Electronic Supplementary Information:

Excited-State Intermolecular Proton Transfer and Competing Pathways in 3-Hydroxychromone: A Non-adiabatic Dynamics Study

Alessandro Nicola Nardi[a] and Morgane Vacher*[a]

[a]Nantes Université, CNRS, CEISAM, UMR 6230, Nantes F-44000, France.
E-mail: morgane.vacher@univ-nantes.fr

Contents

S1 UV-vis absorption spectrum

The UV-vis spectrum is decomposed also in terms of the $\pi\pi^*$ and $n\pi^*$ diabatic states defined at the Franck–Condon (FC) point of the ground state (*cis*-)enol. Adopting the same notation reported by Landry et al.[1], the transformation matrix between the sets of adiabatic and diabatic states is given by

$$\Phi_i = \sum_a U_{ai} \Xi_a \tag{1}$$

with $\{\Phi_i\}$ and $\{\Xi_a\}$ the sets of the adiabatic and diabatic states, respectively. The transformation is represented by the matrix U which is dependent on the molecular configuration of the system and has the property of being unitary[2]. In practice, following the local diabatisation scheme[3], the matrix U is computed for each Wigner sampled initial condition as the Löwdin orthonormalised[4] overlap matrix between a given initial condition and the Franck–Condon (FC) point of the ground state (*cis*-)enol. To obtain the energies and transition dipole moments in the diabatic basis the following transformation was employed:

$$W = U^\dagger X U \tag{2}$$

where W and X are the properties in the diabatic and adiabatic bases, respectively. To compute the spectrum, the Hamiltonian matrix and the x-, y-, z-component transition dipole moment matrices of all the initial conditions were obtained in the diabatic basis and the nuclear ensemble approach was adopted[5].

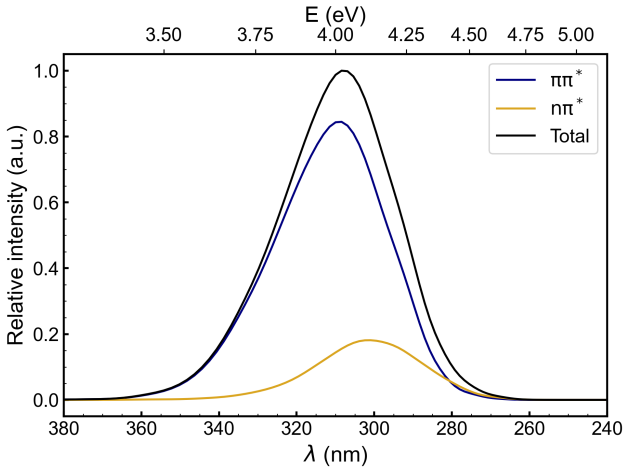


Figure S1: TD-PBE0/cc-pVDZ vertical absorption spectrum of 3-HC obtained via the nuclear ensemble approach. The contributions from the individual $\pi\pi^*$ and $n\pi^*$ diabatic states are highlighted. No energy shift was applied to the computed spectrum.

S2 LIIC along ESIPT coordinate: comparison between cc-pVDZ, cc-pVTZ, aug-cc-pVDZ basis sets

