Supporting Informations for:

Picosecond Quantum-classical Dynamics reveals that the Coexistence of Light-induced Microbial and Animal Chromophore Rotary Motion Modulates the Isomerization Quantum Yield of Heliorhodopsin

Riccardo Palombo[1,2,] Leonardo Barneschi[1], Laura Pedraza-González[1,3], Xuchun Yang[2] and Massimo Olivucci[1,2]

[1] *Dipartimento di Biotecnologie, Chimica e Farmacia, Università di Siena, via A. Moro 2, I-53100 Siena, Siena, Italy.*

[2] *Department of Chemistry, Bowling Green State University, Bowling Green, Ohio 43403 USA.*

[3] *Dipartimento di Chimica e Chimica Industriale, Università di Pisa, Via Giuseppe Moruzzi, 13, I-56124 Pisa, Italy.*

S1 The a-ARM protocol

All the quantum mechanics / molecular mechanics (QM/MM) models reported in this work were generated using the automatic rhodopsin model (*a*-ARM) protocol[1–3]. We constructed congruous QM/MM models for wild-type bovine rhodopsin (Rh) and thermoplasmatales archaeon heliorhodopsin (TaHeR), as well as for some of their mutants, starting from, respectively, the 1U19[4] and 6IS6[4] crystallographic structures. An overview of the *a*-ARM models is provided in Figure 2A. A description of the a-ARM workflow is provided below, however, we intent to redirect the reader to the original protocol publications[1,2,5] for a more in-depth reading.

The *a*-ARM protocol is divided in two phases:

Phase I: input file generation
The first phase of the protocol allows for the automatic or semi-automatic computer-aided preparation (i.e., via a five-step command-line procedure) of a 3D structure in PDB format (without hydrogens) named as PDB[ARM]. This PDB file, which is used as input for the QM/MM model generation phase (see below), contains information on: the monomeric protein structure, including the retinal proton Schiff base (rPSB) chromophore and excluding membrane lipids and non-functional ions; the mutant(s) automatically produced via side-chain replacement using MODELLER[6]; the protonation states for all the ionizable residues based on an algorithm that analyze pKa and partial charges using PROPKA[7], automatically assigned but user-customizable; the positions of Cl^-/Na^+ external counterions needed to neutralize model, based on an energy minimization procedure using PUTION[1], optimized automatically and not user-customizable; and an independent file containing the list of amino acid residues forming the cavity hosting the rPSB, determined automatically with Fpocket[8] but user-customizable. The PDB[ARM] and the cavity files constitute the so-called *a*-ARM input for the second phase.

Phase II: QM/MM model generation

The second phase allows the automatic generation of ground-state (S_0) QM/MM models and the subsequent computation of the maximum absorption wavelength (λ_{max}^a) via vertical excitation energy (ΔE_{S1-S0}) calculations. The procedure is described as follows:

- Classical molecular dynamics simulations

a-ARM input is pre-processed using classical molecular dynamics (MD) simulations. First, the positions of crystallographic (or comparative modelling) waters are optimized and the hydrogens for waters and polar residues are added by using DOWSER[9]. Then, the hydrogens for the rest of the protein and chromophore are added and their positions are optimized by a molecular mechanics (MM) energy minimization using GROMACS[10]. A second MM energy minimization is performed, this time on the side-chains (backbone atoms are fixed at the crystallographic/comparative positions) of the residues belonging to the chromophore cavity sub-system. The resulting structure is employed as an input to generate N=10 independent simulated annealing/MD relaxations at 298 K, each starting with a different randomly chosen seed to warrant independent initial conditions that allow to explore the possible relative conformational phase space of the cavity residue side-chains and chromophore. In the ARM MD approach, that uses GROMACS[10] and AMBER[11] force field, only the rPSB chromophore, the lysine linker, all the residue sidechains constituting the cavity and the cavity waters are relaxed, while the rest of the protein remain fixed. The Lys-QM subsystem (the lysine linker residue plus the rPSB chromophore) is described by using a MM parametrization and partial charges computed as AMBER-like Restrained Electrostatic Potential (RESP) charges, which are specifically parametrized for each employed isomer of the chromophore (e.g., 11-cis, all-trans and 13-cis rPSB)[1]. Moreover, the default heating, equilibration, and production times for the MD are 50, 150, and 800 fs, respectively, for a total length of 1 ns. For each of the N=10 replicas, the frame closest to the average structure of the 1 ns simulation is selected as the starting geometry (i.e., guess structure) for constructing the corresponding QM/MM model.

- QM/MM calculations

Each of the 10 replicas is processed by a particular QM/MM approach implemented into the (Open)Molcas[12]/TINKER[13] interface, where the electrostatic embedding scheme used to describe the interaction between the QM and MM parts of the Lys-QM sub-system involves the electrostatic treatment called Electrostatic Potential Fitted (ESPF)[14]. In the ESPF method, the QM part of the chromophore directly interacts with the MM electrostatic potential through one-electron operators whose expectation values represent the QM charge distribution of the chromophore. In addition, the QM/MM frontier is treated within the link atom approach[15] whose position is restrained according to the Morokuma scheme, and it is placed across the covalently bonded lysine $C\varepsilon$-$C\delta$ bond (where $C\varepsilon$ is a QM atom). The charges of the covalently linked lysine are modified by setting the $C\delta$ charge to zero to avoid hyperpolarization and redistribute the residual fractional charge on the most electronegative atoms of the lysine, thus ensuring a +1 integer charge of the Lys-QM layer. All the 63 Lys-QM atoms (i.e., 62 atoms + linker atom) are free to relax during the QM/MM calculation. By employing such approach, the procedure to obtain an ARM QM/MM model consisting of N=10 replicas, can be described as follows. First, a geometry optimization at the Hartree-Fock (HF) level is performed (HF/3-21G/AMBER). Then, another geometry optimization is carried out this time modeling the QM sub-system with the multi-configurational complete active space self-consistent field (CASSCF) at the 2-roots single-state, (CASSCF(12,12)/6-31G*/AMBER level). This follows an energy correction at the multi-configurational second-order perturbation theory (CASPT2) to recover the missing dynamical electron correlation associated with the CASSCF description. Thus, a 3-roots state-average CASPT2 that uses the three-root (S_0, S_1 and S_2) stage-average CASSCF(12,12)/6-31G*/AMBER is computed (CASTP2(12,12)/6-31G(d)/AMBER). Ultimately, each model replica corresponds to an equilibrated gas-phase and globally uncharged monomer QM/MM model, which is used to evaluate the (ΔE_{S1-S0}). The final a-ARM result is the average of the 10 (ΔE_{S1-S0}) values.
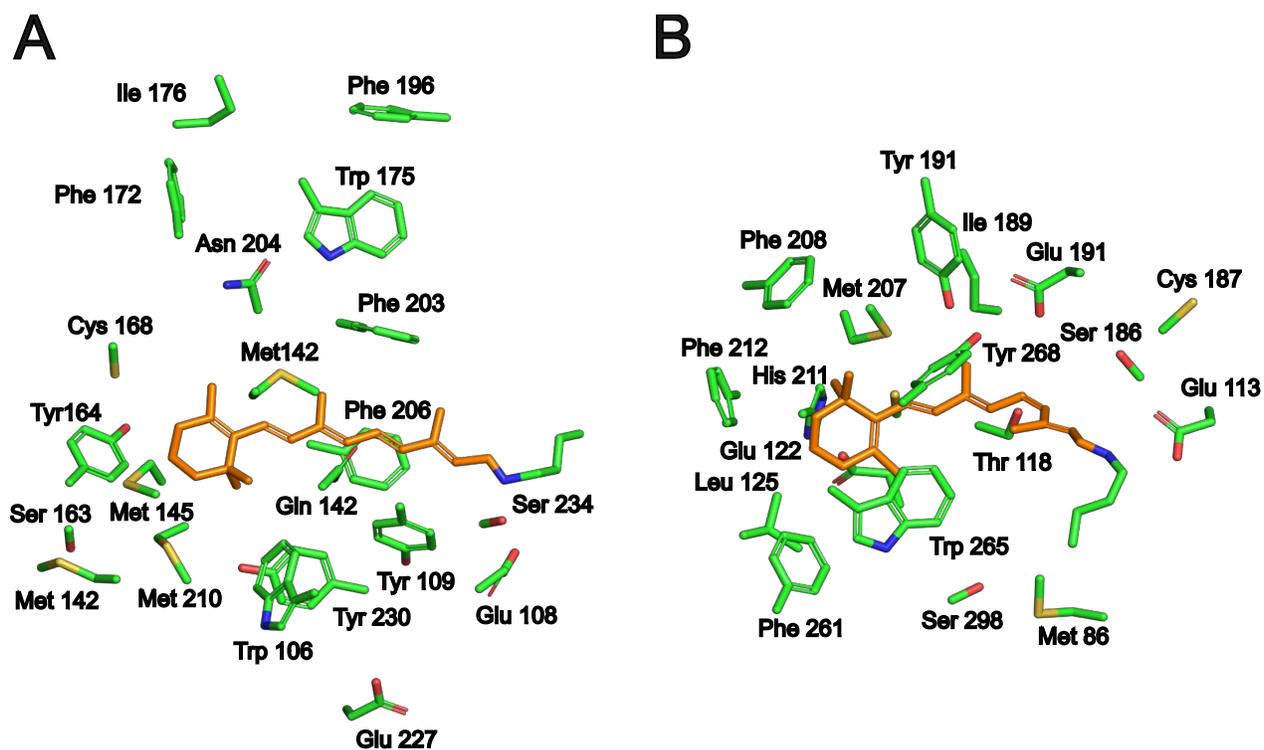
Figure S1. Protein cavity overview. A. Cavity composition in the TaHeR QM/MM model. Residues sidechains (in green) and rPSB chromophore (in orange) are shown. B. Cavity composition in the Rh QM/MM model. Residues sidechains (in green) and rPSB chromophore atoms (in orange) are shown.

