

Electronic Supporting Information for:

Chromophore Interactions Leading to Different  
Absorption Spectra in mNeptune1 and  
mCardinal Red Fluorescent Proteins

Pau Armengol, Ricard Gelabert\*, Miquel Moreno and José M. Lluch

*Departament de Química, and Institut de Biotecnologia i de Biomedicina,  
Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain*

**Contents**

<b>1</b>	<b>RMS of the Full MD Simulation</b>	<b>S2</b>
<b>2</b>	<b>Time Evolution of the Planarity of the Chromophore</b>	<b>S2</b>
<b>3</b>	<b>Changes in Atomic Charges upon Photoexcitation</b>	<b>S3</b>
<b>4</b>	<b>Analysis of the Stabilization of the Photoactive Excited State</b>	<b>S3</b>
<b>5</b>	<b>Benchmarks of System Size and Basis Set</b>	<b>S7</b>
<b>6</b>	<b>Test of QM/MM Excitation Energies</b>	<b>S8</b>
<b>7</b>	<b>CHARMM Force-field Parameters for mNeptune1 and mCardinal Chromophores</b>	<b>S12</b>

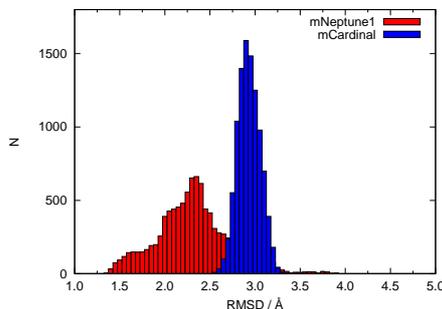


Figure S1: Histograms of the distribution of the RMSD of the protein backbone of each configuration along the MD simulation with respect to the respective X-ray structure

## 1 RMS of the Full MD Simulation

Figure S1 shows the histograms of distribution of the RMS values over the full MD simulation. The spread of this distribution is larger in mNeptune1 than in mCardinal. This can be traced back to the fact that the PDB file of mNeptune1 includes a sequence of 242 amino-acids while mCardinal's only of 228. The extra amino-acids in mNeptune1 constitute a dangling tail from the bottom of the barrel that shows larger mobility in solution. The differences in the mean of the two distributions mean that mCardinal reaches a stable configuration that is slightly more different from the original X-ray structure than in the case of mNeptune1.

## 2 Time Evolution of the Planarity of the Chromophore

Because of the approach used to accumulate MD simulation data, it is only meaningful to extract time-related information from the individual 10 ns stretches of standard simulation. One interesting piece of information is to assess the time-stability of the planarity of the chromophore. We have computed the tilt ( $\tau$ ) and twist ( $\phi$ ) angles (see Figure 1 for a definition), for each of the 10 ns blocks and for each protein.

The values of these two angles are fairly stable over these short stretches of simulation. For mNeptune1, the average tilt is  $\tau=0.57^\circ$ , with an average RMS of  $5.5^\circ$ , while the average twist is  $\phi=-2.8^\circ$  with an average RMS of  $10.1^\circ$ . For mCardinal, the average tilt is  $\tau=-1.9^\circ$ , with an average RMS of  $5.6^\circ$ , while the average twist is  $\phi=-11.0^\circ$  with an average RMS of  $13.3^\circ$ . Figure S2 shows a representation of the time evolution of the tilt and twist angles for both proteins over one of the individual 10 ns blocks of simulation. All these individual blocks have a very similar form to the one shown.

#### 4 ANALYSIS OF THE STABILIZATION OF THE PHOTOACTIVE EXCITED STATE

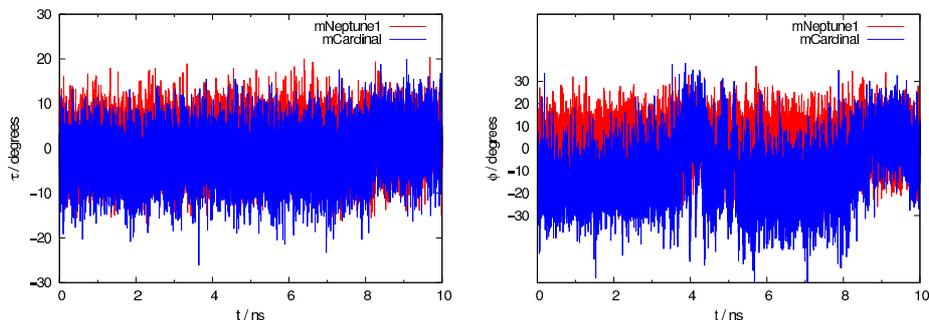


Figure S2: Time evolution along a single 10 ns trajectory of the values of the tilt ( $\tau$ ) and twist ( $\phi$ ) angles, for mNeptune1 (red) and mCardinal (blue).