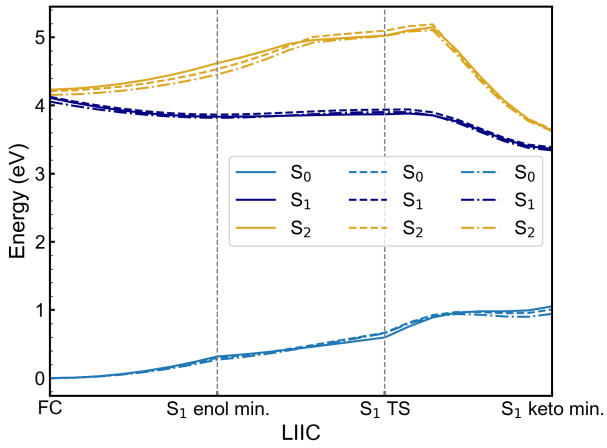
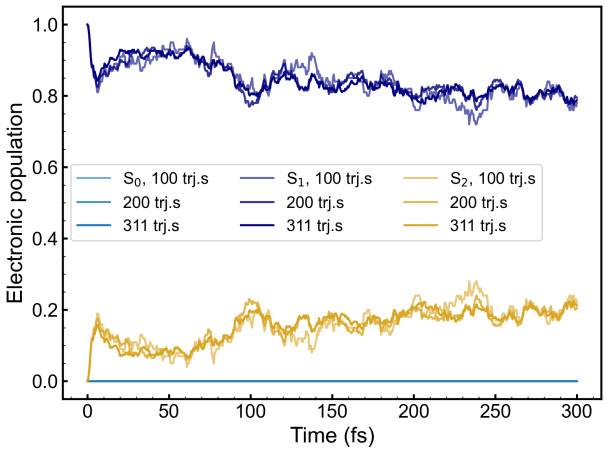
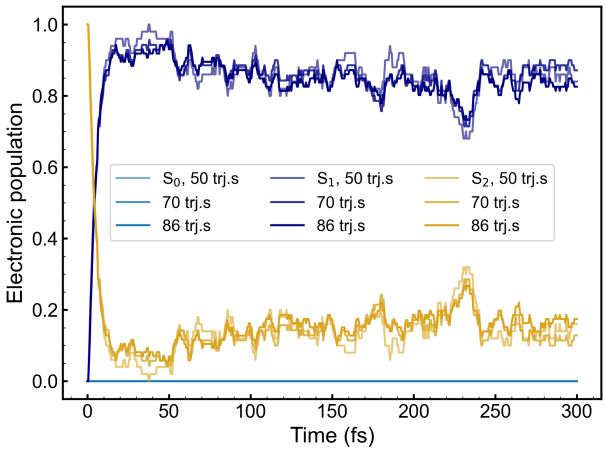


Figure S2: Linear interpolation in internal coordinates (LIIC) along the critical points that characterise the ESIPT coordinate at TD-PBE0/cc-pVDZ (solid lines) compared with the TD-PBE0/cc-pVTZ level of theory (dashed lines), and TD-PBE0/cc-pVTZ level of theory (dash-dotted lines).

S3 Electronic state population convergence



(a) S_1 ensemble.



(b) S_2 ensemble.

Figure S3: Convergence of the electronic state population evolutions upon excitation to the (a) S_1 and (b) S_2 electronic state.

S4 Diabatic electronic state population

The electronic population dynamics can be discussed in terms of the diabatic basis of $\pi\pi^*$ and $n\pi^*$ states at the (*cis*-)enol FC point. To this end, we computed the diabatic electronic populations according to the work of Landry et al.[1] using only the information about the active surface of each trajectory. Let assume, using the same notation reported by Landry et al.[1], that the transformation matrix between the sets of adiabatic and diabatic states is known and it is given by

$$\Phi_i = \sum_a U_{ai} \Xi_a \tag{3}$$

with $\{\Phi_i\}$ and $\{\Xi_a\}$ the sets of the adiabatic and diabatic states, respectively. The matrix U is of course dependent on the position of the system in the configurational space. In this approach, the average diabatic population on the state a, P_a, for an ensemble of trajectories is

$$P_a = \frac{1}{N} \sum_{l=1}^{N} |U_{a\lambda^l}|^2 \tag{4}$$

where N is the number of trajectories in the ensemble, l is the index of the particular trajectory, λ^l is the active state at that particular time step in the l-th trajectory. According to the local diabatisation algorithm[3], the matrix elements of U are obtained as the matrix product of the reference overlap matrix and the time-ordered overlap matrices. The reference overlap matrix is the Löwdin orthonormalised[4] overlap matrix between the Wigner sampled initial condition and the FC geometry. It is necessary to introduce it to provide a common diabatic basis (equivalent to the adiabatic basis at the FC point) to compute meaningful averages across several trajectories.

The diabatic state populations from the full ensemble of trajectories (Figure S4) display a more pronounced change in time compared to the adiabatic ones. The reason for this stems from the fact that the S_1 adiabatic state, once the system crossed the seam, readily accessible from the FC point, changes its character and assumes the $n\pi^*$ character. Indeed, the system's exploration of this region is testified by the increase in the $n\pi^*$ diabatic population. After about 100 fs, the diabatic electronic state populations reach an equilibrium, as well as the adiabatic ones.