Table S1. Comparison between experimental (Exp.) and computed (Comp.) λ^a_{exp} in a set of single and double *Ta*HeR mutants.

	Exp. λ^a_{max} (nm)	Comp. λ^a_{max} (nm)
WT	542[a]	530
A113S	548[a]	535
A207S	538[a]	536
F206A	530[a]	523
S112A	554[a]	545
G146A	528[a]	526
S234A	544[a]	541
M116A	540[a]	532
M142A	528[a]	508
Y164A	540[a]	522
G167A	542[a]	525
F203A	538[a]	515
E108D	542[b]	542
E108D/S112A	559[b]	554
E108D/S234A	540[b]	536

E108D/S78A	532[b]	516

[a] Value found in reference 4
[b] Value found in reference 6

Table S2. Comparison between experimental (Exp.) and computed (Comp.) λ^a_{exp} in a set of single Rh mutants, reported in reference 11.

	Exp. λ^a_{max} (nm)	Comp. λ^a_{max} (nm)
WT	498[a]	499[5]
A269T	514[b]	509[5]
A292S	491[c]	487[5]
E113D	510[a]	516[5]
E122Q	480[c]	477[5]
F261Y	510[b]	509[5]
G90S	489[d]	503[5]
T94S	494[c]	493[5]
W265Y	485[c]	486[5]

[a] Value found in reference 9
[b] Value found in reference 7
[c] Value found in reference 10
[d] Value found in reference 7
[e] Value found in reference 7

S1.1 Choice of the protonation state in the TaHeR model.

The model of TaHeR displays two adjacent histidine residues, His23 and His82, close to the Schiff base moiety of the rPSB chromophore whose protonation state is ambiguous[4] and must be examined in order to assign the correct tautomeric configuration. In fact, this amino acid can display a charge of +1 when both the δ-nitrogen and ε-nitrogen of the imidazole ring are protonated (Hip, [HNδ-His-NεH] configuration), while can be neutral when either the δ-nitrogen (HID, [HNδ-His-Nε] configuration), or the ε-nitrogen (Hie, [NδHis-NεH] configuration) are deprotonated. In this work, for the computation of the TaHeR trajectories, we decided to use the model that yielded the smallest deviation from the experimental $\lambda^a_{exp.}$ For this reason, and leveraging the automatic feature of a-ARM, we built a set of 9 user-customized a-ARM models (see Sec. S1) featuring every possible tautomeric configuration. Figure S2 shows the results in terms of λ^a_{exp} and structural data. Among all, only the configuration with HIP23 and HIE82 (panel H) yielded a good agreement, only -12 nm blue-shifted (λ^a_{exp} = 530 nm) with respect to the experimental value (λ^a_{exp} = 542 nm). This model was therefore selected as operating one and used for trajectory computations (see below, Section S2). All other assessed configurations yielded a poor comparison. Interestingly, our models seem to suggest a net charge of +1 of this histidine couple since when both histidines were set to neutral (panel A,B,D,E) the corresponding models were extremely blue-shifted (~400 nm). Moreover, from the structural point of view all the models displays very similar residue conformations and this suggests a leading electrostatic contribution to the computed λ^a_{exp} shifts found for the different tautomeric configurations.

The His23/His82 couple is conserved among the heliorhodopsin family and a few experimental studies[4,16,17] inferred an active role in the proton transfer pathway of heliorhodopsins (HeR). Furthermore, two recent computational studies[18,19] have investigated the proton transfer pathway in 48C12 HeR, advancing the hypothesis that, after the rPSB photoisomerization, the proton is readily transferred from the rPSB counterion (Glu108 in *Ta*HeR) toward His82. Assuming a similar behavior between 48C12 and TaHeR (43%

of sequence identity), this observation is consistent with our TaHeR model which, indeed, displays a neutral HIE82 residue.
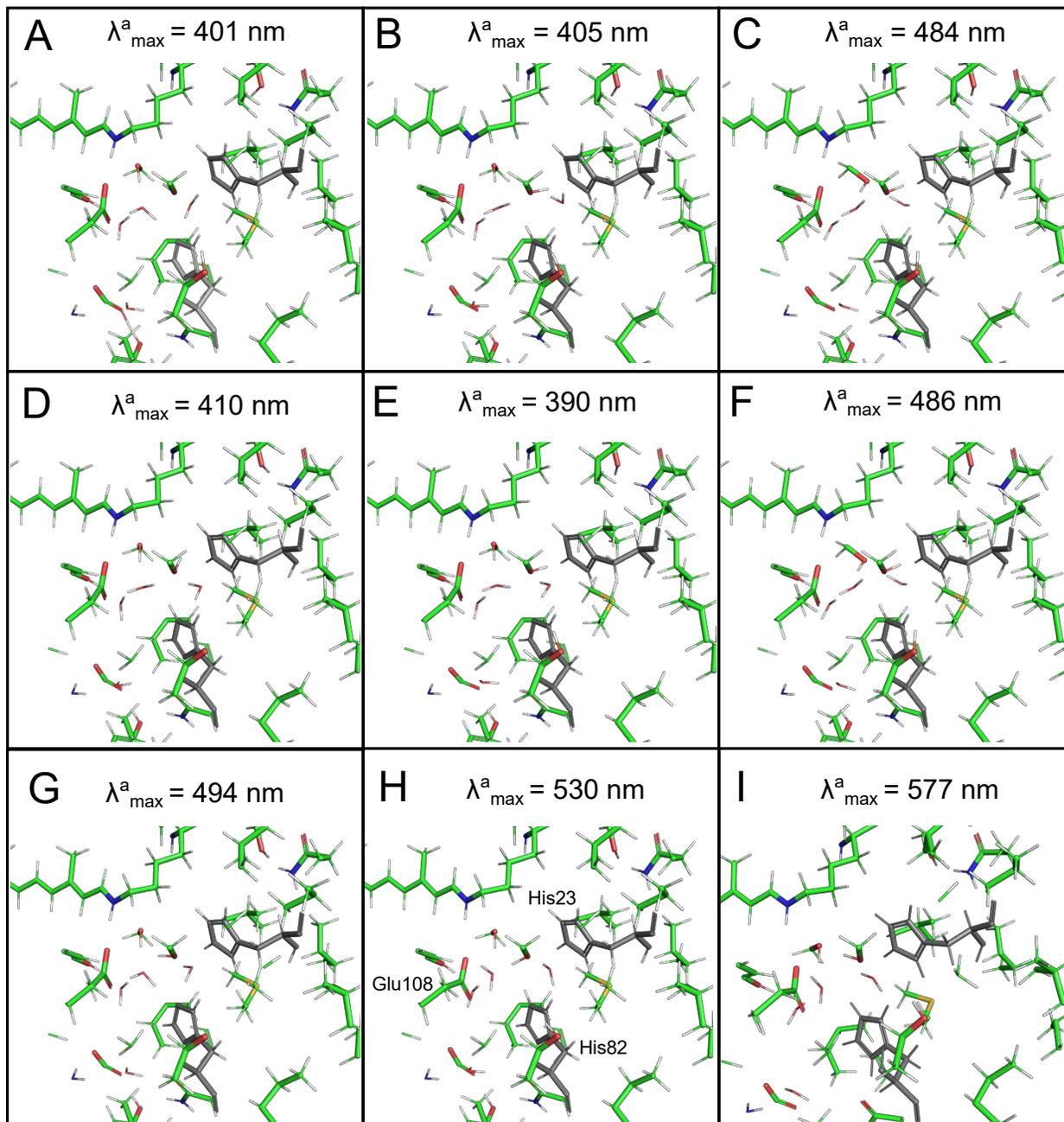
Figure S2. View of different TaHeR *a*-ARM models featuring different His 23 and His 82 tautomeric configurations. (A) 23Hid and 82Hid. (B) 23Hid and 82Hie. (C) 23Hid and 80Hip. (D) 23Hie and 82Hie. (E) 23Hie and 82Hid. (F) 23Hie and 82Hip. (G) 23Hip and 82Hid. (H) 23Hip and 82Hie. (I) 23Hip and 82Hip. In each case the corresponding computed λ^a_{exp} is shown. HID indicates the histidine with [HNδ-His-Nε] configuration, HIE indicates the histidine with [Nδ-His-NεH] configuration and HIP indicates the histidine (+1 charged) with [HNδ-His-NεH] configuration. For panel H, which represents the select configuration for the computation of TaHeR trajectories, rPSB counterion (Glu108), His82 (light gray) and His23 (dark gray) are labelled.