### 3 Changes in Atomic Charges upon Photoexcitation

With the purpose of quantifying the effects of the electronic excitation that gives rise to band I, a random snapshot of the MD simulation of mNeptune1 with excitation energy under the maximum of band I has been selected and on the corresponding structure a QM/MM calculation has been done. Charges have been computed using the Natural Bond Order (NBO) method for this particular snapshot, for the ground state as well as for the excited state contributing to band I. Table S1 details the values of the NBO charges for the ground state ( $Q(S_0)$ ), excited state ( $Q(S_1)$ ), and its difference. In the Table charges of hydrogen atoms have been summed to the charge of the heavy atom they are bound to. Charges are given in atomic units (a.u.) and are given to two significant figures. The Table details the charges on a moiety basis, and the line named “Total” in boldface contains the global change for that moiety. For atom numbering see Figure S3.

The net result (panel A) shows that there is a negative charge flow from the phenolic moiety to the acylimine moiety, with minor changes on the rest of the moieties of the chromophore. A more detailed analysis (based on individual atoms) reveals charge changes with alternate signs on both the phenolic and imidazolinone moieties, as well as the methine bridge.

### 4 Analysis of the Stabilization of the Photoactive Excited State

The discussion in the main text of the paper concerns itself with the number of interactions that stabilize the oxygen of Phe62 with water and, in the case of mCardinal, also with Gln41. Our result establish that in mNeptune1 this atom interacts with 2.2 water molecules, while in the case of mCardinal this figure lowers to 2.0. However, Phe62 is stabilized by Gln41 about 50% of the time along the MD simulation. The “interaction count” favors mCardinal in this sense. That this explains why the excited state of the chromophore is more stable when Gln41 is close requires also that the interactions Phe62-Wat are of equal strength or weaker than interactions Phe62-Gln41.

4 ANALYSIS OF THE STABILIZATION OF THE PHOTOACTIVE  
EXCITED STATE

Table S1: NBO Charges in Selected Parts of the Chromophore in the Ground Electronic State and Excited State Accessed via Band I

Moiety	Atom	$Q(S_0)^a$	$Q(S_1)^{a,b}$	$\Delta Q^a$
Phenolic:	O <sub>1</sub>	-0.83	-0.77	+0.060
	C <sub>2</sub>	+0.39	+0.37	-0.022
	C <sub>3</sub>	-0.079	-0.038	+0.040
	C <sub>4</sub>	-0.019	+0.013	+0.032
	C <sub>5</sub>	+0.058	+0.035	-0.023
	C <sub>6</sub>	+0.050	+0.044	-0.0054
	C <sub>7</sub>	-0.18	-0.11	+0.066
	<b>Total:</b>		<b>-0.61</b>	<b>-0.47</b>
Methine Bridge	C <sub>8</sub>	+0.20	+0.14	-0.059
	<b>Total:</b>	+0.20	+0.14	-0.059
Imidazolinone	C <sub>9</sub>	-0.039	+0.036	+0.074
	C <sub>10</sub>	+0.70	+0.71	+0.013
	O <sub>11</sub>	-0.73	-0.70	+0.033
	N <sub>12</sub>	-0.49	-0.50	-0.0056
	C <sub>13</sub>	+0.34	+0.38	+0.049
	N <sub>14</sub>	-0.44	-0.53	-0.088
	C <sub>15</sub>	+0.22	+0.22	+0.00023
<b>Total:</b>		<b>-0.45</b>	<b>-0.38</b>	<b>+0.077</b>
Acylimine	C <sub>16</sub>	+0.36	+0.27	-0.093
	N <sub>17</sub>	-0.57	-0.59	-0.016
	C <sub>18</sub>	+0.70	+0.68	-0.019
	O <sub>19</sub>	-0.69	-0.70	-0.015
<b>Total:</b>		<b>-0.20</b>	<b>-0.35</b>	<b>-0.14</b>
Methionine	C <sub>20</sub>	+0.068	+0.051	-0.013
	C <sub>21</sub>	-0.097	-0.10	-0.0032
	S <sub>22</sub>	+0.19	+0.20	+0.0021
	C <sub>23</sub>	-0.095	-0.10	-0.0074
	<b>Total:</b>		<b>+0.066</b>	<b>+0.045</b>

