	13.797	4.649	5.350
H	14.554	4.801	3.891
C	14.099	6.678	4.741
O	13.434	7.662	4.260
O	15.169	6.818	5.394
C	17.943	8.972	8.351
H	19.021	8.874	8.525
H	17.529	7.984	8.693
C	17.822	8.954	6.824
H	18.365	9.802	6.411
H	18.291	8.028	6.519
N	16.429	9.014	6.360
H	16.063	8.125	5.951
C	8.692	18.139	9.387
C	8.668	18.686	10.832
C	9.493	17.984	11.883
C	9.037	16.552	11.977
C	8.867	15.889	10.637
C	8.788	16.547	9.422
C	8.912	15.845	8.175
C	9.760	14.817	7.960
C	9.793	13.934	6.836
C	10.852	13.086	6.804
C	11.031	11.992	5.979
C	12.201	11.277	5.821
C	13.496	11.245	6.411
C	14.314	10.172	6.116
C	15.681	10.071	6.490
C	7.240	18.538	8.769
C	9.827	18.738	8.577
C	8.633	14.427	10.680
C	8.687	14.027	5.887
C	14.036	12.302	7.270
H	7.223	19.573	8.454
H	6.380	18.405	9.509
H	7.270	17.944	7.810
H	10.783	18.279	8.896
H	9.838	19.813	8.807
H	9.786	18.396	7.535
H	7.684	18.585	11.209
H	8.936	19.721	10.864
H	9.293	18.558	12.792
H	10.599	17.975	11.733
H	8.115	16.466	12.561
H	9.811	16.068	12.630
H	7.999	14.050	9.870
H	8.035	14.119	11.625
H	9.554	13.842	10.643
H	8.181	16.192	7.430

Continued on next page

Table S17 – Continued from previous page

H	10.521	14.573	8.715
H	8.575	15.088	5.602
H	7.723	13.911	6.371
H	8.751	13.354	5.083
H	11.539	13.261	7.616
H	10.178	11.712	5.340
H	11.967	10.373	5.248
H	15.143	12.407	7.180
H	13.652	13.304	6.892
H	13.713	12.104	8.306
H	14.046	9.436	5.319
H	16.179	10.990	6.780
H[a]	11.455	4.930	4.533
H[a]	17.496	9.764	8.952

[a] Hydrogen capping atoms at the QM/MM boundry.

Table S18: Coordiantes (Å) for Frame 7 from the **Rho-2** run. The coordinates are for the RPSB+Glu113 QM region employed in the excited-state calculations.

C	12.184	4.657	4.417
H	12.289	3.556	4.184
H	12.070	5.075	3.428
C	13.455	5.123	5.113
H	13.397	4.823	6.182
H	14.336	4.530	4.706
C	13.797	6.587	5.009
O	13.125	7.370	4.357
O	14.793	6.992	5.773
C	17.862	9.016	8.797
H	18.976	8.890	8.870
H	17.433	8.164	9.297
C	17.641	9.045	7.306
H	18.100	9.892	6.848
H	18.191	8.175	6.900
N	16.238	8.987	6.786
H	15.836	8.095	6.379
C	8.594	18.393	8.833
C	8.293	19.200	10.054
C	9.172	18.810	11.229
C	8.853	17.305	11.783
C	8.568	16.432	10.498
C	8.652	16.888	9.260
C	8.619	15.922	8.043
C	9.396	14.856	8.017
C	9.532	13.906	6.941
C	10.483	12.909	7.096
C	10.801	11.938	6.057
C	11.949	11.136	5.865
C	13.266	11.277	6.318
C	14.099	10.132	6.452
C	15.471	10.146	6.787
C	7.406	18.561	7.785
C	9.987	18.800	8.202
C	8.200	15.016	10.891
C	8.572	13.974	5.841
C	13.820	12.594	6.719
H	7.432	19.592	7.418
H	6.424	18.503	8.281
H	7.485	17.950	6.869
H	10.756	18.869	8.973
H	9.986	19.741	7.593
H	10.250	17.946	7.591
H	7.238	19.067	10.323
H	8.562	20.257	9.760
H	9.150	19.632	11.943
H	10.249	18.796	10.893
H	7.983	17.218	12.447
H	9.716	16.951	12.342
H	7.797	14.424	10.054
H	7.394	15.077	11.738
H	9.131	14.538	11.237
H	8.065	16.253	7.168