S2 initial condition generation protocol.

The TaHeR and Rh models described above were used to simulate two room temperature Boltzmann-like distributions of 200 initial conditions, where each initial condition consists in one geometry and a set of velocities. In more detail, the following sequential protocol (i-iiii) was adopted [20–23]. i) A room temperature ground state molecular dynamic (MD) of 22 ns was performed at the MM level using AMBER94 force field[11]. The MD consisted in 50 ps of heating phase, of 150 fs of equilibration phase and of 20000 ps of production phase. The MD was carried out in GROMACS[10]. ii) Geometries and velocities corresponding to 200 snapshots were extracted from the production phase every 100 ps. iii) Each snapshot was then used to propagate a 200 fs trajectory at the HF/6-31G*/AMBER94 level of theory followed by iiii) a 50 fs ground state (S_0) trajectory at the two roots SA (i.e S_0 and S_1) CASSCF(12,12)/6-31G*/AMBER94 level of theory. The set of geometries and velocities from the last snapshots of this simulation were then taken as the initial conditions. This set of initial conditions was used to simulate the two absorption bands reported in Figures 1C and 1D. The absorption bands were simulated by the Gaussian convolution of the stick spectra composed by ΔE_{S0-S1} values calculated for each initial condition at the 3 root SA CASPT2/CASSCF(12,12)/6-31G*/AMBER94 level of theory. Oscillator strenghts (f_{S0-S1}) were used to weight the intensity of each $S_0 \rightarrow S_1$ transition. As shown in Figures 1C and 1D the center of the computed absorption band was found very close to the experiments. Moreover, to further test the quality of the initial conditions the 200 corresponding α and α', and bond length alternation (BLA) values were also plotted, and a Gaussian-like shape of the distribution was achieved in all cases (figure S3).

Finally, the two set of 200 initial conditions were propagated on the S_1 using the stochastic Tully surface-hop method[24,25] to account for the S_1 to S_0 hopping events. The corresponding results of the two S_1 population dynamics are documented in the main text, Section B.
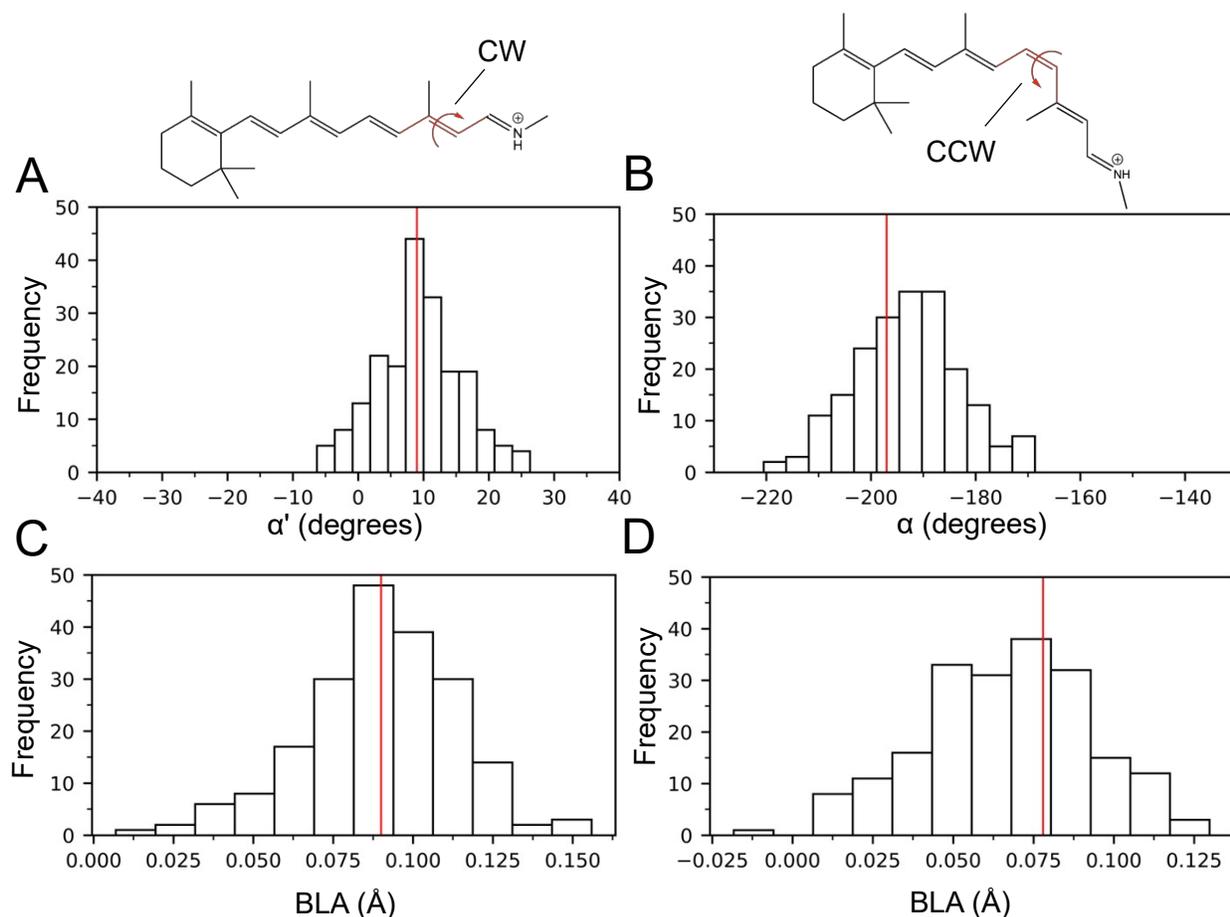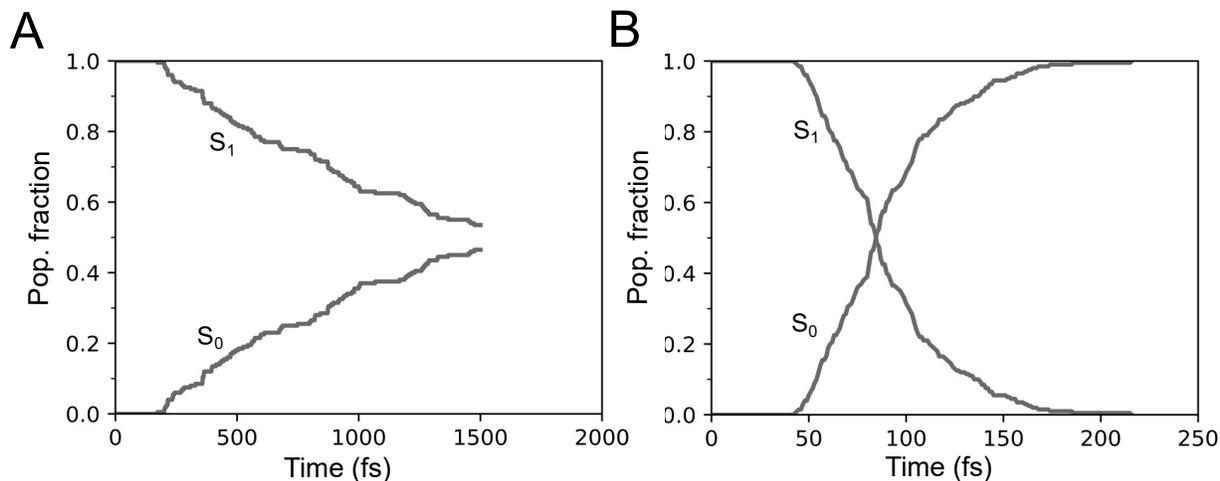
Figure S4. Time evolution of the population fraction propagating along S_0 and S_1 for A) TaHeR and B) Rh. Each starting population is composed by 200 trajectories.
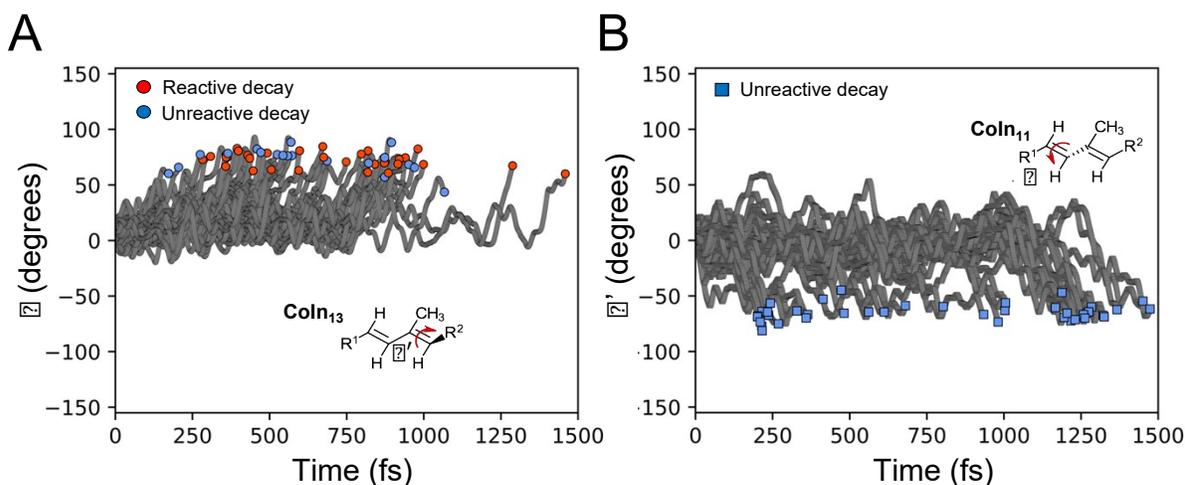
Figure S5. Evolution of the α and α' coordinates (see Fig. 1A) in, respectively, A) TaHeR trajectories undergoing C13=C14 isomerization (i.e. channel CoIn$_{13}$) and in B) TaHeR trajectories undergoing C13=C14 isomerization (i.e. channel CoIn$_{11}$). For the description of the colours and legend see Fig. 3 in main text.