<sup>a</sup> charges in a.u. <sup>b</sup> S<sub>1</sub> is the excited state accessed via Band I

4 ANALYSIS OF THE STABILIZATION OF THE PHOTOACTIVE  
EXCITED STATE

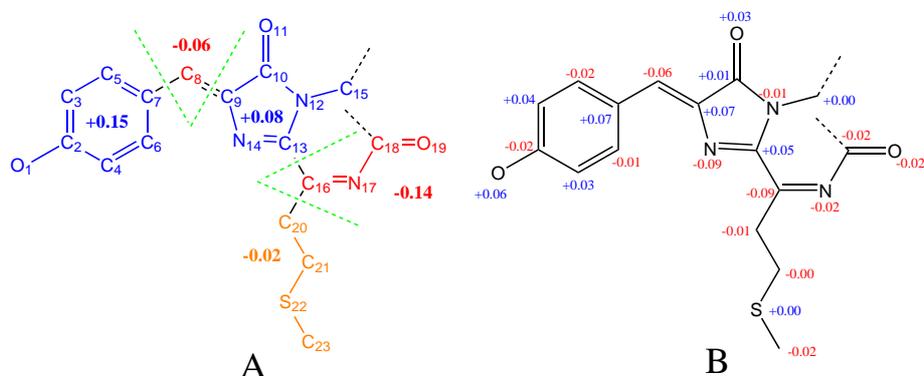


Figure S3: NBO charge changes (in a.u.) undergone in the excitation contributing to band I in a random snapshot of the MD simulation of mNeptune1 with excitation energy under the maximum of band I. A: detail of the structure and the global charge change in each moiety. B: idem, now detailing the charge change of all heavy atoms (hydrogens have been summed to the heavy atom to which they are bonded.). Red and orange figures denote increase of negative charge, and blue ones depletion of negative charge. Green dashed lines separate moieties, and black dashed lines indicate the position of the link atom (H). For the sake of clarity all quantities are given to two decimal figures only.

To analyze this point, we have performed non-polarized QM calculations of the chromophore and neighboring residues of mNeptune1 and mCardinal, and studied the behavior of the excitation energy (computed as the energy difference at a given geometry between the photoactive state and the ground state) as distance to the electrophile varies. Because 1-dimensional potential energy profiles are to be computed we have avoided the QM/MM approach and focused only on the interaction between the chromophore and either water or Gln41. Figure S4 shows the QM model for the calculation, including the distance that is going to be scanned. In a nutshell: the geometries of all atoms except water and glutamine are the same in both systems, and we have computed the excitation energies of the state that describes the same excited state as in the protein (giving rise to band I) for a series of structures with different distances between Phe62 and Wat or Gln41, shown as arrows in Figure S4. For these calculations we have used the same methodology (TDDFT with the same basis set and functional choice) as in the study discussed in the main text. The study of the resulting 1-dimensional potential energy scans permits to analyze the effect of the presence of each electrophile on the excited state energy.

The results of the (unrelaxed) potential energy scans for both models are depicted in Figure S5. The excitation energies of the chromophore in presence of an electrophile behave similarly for both water and glutamine at distances of 3.3 Å and above. Distances below this value show that the excited state of the chromophore lies somewhat lower in energy when the chromophore is interacting with glutamine (when the distance is 2.5 Å this amounts to an extra stabilization of about 0.005 eV, or 0.12 kcal mol<sup>-1</sup>). A maximum difference of this magnitude means that in effect, water and glutamine can be considered to all practical ends to exert interactions that are piecewise of the same intensity