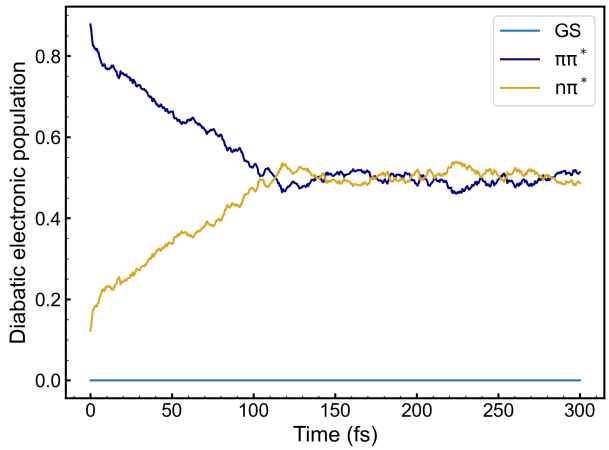
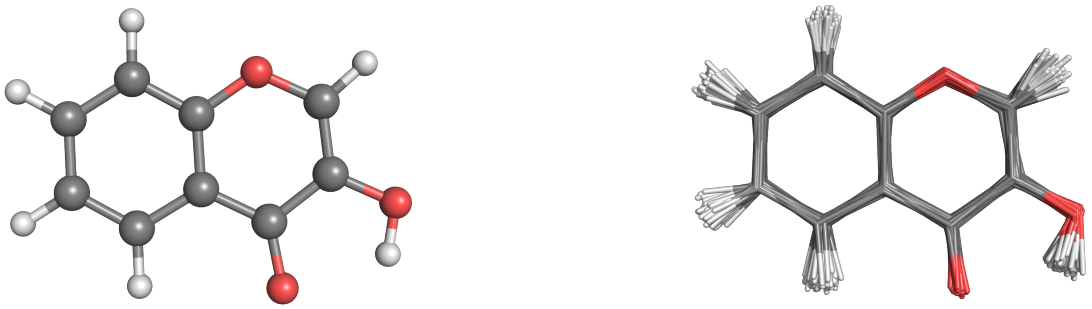


Figure S4: Diabatic electronic state population evolution from the full trajectory ensemble.

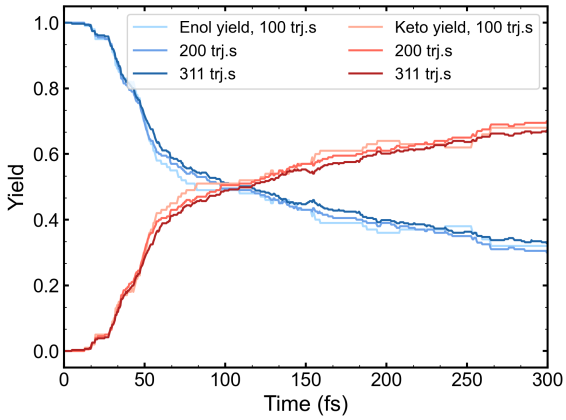
S5 Enol-type MECI and hopping geometries



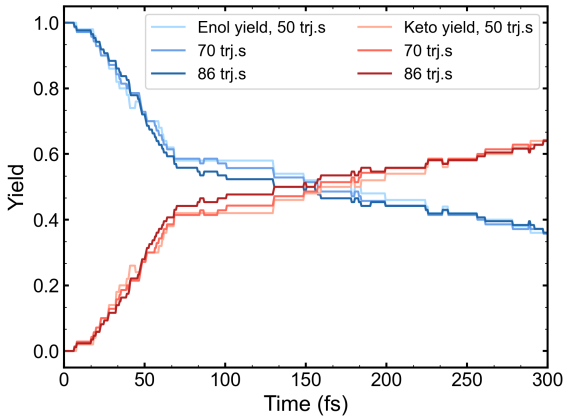
(a) (b)

Figure S5: Ball and stick representation of the (a) planar (C_s symmetry) S_2/S_1 enol-type MECI and (b) stick representation of the hopping geometries between the S_1 and S_2 electronic states within the first 25 fs of the mixed quantum-classical trajectories.