S3 Minimum Energy Path calculation in *Ta*HeR.

Figure S6 shows the S_1 minimum energy paths (MEP's) computed for the TaHeR rPSB chromophore. The first (figure S6A) describes the clockwise (CW) photoisomerization around the C13=C14 double bond while the second (figure S6B) describes the counterclockwise (CCW) photoisomerization around the C11=C12 double bond. Each path is computed by relaxing the Franck-Condon (FC) rPSB structure (i.e the first point in the MEP) and then proceed by a series of constrained geometry optimizations characterized by progressive rotation of 5 degrees of the selected double bond. The choice of the rotations was decided based on the results of the trajectories computations where the CW and CCW isomerization motions were observed for, respectively, the C13=C14 and C11=C12 double bonds (see section B, main text). Finally, in

order to locate the two corresponding minimum energy conical intersections (Min CoIn_A and Min CoIn_B) we carried out geometry optimization without the geometrical constrains, but using the energetical constrain of the energy difference between S_0 and S_1 equal to zero. Both MEP's were computed with the necessary S_1 energy gradients calculated at the 2 root (i.e. S_0 and S_1) SA CASSCF(12,12)/6-31G*/AMBER94 level of theory. However, to account for the missing dynamic correlation effect of the CASSCF method, we have re-evaluated (i.e. single point calculation) the energies of the CASSCF optimized geometries via the more expensive 3 root SA CASPT2/CASSCF(12,12)/6-31G*/AMBER94 level of theory. Fig. S6 shows the corresponding energy profiles for S_0, S_1 and S_2 states.
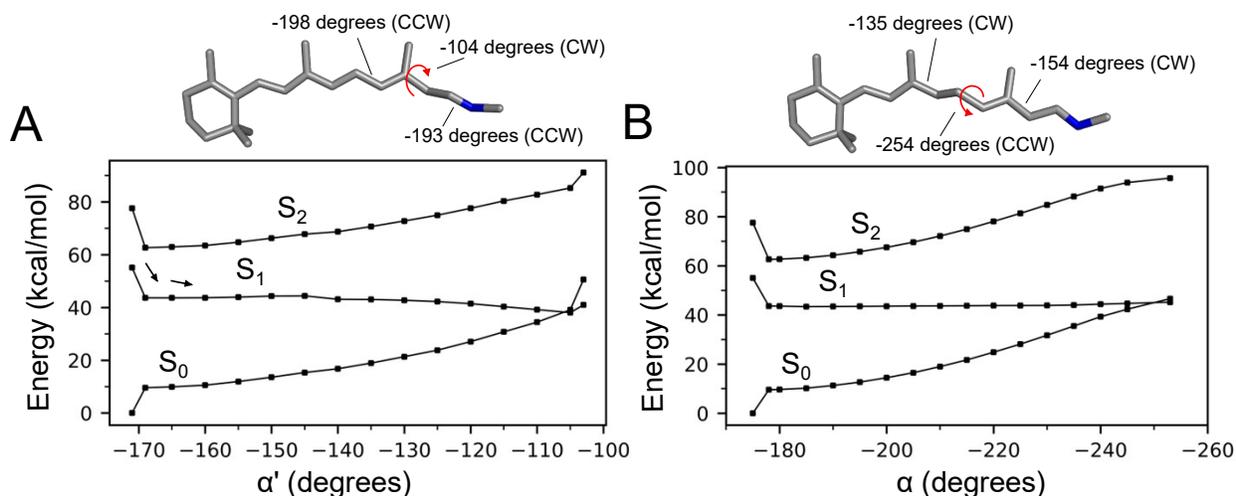
Figure S6. Minimum energy path describing A. the clockwise (CW) photoisomerization around the C13=C14 double bond (α') and B) the counterclockwise (CCW) photoisomerization around the C11=C12 double bond (α). Each profile shows S_0, S_1 and S_2 energy profiles computed at the 3 root SA CASPT2/CASSCF(12,12)/6-31G*/AMBER94 energy profiles. On top, view of the two located minimum energy CoIns (S4A, Min. CoIn_{13} and S4B, Min. CoIn_{11}) displaying the relevant torsional deformations.
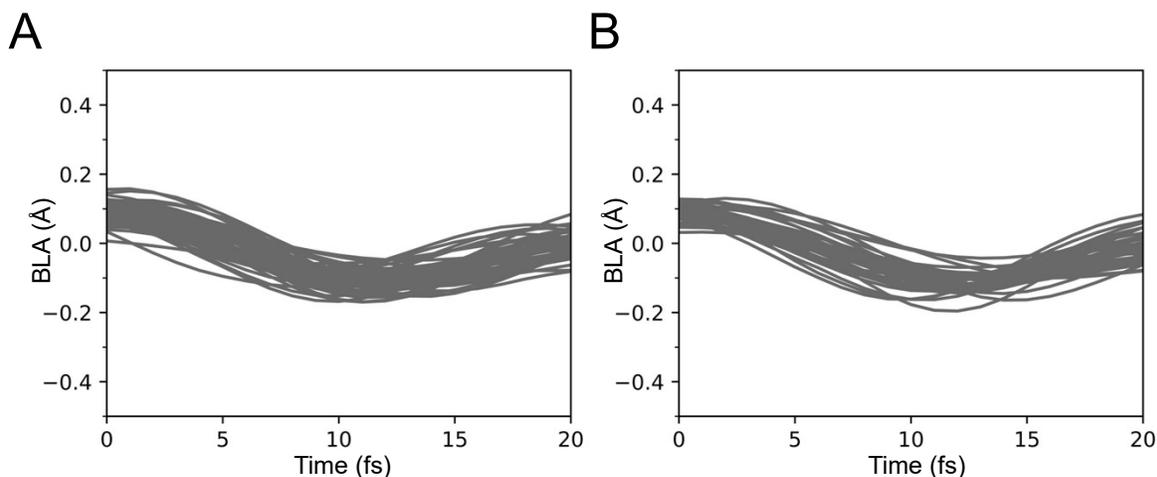
Figure S7. Early time evolution of the bond length alternation (BLA) in A) CoIn_{13} trajectories and in B) CoIn_{11} trajectories. BLA is defined as the difference between the average single bond length minus the average double bond length along the C5 to N of the rPSB conjugated chain. To notice how, in both populations, at