#### 4 ANALYSIS OF THE STABILIZATION OF THE PHOTOACTIVE EXCITED STATE

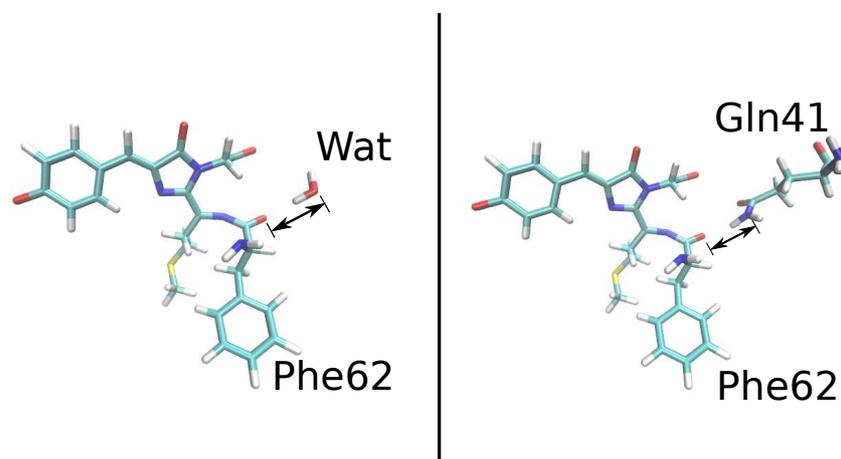


Figure S4: QM cluster models used to analyze the effect of approximation of a water molecule (left) and a glutamine (right) to the oxygen of Phe62 in mNeptune1 and mCardinal proteins.

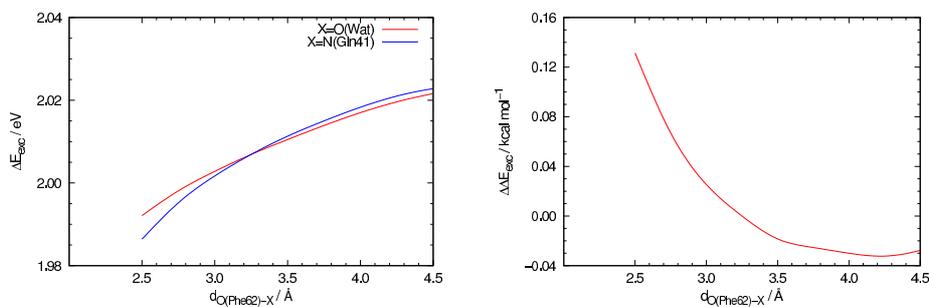


Figure S5: Left panel: Excitation energy along the potential energy scan of approximation of a water molecule (red) or of a glutamine (blue) to the oxygen of Phe62 in mNeptune1 and mCardinal proteins. The excitation energy is computed as the energy difference of the photoactive state and the ground state at the same configuration. Right panel: Excitation energy difference ( $\Delta \Delta E_{\text{exc}}$ ) along the same potential energy scan, computed as the difference between excitation energy for water complex *minus* the excitation energy of glutamine complex.

with Phe62. Thus, it is legitimate to consider that mCardinal has a preferential stabilization of the excited state because the number of interactions (which are now seen to be of almost equal strength) is larger than in mNeptune1.

The water radial distribution function peaks for both proteins at 2.75 Å, and in the case of mCardinal, the peak of the distribution of Phe62-Gln41 distances appears at 3.0 Å. A glance at the right panel in Figure S5 shows that for these short distances, slightly lower excitation energies are rendered by the interaction with glutamine, that is, the interaction with glutamine is a little bit stronger at short distances than with water. The reason behind this slight preferential stabilization by glutamine might be connected to the fact that at short distances glutamine is better able to stabilize charge that is transferred from Phe62 through delocalization over its more extensive expanse, a situation that cannot be accommodated in the case of water.

## 5 Benchmarks of System Size and Basis Set

Besides the lengthy MD simulations, a sizable number of QM/MM calculations have to be done to determine excitation energies and oscillator strengths. Therefore it is necessary to pick an affordable combination of methodology and basis set that makes it possible to obtain reasonably correct results in an affordable length of time. To this end we have performed simple benchmarking testing size of the QM part and basis set, on a single snapshot of the MD simulation of mNeptune1 protein. The calculations described below are QM/MM in nature, and consisted of determining 10 excited states using TDDFT with the CAM-B3LYP functional. We focus specifically on the description of the excitation energy of the excited state of interest, that is, the one that will give rise to band I.

The basis set chosen is of concern because the QM part is in fact anionic. Common wisdom says that diffuse functions might be important to correctly describe anionic systems. We have selected several basis set, from valence double- $\zeta$  to triple- $\zeta$  quality, with polarization functions on all atoms and also including diffuse functions.

