Enhanced Sampling Aided Design of Molecular Photoswitches

Umberto Raucci,\textsuperscript{1} David M. Sanchez,\textsuperscript{2,3,*} Todd J. Martínez,\textsuperscript{2,3,*} Michele Parrinello\textsuperscript{1*}

\textsuperscript{1}Italian Institute of Technology, Genova GE, Italy
\textsuperscript{2}Department of Chemistry, Stanford University, Stanford, CA 94305
\textsuperscript{3}Stanford PULSE Institute, SLAC National Accelerator Laboratory, Menlo Park, CA 94025

Abstract

Advances in the evolving field of atomistic simulations promise important insights for the design and fundamental understanding of novel molecular photoswitches. Here, we use state-of-the-art enhanced simulation techniques to unravel the complex, multistep chemistry of donor-acceptor Stenhouse adducts (DASAs). Our reaction discovery workflow consists of enhanced sampling for efficient chemical space exploration, refinement of newly observed pathways with more accurate ab initio electronic structure calculations, and structural modifications to introduce design principles within future generations of DASAs. We showcase our discovery workflow by not only recovering the full photoswitching mechanism of DASA, but predicting a plethora of new plausible thermal pathways and suggesting a way for their experimental validation. Furthermore, we illustrate tunability of these newly discovered reactions, leading to a potential avenue for controlling DASA dynamics through multiple external stimuli. Overall, these insights could offer alternative routes to increase the efficiency and control of DASA’s photoswitching mechanism, providing new elements to design more complex light-responsive materials.

\textsuperscript{*}Present address: Design Physics Division, Lawrence Livermore National Laboratory, Livermore, USA.

Raucci, U and Sanchez, D. M et al. – Enhanced Sampling Aided Design
Smart materials that change shape in response to an external stimulus (e.g., heat, pH, electricity, pressure, or light) hold promise in the areas of robotics, artificial muscles, motors, and sensors. Among the many emerging frontiers in this field, light-responsive systems have distinguished themselves due to significant advances in wireless actuation and remote control. These materials can convert the dynamical transduction of light energy at the molecular level into macroscopic responses via coupled mechanical motion between neighboring light-sensitive molecules known as molecular photoswitches. Prototypical photoswitches (e.g., spiropyran, diarylethene, and azobenzene) contain just one reactive moiety that undergoes photoinduced pericyclic or Z/E isomerization reactions, effectively leading to the ability to “turn on/off” the photo-response of materials.

Connecting individual photoswitches or adopting multi-stimuli approaches (e.g., combining pH changes or electrochemical regulation with light) are some approaches proposed for moving beyond binary systems and achieving more complex material responses. Another way is to increase the chemical complexity of next generation photoswitches by introducing multiple reactive moieties that are capable of displaying a rich and diverse range of chemistry.

Read de Alaniz and co-workers have proposed a novel class of photochromes named donor-acceptor Stenhouse adducts (DASAs), consisting of a donor-group connected to an acceptor moiety through a hydroxylated hexatriene bridge (Fig. 1). Unlike previous photoswitches, DASAs can switch from a colored conjugated form to a colorless ring-closed structure via a rather complex multistep mechanism initiated by light with subsequent thermal steps governing the overall switching behavior. In particular, the photoswitching mechanism of DASA-based systems starts with an actinic step (E/Z photoisomerization) but is followed by a multitude of competing thermal pathways involving 4π electrocyclization, bond rotation and proton transfer. In our

Raucci, U and Sanchez, D. M et al. – Enhanced Sampling Aided Design
recent work\textsuperscript{29-30} we showed for the first time the natural evolution of the nuclear wavepacket relaxing through conical intersections via \textit{ab initio} nonadiabatic and adiabatic quantum molecular dynamics (Fig. 1).