Continued on next page

Table S18 – Continued from previous page

H	10.038	14.687	8.885
H	8.453	14.980	5.392
H	7.578	13.660	6.214
H	8.817	13.233	5.102
H	11.091	13.044	8.036
H	9.918	11.697	5.414
H	11.775	10.269	5.241
H	14.794	12.728	6.291
H	13.135	13.381	6.402
H	13.935	12.526	7.863
H	13.772	9.116	6.108
H	15.940	11.094	7.155
H[a]	11.262	4.802	4.980
H[a]	17.556	9.877	9.392

[a] Hydrogen capping atoms at the QM/MM boundry.

Table S19: Coordiantes (Å) for Frame 8 from the **Rho-2** run. The coordinates are for the RPSB+Glu113 QM region employed in the excited-state calculations.

C	12.496	4.387	4.513
H	12.545	3.429	4.036
H	12.189	5.090	3.768
C	13.832	4.737	5.149
H	13.936	4.417	6.190
H	14.708	4.233	4.653
C	14.187	6.261	5.128
O	13.482	7.101	4.549
O	15.188	6.507	5.896
C	17.756	8.583	8.568
H	18.863	8.543	8.670
H	17.322	7.659	8.945
C	17.443	8.900	7.138
H	17.752	9.911	6.930
H	18.047	8.244	6.520
N	16.062	8.774	6.808
H	15.805	7.875	6.252
C	8.509	18.313	8.270
C	8.130	19.302	9.355
C	8.860	19.033	10.689
C	8.418	17.671	11.235
C	8.447	16.580	10.203
C	8.545	16.811	8.861
C	8.511	15.780	7.870
C	9.240	14.672	7.835
C	9.271	13.621	6.897
C	10.322	12.737	7.067
C	10.594	11.563	6.319
C	11.736	10.843	6.300
C	13.100	10.992	6.768
C	13.898	9.885	6.520
C	15.235	9.829	6.864
C	7.372	18.357	7.253
C	9.821	18.698	7.619
C	8.248	15.252	10.909
C	8.192	13.518	5.878
C	13.661	12.226	7.455
H	6.915	19.347	7.031
H	6.537	17.835	7.767
H	7.671	17.840	6.299
H	10.625	18.856	8.401
H	9.641	19.678	7.145
H	10.125	17.843	6.968
H	7.069	19.266	9.576
H	8.266	20.383	9.036
H	8.532	19.804	11.392
H	9.941	19.259	10.505
H	7.326	17.946	11.370
H	8.945	17.323	12.108
H	8.250	14.501	10.136
H	7.284	15.228	11.478
H	9.201	15.004	11.466
H	7.891	16.009	7.018

Continued on next page

Table S19 – Continued from previous page

H	9.966	14.469	8.688
H	7.890	14.520	5.435
H	7.220	13.221	6.341
H	8.351	12.846	5.067
H	10.890	13.022	7.977
H	9.867	11.302	5.563
H	11.602	9.870	5.712
H	12.960	13.066	7.291
H	13.649	12.093	8.553
H	14.646	12.430	7.050
H	13.540	9.090	5.808
H	15.689	10.763	7.261
H[a]	11.685	4.397	5.241
H[a]	17.473	9.433	9.190

[a] Hydrogen capping atoms at the QM/MM boundry.

Table S20: Coordiantes (Å) for Frame 9 from the **Rho-2** run. The coordinates are for the RPSB+Glu113 QM region employed in the excited-state calculations.

