Excited-state distortions control the reactivities and regioselectivities of photochemical 4π-electrocyclizations of fluorobenzenes

Jingbai Li and Steven A. Lopez*
Department of Chemistry and Chemical Biology, Northeastern University, Boston, MA 02115, USA

*Correspondence to s.lopez@northeastern.edu

Abstract

The photochemistry of benzene is complex and non-selective because numerous mechanistic pathways are accessible in the ground- and excited-states. Fluorination is a known strategy to increase the chemoselectivities for Dewar-benzenes via 4π-disrotatory electrocyclization. However, the origin of the chemo- and regioselectivities of fluorobenzenes remains unexplained because of experimental limitations in resolving the excited-state structures on ultrafast timescales. The computational cost of multiconfigurational nonadiabatic molecular dynamics simulations is also generally prohibitive. We now provide high-fidelity structural information and reaction outcome predictions with machine-learning-accelerated photodynamics simulations of a series of fluorobenzenes, $C_6F_{6-n}H_n$, n=0–3 to study their $S_1 \rightarrow S_0$ decay in 4 ns. We trained neural networks with XMS-CASPT2(6,7)/aug-cc-pVDZ calculations, which reproduced the S_1 absorption features with mean absolute errors of 0.04 eV (< 2 nm). The predicted S_1 excited-state lifetimes for $C_6F_4H_2$, C_6F_6, C_6F_5H, and $C_6F_3H_3$ are 64, 40, 18, and 8 ps, respectively. The trend is in excellent agreement with the experimental lifetimes. Our calculations show that the pseudo Jahn-Teller distortions create the S_1 minimum region that prolongs the excited-state lifetime of fluorobenzenes. The pseudo Jahn-Teller distortions reduce when fluorination decreases. Characterization of the surface hopping structures suggests that the S_1 relaxation first involves a *cis-trans* isomerization of a π_{C-C}-bond in the benzene ring, promoted by the pseudo-Jahn-Teller distortions. A branching plane analysis revealed that the conical intersections favoring 4π-electrocyclization are less energetically accessible through the S_1 relaxation; lower-energy conical intersections resemble the reactant and favor reversion.
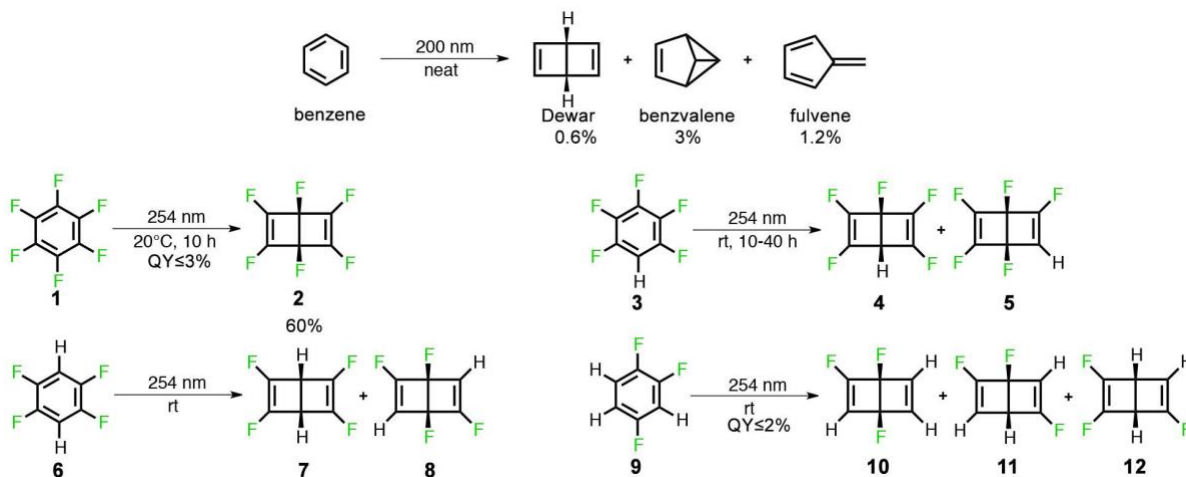
Introduction

Photochemistry is leading a sustainable future of industrial chemical syntheses. The use of light-promoted chemical reactions avoids harsh conditions (e.g., prolonged heating, high pressure, and or extreme pH). Organic photoreactions are of great interest because their reactivity and selectivity can depend simply on the wavelength of light absorption and are free of typically expensive organometallic catalysts. Photoreactions such as 4π-electrocyclization[1] and [2+2]-cycloaddition[2-3] feature shortened synthetic routes with high atom economy. Recent innovations in chemical biology[4] and material science have made significant progress on novel light-harnessing materials. Solar energies are converted to electricity using organic photovoltaics,[5] and stored in solar-thermal storage[6-7] and various strained organic compounds, such as quadricyclanes,[8] bicyclobutanes,[9] cubanes,[10] and ladderenes.[11]

Molecules undergo electronic excitations upon photon absorption. The molecular excited states convert excess energy into enhanced vibrational energy, inducing ultrafast molecular

transformations. Benzene is a well-known molecule with complex excited-state reaction mechanisms leading to several photo products; the photochemistry of benzene also depends on the wavelength of the irradiating light (Scheme 1). Wishnok and co-workers found the Dewar-benzene can only form when exciting benzene to S_2 with λ = 200 nm.[12] The photochemical reaction pathways have been mapped by quantum chemical calculations[13] and dynamics modeling.[14] Fluorination significantly changes the photochemical reactivity and chemoselectivity of benzene. Haller[15] and Camaggi[16] independently discovered that hexafluorobenzene (**1**) exclusively forms Dewar-fluorobenzenes (**2**, Scheme 1) from S_1. A low quantum yield (≤3%) of Dewar-hexafluorobenzene (**2**) has been measured under pressures ranging from 44 to 1052 Torr.[17] Nevertheless, this reaction provides a convenient and chemoselective route to highly strained molecular structures.[11]

Other fluorinated benzenes ($C_6F_{6-n}H_n$, n=1–3) form various Dewar-fluorobenzenes depending on the fluorination positions and numbers (Scheme 1). Ratajczak observed the coexistence of two Dewar-isomers (**4** and **5**) of pentafluorobenzene (**3**), although their quantum yields have not been determined.[18] Ratajczak also reported the formation of Dewar-1,2,4,5-tetrafluorobenzene, **8** in the two possible products of **6** while the quantitative measurement is unavailable.[19] Semeluk performed vapor phase photolysis of 1,2,4-trifluorobenzene (**9**) and identified **11** and **12** in the three possible products, where the overall quantum yields of **11** and **12** are ≤2%.[20] These experiments imply that fluorine plays a significant role in controlling the chemo- and regioselectivity of benzene photoreaction. However, due to the lack of structural information and low time-resolution in the photolysis experiments, the role of fluorination in the regioselectivity of benzene photoreaction remained unresolved.
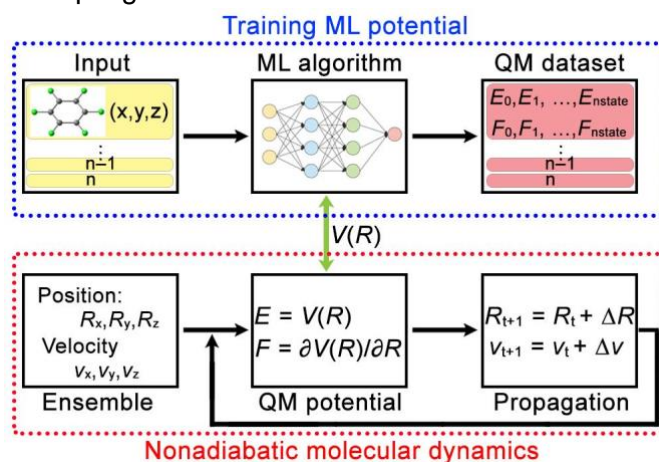
Scheme 1. Photochemical isomerization of benzene and fluorobenzenes.