![Enhanced Sampling to Explore DASA's Multistep Chemistry](image)

**Figure 1. Enhanced Sampling to Explore DASA's Multistep Chemistry.** DASA’s photoswitching mechanism is initiated upon absorbing a 530nm photon, promoting the nuclear wave packet (blue) to the excited $S_1$ electronic state. We previously simulated the excited state relaxation dynamics using nonadiabatic AIMD (yellow), which correctly captures the propagation of the nuclear wavepacket through conical intersections over several ps in DASA’s actinic step.\textsuperscript{25} Furthermore, we explored ground-state thermal pathways by continuing the nonadiabatic AIMD trajectories with adiabatic AIMD (purple), providing insight into the natural evolution of the nuclear wavepacket across multiple electronic states and intermediates. Lastly, in this contribution we use biased AIMD (green/blue) to move beyond the limitations of bias free AIMD, making it possible to sample “rare events” that take place over several nano/microseconds. The different DASA generations are shown above the potential energy diagram while the general scheme for enhanced sampling via bias potentials is shown at the bottom.

Our dynamics suggested that several pathways are accessible on the electronic ground state beyond the main photoswitching mechanism. Up to this point, these pathways have been essentially unexplored and their inhibition can improve control of photoswitching in DASA increasing its overall efficiency.

The difficulty in modeling these alternate thermal pathways lies in the inability of \textit{ab initio} molecular dynamics (AIMD) to efficiently sample rare events that may proceed over long
timescales (i.e., ns and greater). One way to accelerate this sampling is by applying an external bias potential to a set of carefully selected collective variables (CVs) that encode the generally undersampled and slow degrees of freedom of a system.\textsuperscript{36-40} When proper CVs are identified, this procedure allows for the efficient exploration of reaction space and subsequent discovery of novel chemical pathways otherwise inaccessible on the timescale of AIMD (Fig. 1). Recently, we proposed an automatic workflow for reaction discovery that relies on a generic CV derived from spectral graph theory\textsuperscript{41} and on the explore variant of the on-the-fly probability enhanced sampling method (OPES\textsubscript{E}).\textsuperscript{42-43} With our approach, a molecule is represented as a graph whose vertices and edges are its atoms and chemical bonds, respectively, and the CV is the maximum eigenvalue ($\lambda_{\text{max}}$) of the symmetric adjacency matrix associated to the graph.\textsuperscript{41} In the OPES\textsubscript{E} method, $\lambda_{\text{max}}$ fluctuations are enhanced by introducing a bias from an on-the-fly estimation of the CV probability distribution. Furthermore, in OPES\textsubscript{E}, the bias deposited is controlled by a barrier cutoff, which sets a limit to the maximum value of the bias and allows control of the discovery phase. Thus, the task of finding new chemical pathways is completely left to the molecular dynamics engine, allowing for a more thorough exploration of the reaction space independent of a user’s chemical intuition.

Numerous other approaches to reaction discovery have been suggested, for both ground\textsuperscript{44-50} and excited states.\textsuperscript{51-52} Here, we complement our previous findings obtained from traditional bias-free molecular dynamics (to describe the evolution on the excited state and the first steps after reaching the ground state) with OPES\textsubscript{E} guided reaction discovery to shed light on the complex chemistry that governs DASA-based photoswitches.

We apply OPES\textsubscript{E} as part of a general three-step workflow consisting of: a) a discovery phase spanning chemical space (using a semiempirical potential (PM6)), b) a refinement stage of the determined pathways based on more accurate ab initio electronic structure calculations (\omega B97x-
D3/def2-TZVP(-f)/COSMO(ε=2.38)), and c) a design step where pathways are tuned via structural changes in DASA. This workflow has assisted in unraveling the complexity of DASA chemistry, revealing a plethora of novel thermal pathways that were previously inaccessible using bias free AIMD. Furthermore, we compare the stability and energetic barriers of these newly determined species with those involved in the main photoswitching mechanism. Lastly, we underline the potential tunability of these newly discovered pathways via 2\textsuperscript{nd} and 3\textsuperscript{rd} generation DASAs (Fig. 1).
Figure 2. Discovered Pathways from Nonadiabatic, Adiabatic, and Biased AIMD. The full 4π-electrocyclization pathway along with all alternate pathways discovered during AIMS (yellow), AIMD (purple), and OPSE (blue/green). In total 17 species have been identified in the discovery phase, and 8 of these are proposed for the first time. Labels for dihedral angles and carbon atoms are shown along the hexatriene bridge for A. Each “open” structure is labeled according to the E/Z configuration assumed by the ω, δ, α, β bonds, respectively. All protons are omitted for clarity except for H₄ (the H atom involved in the initial intramolecular H-bond).
The discovery phase was initiated from the open-form species A and A' (Fig. 2) which consisted of performing 16 molecular dynamics simulations (3ns each) at the PM6 level of theory using barrier values between 10-50 kcal/mol in OPES_E (a summary of the discovery phase results is reported in Table S1 and S2). Following the initial discovery phase, species and the corresponding transition states (TS) between them were further refined using the wB97x-D3/def2-TZVP(-f)/COSMO(ε=2.38) level of theory. Lastly, aromatic substituents and heteroatoms were introduced into the donor and acceptor groups in the design phase revealing our ability to tune the energetics and subsequent reactivity of these newly discovered thermal pathways.