the FC point the BLA evolves from positive to negative values, reflecting the different charge transfer electronic character of S_1 with respect to S_0.
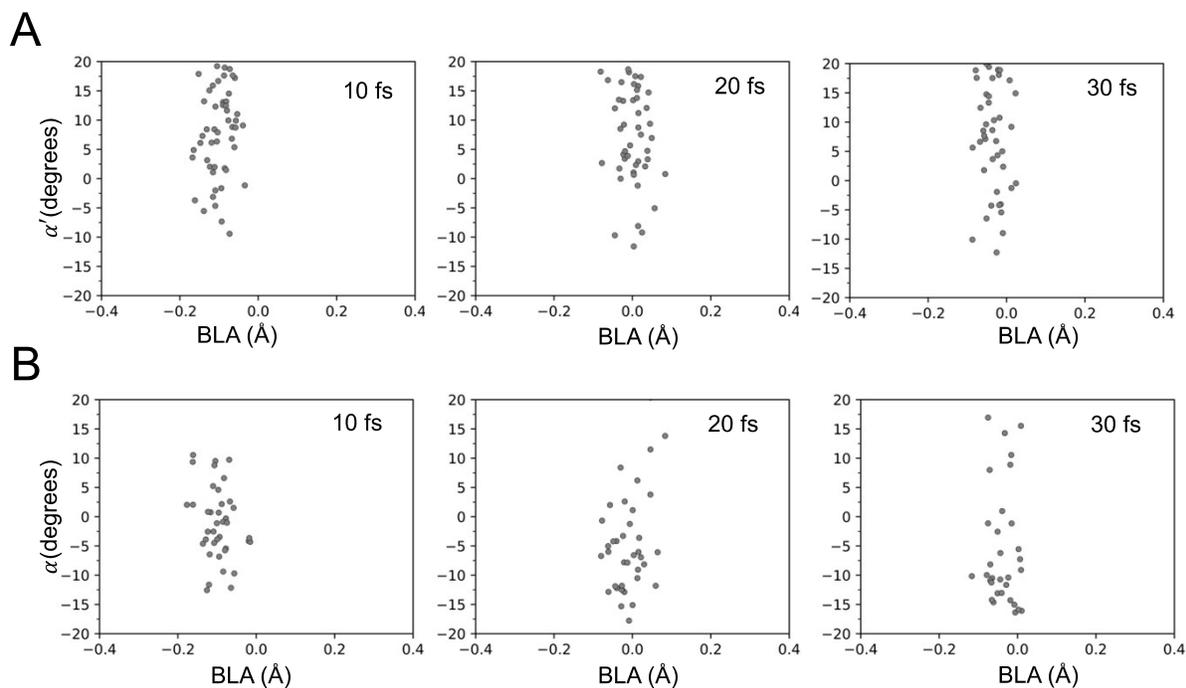
Figure S8. Initial change in α, α' and bond length alternation (BLA) values computed after, from left to right, 10 fs, 20 fs and 30 fs in A) $CoIn_{13}$ trajectories and in B) $CoIn_{11}$ trajectories. α and α' indicate the dihedral angle corresponding to the isomerization coordinate (see Fig. 1A). BLA is defined as the difference between the average single bond length minus the average double bond length along the C5 to N of the rPSB conjugated chain.
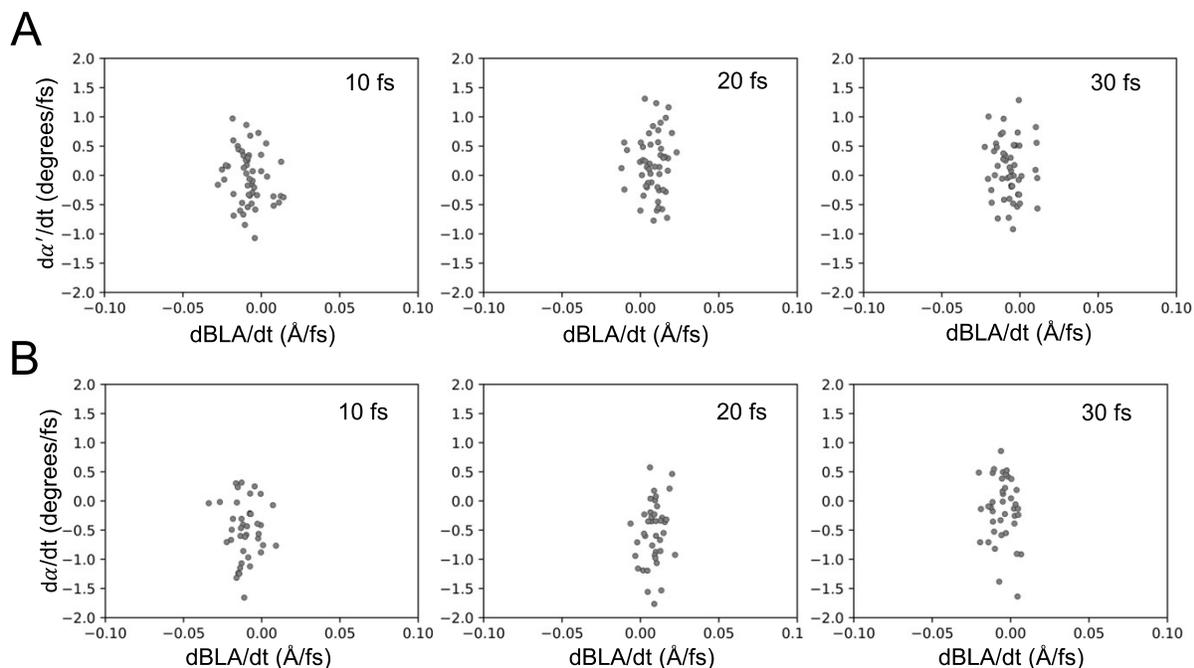
Figure S9. Initial change in α, α' and bond length alternation (BLA) velocities computed after, from left to right, 10 fs, 20 fs and 30 fs in A) $CoIn_{13}$ trajectories and in B) $CoIn_{11}$ trajectories. α and α' indicates the dihedral angle corresponding to the isomerization coordinate (see Fig. 1A). BLA is defined as the difference between the average single bond length minus the average double bond length along the C5 to N of the rPSB conjugated chain.

S4 Analysis of the S2/S1 mixing in _Ta_HeR

In Figures S10, S11, S12 we report the results of 3 root (i.e. S_0, S_1 and S_2) SA CASPT2/CASSCF(12,12)/6-31G*/AMBER94 single point calculations along 3 selected trajectories that belong to the $CoIn_{11}$ (Figure S10, namely geom188, geom204 and geom105) and $CoIn_{13}$ (Figure S11, namely geom207, geom036 and geom021) isomerization decay channels and 3 selected trajectories that have never displayed $S_1 \rightarrow S_0$ decay within the adopted simulation time of 1.5 ps (Figure S12, namely geom075, geom087 and geom206). For each case, the S_1/S_2 energy gap, QM charge (Mulliken analysis) and BLA profiles of the rPSB chromophore are shown. This 3 root state energy re-evaluation has allowed us to observe the involvement of an electronic S_1/S_2 state mixing along each type of trajectory. Besides the S_1/S_2 energy gap, the mixing can be detected from either the charge or the BLA evolution. In fact, the electronic character of S_1 and S_2 states can be described by, respectively, a charge transfer (CT) and diradical (DIR) characters associated with the limiting resonance structures of Fig. S12A. CT and DIR display a different positive charge distribution and, consequently, a different C-C/C=C bond length pattern of the polyene chain. Compared to CT, the DIR exhibits closer C-C/C=C bond lengths and so smaller BLA values. As a result, along the trajectory the occurrence of the S_1/S_2 mixing can be detected by decreasing BLA values approaching the 0. Notice how the mixing is particularly evident in geom206 (Figure S12) where it lasts for nearly 0.5 ps.