Experimental and computational studies primarily focused on the electronic spectroscopy of fluorobenzenes to determine the effects of fluorination on benzene photochemistry. Philis *et al.* compared the absorption spectra of **1**, **3**, **6,** and **9** with benzene, suggesting that the fluorination lowers the energy of the $\pi\pi^*$ ($^1B_{2u}$) state. The electronegative fluorine atoms lower the energy of the σ^*_{C-F} orbital to create a relatively low-lying $\pi\sigma^*$ ($^1E_{1g}$) state.[21] Our group used multiconfigurational calculations to determine that the $\pi\pi^*$ state is S_1 and the $\pi\sigma^*$ state is S_2 for **1**. Bradforth and co-workers measured a 2.29 ns fluorescence time constant of **1**.[22] It is consistent with the early fluorescence study by Loper *et al.*, where the reported time-constants of **1** and **3** at

254–280 nm were 1.7–3.6 and 1.6–1.7 ns, respectively.[23] Loper explains that the symmetry reduction from D_{6h} in **1** to C_{2v} in **3** favors the electronic transition, thus shortening the excited-state lifetime of **3**. The nonradiative $S_1{\rightarrow}S_0$ decay measured by Studzinski *et al.* showed that the lifetime of **6** was 2.7 times longer than **1**.[24] The decay time of **9** has not been reported in the literature. Despite these spectroscopic studies, the mechanisms of fluorobenzene photoreactions remained largely unexplored. Tracing the excited-state 4π-electrocyclization of fluorobenzenes is almost impossible due to the low quantum yields. Theoretical works only implied a hypothetical pathway toward the 4π-electrocyclization based on static potential energy surfaces (PESs).[22] However, the long excited-state lifetimes have prevented quantum chemical simulations from discovering the mechanistic pathways toward Dewar fluorobenzene.

We now apply our recently developed machine-learning (ML) accelerated nonadiabatic molecular dynamics (NAMD) code PyRAI²MD[25] to predict the reactivities and regioselectivities of fluorobenzenes (**1**, **3**, **6**, and **9**) at the nanosecond timescale. ML techniques (e.g., Kernel Ridge Regression[26] and Neural Networks[27]) have recently emerged to circumvent the computational bottleneck by reproducing the resource-intensive quantum chemical (QC) calculations with drastically decreased computational cost. A ML-based molecular dynamics simulation first trains a (ground/excited-state) potential model from a big set of precomputed QC data, then uses it for on-the-fly trajectories propagations (Figure 1). To date, ML potentials have achieved cost-efficient molecular dynamics simulations with high-fidelity, such as coupled-cluster with single, double, and perturbative triple excitations (CCSD(T)) by Bogojeski *et al.*[28] and multireference configuration interaction (MRCI) by Westermyr *et al.*[29] Our recent adaptation of neural networks (NNs) based on multiple perception layers to NAMD successfully simulated several organic photoreactions of unprecedented complexity in nanoseconds.[25] It enables a heuristic discovery of photoreaction mechanisms by interrogating all possible pathways in excited-state, which are currently prohibitive due to the requisite multiconfigurational QC calculations for excited-state PESs and nonadiabatic couplings.

Figure 1. A schematics of training machine learning potential for nonadiabatic molecular dynamics simulation.

Our NN modes are trained with the extended multistate complete active space second-order perturbation (XMS-CASPT2) calculations to account for static and dynamical electron correlation. This unprecedented level of theory is needed to generate high-fidelity training data for 4 ns
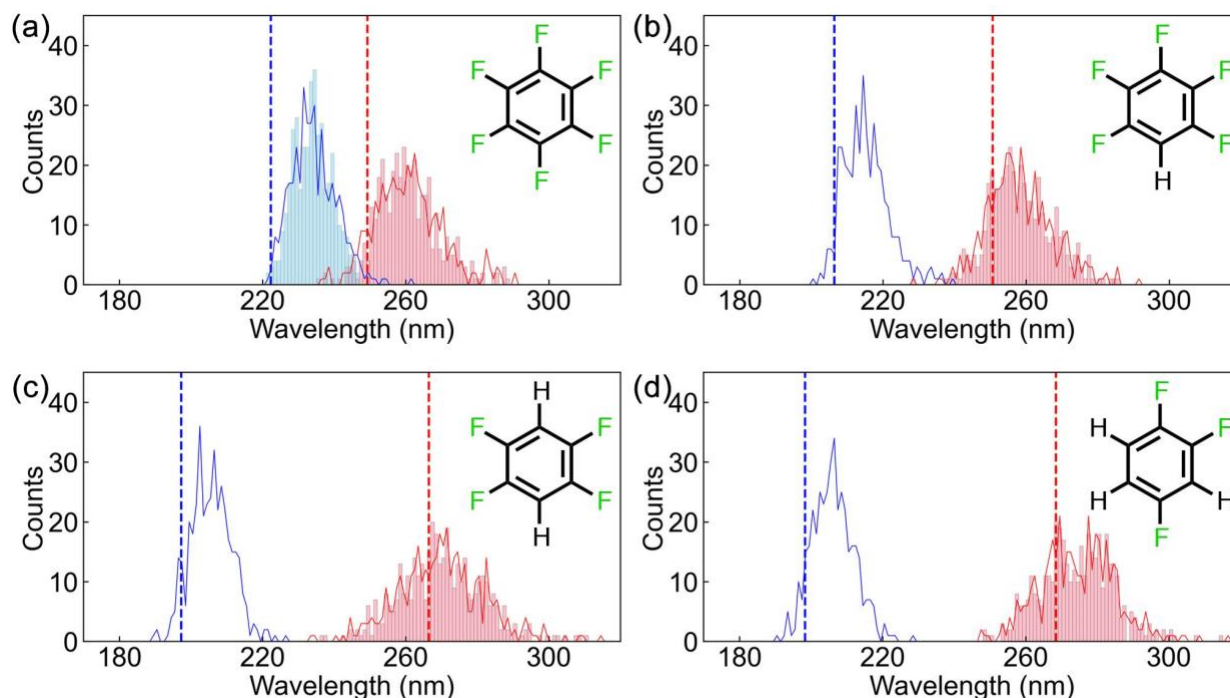
simulations and ensure robust explorations of the possible excited-state reaction pathways. The following discussion focuses on 1) NNs validation by examining the predicted absorption wavelengths and excited-state lifetimes; 2) analysis of the ML-NAMD trajectories to understand how fluorination affects the quantum yields and regioselectivities; 3) characterization of the conical intersections to explain the interplay of the ground- and excited-state topologies. This work demonstrates the value of the ML-NAMD technique for disentangling elusive structure-property relationships and dynamical effects.