S6 Time-dependent yields convergence



(a) S_1 ensemble.



(b) S_2 ensemble.

Figure S6: Convergence of the time-dependent yields of the reactant (enol, in blue) and product (keto, in red) forms upon excitation to the (a) S_1 and (b) S_2 electronic states.

S7 Comparison between the mono- and bi-exponential fit of the yields

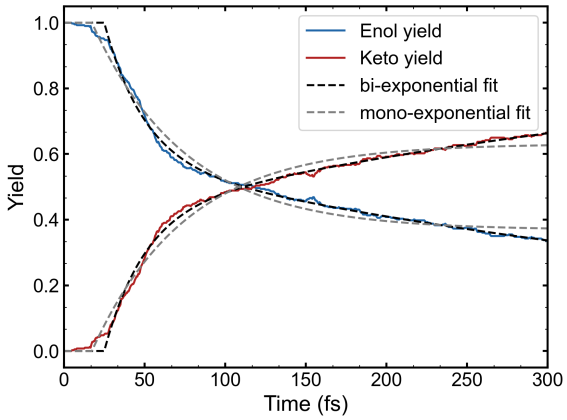


Figure S7: Time-dependent yield of the reactant (enol, in blue) and product (keto, in red) forms. The black dashed line represents the bi-exponential fit of the reactant and product populations in time. The grey dashed line represents the mono-exponential fit of the reactant and product populations in time.

Fit parameters (bi-exponential)						
t_0 (fs)	δ_1	δ_2	τ_1 (fs)	τ_2 (fs)	C	RSS
24.97 (0.20)	0.42 (0.01)	0.58 (0.14)	24.82 (0.84)	510.92 (162.79)	0.00 (0.13)	0.08
Fit parameters (mono-exponential)						
t_0 (fs)	δ_1	δ_2	τ_1 (fs)	τ_2 (fs)	C	RSS
17.32 (0.42)	1.00	-	58.02 (0.85)	-	0.37 (0.01)	0.31

Table S1: Optimised parameters and the residual sum of squares (RSS) from the fit of the enol yield in time with a mono-exponential or bi-exponential decay function. In parentheses one standard deviation errors on the fit parameters are reported.

S8 Influence of the methylcyclohexane solvent on the LIIC along ESIPT coordinate

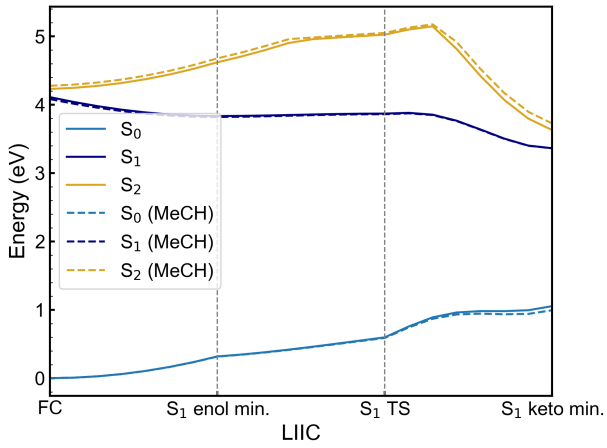


Figure S8: Linear interpolation in internal coordinates (LIIC) along the critical points that characterise the ESIPT coordinate in the gas phase (solid lines) and in methylcyclohexane (MeCH) (dashed lines) at TD-PBE0/cc-pVDZ. The solvent was described as a continuum within the IEFPCM model.

S9 Yields from S_1 and S_2 ensembles

The enol and keto yields in time were obtained for the ensembles of trajectories that started from the excited-states S_1 and S_2, separately.