C	12.310	4.639	4.058
H	12.491	3.575	3.808
H	11.856	5.183	3.267
C	13.653	5.190	4.576
H	14.124	4.606	5.366
H	14.354	5.116	3.742
C	13.787	6.694	4.863
O	13.002	7.531	4.325
O	14.783	7.034	5.680
C	17.886	9.111	8.422
H	19.010	9.214	8.533
H	17.588	8.086	8.872
C	17.562	9.335	6.955
H	17.802	10.355	6.677
H	18.177	8.725	6.250
N	16.106	9.250	6.516
H	15.687	8.296	6.192
C	8.385	18.075	8.674
C	7.928	19.262	9.499
C	8.732	19.237	10.819
C	8.343	17.981	11.573
C	8.415	16.721	10.747
C	8.390	16.722	9.394
C	8.357	15.550	8.551
C	9.050	14.367	8.562
C	9.085	13.464	7.485
C	10.218	12.654	7.470
C	10.469	11.749	6.389
C	11.627	11.152	5.992
C	13.026	11.266	6.421
C	13.914	10.177	6.227
C	15.312	10.279	6.519
C	7.406	17.969	7.463
C	9.856	18.205	8.196
C	8.481	15.501	11.626
C	8.103	13.622	6.356
C	13.473	12.580	7.103
H	7.448	19.038	6.973
H	6.391	17.861	7.863
H	7.728	17.217	6.750
H	10.507	18.141	9.109
H	10.071	19.132	7.663
H	10.217	17.332	7.610
H	6.839	19.234	9.731
H	8.157	20.237	8.945
H	8.476	20.144	11.462
H	9.828	19.185	10.580
H	7.402	18.004	12.003
H	8.947	17.908	12.488
H	8.026	14.586	11.098
H	8.062	15.735	12.609
H	9.590	15.232	11.844
H	7.823	15.665	7.608

Continued on next page

Table S20 – Continued from previous page

H	9.775	14.178	9.380
H	8.329	14.539	5.790
H	7.070	13.730	6.833
H	7.988	12.718	5.734
H	10.824	12.780	8.349
H	9.576	11.560	5.803
H	11.431	10.207	5.446
H	12.535	13.031	7.307
H	14.048	12.353	8.036
H	13.909	13.342	6.473
H	13.493	9.273	5.918
H	15.711	11.273	6.762
H[a]	11.536	4.616	4.825
H[a]	17.442	9.797	9.143

[a] Hydrogen capping atoms at the QM/MM boundry.

Table S21: Coordiantes (Å) for Frame 10 from the **Rho-2** run. The coordinates are for the RPSB+Glu113 QM region employed in the excited-state calculations.

C	12.338	4.721	4.270
H	12.274	3.747	3.757
H	12.106	5.473	3.553
C	13.805	4.769	4.760
H	13.957	4.147	5.624
H	14.504	4.445	3.998
C	14.206	6.258	5.023
O	13.641	7.158	4.418
O	15.256	6.449	5.790
C	17.956	8.605	8.356
H	19.052	8.489	8.450
H	17.517	7.677	8.861
C	17.664	8.744	6.877
H	18.301	9.523	6.457
H	17.968	7.780	6.413
N	16.212	8.877	6.509
H	15.706	8.002	6.154
C	8.324	17.945	8.559
C	8.037	18.961	9.669
C	9.008	18.808	10.820
C	8.789	17.440	11.518
C	8.729	16.321	10.516
C	8.553	16.544	9.146
C	8.596	15.475	8.124
C	9.353	14.341	8.076
C	9.397	13.360	7.022
C	10.533	12.622	6.997
C	10.790	11.589	6.093
C	11.917	10.763	5.903
C	13.267	10.995	6.267
C	14.172	9.946	6.138
C	15.568	10.014	6.367
C	7.220	18.003	7.513
C	9.629	18.232	7.782
C	8.688	14.942	11.149
C	8.313	13.272	6.008
C	13.677	12.373	6.870
H	6.917	19.017	7.291
H	6.349	17.394	7.788
H	7.586	17.542	6.554
H	10.511	18.057	8.443
H	9.656	19.336	7.482
H	9.629	17.531	6.932
H	6.984	18.743	10.068
H	8.074	20.066	9.394
H	8.985	19.605	11.558
H	10.022	18.781	10.356
H	7.819	17.492	11.976
H	9.556	17.209	12.323
H	7.992	14.172	10.626
H	8.219	15.085	12.107
H	9.682	14.458	11.318
H	7.828	15.532	7.374

Continued on next page

Table S21 – Continued from previous page

H	10.161	14.350	8.802
H	8.117	14.225	5.501
H	7.300	13.111	6.453
H	8.460	12.435	5.303
H	11.395	12.755	7.694
H	9.865	11.225	5.697
H	11.756	9.847	5.251
H	12.971	13.183	6.486
H	13.609	12.313	7.977
H	14.688	12.720	6.562
H	13.904	8.967	5.649
H	16.010	10.994	6.546
H[a]	11.474	4.699	4.935
H[a]	17.646	9.426	9.002

[a] Hydrogen capping atoms at the QM/MM boundry.