Results and discussion
NN predicted absorption wavelengths of fluorobenzenes.
Constructing reliable NNs potential to simulate fluorobenzene photodynamics requires adequate QC methods to generate the training data. We first benchmarked vertical excitation energies with various QC methods against experiment data (Tabe S3). The best agreement was achieved by XMS-CASPT2(6,7)/aug-cc-pVDZ for **1**, **3**, **6,** and XMS-CASPT2(6,6)/aug-cc-pVDZ for **9**. We plotted the absorption spectra to select an appropriate number of low-lying states to build the NN potential (Figure S6). We noted an unambiguous overlap between the S_1 and S_2 in **1**, which suggests $S_0{\rightarrow}S_1$ and $S_0{\rightarrow}S_2$ electronic transitions and possible S_2/S_1 intersections. As such, we reasoned that the NN for **1** should at least include the S_0, S_1, and S_2 states. In **3**, **6,** and **9**, the higher-lying excited states show negligible interactions with S_1 because of the larger energy gaps (Figure S6). Detailed information about NN training is available in the Computational Methods section of the manuscript and Supporting Information.

We first validated the NN predictions by comparing the predicted excited-state energies to the reference QC data. Figure 2 illustrates the NN predicted absorption spectra of **1**, **3**, **6,** and **9** overlaid with the QC results.
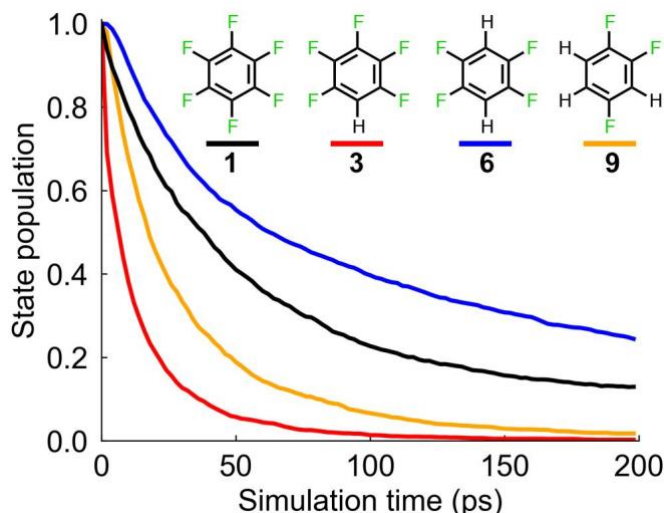

Figure 2. Quantum chemical and neural networks predicted absorption wavelength distributions

of (a) C_6F_6, (b) C_6F_5H, (c) $C_6F_4H_2$ and (d) $C_6F_3H_3$. The dashed lines mark the vertical excitation energies. The solid curves depict the wavelength distributions obtained from XMS-CASPT2(6,7)/aug-cc-pVDZ for **1**, **3**, **6,** and XMS-CASPT2(6,6)/aug-cc-pVDZ for **9**. The S_2 and S_1 wavelengths are in blue and red, respectively. The histograms illustrate the NN predicted wavelength distributions of 400 Wigner sampled configurations.

The S_1 and S_2 of **1** are $\pi\pi^*$ and $\pi\sigma^*$ states centered at 249 nm (4.97 eV) and 222 nm (5.58 eV). The $\pi\sigma^*$ nature of S_2 is contributed by a low-lying σ_{C-F}^*-orbital (Figure S3). Our previous experimental work with Bradforth and co-workers reported red-shifted S_1 and S_2 absorptions due to the breaking of the D_{6h} symmetry.[22] The NN predictions show an excellent agreement with the QC results. It reproduces the red-shift (QC: 263 and 232 nm *vs.* NN: 259 and 235 nm) and the overlap between S_1 and S_0 (QC: 237–262 nm *vs.* NN: 236–254 nm, Figure 2a). The mean absolute error of predicted wavelengths (MAE) is 2 nm (0.04 eV). Figure 2b shows the S_1 and S_2 of **3** at 251 nm (4.95 eV) and 206 nm (6.01 eV), and the band centers shift to 258 and 214 nm, respectively. The S_2 energy increases because the pentafluorination stabilizes the $\pi\sigma^*$ configuration less than hexafluorination. The NN predicted S_1 wavelength distribution is consistent with QC calculations showing a shifted band center at 255 nm (Figure 2b). The MAE of the predicted wavelengths is nearly 2 nm (0.03 eV). In **6** and **9,** the S_2 is transforming from a $\pi\sigma^*$ to $\pi\pi^*$ state with decreasing number of the fluorines. The QC calculations show that the S_1 and S_2 of **6** are centered at 266 nm (4.65 eV) and 197 nm (6.28 eV) in FIgure 2c; the S_1 and S_2 of **9** are centered at 268 nm (4.62 eV) and 198 nm (6.25 eV) in Figure 2d. The MAE of predicted S_1 wavelengths of **6** and **9** is 1 nm (0.02 eV). The NN also captures the features of S_1 distributions - **6** resembles a normal distribution at 266 nm (Figure 2c); **9** exhibits two dominant peaks at 268 and 277 nm (Figure 2d).
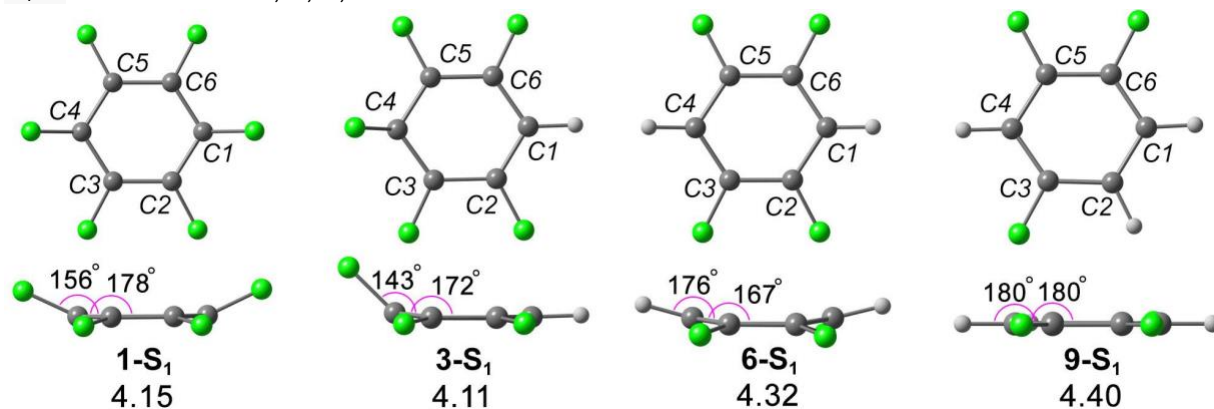
NN predicted excited-state lifetimes of fluorobenzenes.
Given the remarkable accuracy in predicting the spectra, we turned our attention to understand the photochemistry of fluorobenzenes with photodynamics simulations. The initial conditions of **1**, **3**, **6**, and **9** are prepared by Wigner sampling at the zero-point energy level. The ML-NAMD simulations propagated the trajectories from the S_1-Franck-Condon (FC) points in 4 ns with a 0.5 fs time step. Detailed information about the ML-NAMD simulations is available in the Computational Methods section of the manuscript. We performed 1878, 1993, 1797, and 1830 trajectories of **1**, **3**, **6**, and **9**, respectively. The S_1 population dynamics are shown in Figure 3.