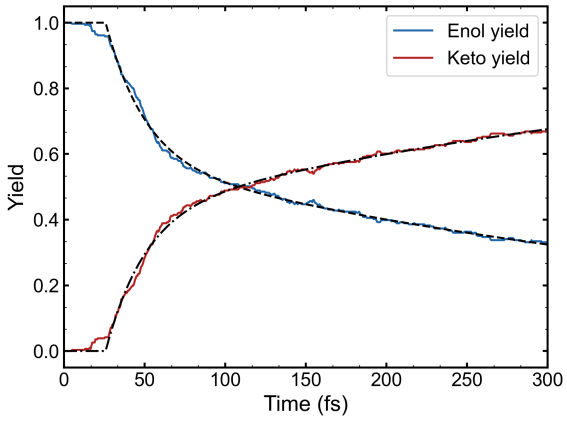


Figure S9: Time-dependent yield of the reactant (enol, in blue) and product (keto, in red) forms, from the ensemble of trajectories that started in the S_1 state. The black dashed line represents the bi-exponential fit of the reactant and product populations in time.

Fit parameters (bi-exponential)					
t_0 (fs)	δ_1	δ_2	τ_1 (fs)	τ_2 (fs)	C
25.82 (0.18)	0.42 (0.01)	0.58 (0.10)	24.43 (0.84)	475.92 (120.55)	0.00 (0.09)

Table S2: Optimised parameters from the fit of the enol yield in time with a bi-exponential decay function. In parentheses one standard deviation errors on the fit parameters are reported.

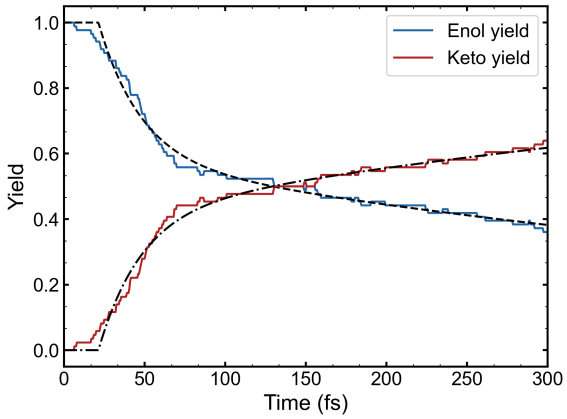


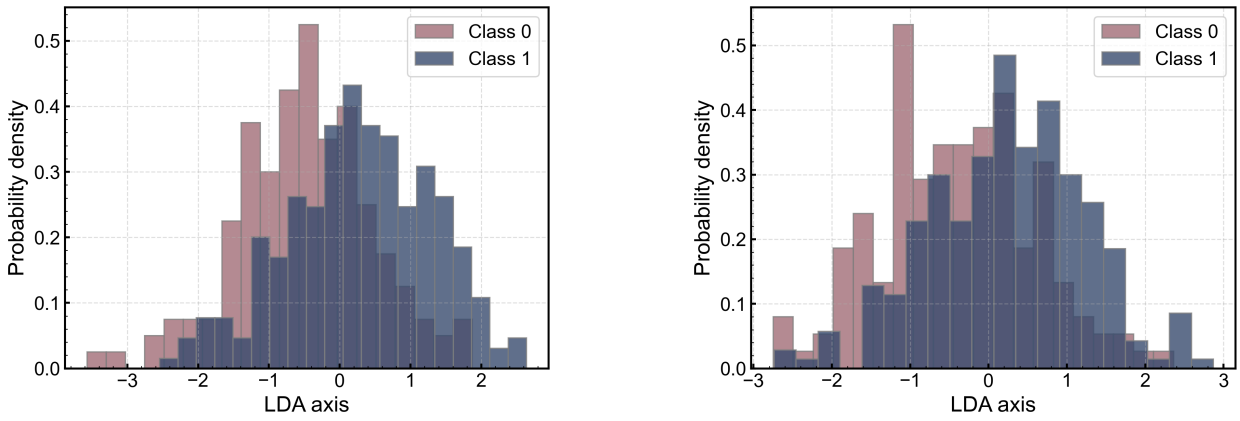
Figure S10: Time-dependent yield of the reactant (enol, in blue) and product (keto, in red) forms, from the ensemble of trajectories that started in the S_2 state. The black dashed line represents the bi-exponential fit of the reactant and product populations in time.