Concerning the size of the QM system, we have studied two possibilities. “Large” system consists of the chromophore (phenolic, imidazolinone and acylimine moieties) including also the full methionine residue. Results for this model are presented in Table S2. The “Small” system has the methionine portion treated as a MM appendage to the chromophore and as such has three (not two) link atoms. Results for this model are presented in Table S3. Enlarging the size of the selection radius of the QM part presents a rapid increase (it scales as  $R^3$ ) in the number of atoms.

The results show that there is no real difference in quality, but also not in time, in leaving the methionine out of the QM part. Because that also makes it easier to prepare the input files, we have opted to keep the “Large” model. As for the basis sets, and taking into account the known error attributed to TDDFT to compute excitation energies, all of them rendered values that are quite close together. In view of this, and of the large increase in cost incurred with the larger basis sets, we have decided to stay with the most economic level of calculation.

## 6 TEST OF QM/MM EXCITATION ENERGIES

Table S2: Excitation Energies, Oscillator Strengths & CI coefficients for the “Large” Model

Basis Set	Excitation Energy / eV	$\lambda$ nm	Oscillator Strength	Largest CI coeff.	$t^a$
6-31G(d,p)	2.3302	532.07	0.836	HOMO→LUMO: 0.699	1.0
6-31+G(d,p)	2.3167	535.17	0.866	HOMO→LUMO: 0.695	1.7
6-311G(2d,2p)	2.3141	535.77	0.838	HOMO→LUMO: 0.697	2.0
6-311+G(2d,2p)	2.3056	537.75	0.863	HOMO→LUMO: 0.694	6.5
cc-pVTZ	2.3194	534.56	0.843	HOMO→LUMO: 0.696	6.7

<sup>a</sup> calculation time relative to the calculation with the smallest basis set.

Table S3: Excitation Energies, Oscillator Strengths & CI coefficients for the “Small” Model

Basis Set	Excitation Energy / eV	$\lambda$ nm	Oscillator Strength	Largest CI coeff.	$t^a$
6-31G(d,p)	2.3440	528.94	0.827	HOMO→LUMO: 0.700	1.0
6-31+G(d,p)	2.3201	534.38	0.866	HOMO→LUMO: 0.695	1.8
6-311G(2d,2p)	2.3302	532.08	0.829	HOMO→LUMO: 0.698	1.6
6-311+G(2d,2p)	2.3091	536.94	0.861	HOMO→LUMO: 0.695, HOMO-3→LUMO: 0.100	3.2
cc-pVTZ	2.3299	532.14	0.8546	HOMO→LUMO: 0.697	4.0

<sup>a</sup> calculation time relative to the calculation with the smallest basis set.

## 6 Test of QM/MM Excitations Against Reference Method

To rule out or identify the possibility that errors concur in our determination of excited states, we have selected a snapshot of the MD simulation of mNeptune1 and done QM/MM calculations of the excitation energies using our methodology (TDDFT with the CAM-B3LYP functional) and with a more robust method such as SOS-CIS(D). The calculations have been done on the same geometry, using the 6-31G(d,p) basis set and converging 10 states with TDDFT/CAM-B3LYP, as well as CIS. Later, on the CIS roots the SOS-CIS(D) energy was computed using the auxiliary basis RIMP2-ccpVDZ. The resulting excitation energies, oscillator strengths and CI coefficient details are presented in Tables S4 and S5.

We want to establish a 1:1 correspondence, when possible, between the TDDFT/CAM-B3LYP and their SOS-CIS(D) counterparts. This should be done comparing the nature of the states, which means analyzing the description of the excitation in terms of the MOs involved and their respective nature. This is lengthy and difficult to carry out when, as above, many of the CI expansions include more than two significant excitation amplitudes. In general it is advisable to use the natural transition orbital (NTO) transformation to reduce the CI expansion to a compact set of orbitals to interpret the nature of the excitation.[1] Hence we have done this NTO transformation on the two sets of excited states, which are presented in Figures S6 and S7.