Figure 3. NN predicted S_1 population dynamics of C_6F_6, C_6F_5H, $C_6F_4H_2$, and $C_6F_3H_3$.

The predicted S_1 half-lives for **6**, **1**, **9**, and **3** are 64 ps, 40 ps, 18 ps, and 8 ps, respectively. This trend agrees with the literature reported $S_1 \rightarrow S_0$ decay times, where **6** > **1** > **3**. [23-24]The decay time of **9** has not yet been reported; our calculations predict that it will be between **1** and **3**. The excited-state lifetimes of fluorobenzenes are longer than 1 picosecond because they quickly locate an S_1-minimum region near the S_1-FC region, where they can radiatively decay. We sought to generalize the role of fluorination role on photochemical 4π-electrocyclizations and identify a connection with the extent of PJT distortion, previously identified by our group.[22] Figure 4 shows the optimized S_1 local minima for **1**, **3**, **6**, and **9**.

Figure 4. Top and side views of the S_1 minimum geometries of C_6F_6, C_6F_5H, $C_6F_4H_2$, and $C_6F_3H_3$, optimized with XMS-CASPT2(6,7)/aug-cc-pVDZ for **1-S₁**, **3-S₁**, **6-S₁** and XMS-CASPT2(6,6)/aug-cc-pVDZ for **9-S₁**. The relative energies are reported in eV.
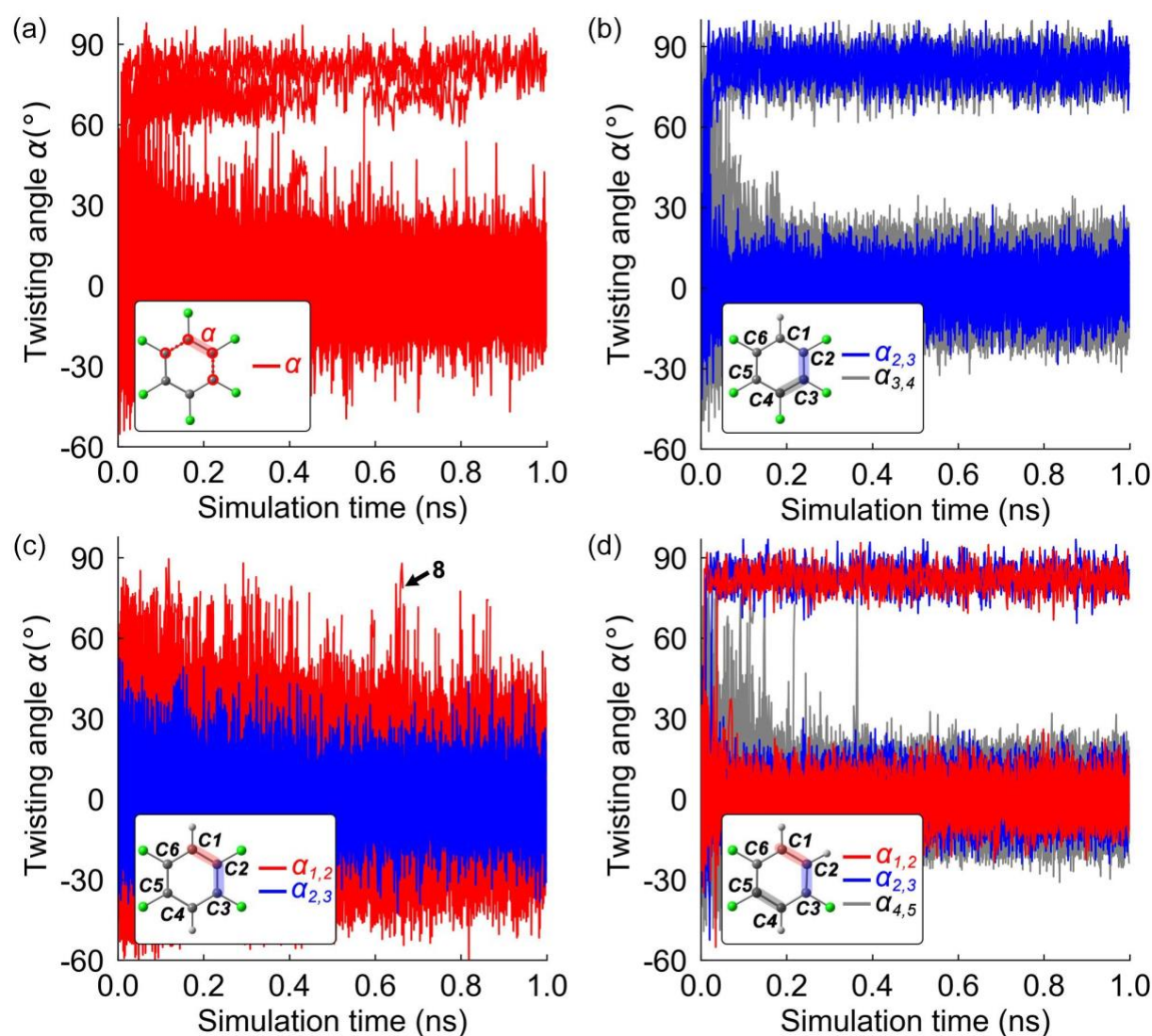
The S_1 minimum of **1** has a non-planar geometry that breaks the D_{6h} symmetry of the reactant. The C-F bonds in a 1,4 relationship feature an out-of-plane angle of 156° because vibronic mixing of the $\pi\pi^*$ and $\pi\sigma^*$ states with a small S_1-S_2 energy difference of 0.60 eV induces the PJT distortions, originating from the perfluoro effect.[22] Similar structures have been reported in **3** by Temps[30] and Sala[31] with XMSCQDPT2/cc-pVDZ calculations. The reduced fluorination in **3** increases the S_1-S_2 energy gap to 1.06 eV, which disfavors the vibronic couplings between the

6

ππ* and πσ* states. As such, the geometry of **3-S$_1$** shows a 143° out-of-plane bending of the C4-F bond, whereas the C1-H bond lies on the benzene plane. The PJT distortions become even weaker in **6** because the energy gap between the πσ* (now S$_3$) and ππ* state (S$_1$) substantially increases to 2.51 eV. As a result, the S$_1$ minimum geometry **6-S$_1$** only deviates from planarity with a ring and C-H bending angles of 167° and 176°, respectively. It agrees with the so-called "butterfly" motions observed in the experiments by Ito and co-workers.[32] The energy of **6-S$_1$** is 0.2 eV higher than those of **1-S$_1$** and **3-S$_1$**, confirming that the PJT distortions proportionally increase with the number of fluorine substituents. The PJT distortions disappear in **9** due to the highly disfavored mixing of the ππ* and πσ* states, where the energy gap is larger than 3 eV. Thus, the S$_1$ minimum **9-S$_1$** retains a planar geometry, and the relative electronic energy is 0.1–0.3 eV higher than **1-S$_1$**, **3-S$_1$**, and **6-S$_1$**.