In order to further investigate this S_2/S_1 mixing of TaHeR we have also located the corresponding minimum energy CoIn (Min. $CoIn_{S2/S1}$) which was found 7 kcal/mol lower in energy than the FC (Figure S13B). The optimization was carried out at 2 root SA (i.e. S_1 and S_2) CASSCF(12,12)/6-31G*/Amber94 level of theory. Figure S13C shows the Min $CoIn_{S2/S1}$ structure compared with the FC. Consistently with the behavior observed along the trajectories (see above), the Min $CoIn_{S2/S1}$ displays an almost null BLA (0.011 Å).
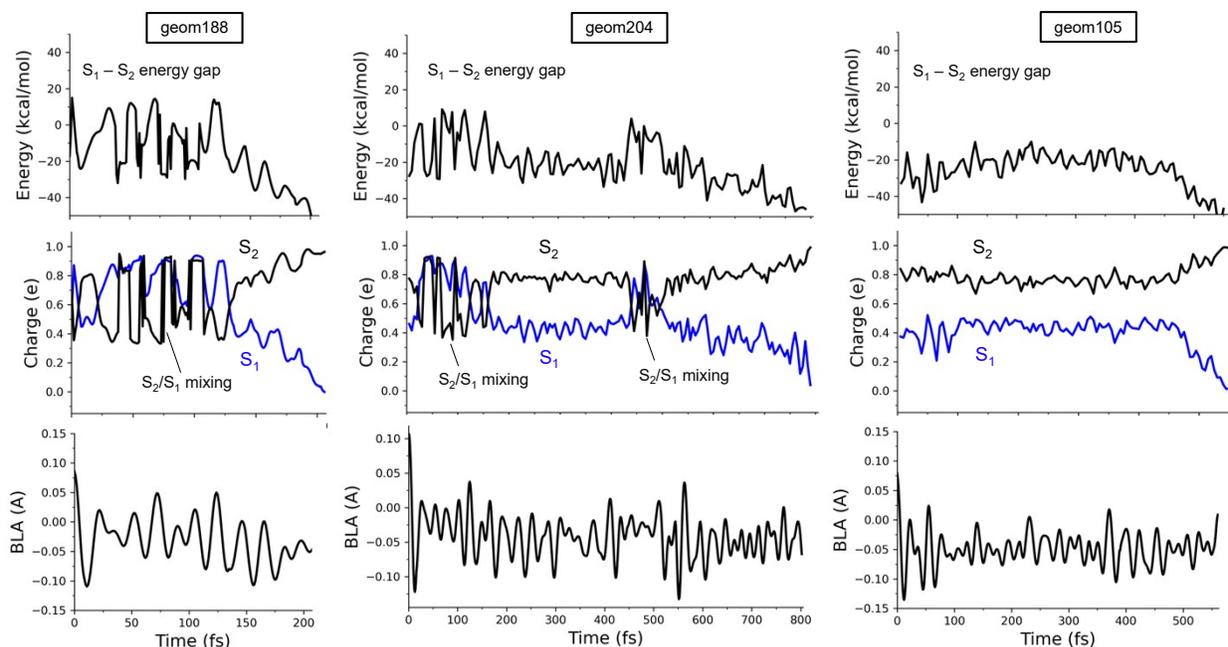
Figure S10. Energy, charge and bond length alternation (BLA) profiles along three TaHeR trajectories belonging to the $CoIn_{11}$ channel (namely geom188, geom204 and geom105). Top row, S_1/S_2 CASPT2/CASSCF(12,12)/6 31G*/Amber94 energy gap evolution. Middle row, QM (Mulliken) charge evolution of the -C12-C13-C14-C15-N rPSB fragment. Bottom row, BLA evolution. BLA is defined as the difference between the average of the double-bond lengths and the average of the single-bond lengths of the rPSB chromophore.
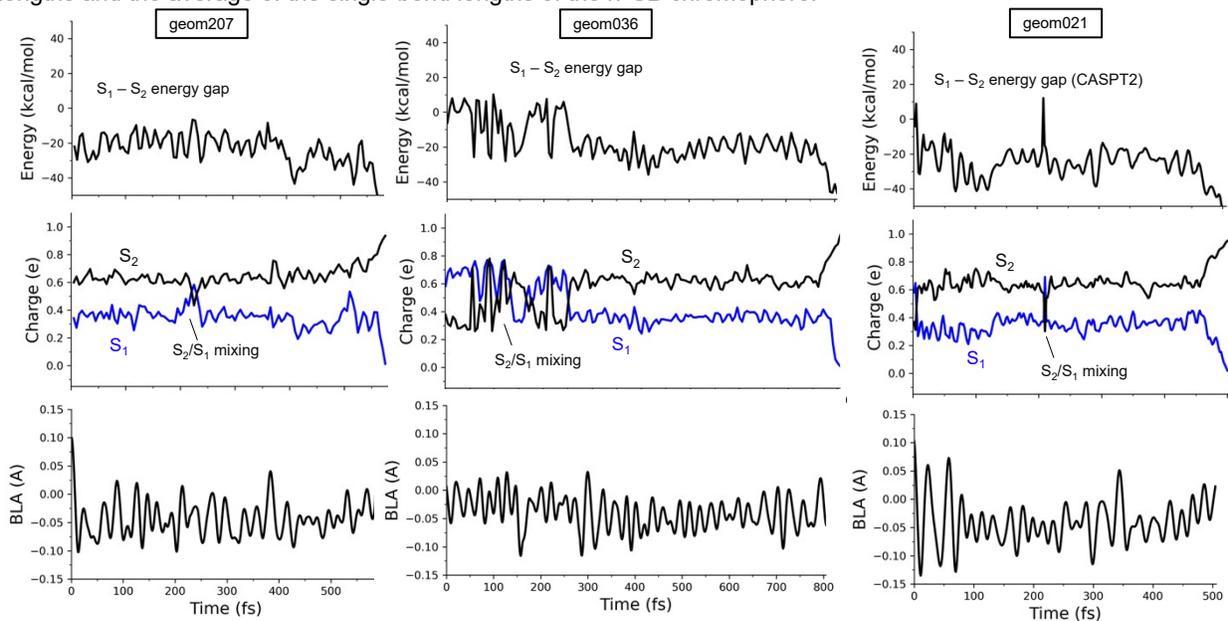
Figure S11. Energy, charge and bond length alternation (BLA) profiles along three TaHeR trajectories belonging to the $CoIn_{13}$ channel (namely geom207, geom036 and geom021). Top row, S_1/S_2 CASPT2/CASSCF(12,12)/6-31G*/Amber94 energy gap evolution. Middle row, QM (Mulliken) charge evolution of the -C14-C15-N rPSB fragment. Bottom row, BLA evolution. BLA is defined as the difference between the average of the double-bond lengths and the average of the single-bond lengths of the rPSB chromophore.
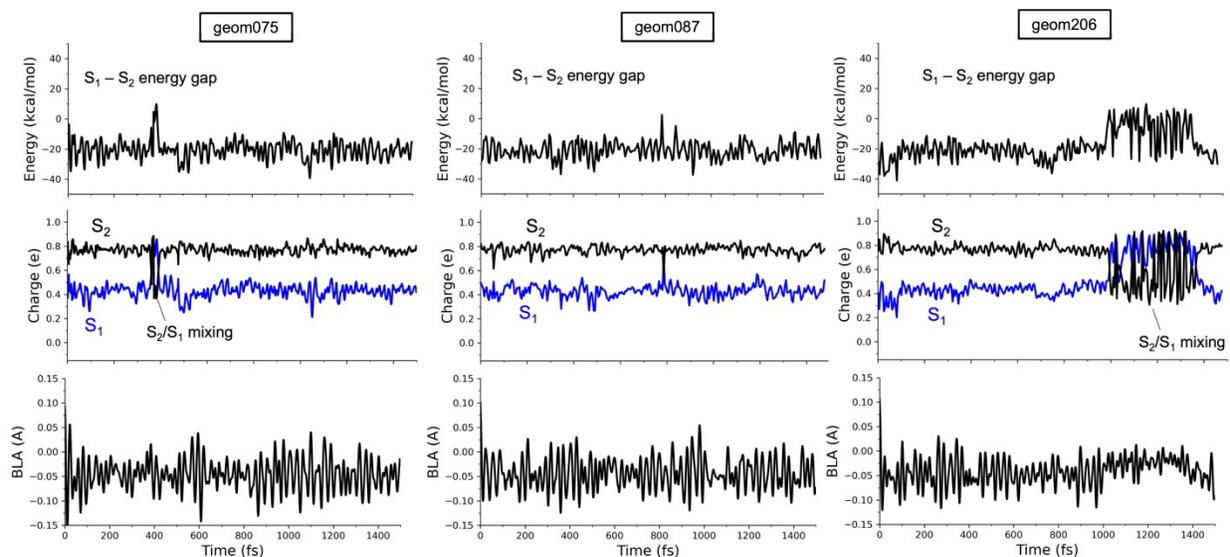
Figure S12. Energy, charge and bond length alternation (BLA) profiles along three TaHeR trajectories that have never displayed $S_1 \rightarrow S_0$ decay within 1.5 ps (namely geom075, geom087 and geom206). Top row, S_1/S_2 CASPT2/CASSCF(12,12)/6-31G*/Amber94 energy gap evolution. Middle row, QM (Mulliken) charge evolution of the -C12-C13-C14-C15N rPSB fragment. Bottom row, BLA evolution. BLA is defined as the difference between the average of the double-bond lengths and the average of the single-bond lengths of the rPSB chromophore.
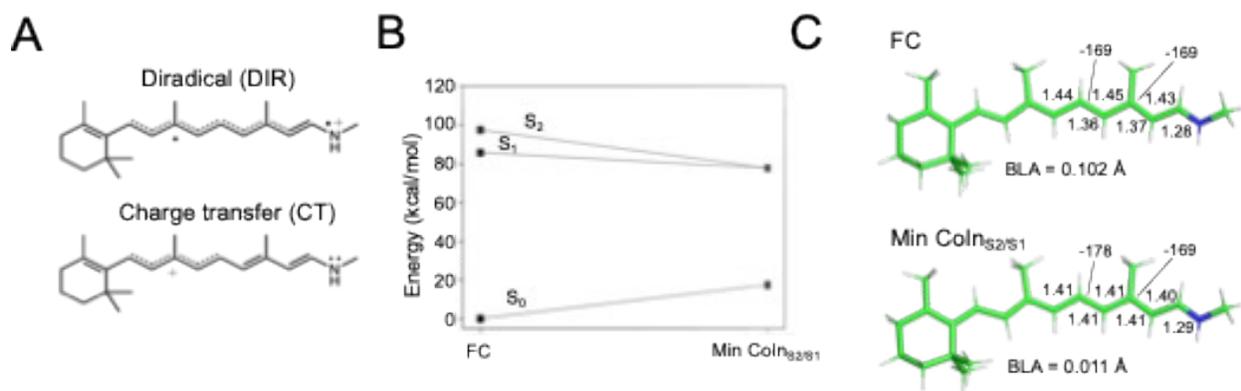
Figure S13. Energetical and geometrical properties of the Min CoIns$_{S2/S1}$ of TaHeR. A) Representation of the two limiting resonance formulas adopted to describe the charge transfer (CT) and diradical (DIR) electronic character of the rPSB chromophore. B) S$_0$, S$_1$ and S$_2$ energy values at the FC and Min CoIns$_{S2/S1}$ structures. C) Representation of the FC and MinCoIns$_{S2/S1}$ structures. Relevant bond lengths (in Å) and dihedral angles (in degrees) of the rPSB chromophore are shown.