Table S22: Coordiantes (Å) for Frame 11 from the **Rho-3** run. The coordinates are for the RPSB+Glu113 QM region employed in the excited-state calculations.

C	12.514	4.397	4.249
H	12.570	3.304	3.987
H	12.136	4.987	3.431
C	13.819	4.723	4.986
H	13.850	4.128	5.913
H	14.712	4.504	4.322
C	14.015	6.278	5.329
O	13.406	7.198	4.762
O	14.952	6.420	6.244
C	17.778	8.770	8.828
H	18.851	8.756	8.997
H	17.372	7.901	9.293
C	17.431	8.935	7.361
H	17.843	9.911	6.896
H	17.974	8.127	6.787
N	15.971	8.910	7.121
H	15.551	7.973	6.851
C	8.723	18.433	8.882
C	8.938	19.527	9.969
C	10.109	19.111	10.888
C	9.541	17.917	11.706
C	9.026	16.813	10.810
C	8.745	17.006	9.499
C	8.594	15.924	8.563
C	9.417	14.864	8.565
C	9.297	13.770	7.586
C	10.355	12.870	7.585
C	10.570	11.807	6.645
C	11.699	11.014	6.327
C	13.027	11.058	6.794
C	13.847	9.916	6.656
C	15.155	10.014	7.117
C	7.349	18.775	8.197
C	9.837	18.525	7.829
C	8.758	15.567	11.639
C	8.077	13.603	6.714
C	13.645	12.376	7.264
H	7.296	19.859	8.100
H	6.549	18.568	8.908
H	7.121	18.367	7.243
H	10.848	18.480	8.265
H	9.780	19.420	7.246
H	9.721	17.713	7.138
H	8.119	19.610	10.740
H	9.141	20.509	9.502
H	10.409	19.829	11.576
H	11.097	18.879	10.354
H	8.731	18.291	12.361
H	10.358	17.576	12.444
H	8.262	14.766	11.149
H	8.081	15.975	12.385
H	9.624	15.212	12.177
H	7.901	16.031	7.768

Continued on next page

Table S22 – Continued from previous page

H	10.178	14.721	9.325
H	7.934	14.507	6.129
H	7.221	13.275	7.242
H	8.291	12.825	5.999
H	11.046	13.020	8.355
H	9.739	11.533	5.991
H	11.533	10.262	5.604
H	12.919	13.153	7.553
H	14.287	12.129	8.112
H	14.412	12.735	6.543
H	13.546	8.965	6.169
H	15.563	10.922	7.505
H[a]	11.710	4.349	4.983
H[a]	17.379	9.632	9.363

[a] Hydrogen capping atoms at the QM/MM boundry.

References

(S1) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09 Revision A.02. Gaussian Inc. Wallingford CT 2009.

(S2) Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51–57.

(S3) Vydrov, O. A.; Heyd, J.; Krukau, A. V.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 074106.

(S4) Vydrov, O. A.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 234109.

(S5) TURBOMOLE V5.1 2008, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from `http://www.turbomole.com`.

(S6) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240*, 283–289.

(S7) Widmark, P.; Malmqvist, P.; Roos, B. O. *Theor. Chem. Acc.* **1990**, *77*, 291–306.

(S8) Dunning Jr, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(S9) Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175.

(S10) Send, R.; Valsson, O.; Filippi, C. *J. Chem. Theory Comput.* **2011**, *7*, 444–455.

(S11) Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Roos, B. O.; Ryde, U.; Veryazov, V.; Widmark, P.-O.; Cossi, M.; Schimmelpfennig, B.; Neogrady, P.; Seijo, L. *Comput. Mater. Sci.* **2003**, *28*, 222–239.

(S12) Ghigo, G.; Roos, B. O.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **2004**, *396*, 142–149.

(S13) Forsberg, N.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **1997**, *274*, 196–204.

(S14) Aquilante, F.; Malmqvist, P.-Å.; Pedersen, T. B.; Ghosh, A.; Roos, B. O. *J. Chem. Theory Comput.* **2008**, *4*, 694–702.

(S15) Ferré, N.; Ángyán, J. G. *Chem. Phys. Lett.* **2002**, *356*, 331 – 339.