A complete picture of DASA’s reaction space is reported in Fig. 2, where the E/Z and atom labels on A refer to the stereochemical configuration of the ω, δ, α, and β dihedral angles and the carbon atoms along the hexatriene backbone, respectively. The computed energies of the ground-state minima and transition states along all the discovered pathways are reported in Table S3. Using our reaction discovery workflow, we recover the full productive photoswitching pathway (A→B‴ transformation in Fig. 2) previously shown with unbiased nonadiabatic/adiabatic molecular dynamics.30 This pathway consists of multistep reactions: the isomerization around α leads to the formation of A' (EEEZ), which in turn leads to the EZEZ intermediate (A″) after subsequent rotation around δ. A″ undergoes 4π-electrocyclization resulting in the formation of the closed cyclopentenone form B, which evolves into B‴ after a rotation around the γ angle. Finally, the zwitterionic species B‴ is formed after a proton transfer reaction between the hydroxyl and amino groups (Fig. 2).

Additionally, we observe several competing pathways beyond the main photoswitching mechanism leading to species D, E, F, G, and H. Inhibiting the accessibility of these side reactions,
should improve the efficiency of the main \( \text{A} \rightarrow \text{B} ''' \) transformation and, in turn, its overall photoswitching performance. In particular, \( \text{A}''' \) can undergo further cyclization upon bond formation between \( \text{C}_2-\text{C}_5 \), resulting in the cyclobutenol form \( \text{D} \). This zwitterionic species can either experience nucleophilic attack from the carbonyl oxygen \( \text{O}_b \) onto \( \text{C}_6 \) forming \( \text{D}' \) or bond formation between \( \text{C}_1-\text{C}_6 \) leading to \( \text{D}'' \) (Fig. 2). Other side-pathways are accessible from \( \text{A}' \) proceeding through the \( \text{EEEE} \) species resulting after rotation around the \( \beta \) dihedral angle. In particular, \( \text{EEEE} \) has also been observed in our previous unbiased non-adiabatic molecular dynamics simulations, and it represents the main side-product of the photorelaxation dynamics (\( \sim10\% \) of the total wavepacket).\(^{30}\) By introducing a bias into the graph CV, we reveal two possible routes accessible from \( \text{EEEE} \): an intramolecular cyclization leading to \( \text{E} \) upon nucleophilic attack \( \text{O}_b-\text{C}_4 \) or another rotation around \( \delta \) resulting in \( \text{EZEE} \), which is ideally arranged to form the spiro compound \( \text{F} \) (Fig. 2). The species \( \text{E} \) can also be formed from \( \text{EEZE} \), which results from the open form \( \text{A} \) after isomerization around the \( \beta \) bond. The energy barrier for the \( \text{EEZE} \rightarrow \text{E} \) reaction is quite similar to the one for \( \text{EEEE} \rightarrow \text{E} \) (\( \sim18\text{kcal/mol} \) and \( \sim16\text{kcal/mol} \), respectively). In addition, Fig. 2 shows the formation of \( \text{EZZZ} \) from isomerization around the \( \delta \) bond in \( \text{A} \). \( \text{EZZZ} \) is an attractive target for design because 1) it is a relatively abundant side-product according to both our photorelaxation dynamics\(^{30}\) and Feringa and co-workers’ characterization using \(^1\text{H NMR}\),\(^{22} \) 2) it offers a potential avenue for the formation of \( \text{G} \) via proton transfer and nucleophilic attack of \( \text{O}_a \) from \( \text{C}_6 \), and 3) \( \text{G} \) is the most stable product among the newly discovered species (Table S3). Lastly, we show the formation of two other ground-state DASA species (\( \text{G}' \) and \( \text{H} \)) resulting from the proton transfer from the adjacent amino group followed by the subsequent dissociation into a secondary amine and Meldrum acid species, respectively (Fig. 2).
Overall, these newly discovered pathways are rather high in energy when compared to the main photoswitching mechanism in Meldrum’s acid first-generation DASAs (Fig. 3, Table S3). However, these pathways could become quite important to DASA’s switching behavior due to its sensitivity to structural modifications of its donor and acceptor moieties. Indeed, a dissociative side reaction resulting in the release of an amine and activated furan has been recently reported in 5-methyl substituted DASAs.\textsuperscript{31} However, a full mechanistic characterization of this dissociative side reaction has yet to be reported. We find that their experimental findings could match our discovered reaction $\text{A} \rightarrow \text{EZZZ} \rightarrow \text{G} \rightarrow \text{G'} \rightarrow \text{H}$ (Fig. 2). To determine whether the newly discovered species are general to other DASA derivatives, we investigated the stability of $\text{A'}$, $\text{EZZZ}$, $\text{B}$, $\text{E}$, $\text{F}$, $\text{G}$, and $\text{H}$ relative to $\text{A}$ in Fig. 3 and Table S4 with 1\textsuperscript{st}, 2\textsuperscript{nd}, and 3\textsuperscript{rd} generation donor and acceptor groups.
We explored the relative stability of these newly discovered species by systematically changing donor and acceptor groups: the comparison of 1a/2a and 1b/2b highlights the effect of a secondary aniline donor group with the same acceptor moiety. Compounds 3a, 3b, 3c and 3d show the impact of strong carbon acid acceptors with the same indoline donor, while in 3d, 3e and 3f we fixed the acceptor group and analysed the effects of increasing the electron-withdrawing character of the donor moiety.