With the NTO descriptions in Figures S6 and S7, it is possible to try to connect TDDFT/CAM-B3LYP and CIS states. Table S6 shows the tentative equivalences between both sets. An asterisk has been included in the Table to

6 TEST OF QM/MM EXCITATION ENERGIES

Table S4: Excitation Energies Obtained at TDDFT/CAM-B3LYP

Excited State Number	Excitation Energy / eV	Oscillator Strength	Largest CI coeffs.
1	2.3027	0.7847	0.70 (87→88)
2	2.7987	0.0002	0.64 (86→88)
3	2.9377	0.0028	0.64 (82→88)
4	3.5040	0.0063	0.68 (85→88)
5	3.5886	0.0198	0.60 (84→88), 0.24 (87→89)
6	3.7101	0.2858	0.50 (87→89), 0.46 (83→88)
7	3.8862	0.1371	0.44 (83→88), -0.40 (87→89)
8	4.0994	0.0005	0.62 (80→88)
9	4.6953	0.0249	0.49 (79→88), -0.33 (81→88), -0.26 (79→90)
10	4.7183	0.0477	0.62 (87→91), 0.20 (86→89)

Table S5: Excitation Energies Obtained at CIS and SOS-CIS(D) Levels

CIS Excited State Number	CIS Excitation Energy / eV	SOS-CIS(D) Exc. Energy / eV	Oscillator Strength	Largest CI coeffs.
1	3.3179	1.6193	1.2855	0.63 (87→88), -0.20 (85→88)
2	3.5757	2.5704	0.0229	0.60 (81→88), -0.26 (81→89)
3	4.8824	3.5862	0.0475	0.57 (86→88), 0.31 (86→89)
4	5.1529	3.5491	0.2904	0.54 (87→89), 0.28 (85→88)
5	5.5354	3.6554	0.1977	0.50 (85→88), -0.26 (87→89), 0.25 (87→88), 0.22 (82→88)
6	5.6633	2.6589	0.0007	0.47 (83→88), 0.34 (83→89), -0.27 (83→94)
7	5.8056	4.5759	0.0112	0.31 (80→91), 0.28 (77→88), -0.26 (84→91), 0.24 (84→88)
8	5.8663	5.1550	0.0011	-0.51 (84→91), 0.24 (84→88)
9	6.1524	4.8445	0.0039	0.52 (74→90), 0.26 (76→90)
10	6.2313	4.6085	0.1573	0.57 (87→92), -0.29 (86→88)