Fit parameters (bi-exponential)					
t_0 (fs)	δ_1	δ_2	τ_1 (fs)	τ_2 (fs)	C
21.33 (0.33)	0.42 (0.02)	0.58 (0.32)	26.39 (1.39)	669.42 (521.69)	0.00 (0.30)

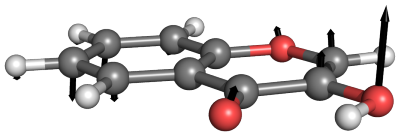
Table S3: Optimised parameters from the fit of the enol yield in time with a bi-exponential decay function. In parentheses one standard deviation errors on the fit parameters are reported.

S10 Linear discriminant analysis (LDA)

Linear Discriminant Analysis (LDA)[6,7] is a classification method that constructs linear decision boundaries. It is also commonly applied for supervised dimensionality reduction, where data are projected onto a linear subspace defined by directions that best separate the different classes. Regarding the mathematical foundation of the method, further details can be found in the literature[6,8]. In this context, LDA was applied to have insights about the separability in terms of structural features of the initial conditions into two classes. Specifically, class 0/1 corresponds to the initial conditions (expressed in Cartesian coordinates) set from which the trajectories explore a displacement along $|\phi_1|$ smaller/greater than 21.6 degrees, respectively.



(a) (b)



(c)

Figure S11: Linear discriminant analysis (LDA) on the (a) initial geometries and (b) velocities. The initial conditions were divided into two classes. The 0-/1-labelled class identifies the initial conditions from which the trajectories explore a displacement along $|\phi_1|$ smaller/greater than 21.6 degrees. (c) LDA axis of the initial geometries in Cartesian space.

S11 Principal component analysis (PCA)

Previous analyses linked the slower ESIPT time constant to out-of-plane hydrogen torsion and sampling of the S_1 torsional minimum. To better capture the complex excited-state dynamics and identify the dominant collective motions underlying this evolution, we performed a principal component analysis (PCA)[9] on the full trajectory ensemble. The projection of the sampled geometries, together with the previously identified critical points (*cis*-enol, *tor*-enol, *trans*-enol, and keto), onto the first two principal components (PCs), which account for 65% of the total variance (see Figure S12), is shown in Figure S13. This two-dimensional subspace effectively captures the essential dynamics of the system, illustrating how the trajectories explore the conformational space connecting the key tautomers and intermediate states. The other PCs have a significantly smaller individual explained variance ($< 4\%$) which sum up to 35%, accounting for the fluctuations of the individual less important degrees of freedom that does not characterise the dynamics in terms of the ESIPT and/or competitive pathways. The motion along the positive direction of first principal component (PC0) describes the collective coordinate associated with ESIPT, while the second principal component (PC1) captures the out-of-plane motion of the hydrogen atom, connecting the *cis*- and *tor*-enol conformers. Displacements along the negative direction of PC0 from the *tor*-enol region lead the system into the *trans*-enol conformation. This analysis confirms that the previously identified critical structures represent the main basins/critical points of the excited-state dynamics.

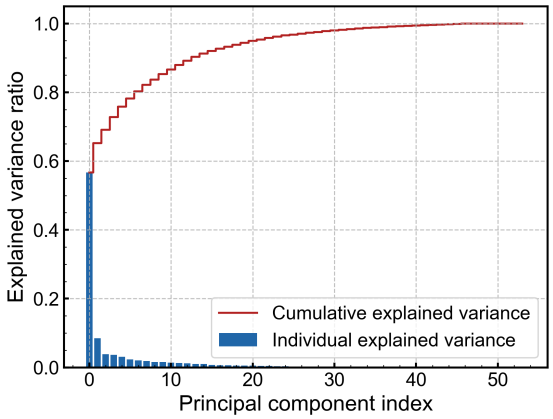
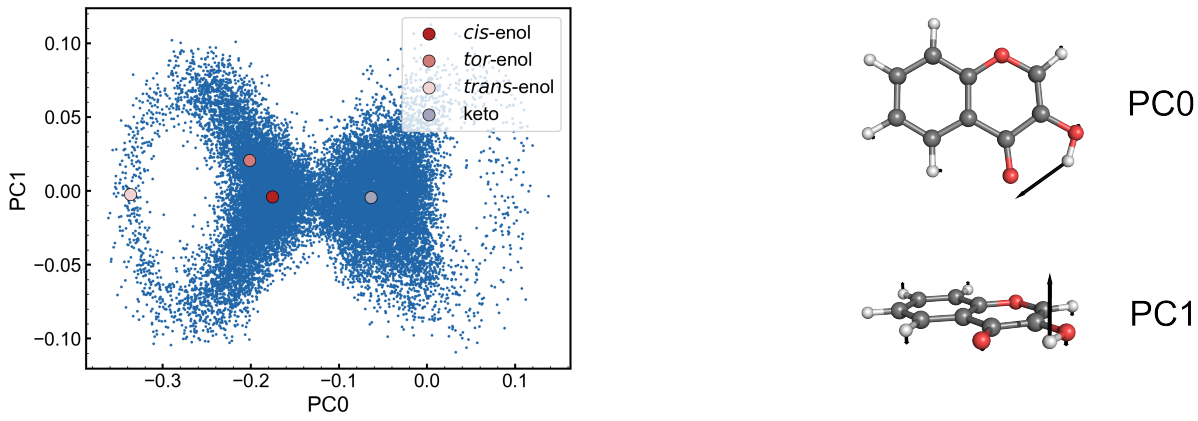


