UC Davis

Title

Mechanistic Investigations of Chemical and Photochemical Reactions with Quantum Chemistry and Machine-Learning Tools

Permalink

https://escholarship.org/uc/item/9hh0w8c7

Author

Feng, ZHITAO

Publication Date

2023

Peer reviewed|Thesis/dissertation

Mechanistic Investigations of Chemical and Photochemical Reactions with Quantum Chemistry and Machine-Learning Tools

By

ZHITAO FENG
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Chemistry

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

_____
Dean J. Tantillo, Chair

_____
Davide Donadio

_____
Lee-Ping Wang

Committee in Charge

2023

i

Dedicated to my family and friends who have supported me.

Acknowledgment

I would like to thank many people who have supported me through Ph.D. career.

Most importantly, I would like to acknowledge and thank my supervisor, Dean Tantillo, for his careful mentorship, support, and always insightful suggestions. I remember there were numerous days that I felt my research is going nowhere, but you would always give me useful suggestions that will alleviate my anxiety and help me back on track. The knowledge I learnt from you is not just computational chemistry, but also how to be a curious scientist and patient mentor. Your passion and philosophy towards education have also enlightened me and make me constantly rethink what education and science really mean. I wouldn't have made such accomplishments without your support.

I am very grateful to everyone in the Tantillo group as well. Specifically I would like to thank Stephanie Hare, Croix Laconsay, Yue Zhang, Amy Merrill, Shunyang Wang, Jesi Lee, Volga Kojasoy, Gil Fernandez, Phil Gingrich, Wentao Guo, Dongjie Chen, Yusef Ahmed, Kevin Kong, Will Desnoo, Ian Torrence, Alicia Ross, Yumeng Cao, Nutthakarn Soisang, Yibo Wang, Nico Havenner, Amy Cheung, Shusen Chen, Yangyang Xing, Mengna Bai, Yoshimitsu Hashimoto. I also would like to thank other people from the chemistry department at UC Davis, Nathaniel Troup, Zekun Chen, Zimin Zheng, Zhuoqi Yang, Chengyang Ji, Yingxin Su, Zhecheng He, Shuhao Yang.

I would like to acknowledge the help from many professors as well. First and foremost I would like to thank my dissertation committee, Prof. Lee-Ping Wang and Prof. Davide Donadio. I have learned a lot from your classes and your insightful suggestions are very helpful. Additionally, I would like to thank Prof. Dan Singleton at Texas A&M for helpful discussions about Progdyn and non-statistical effects in organic reactions. Discussions on photochemistry

Table of Contents

ABSTRACT

This dissertation is a collection of computational chemistry works on chemical and photochemical reaction mechanisms. Multiple theoretical tools have been utilized to tackle mechanistic questions in organic chemistry and photochemistry. They include density functional theory, machine learning and *ab initio* molecular dynamics and they are briefly discussed in chapter 1.

Chapter 2 is about the migratory aptitude of carbocation rearrangement from the perspective of dynamic effects. Carbocation rearrangement reactions are of great significance to synthetic and biosynthetic chemistry. In pursuit of a scale of inherent migratory aptitude that takes into account dynamic effects, both uphill and downhill ab initio molecular dynamics (AIMD) simulations were used to examine competing migration events in a model system designed to remove steric and electronic biases. The results of these simulations were combined with detailed investigations of potential energy surface topography and variational transition state theory calculations to reveal the importance of non-statistical dynamic effects on migratory aptitude.

Chapter 2 formulates a selectivity model based on the widths of pathways to competing products, rather than barrier heights, for a butadiene + allyl cation reaction. This model was arrived at via analysis of stationary points, intrinsic reaction coordinates, potential energy surface shapes and direct dynamics trajectories, all determined using quantum chemical methods.

Chapter 3 is on photochemistry. Selectivity for photochemical reactions where crossing between excited state and ground state surfaces occurs near ground state transition structures that interconvert competing products should be controlled by the momentum of the reacting molecules as they return to the ground state and the shape of the potential energy surfaces

involved. The roles of these factors are revealed here for a classic photochemical reaction: deazetization of 2,3-diazabicyclo[2.2.2]oct-2-ene. The utility of analogies between such photochemical reactions and ground state reactions with post-transition state bifurcations for forward design is suggested.