6 TEST OF QM/MM EXCITATION ENERGIES

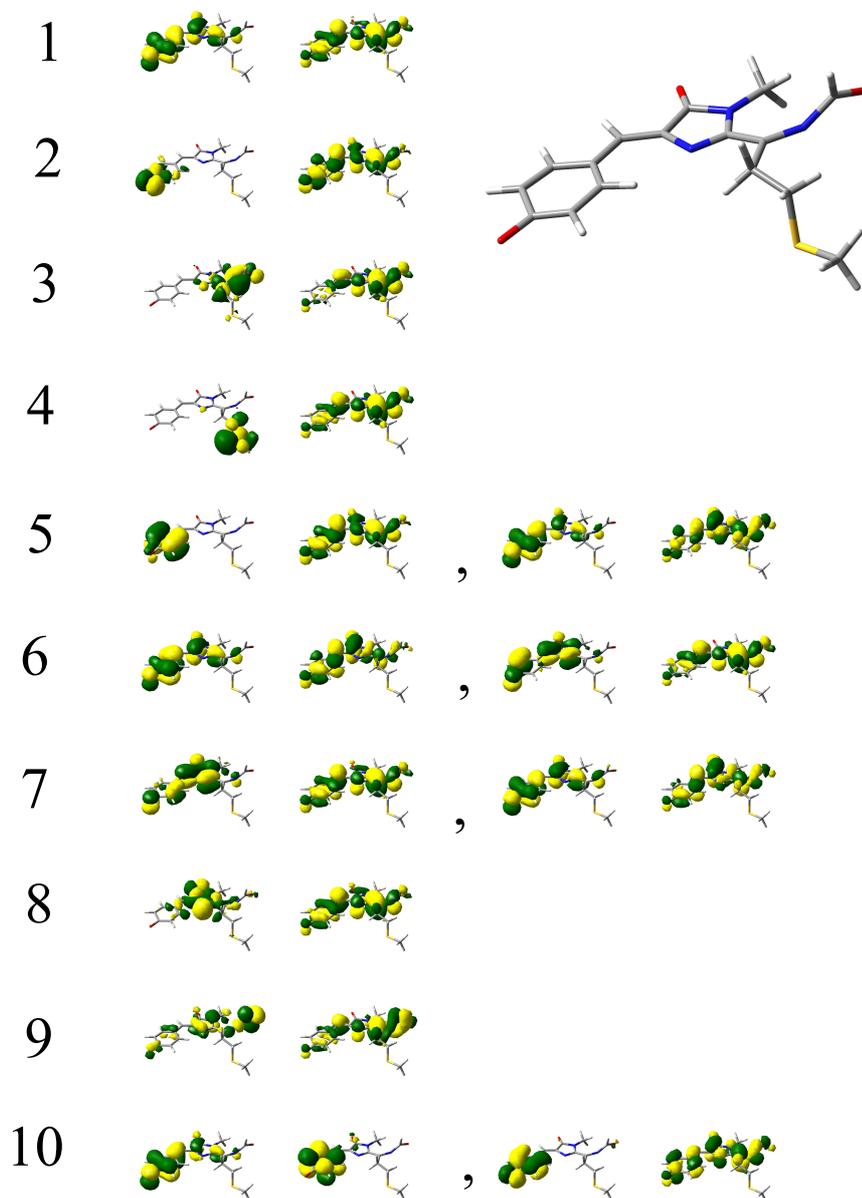


Figure S6: NTO description of the excited states obtained with TDDFT/CAM-B3LYP. The numbers indicate the ordinal number of the TDDFT/CAM-B3LYP excited state, as described in Table S4. Each pair of orbitals indicates the “hole” and the “particle” parts of the pair, respectively. When two pairs are given, the first corresponds to the largest  $\lambda$  eigenvalue and the one after the comma to the second largest eigenvalue.