Figure S12: Cumulative and individual explained variance for different principal components.



(a) (b)

Figure S13: (a) Projection of the geometries visited along the trajectories on the first two principal components. (b) Cartesian directions associated to the first two principal components.

S12 Transitions between states

All structures visited along the trajectories were clustered into four categories based on their geometry:

- *cis*-enol ($d(\mathrm{O_d} - \mathrm{H}) \leq d(\mathrm{H} - \mathrm{O_a})$ and $|\phi_1| < 21.6$ degrees)
- *tor*-enol ($d(\mathrm{O_d} - \mathrm{H}) \leq d(\mathrm{H} - \mathrm{O_a})$ and 21.6 degrees $\leq |\phi_1| < 130.0$ degrees)
- *trans*-enol ($d(\mathrm{O_d} - \mathrm{H}) \leq d(\mathrm{H} - \mathrm{O_a})$ and $|\phi_1| \geq 130.0$ degrees)
- keto ($d(\mathrm{O_d} - \mathrm{H}) > d(\mathrm{H} - \mathrm{O_a})$)

Since each conformer/tautomer can be populated in either the S_1 or S_2 electronic state, this leads to a total of eight distinct states (see Table S4 for the numerical labels). Every frame of the ensemble of trajectories was assigned to one of these states, and the transitions between them were counted. The resulting transition matrix (with normalised rows) is reported in Figure S14.

Label	Conformer/Tautomer	Electronic state
1	*cis*-enol	S_1
2	*cis*-enol	S_2
3	*tor*-enol	S_1
4	*tor*-enol	S_2
5	*trans*-enol	S_1
6	*trans*-enol	S_2
7	keto	S_1
8	keto	S_2

Table S4: Labels associated with the conformational/tautomeric forms in their corresponding electronic state.