ML-NAMD trajectories analysis of fluorobenzenes.
We identified two stages for the photodynamics of fluorobenzenes: the S$_1$ relaxation that moves to the S$_1$/S$_0$ seam and the subsequent S$_0$ bifurcation that controls the reactant and product distribution. During nonadiabatic relaxation along the S$_1$ surface, the PJT distortions facilitate a *cis-trans* isomerization of a π$_{C-C}$ bond in the fluorobenzenes. We introduce a geometrical angular parameter, α, which quantifies the extent of twisting over four carbons during the trajectories (Figure 5a). Figure 5a–5d plots the changes of twisting angle in the first 1 ns of the trajectories of **1**, **3**, **6**, and **9**.

Figure 5. Plots of the twisting angle of (a) C_6F_6, (b) C_6F_5H, (c) $C_6F_4H_2$, and (d)$C_6F_3H_3$ in 1 ns ML-NAMD simulations. The plots illustrate the first 1 ns of 400 trajectories of **1**, **3**, **6**, and **9** for clarity. The twisting angle is the dihedral of four adjacent carbons, measured in a 2 ps interval. The trajectories are color-coded according to the middle C-C bonds highlighted in the legends.
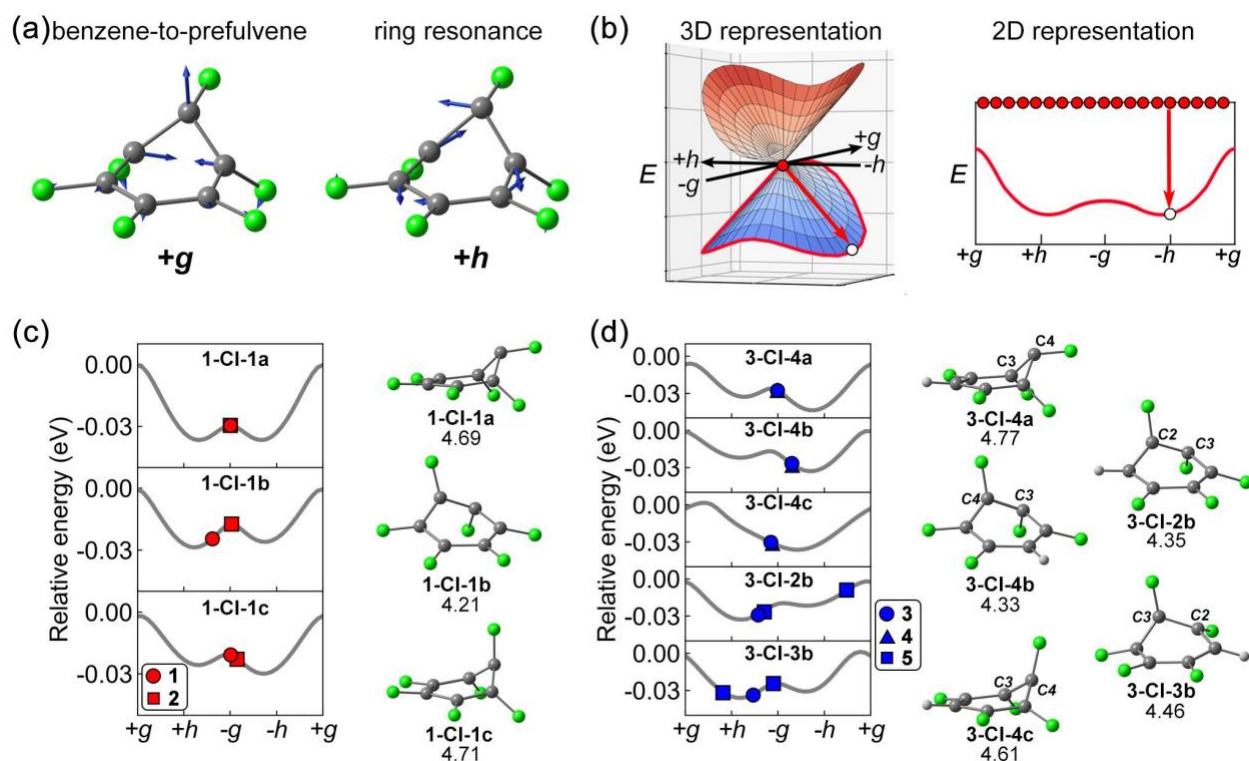
In the first 100 ps, the twisting angle in the trajectories of **1** increases more than 60° as it approaches the S_1 minimum region (Figure 5a). The trajectories immediately bifurcate after hopping to the ground-state, where 99.6% revert to **1** with α approaching 30°. 0.4% of the trajectories correspond to the 4π-electrocyclization reaction toward **2** with α approaching 80°. The trajectories of **3**, **6**, and **9** show different C-C twisting motions along the S_1 relaxations. We grouped the trajectories based on the twisting angle in the surface hopping points. The trajectories of **3** undergo twisting along the C2-C3 (C4-C5) and C3-C4 (C4-C5) bonds (Figure 5b). 92.7% of the trajectories feature a C3-C4 twisting motion, while just 7.1% of the trajectories show the C2-C3 twisting. The C3-C4 twisting motion is significantly preferred than C2-C3 because the fluorinations on C3-C4 bond facilitates a symmetry-breaking C4-F bond bending, which promotes the PJT distortions reported by Sala,[31] leading to the S_1 minimum region (**3-S₁**, Figure 4). When

8

the trajectories pass through the S_1/S_0 seam, the C3-C4 twisting produces **4,** and the C2-C3 twisting leads to **5**. The predicted yields of **4** and **5** are 0.5% and 0.3%. Figure 5c shows that 94.3% of the trajectories of **6** hop to the ground-state via the C1-C2 twisting, and 3.7% of the trajectories are initiated by the C2-C3 twisting. These are in line with the PJT distortions in **6** that the C1-C2 twisting favors the geometry of the S_1 minimum (**6-S_1**, Figure 4). Of the 1797 trajectories of **6**, only 1 trajectory (0.1%) forms **8** through the C1-C2 pathway, and none forms **7**. The trajectories of **9** went through all six C=C bond twisting pathways. Figure 5d illustrates the dominant C4-C5 (85.0%), C2-C3 (7.8%), and C1-C2 (3.1%) twisting. The trajectories in the other pathways are <2.0%. Due to the planar structure of the S_1 minimum, the C=C twisting of **9** leads to two regioisomers of Dewar-fluorobenzenes depending on the twisting site. 0.4% of the trajectories lead to **10** via C3-F bending in the C2-C3 twisting and 0.1% lead to **12** via C1-H bending in the C1-C2 bond twist. Both C1-C2 and C2-C3 twisting formed **11** by the C2-H bending, which gives a total yield of 0.3%, however none of the 4π-electrocyclization was initialized by C4-C5 twisting. Collectively, our ML-NAMD simulations indicate that the regioselectivity of 4π-electrocyclization does not correlate to the preference of the S_1 excited-state minimum of fluorobenzene.