(S16) Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. *J. Chem. Phys.* **2001**, *114*, 10252–10264.

(S17) Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. *Chem. Phys. Lett.* **2001**, *350*, 297–305.

(S18) Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. *J. Chem. Phys.* **2002**, *117*, 9138–9153.

(S19) Angeli, C.; Pastore, M.; Cimiraglia, R. *Theor. Chem. Acc.* **2007**, *117*, 743–754.

(S20) Neese, F. ORCA – an ab initio, Density Functional and Semiempirical program package, Version 2.8. Max-Planck-Insitut für Bioanorganische Chemie, Mülheim an der Ruhr, 2011.

(S21) Neese, F. *WIREs Comput Mol Sc* **2012**, *2*, 73–78.

(S22) Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. *Chem. Phys.* **2009**, *356*, 98–109.

(S23) Dyall, K. G. *J. Chem. Phys.* **1995**, *102*, 4909–4918.

(S24) Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. *Chem. Phys. Lett.* **2000**, *317*, 472–480.

(S25) CHAMP is a quantum Monte Carlo program package written by C. J. Umrigar, C. Filippi and collaborators.

(S26) Burkatzki, M.; Filippi, C.; Dolg, M. *J. Chem. Phys.* **2007**, *126*, 234105.

(S27) Filippi, C.; Umrigar, C. J. *J. Chem. Phys.* **1996**, *105*, 213–226, As Jastrow correlation factor, we use the exponential of the sum of three fifth-order polynomials of the electron-nuclear (e-n), the electron-electron (e-e), and of pure 3-body mixed e-e and e-n distances, respectively. The Jastrow factor is adapted to deal with pseudo-atoms, and the scaling factor κ is set to 0.6 a.u.

(S28) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Jr, J. A. M. *J. Comput. Chem.* **1993**, *14*, 1347–1363.

(S29) Schautz, F.; Filippi, C. *J. Chem. Phys.* **2004**, *120*, 10931–10941.

(S30) Casula, M. *Phys. Rev. B* **2006**, *74*, 161102.

(S31) We add one s and one p diffuse function on the carbon and the nitrogen using exponents from the aug-cc-pVDZ basis set, taken from EMSL Basis Set Library (http://bse.pnl.gov, accessed Sept 1, 2012).

(S32) Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Chem. Phys.* **2002**, *116*, 6941.

(S33) Strambi, A.; Coto, P. B.; Ferré, N.; Olivucci, M. *Theor. Chem. Acc.* **2007**, *118*, 185–191.

(S34) Melaccio, F.; Olivucci, M.; Lindh, R.; Ferré, N. *Int. J. Quantum Chem.* **2011**, *111*, 3339–3346.

(S35) see `http://sites.univ-provence.fr/lcp-ct/ferre/nf_tinker_qmmm.html` (accessed Sept 1, 2012).

(S36) Ponder, J. W.; Richards, F. M. *J. Comput. Chem.* **1987**, *8*, 1016–1024.

(S37) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(S38) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

(S39) Okada, T.; Sugihara, M.; Bondar, A.-N.; Elstner, M.; Entel, P.; Buss, V. *J. Mol. Biol.* **2004**, *342*, 571–583.

(S40) Valsson, O.; Angeli, C.; Filippi, C. *Phys. Chem. Chem. Phys.* **2012**, *14*, 11015.

(S41) Wanko, M.; Hoffmann, M.; Strodel, P.; Koslowski, A.; Thiel, W.; Neese, F.; Frauenheim, T.; Elstner, M. *J. Phys. Chem. B* **2005**, *109*, 3606–3615.

(S42) Valsson, O.; Filippi, C. *J. Chem. Theory Comput.* **2010**, *6*, 1275–1292.

(S43) Cembran, A.; Bernardi, F.; Olivucci, M.; Garavelli, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6255–6260.

(S44) Valsson, O.; Filippi, C. *J. Phys. Chem. Lett.* **2012**, *3*, 908–912.

(S45) Pastore, M.; Angeli, C.; Cimiraglia, R. *Theor. Chem. Acc.* **2007**, *118*, 35–46.

(S46) Angeli, C.; Cimiraglia, R.; Cestari, M. *Theor. Chem. Acc.* **2009**, *123*, 287–298.

(S47) Angeli, C.; Pastore, M. *J. Chem. Phys.* **2011**, *134*, 184302.