Figure 3. Relative Free Energies of A’, EZZZ, B, E, F, G and H Minima between the 1st, 2nd and 3rd Generation DASAs. The relative free energies of A’, EZZZ, B, E, F, G, and H ground-state intermediates relative to A, computed with oB97x-D3/def2-TZVP(-f)/COSMO(ε=2.38) and using the harmonic approximation for vibrational contributions to the free energy. All open-form structures (A) of all three generations are shown on the bottom. The relative energy of intermediate G is shown in the box for 3f.
Along the main pathway, \( A' \) is \(~5 \text{ kcal/mol} \) higher in energy than \( A \) for all the derivatives, while the closed cyclopentenone form \( B \) is more stable than \( A \) for all the compounds of 2\(^{nd} \) and 3\(^{rd} \) generation (compounds 2\( a-2b \) and 3\( a-3f \), respectively). \textbf{Fig. 3} shows that changing the Meldrum acid into a Barbituric group (e.g., 1\( a \) vs 1\( b \), 2\( a \) vs 2\( b \)) increases the nucleophilicity of the acceptor group and stabilizes species \( E, F, G \). In contrast, the ability of the donor group to stabilize the newly discovered intermediates is limited to only \( F \) and has little effect on the stability of \( E \) and \( G \) (1\( a \) vs 2\( a \), 1\( b \) vs 2\( b \)).