Characterization of the conical intersections of the fluorobenzenes.
We hypothesized that the low quantum yields were due to dynamical effects resulting from a slanted ground-state PES gradient towards the reactants **1**, **3**, **6**, and **9**. To determine the ground-state PES topology, we computed the branching planes of the S_1/S_0 minimum energy conical intersections. A branching plane spans the gradient difference vector, ***g,*** and the derivative coupling vector, ***h*** resembling all directions that lift the S_1/S_0 degeneracy. Figure 6a shows the ***g*** and ***h*** vectors of a conical intersection (CI), **1-CI-1a**. The ***g*** vector corresponds to the pyramidalization of benzene toward prefulvene, and the ***h*** vector represents in-plane distortions along the π_{C-C} bond length alternation (i.e., resonance of two Kekulé benzene structures). The initial relaxation direction toward reactants and products can be determined by projecting the Euclidean distance vectors onto the branching plane (Figure 6b). Our CI optimizations searched all possible C=C bond twisting in **1**, **3**, **6**, and **9** (Figure S5-S7). Figure 6c–6d illustrates the 2D representations of the branching planes and the geometries of the CIs optimized from the most frequently detected surface hopping structures in the trajectories of **1** and **3**.
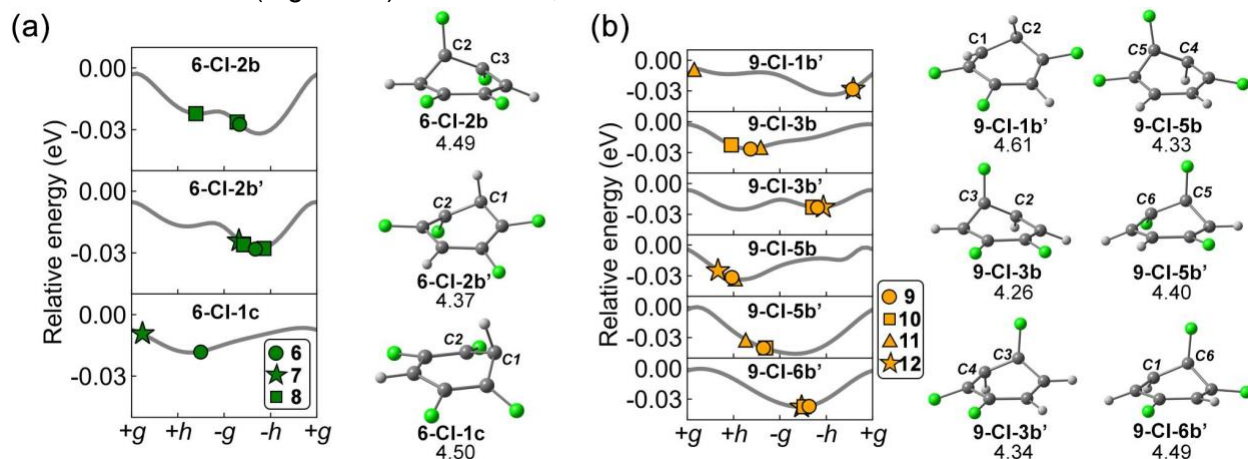
Figure 6. (a) g and h vector of **1-Cl-1a** computed at XMS-CASPT2(6,7)/cc-pVDZ. (b) 3D and 2D representation of the branching plane spanning over the g and h vectors of **1-Cl-1a**. The red vectors illustrate the same relaxation path from **1-Cl-1a** (red dot) toward **-h** direction (white dot) in the 2D and 3D representation. 2D representations of the selected branching planes of (c) C_6F_6 and (d) C_6F_5H, with the corresponding S_1/S_0 conical intersections optimized with XMS-CASPT2(6,7)/aug-cc-pVDZ. The S_0 PES energies are computed at a radius of 0.004 Å to the CI. The relative energies of the CIs are set to 0 eV. The markers are red dot: **1**, red star: **2** in (c); blue dot: **3**, blue triangle: **4**, blue square: **5** in (d).

1-Cl-1a and **1-Cl-1c** have two adjacent twisting C=C bonds. We adapted the *exo* and *endo* terminology from the cycloaddition literature[33] to describe the structural differences between **1-Cl-1a** and **1-Cl-1c**. **1-Cl-1a** bends the C-F bond to the *exo* position in the benzene ring, but **1-Cl-1c** bends it to the *endo* position (Figure 6c). **1-C1-1b** shows the *cis-trans* isomerization of the C=C bond. The trajectories of **1** find 41.6% and 54.9% surface hopping events near **1-C1-1b** and **1-C1-1c**, respectively. The **1-Cl-1a** is less favored (3.4%) because the C-F bond at the *exo* position significantly deviates from the S_1 minimum (**1-S$_1$**, Figure 4). From **1-Cl-1a**, the 4π-electrocyclization and reversion pathways overlap in the **-g** direction, and all trajectories revert to **1**, representing the MECI for a deactivation channel. **1-C1-1b** shows separate 4π-electrocyclization (**-g** direction) and reversion (**+h** and **-g** directions) pathways, but the 4π-electrocyclization pathway is higher than the reversion pathway (Figure 6c). Thus, none of the trajectories yield **2** through this crossing region. The trajectories formed **2** come from the crossing structures resembling **1-C1-1c**. The branching plane shows a slightly steeper 4π-electrocyclization pathway than the reversion near the **-g** direction (Figure 6c). **1-C1-1c** (4.71 eV) is 0.50 eV higher than **1-C1-1b** (4.21 eV). It suggests **1-C1-1c** is less energetically accessible, which explains the minor quantum yields of the 4π-electrocyclization.

The dominant C3-C4 twisting of **3** leads to three CIs, **3-CI-4a**, **3-CI-4b**, and **3-CI-4c**. The branching planes generally show a blocked 4π-electrocyclization pathway; they instead proceed to the **-g** direction adjacent to the reversion pathways (Figure 6d). **3-CI-4a** (4.77 eV) is higher than **3-CI-4b** (4.33 eV) and **3-CI-4c** (4.61 eV). We only observed minor formations of **4** from surface hopping points near **3-CI-4b** and **3-CI-4c**. The less favored C2-C3 twisting of **3** generates two accessible CIs, **3-CI-2b** (4.35 eV) and **3-CI-3b** (4.46 eV). The branching plane shows unambiguous pathways to **5** following the **-g** and **+g** directions from **3-CI-2b** and the **-g** and **+h** directions from **3-CI-3b** (Figure 6d). However, these pathways are energetically less accessible due to higher energies than the reversion pathways.