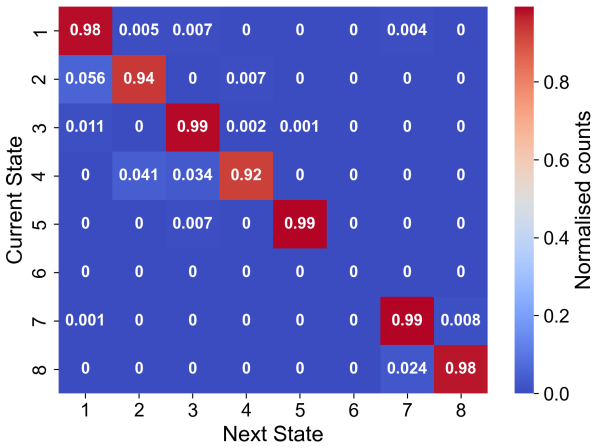
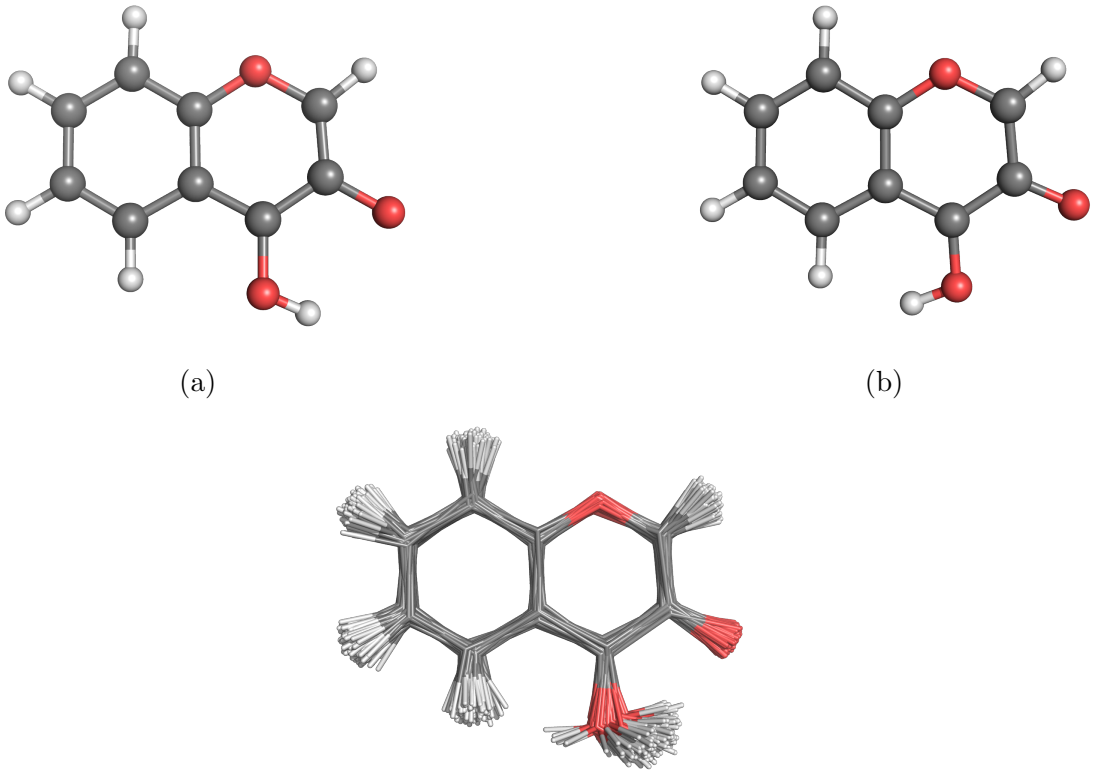


Figure S14: Counts of the transitions between one (conformational/tautomeric + electronic) state to the other along the full ensemble of trajectories. Refer to the Table S4 for the labels.

S13 Keto-type MECIs and hopping geometries



(a) (b)

(c)

Figure S15: Ball and stick representation of the (a) S_2/S_1 keto-type MECI, (b) S_2/S_1 *trans*-keto-type MECI and (c) stick representation of the keto-type hopping geometries between the S_1 and S_2 electronic states from the mixed quantum-classical trajectories.

Bibliography

[1] B. R. Landry, M. J. Falk and J. E. Subotnik, *The Journal of Chemical Physics*, 2013, **139**, 211101.

[2] M. Baer, *Chemical Physics Letters*, 1975, **35**, 112–118.

[3] G. Granucci, M. Persico and A. Toniolo, *The Journal of Chemical Physics*, 2001, **114**, 10608–10615.

[4] P.-O. Löwdin, *Advances in quantum chemistry*, Elsevier, 1970, vol. 5, pp. 185–199.

[5] R. Crespo-Otero and M. Barbatti, *Marco Antonio Chaer Nascimento: A Festschrift from Theoretical Chemistry Accounts*, Springer, 2012, pp. 89–102.

[6] R. A. Fisher, *Annals of Eugenics*, 1936, **7**, 179–188.

[7] T. Hastie, R. Tibshirani, J. Friedman *et al.*, *The Elements of Statistical Learning*, 2009.

[8] P. E. Hart, D. G. Stork and R. Duda, *Pattern classification*, Wiley Hoboken, 2001.

[9] K. Pearson, *The London, Edinburgh, and Dublin philosophical magazine and journal of science*, 1901, **2**, 559–572.