S5 Steric effect analysis.

In the attempt to gain a deeper insight into the lack in reactivity of the CoIn$_{11}$ channel, we repeated the calculation of the CoIn$_{11}$ trajectories starting 20 fs before each S$_1$→S$_0$ decay event. Crucially, we performed this calculation after setting the MM charges of the residues in the protein cavity to zero. The simulation allowed us to exclude the impact of the electrostatic of the cavity from the observed restraining force that systematically inverts the motion from CCW to CW for each trajectory toward the approach to the CoIn$_{11}$ decay (see Fig. 5B). These data are reported in Fig. S14. Interestingly, we observed the same inversion for either the set of trajectories computed with the full protein charge or in the absence of the cavity charge (see yellow curves in Fig. S14A and Fig. S14B), suggesting a merely steric origin of the restraining force and, in turn, of the null computed value of $\varphi_{iso,11}$.
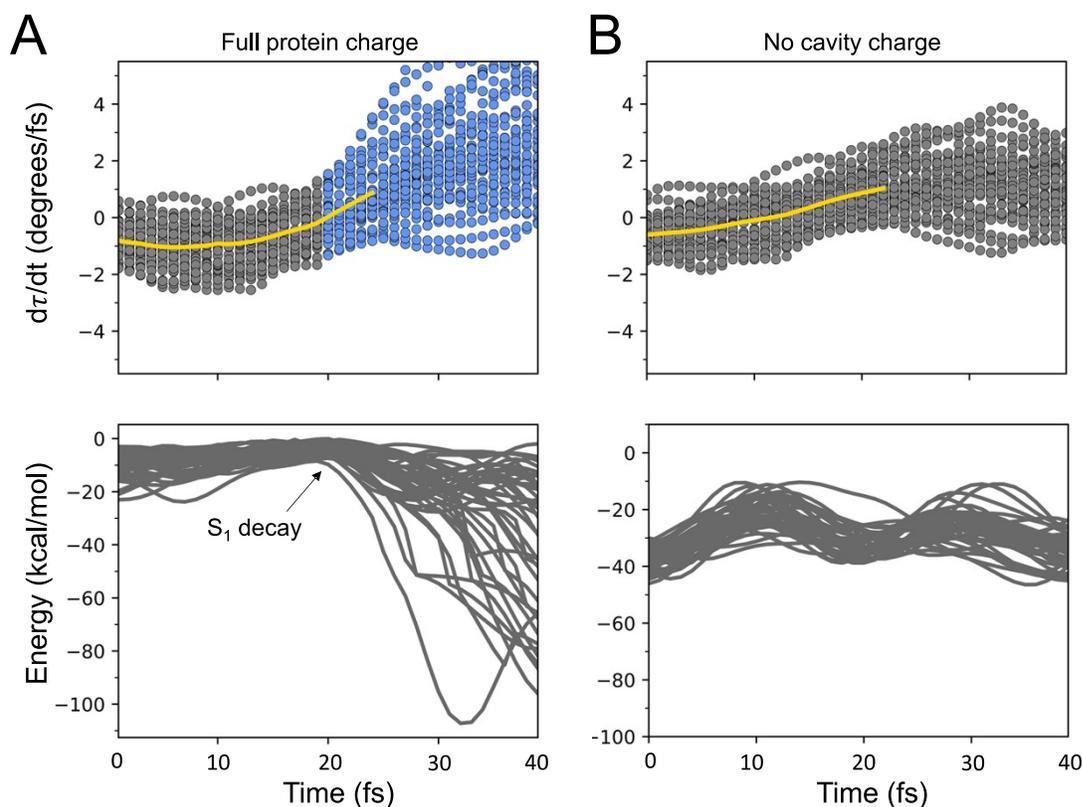
Figure S14. Demonstration of a steric effect along the CoIn$_{11}$ channel by testing the electrostatic effects. A) Top, time evolution of the overlap index dτ/dt towards the CoIn$_{11}$ channel for the full TaHeR model. For each trajectory, the plot starts 20 fs before the S$_1 \rightarrow$ S$_0$ decay time. Accordingly, gray data points refer to S$_1$ while blue data points refer to S$_0$. The yellow curve represents the average value. Bottom, the corresponding time evolution of the S$_1$-S$_0$ energy gap between S$_1$ and S$_0$. B. Same as A but with all charges of the protein cavity set to 0. Top. The dτ/dt slope is very similar to the one of the "charged system", showing a positive slope, revealing that the repulsive force is not due to an electrostatic effect and must thus be due to a steric effect. Bottom. Demonstration that the charges have an effect on the energy S$_1$-S$_0$ energy gap, which become slightly larger and, as a consequence, no S$_1$ decay is observed.

REFERENCES

(1)   Melaccio, F.; del Carmen Marín, M.; Valentini, A.; Montisci, F.; Rinaldi, S.; Cherubini, M.; Yang, X.; Kato, Y.; Stenrup, M.; Orozco-Gonzalez, Y.; Ferré, N.; Luk, H. L.; Kandori, H.; Olivucci, M. Toward Automatic Rhodopsin Modeling as a Tool for High-Throughput Computational Photobiology. *J Chem Theory Comput* **2016**, *12* (12), 6020–6034. https://doi.org/10.1021/acs.jctc.6b00367.

(2)   Pedraza-González, L.; Barneschi, L.; Padula, D.; De Vico, L.; Olivucci, M. Evolution of the Automatic Rhodopsin Modeling (ARM) Protocol. *Top Curr Chem* **2022**, *380* (3), 21. https://doi.org/10.1007/s41061-022-00374-w.