Figure 7a illustrates the three lowest CIs of **6**. **6-CI-2b** originates from the C2-C3 twisting; **6-CI-2b'** and **6-CI-1c** resemble the C1-C2 twisting. **6-CI-2b** can proceed to **8** in the **-g** or **+h** direction (Figure 7a). These pathways lie higher than the reversion pathways, thus disfavoring the 4π-electrocyclization. **6-CI-2b'** is the global minimum CIs (4.37 eV). The branching plane shows that the 4π-electrocyclization along the **-h** direction is separate and nearly degenerate to the reversion pathway (Figure 7a). It is consistent with the single trajectory formed **8** near **6-CI-2b'**. A possible reaction channel to **7** is near the **-g** direction, but it points to a higher energy pathway. **6-CI-1c** (4.50 eV) is higher than **6-CI-2b'** (4.37 eV) because it bends a C-H bond with two twisting C=C bonds. The 4π-electrocyclization follows the **+g** direction and is higher than the reversion pathway in the **+h** direction (Figure 7a). As a result, **6-CI-1c** deactivates the formation of **7**.

Figure 7. 2D representations of the branching planes of (a) $C_6F_4H_2$ and (b) $C_6F_3H_3$ with the S_1/S_0 conical intersections optimized with XMS-CASPT2(6,7)/aug-cc-pVDZ for $C_6F_4H_2$ and XMS-CASPT2(6,6)/aug-cc-pVDZ for $C_6F_3H_3$. The S_0 PES energies are computed at a radius of 0.004 Å to the CI. The relative energies of CIs are set to 0 eV. The markers are green dot: **6**, green star: **7**, green square: **8** in (a); orange dot: **9** orange square: **10**, orange triangle: **11**, and orange star: **12** in (b).

The C=C twisting of **9** produces multiple CIs (Figure S7). Here, we focus on the lowest CI corresponding to each unique twisting (Figure 7b). The CI energies range from 4.26 to 4.61 eV, close to the S_1 minimum region (**9-S₁,** 4.40 eV). 73% of the trajectories hopped near **9-CI-5b** (C4-C5). The branching plane shows the pathway to **11** moves in the same **+h** direction as the reversion, and the pathway to **12** is disfavored. None of the trajectories formed **11** or **12** following these pathways. Thus, the pathway involving **9-CI-5b** significantly deactivates the 4π-electrocyclization. The C1-C2 and C2-C3 twisting lead to **9-CI-1b'** and **9-C1-3b. 9-CI-1b'**, the

formation of **12** is merged with the reversion pathway, and the pathway to **11** has higher energy (Figure 6d). **9-C1-3b** branches to **10** and **11**. Their pathways follow a similar energy descent to the reversion, explaining the predicted minor yield of **10** (0.4%) and **11** (0.3%). The rest of the CIs, **9-CI-5b'** (C5-C6), **9-CI-3b'** (C3-C4), and **9-CI-6b'** (C6-C1) do not contribute to the 4π-electrocyclization. Other pathways are either energetically unfavorable or higher in energy than the reversion pathway (Figure 6d).

Conclusions

We have uncovered the origin of fluorination effects on 4π-electrocyclization of fluorobenzenes towards Dewar fluorobenzenes with unprecedented 4ns ML-NAMD simulations. Our ML-NAMD trajectories provided detailed time-resolved structural information that elucidates the low quantum yield and regioselectivity of Dewar-fluorobenzenes. We trained the NNs with XMS-CASPT2 for the first time. The NNs reproduced the S_1 absorption features with an MAE < 0.04 eV (2nm) and showed excellent agreement with the trend of the experimentally measured time constants (**6**, **1**, **9**, and **3** are 64, 40, 18, and 8 ps, respectively). The excited-state lifetimes of fluorobenzenes are longer than 1 ps because they quickly relax to an S_1 minimum region created by the PJT distortions. The PJT distortions reduce along with the decreasing number of fluorination.

Characterization of the surface hopping structures suggests that the S_1 relaxation of fluorobenzene resembles the *cis-trans* isomerization of C=C bonds. The fluorination effects control the S_1 relaxations by regioselective activation of the π_{C-C} twisting that can facilitate in the PJT distortions. The most selective π_{C-C} bonds in **3**, **6**, and **9** are C3-C4 (92.7%), C1-C2 (94.3%), and C4-C5 (85.0%). The predicted quantum yields of **2** (0.4%), **4** (0.5%), **5** (0.3%), **7** (0%), **8** (0.1%), **10** (0.4%), **11** (0.3%), and **12** (0.1%) suggest an intrinsically inefficient 4π-electrocyclization of fluorobenzenes. We determined that the CI accessibility during S_1 relaxation and the following S_0 bifurcation synergetically govern the 4π-electrocyclization regioselectivities of the fluorobenzenes. The branching plane analysis showed the S_1 relaxation of fluorobenzene prefers the CIs that deactivate the 4π-electrocyclization, while the CIs favoring 4π-electrocyclization are less energetically accessible through the S_1 relaxation.

.

Computational Methods

Machine learning methods.

The NN training and ML-NAMD simulation use the Python Rapid Artificial Intelligence Ab Initio Molecular Dynamics program (PyRAI²MD).[25] The NN consists of fully connected feedforward multiple perception layers. The input layer takes a flattened inverse distance matrix computed from Cartesian coordinates. The neuron activations employ a leaky softplus function. The output layer fits the energies and forces for all states. The energy and force are trained together in a 40:1 loss weight. The NN implementation uses TensorFlow/Keras (v2.3) API[34] for Python.

The training data generation includes two procedures. 1) We prepared a chemical intuitive initial set that samples the configurations near the reactants and products then interpolate the intermediate configurations via possible conical intersection(s). 2) We performed adaptive sampling[35] to collect undersample configurations in all electronic states. It propagates 200 trajectories to iteratively search undersample configurations using a committee of two NNs. The energies and forces of the uncertain geometries are recomputed with quantum chemical calculations and added to the initial set for retraining the NN potential. The final sets have 3051,

4037, 4710, 10204 data points for **1**, **3**, **6**, and **9**. The validation MAE in energy predictions ranges from 0.020–0.036eV, satisfy the chemical accuracy, 0.043eV (1 kcal·mol^{-1}). More details about NN hyperparameters, training data generation, and adaptive sampling are available in Supporting Information.

ML-NAMD simulations.

The initial nuclear positions and velocities are sampled from the vibrational modes in the Wigner distribution at the zero-point energy level. The simulation timestep is 0.5 fs. The surface hopping probability calculations apply the Zhu-Nakamura theory[36] with an energy difference threshold of 1.0 eV. The trajectories employed start in a microcanonical ensemble (NVE) to model the nonadiabatic transitions. In 100 fs, after arriving at the ground state, they switch to a canonical ensemble (NVT) at 300 K to include the thermalization in pico- and nanoseconds. In the 4 ns simulations, the translation and rotation velocities are removed every 4000 steps (2 ps). With an approximately $7\cdot10^5$-fold acceleration, the simulations finished in 12–27 hours on a single CPU.