6 TEST OF QM/MM EXCITATION ENERGIES

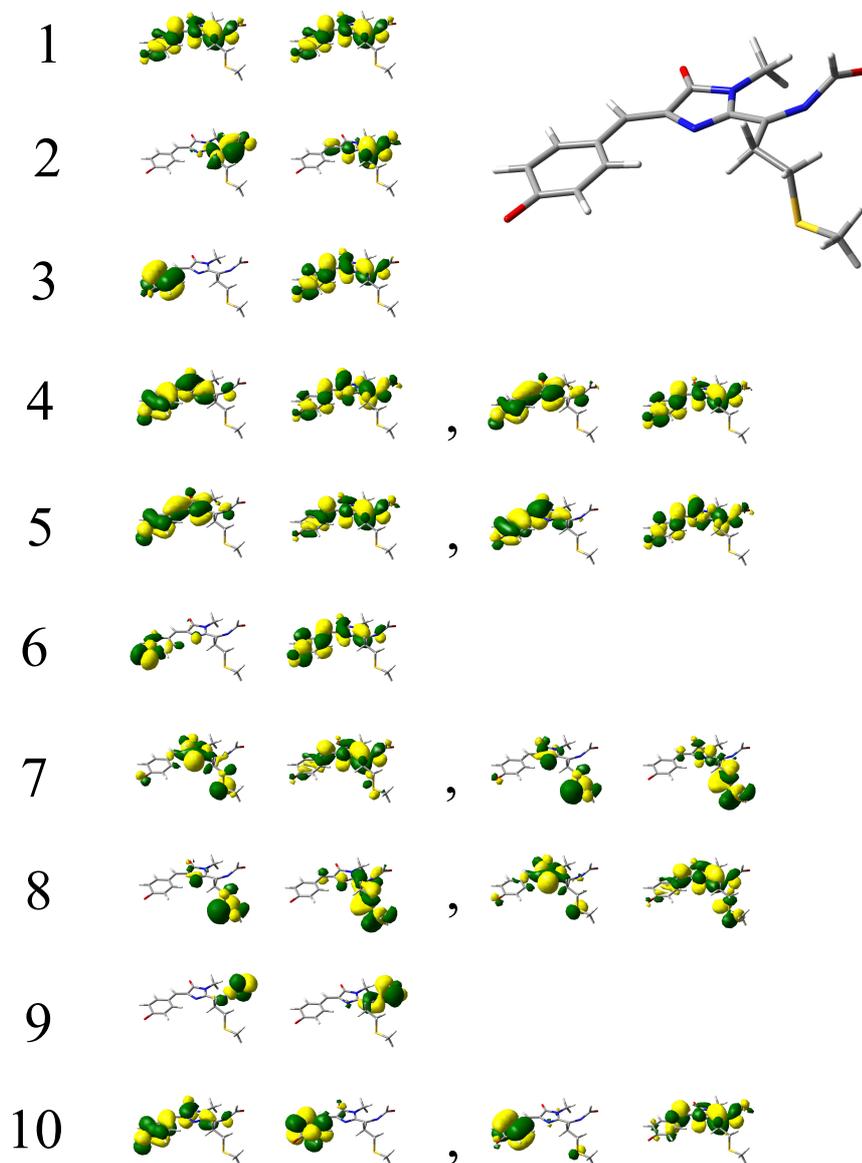


Figure S7: NTO description of the excited states obtained with CIS. The numbers indicate the ordinal number of the CIS excited state, as described in Table S5. Each pair of orbitals indicates the “hole” and the “particle” parts of the pair, respectively. When two pairs are given, the first corresponds to the largest  $\lambda$  eigenvalue and the one after the comma to the second largest eigenvalue.

7 CHARMM FORCE-FIELD PARAMETERS FOR MNEPTUNE1 AND  
MCARDINAL CHROMOPHORES

Table S6: Correspondence between States Computed with SOS-CIS(D) and TDDFT/CAM-B3LYP

State	TDDFT/CAM-B3LYP		CIS State	CIS & SOS-CIS(D)	
	Excitation Energy / eV	$f$		SOS-CIS(D) Excitation Energy / eV	$f$
1	2.3027	0.7847 *	1	1.6193	1.2855
2	2.7987	0.0002	6	2.6589	0.0007
3	2.9377	0.0028	2	2.5704	0.0229
4	3.5040	0.0063	mix 7&8		
5	3.5886	0.0198	3	3.5862	0.0475
6	3.7101	0.2858 *	4	3.5491	0.2904
7	3.8862	0.1371 *	5	3.6554	0.1977
8	4.0994	0.0005	7	4.5759	0.0112
9	4.6953	0.0249	9	4.8445	0.0039
10	4.7183	0.0477	10	4.6085	0.1573

those states that have large oscillator strength and contribute to bands I and II in the computed spectrum.

While the equivalence between states computed with both methods has been done visually assessing the shape of the NTOs involved, it is quite remarkable that a good correspondence exists. Once the connections are included, the SOS-CIS(D) results are also almost ordered energetically (exceptions being state pair 2 and 3, 5 and 6, and 9 and 10). There is also reasonable concordance between the oscillator strengths. The values of the TDDFT/CAM-B3LYP and SOS-CIS(D) energies are in reasonable agreement, with the exception of state 1, where the difference between both is  $\sim 0.7$  eV.

Summarizing: a good correspondence has been found between the series of 10 excited states computed with SOS-CIS(D) and TDDFT/CAM-B3LYP. This confirms that SOS-CIS(D) predicts the existence, for this snapshot, of the three “bright” states, as well as basically all the dark states. Because of this, we can conclude that our method for this study is not affected by the existence of self-interaction error (SIE).

## 7 CHARMM Force-field Parameters for mNeptune1 and mCardinal Chromophores

The values of the CHARMM27 force-field parameters for the mKeima chromophore were published some time ago [2], based on the parameter set developed by Reuter et al. for the GFP chromophore [3], modified to accommodate the new C=N bond present in these red fluorescent protein chromophores.

The parameters we have used in the current study include an improvement in the dihedral section to help maintain the C=N double bond on the plane of the chromophore in mNeptune1 and mCardinal. The adjusted parameter was computed by fitting against QM data derived from MP2-quality potential energy scans.

Figure S8 explains the nomenclature of the affected atoms in the force-field used in Tables S7, S8, S9, S10 and S11 contain the full set of CHARMM27 force-field parameters. A full description of the chromophore requires use of the



Dihedral Angle	$k_\chi / \text{kcal mol}^{-1} \text{ rad}^{-2}$	n	$\delta / \text{degrees}$
X-CA1-C1-X	4.50	2	180.000

Table S9: Torsion force-field parameters

Improper Torsion	$k_\phi / \text{kcal mol}^{-1} \text{ rad}^{-2}$	$\phi_0 / \text{degrees}$
C1-CA1-N2-N3	60.37	-2.97
CA1-N-CB1-C1	103.33	-2.41

Table S10: Improper force-field parameters

Atom Type	Charge / au	$\epsilon / \text{kcal mol}^{-1}$	$R_{\text{min}}/2 / \text{\AA}$
N	-0.8651	-0.09649	1.9352
CA1	0.1339	-0.00043	2.0835

Table S11: Non-bonded force-field parameters