(3)   Pedraza-González, L.; Barneschi, L.; Marszałek, M.; Padula, D.; De Vico, L.; Olivucci, M. Automated QM/MM Screening of Rhodopsin Variants with Enhanced Fluorescence. *J Chem Theory Comput* **2023**, *19* (1), 293–310. https://doi.org/10.1021/acs.jctc.2c00928.

(4)   Shihoya, W.; Inoue, K.; Singh, M.; Konno, M.; Hososhima, S.; Yamashita, K.; Ikeda, K.; Higuchi, A.; Izume, T.; Okazaki, S.; Hashimoto, M.; Mizutori, R.; Tomida, S.; Yamauchi, Y.; Abe-Yoshizumi, R.; Katayama, K.; Tsunoda, S. P.; Shibata, M.; Furutani, Y.; Pushkarev, A.; Béjà, O.; Uchihashi, T.; Kandori, H.; Nureki, O. Crystal Structure of Heliorhodopsin. *Nature* **2019**, *574* (7776), 132–136. https://doi.org/10.1038/s41586-019-1604-6.

(5)   Pedraza-González, L.; De Vico, L.; Marín, M. del C.; Fanelli, F.; Olivucci, M. *A* - ARM: Automatic Rhodopsin Modeling with Chromophore Cavity Generation, Ionization State Selection, and External Counterion Placement. *J Chem Theory Comput* **2019**, *15* (5), 3134–3152. https://doi.org/10.1021/acs.jctc.9b00061.

(6)   Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Bioinformatics* **2016**, *54* (1). https://doi.org/10.1002/cpbi.3.

(7)   Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical p K_a Predictions. *J Chem Theory Comput* **2011**, *7* (2), 525–537. https://doi.org/10.1021/ct100578z.

(8)   Le Guilloux, V.; Schmidtke, P.; Tuffery, P. Fpocket: An Open Source Platform for Ligand Pocket Detection. *BMC Bioinformatics* **2009**, *10* (1), 168. https://doi.org/10.1186/1471-2105-10-168.

(9)   Zhang, L.; Hermans, J. Hydrophilicity of Cavities in Proteins. *Proteins: Structure, Function, and Genetics* **1996**, *24* (4), 433–438. https://doi.org/10.1002/(SICI)1097-0134(199604)24:4<433::AID-PROT3>3.0.CO;2-F.

(10)   Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. *Bioinformatics* **2013**, *29* (7), 845–854. https://doi.org/10.1093/bioinformatics/btt055.

(11)   Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J Am Chem Soc* **1995**, *117* (19), 5179–5197. https://doi.org/10.1021/ja00124a002.

(12)   Aquilante, F.; Autschbach, J.; Baiardi, A.; Battaglia, S.; Borin, V. A.; Chibotaru, L. F.; Conti, I.; De Vico, L.; Delcey, M.; Fdez. Galván, I.; Ferré, N.; Freitag, L.; Garavelli, M.; Gong, X.; Knecht, S.; Larsson, E. D.; Lindh, R.; Lundberg, M.; Malmqvist, P. Å.; Nenov, A.; Norell, J.; Odelius, M.; Olivucci, M.; Pedersen, T. B.; Pedraza-González, L.; Phung, Q. M.; Pierloot, K.; Reiher, M.; Schapiro, I.; Segarra-Martí, J.; Segatta, F.; Seijo, L.; Sen, S.; Sergentu, D.-C.; Stein, C. J.; Ungur, L.; Vacher, M.; Valentini, A.; Veryazov, V. Modern Quantum Chemistry with [Open]Molcas. *J Chem Phys* **2020**, *152* (21). https://doi.org/10.1063/5.0004835.

(13)   Rackers, J. A.; Wang, Z.; Lu, C.; Laury, M. L.; Lagardère, L.; Schnieders, M. J.; Piquemal, J.-P.; Ren, P.; Ponder, J. W. Tinker 8: Software Tools for Molecular Design. *J Chem Theory Comput* **2018**, *14* (10), 5273–5289. https://doi.org/10.1021/acs.jctc.8b00529.

(14)   Ferré, N.; Ángyán, J. G. Approximate Electrostatic Interaction Operator for QM/MM Calculations. *Chem Phys Lett* **2002**, *356* (3–4), 331–339. https://doi.org/10.1016/S0009-2614(02)00343-3.

(15) Singh, U. C.; Kollman, P. A. A Combinedab Initio Quantum Mechanical and Molecular Mechanical Method for Carrying out Simulations on Complex Molecular Systems: Applications to the CH3Cl + Cl? Exchange Reaction and Gas Phase Protonation of Polyethers. *J Comput Chem* **1986**, *7* (6), 718–730. https://doi.org/10.1002/jcc.540070604.

(16) Pushkarev, A.; Inoue, K.; Larom, S.; Flores-Uribe, J.; Singh, M.; Konno, M.; Tomida, S.; Ito, S.; Nakamura, R.; Tsunoda, S. P.; Philosof, A.; Sharon, I.; Yutin, N.; Koonin, E. V.; Kandori, H.; Béjà, O. A Distinct Abundant Group of Microbial Rhodopsins Discovered Using Functional Metagenomics. *Nature* **2018**, *558* (7711), 595–599. https://doi.org/10.1038/s41586-018-0225-9.

(17) Kovalev, K.; Volkov, D.; Astashkin, R.; Alekseev, A.; Gushchin, I.; Haro-Moreno, J. M.; Chizhov, I.; Siletsky, S.; Mamedov, M.; Rogachev, A.; Balandin, T.; Borshchevskiy, V.; Popov, A.; Bourenkov, G.; Bamberg, E.; Rodriguez-Valera, F.; Büldt, G.; Gordeliy, V. High-Resolution Structural Insights into the Heliorhodopsin Family. *Proceedings of the National Academy of Sciences* **2020**, *117* (8), 4131–4141. https://doi.org/10.1073/pnas.1915888117.

(18) Tsujimura, M.; Chiba, Y.; Saito, K.; Ishikita, H. Proton Transfer and Conformational Changes along the Hydrogen Bond Network in Heliorhodopsin. *Commun Biol* **2022**, *5* (1), 1336. https://doi.org/10.1038/s42003-022-04311-x.

(19) Chiba, Y.; Tsujimura, M.; Saito, K.; Ishikita, H. PH-Dependent Binding and Releasing Mechanism of Acetate in the Inner Water Cavity of Heliorhodopsin. *Biochemistry* **2023**, *62* (16), 2363–2370. https://doi.org/10.1021/acs.biochem.3c00193.

(20) Yang, X.; Manathunga, M.; Gozem, S.; Léonard, J.; Andruniów, T.; Olivucci, M. Quantum–Classical Simulations of Rhodopsin Reveal Excited-State Population Splitting and Its Effects on Quantum Efficiency. *Nat Chem* **2022**, *14* (4), 441–449. https://doi.org/10.1038/s41557-022-00892-6.

(21) Schnedermann, C.; Yang, X.; Liebel, M.; Spillane, K. M.; Lugtenburg, J.; Fernández, I.; Valentini, A.; Schapiro, I.; Olivucci, M.; Kukura, P.; Mathies, R. A. Evidence for a Vibrational Phase-Dependent Isotope Effect on the Photochemistry of Vision. *Nat Chem* **2018**, *10* (4), 449–455. https://doi.org/10.1038/s41557-018-0014-y.

(22) Manathunga, M.; Yang, X.; Luk, H. L.; Gozem, S.; Frutos, L. M.; Valentini, A.; Ferrè, N.; Olivucci, M. Probing the Photodynamics of Rhodopsins with Reduced Retinal Chromophores. *J Chem Theory Comput* **2016**, *12* (2), 839–850. https://doi.org/10.1021/acs.jctc.5b00945.

(23) Blanco-Gonzalez, A. and M. M. and Y. X. and O. M. Comparative Quantum-Classical Dynamics of Natural and Synthetic Molecular Rotors Show How Vibrational Synchronization Modulates the Photoisomerization Quantum Efficiency. *Chemrxiv* **2023**.

(24) Granucci, G.; Persico, M. Critical Appraisal of the Fewest Switches Algorithm for Surface Hopping. *J Chem Phys* **2007**, *126* (13). https://doi.org/10.1063/1.2715585.

(25) Tully, J. C. Molecular Dynamics with Electronic Transitions. *J Chem Phys* **1990**, *93* (2), 1061–1071. https://doi.org/10.1063/1.459170.