Multiconfigurational calculations.

The multiconfigurational quantum chemical calculations use BAGEL.[37-38] We choose a (6,7) active space for **1**, **3**, and **6**. It contains 6 π-electrons and 6 π-orbitals of the benzene ring and a low-lying in-phase σ$^*_{C-F}$ orbital. The σ$^*_{C-F}$ orbital is removed in the (6,6) space for **9** because it no longer contributes to the low-lying excited-states. The calculations use the aug-cc-pVDZ[39] basis set and a real shift of 0.5 Hartree in the XMS-CASPT2 calculation.

Supporting information

NN hyperparameters optimizations; training data generation; active learning; supplementary trajectory results; active spaces; optimized geometries of S$_1$ minimum and S$_1$/S$_0$ conical intersections; conical intersection *g* and *h* vectors; simulated spectra; electronic energies; Cartesian coordinates; data availability

Reference

[1]   S. C. Coote, *Eur. J. Org. Chem.* **2020**, 1405-1423.
[2]   S. Poplata, A. Troster, Y. Q. Zou, T. Bach, *Chem. Rev.* **2016**, *116*, 9748-9815.
[3]   D. Sarkar, N. Bera, S. Ghosh, *Eur. J. Org. Chem.* **2020**, 1310-1326.
[4]   B. D. Fairbanks, L. J. Macdougall, S. Mavila, J. Sinha, B. E. Kirkpatrick, K. S. Anseth, C. N. Bowman, *Chem. Rev.* **2021**, *121*, 6915-6990.
[5]   Y. Yang, K. Börjesson, *Trends Chem.* **2021**, DOI: 10.1016/j.trechm.2021.1010.1007.
[6]   A. K. Saydjari, P. Weis, S. Wu, *Adv. Energy Mater.* **2017**, *7*, 1601622.
[7]   S. Wu, H. J. Butt, *Macromol. Rapid Commun.* **2020**, *41*, 1900413.

[8]   V. A. Bren, A. D. Dubonosov, V. I. Minkin, V. A. Chernoivanov, *Russ. Chem. Rev.* **1991**, *60*, 451-469.
[9]   M. H. Chang, D. A. Dougherty, *J. Am. Chem. Soc.* **2002**, *104*, 2333-2334.
[10]   K. F. Biegasiewicz, J. R. Griffiths, G. P. Savage, J. Tsanaktsidis, R. Priefer, *Chem. Rev.* **2015**, *115*, 6719-6745.
[11]   B. R. Boswell, C. M. F. Mansson, J. M. Cox, Z. Jin, J. A. H. Romaniuk, K. P. Lindquist, L. Cegelski, Y. Xia, S. A. Lopez, N. Z. Burns, *Nat. Chem* **2021**, *13*, 41-46.
[12]   H. R. Ward, J. S. Wishnok, *J. Am. Chem. Soc.* **2002**, *90*, 1085-1086.
[13]   I. J. Palmer, I. N. Ragazos, F. Bernardi, M. Olivucci, M. A. Robb, *J. Am. Chem. Soc.* **2002**, *115*, 673-682.
[14]   E. Pieri, D. Lahana, A. M. Chang, C. R. Aldaz, K. C. Thompson, T. J. Martinez, *Chem Sci* **2021**, *12*, 7294-7307.
[15]   I. Haller, *J. Am. Chem. Soc.* **1966**, *88*, 2070-2071.
[16]   G. Camaggi, F. Gozzo, G. Cevidalli, *Chem. Commun.* **1966**, 313–314.
[17]   I. Haller, *J. Chem. Phys.* **1967**, *47*, 1117-1125.
[18]   P. Cadman, E. Ratajczak, A. F. Trotman-Dickenson, *J. Chem. Soc. A* **1970**.
[19]   B. Sztuba, E. Ratajczak, *J. Chem. Soc. Perkin Trans. 2* **1982**.
[20]   G. P. Semeluk, R. D. S. Stevens, *J. Chem. Soc. D* **1970**.
[21]   J. Philis, A. Bolovinos, G. Andritsopoulos, E. Pantos, P. Tsekeris, *J. Phys. B: At. Mol. Opt. Phys.* **1981**, *14*, 3621-3635.
[22]   J. M. Cox, M. Bain, M. Kellogg, S. E. Bradforth, S. A. Lopez, *J Am Chem Soc* **2021**, *143*, 7002-7012.
[23]   G. L. Loper, E. K. C. Lee, *Chem. Phys. Lett.* **1972**, *13*, 140-143.
[24]   H. Studzinski, S. Zhang, Y. Wang, F. Temps, *J Chem Phys* **2008**, *128*, 164314.
[25]   J. Li, P. Reiser, B. R. Boswell, A. Eberhard, N. Z. Burns, P. Friederich, S. A. Lopez, *Chem. Sci.* **2021**, *12*, 5302-5314.
[26]   J. Westermayr, F. A. Faber, A. S. Christensen, O. A. von Lilienfeld, P. Marquetand, *Mach. Learn.: Sci. Technol.* **2020**, *1*.
[27]   J. Westermayr, P. Marquetand, *Chem Rev* **2021**, *121*, 9873-9926.
[28]   M. Bogojeski, L. Vogt-Maranto, M. E. Tuckerman, K. R. Muller, K. Burke, *Nat. Commun.* **2020**, *11*, 5223.
[29]   J. Westermayr, M. Gastegger, P. Marquetand, *J. Phys. Chem. Lett.* **2020**, *11*, 3828-3834.
[30]   O. Huter, M. Sala, H. Neumann, S. Zhang, H. Studzinski, D. Egorova, F. Temps, *J. Chem. Phys.* **2016**, *145*, 014302.
[31]   M. Sala, D. Egorova, *Photochem Photobiol Sci* **2018**, *17*, 1036-1048.
[32]   K. Okuyama, T. Kakinuma, M. Fujii, N. Mikami, M. Ito, *J. Phys. Chem.* **2002**, *90*, 3948-3952.
[33]   K. N. Houk, *Tetrahedron Lett.* **1970**, *11*, 2621-2624.
[34]   A. A. Martín Abadi, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mane, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viegas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng, **2016**, Software available from tensorflow.org.
[35]   J. Westermayr, M. Gastegger, M. Menger, S. Mai, L. Gonzalez, P. Marquetand, *Chem. Sci.* **2019**, *10*, 8100-8107.
[36]   T. Ishida, S. Nanbu, H. Nakamura, *Int. Rev. Phys. Chem.* **2017**, *36*, 229-285.
[37]   T. Shiozaki, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2017**, *8*.

[38]   BAGEL, Brilliantly Advanced General Electronic-structure Library. http://www.nubakery.org under the GNU General Public License.

[39]   R. A. Kendall, T. H. Dunning, R. J. Harrison, *J. Chem. Phys.* **1992**, *96*, 6796-6806.