Additionally, \textbf{Fig. 3} reveals that \textit{EZZZ} is a very stable intermediate in all DASA generations (the \( \Delta E \) (A-\textit{EZZZ}) ranges in an interval \( 1.6/-0.2 \text{ kcal/mol} \) being even slightly more stable than \( A \) for compound 3\( f \). Increasing the electron-withdrawing character of the donor group leads to an enhanced stabilization of \textit{EZZZ} (3\( d \), 3\( e \), and 3\( f \)). Furthermore, the inclusion of a strong carboxylic acid on the acceptor side (3\( a-3f \)) stabilizes \( G \) and \( H \) by promoting the nucleophilic attack of the hydroxyl and subsequent proton transfer to the acceptor group. \textbf{Fig. 3} shows that the relative energy between \( G \) and the open form \( A \) can be systematically decreased by enhancing the electron-withdrawing character of the donor group (3\( d-3f \)). In particular, \( G \) is slightly more stable than \( A \) and \textit{EZZZ} (-0.8 and -0.5 \text{ kcal/mol}, respectively) for compound 3\( f \). This opens a potential avenue where two orthogonal competing pathways could be achieved: the \( A\rightarrow B \) transformation initiated by light irradiation or the thermal \( A\rightarrow G \) route observed using heat to cross barriers (relative barriers are reported in \textbf{Table S5}). \textbf{Fig. S2} highlights the effect of the donor group on the \( A\rightarrow B \) and \( A\rightarrow G \) energy barriers for 3\( d \) and 3\( f \), which show that enhancing the electron-withdrawing character on the donor side lowers the barriers along the \( A\rightarrow G \) pathway while the \( A\rightarrow A' \) barrier is essentially unaffected.
These two different transformations could be experimentally followed via UV-vis absorption spectroscopy since $A$, $A'$, $B$, and $G$ absorb in different spectral windows. The simulated UV-vis absorption spectrum of $A$, $A'$, $B$, and $G$ is reported in Fig. 4 for compound $3d$. Along the $A \rightarrow B$ photoinduced pathway, the formation of $A'$ appears as a red-shifted shoulder ($\sim 0.2$ eV) on the $A$ spectrum, while the cyclopentenone closed form $B$ does not absorb in the visible range showing an absorption maximum in the UV-region $\sim 3$ eV blue-shifted with respect to $A$. On the other hand, $G$ has a $\sim 2$ eV blue-shifted absorption with respect to $A$ which could be used as a spectroscopic fingerprint to track its formation. This blue-shifted peak is a common feature of $G$ also present in the absorption spectrum of compound $1a$ and $2b$ as reported in Fig. S3.

In conclusion, our approach has unraveled alternative possibilities of DASA’s ground state chemical transformations bridging the gap between short fs/ps relaxation dynamics and long ns/μs thermal reaction regimes. Using state-of-the-art enhanced simulation techniques to drive reaction discovery runs, we revealed a plethora of thermal pathways which were uncharacterized until now.
In total 17 species have been identified in the discovery phase, and 8 of these (D, D’, D’’, E, F, G, G’, H) are proposed for the first time. Furthermore, we showed the potential stabilization of the newly discovered species via systematic structural changes in DASAs donor and acceptor moieties. These novel atomistic insights could be exploited to maximize the efficiency of the photoswitching process by limiting these newly discovered side-pathways. Furthermore, appropriately tuning the structure of compound 3f could lead to a designed multi-controllable DASA switch where two orthogonal pathways could be achieved and tuned using two different types of external stimuli (heat or light). Work in this direction is currently underway.
ASSOCIATED CONTENT

Supporting Information

Computational Details; reaction discovery workflow; summary of the discovered phase starting from A and A’; energies of the ground-state minima and transition states for Meldrum’s acid first-generation DASAs; relative energies of A’, EZZZ, B, E, F, G and H minima between the 1\textsuperscript{st}, 2\textsuperscript{nd} and 3\textsuperscript{rd} generation DASAs; energies of the ground-state minima and transition states for compound 3d and 3f; atoms included in the definition of the collective variable $\lambda^{\text{max}}$; reaction profile for DASA 3d and 3f; simulated UV electronic absorption spectra of compound 1a and 2b.

AUTHOR INFORMATION

Corresponding Author

michele.parrinello@iit.it; todd.martinez@stanford.edu

Present Addresses

#Design Physics Division, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

ACKNOWLEDGMENTS

The authors would like to thank Valerio Rizzi, Enrico Trizio and Friedrich Stricker for helpful discussions. This work was supported in part by the Office of Naval Research (N00014-18-1-2624 and N00014-18-1-2659). Lawrence Livermore National Laboratory is operated by Lawrence Livermore National Security, LLC, for the U.S. Department of Energy, National Nuclear Security Administration under Contract DE-AC52-07NA27344.
References


31. Peterson, J. A.; Stricker, F.; Read de Alaniz, J., Improving the kinetics and dark equilibrium of donor–acceptor Stenhouse adduct by triene backbone design. *Chemical Communications* 2022, 58 (14), 2303-2306.