Chapter 4 is a collaborative project with the Sarpong group at UC Berkley. They have made synthetic efforts towards the pupukeanane natural products. The 10-step enantiospecific synthetic route to 2-isocyanoallopupukeanane facilitates an unprecedented bio-inspired 'contra-biosynthetic' rearrangement, providing divergent access to the pupukeanane core. The computational studies provide insights into the nature of this novel rearrangement.

Chapter 5 is a collaborative project with the Zheng group at the University of Arkansas. Distonic radical cations generated from the ring-opening of cyclopropylamines have shown distinct reactivities and TMSCN has been shown to be able to make the ring-opening happen. Our computational study has pointed out that the change in the photoredox potential and the quenching of the distonic radical cations by TMSCN are critical to the reactivities.

Chapter 6 is a collaborative project with the Lectka group at Johns Hopkins University on a through-space arene activations with halogens, tetrazoles and achiral esters and amides. Contrary to previously assumed direct activation through σ-complex stabilization, computational and experimental results suggest that these reactions proceed by a relay mechanism wherein the lone pair-containing activators form exothermic π-complexes with electrophilic nitronium ion before transferring it to the probe ring through low barrier transition states.

Publication List

Publication at UC Davis

[1]   **Zhitao Feng**, Wentao Guo, Wang-Yeuk Kong, Dongjie Chen, Shunyang Wang and Dean J. Tantillo*. Analogies between Photochemical Reactions and Ground State Reactions with Post-Transition State Bifurcations – Prospects for Rational Design. (Submitted)

[2]   **Zhitao Feng** and Dean J. Tantillo*. Dynamic Effects on Migratory Aptitudes in Carbocation Reactions. *J. Am. Chem. Soc.* **2021**, *143*, 1088-1097.

[3]   Mengna Bai†, **Zhitao Feng**†, Jun Li, Dean J. Tantillo*. Bouncing off Walls – Widths of Exit Channels from Shallow Minima Can Dominate Selectivity Control. *Chem. Sci.* **2020**, *11*, 9937-9944. (†: these authors contributed equally)

[4]   Melissa A. Hardy, **Zhitao Feng**, Isabel Kerschgens, Lynée A. Massey, Dean J. Tantillo*, Richmond Sarpong*. Unifying Synthesis of 2-isocyanoallopupukeanane and the Pupukeanane Core through a Bio-Inspired Rearrangement (Submitted)

[5]   Muhmmad Kazim, **Zhitao Feng**, Srini Vemulapalli, Maxime A. Siegler, Anant Chopra,1 Phuong Minh Nguyen, Maxwell Gargiulo Holl, Liangyu Guan, Dean J. Tantillo*, Travis Dudding*, Thomas Lectka* Through-Space, Lone-Pair Promoted Aromatic Substitution: A Relay Mechanism Beats Out Direct Activation. *Chem. Eur. J.*

[6]   Ran Fang*, **Zhitao Feng**, Alexander M. Kirillov, Lizi Yang*. Density Functional Theory of the Metal-Catalyzed Cycloaddition of Indolyl-Allenes: Possible Reaction Pathways, Stereoselectivity, and Regioselectivity. *Organometallics*. **2020**, 39, 1782-1789.

[7]   Chuan-Gang Zhao, **Zhi-Tao Feng**, Guo-Qiang Xu*, Ang Gao, Jing-Wei Chen, Zhu-Yin Wang, Peng-Fei Xu*. Highly Enantioselective Construction of Strained Spiro[2,3]hexanes through a Michael Addition/Ring Expansion/ Cyclization Cascade. *Angew. Chem. Int. Ed.* **2020**, 132, 3082-3086

[8]   Guo-Qiang Xu, **Zhi-Tao Feng,** Ji-Tao Xu, Zhu-Yin Wang, Yong Qin*, Peng-Fei Xu*. Transition-Metal-Free Selective C−H Benzylation of Tertiary Arylamines by a Dearomatization-Aromatization Sequence. *Chem. Eur. J.* **2018**, 24(52), 13778-13782

[9]   Wanjia Zhang, **Zhitao Feng**, Happy Mulenga, Wei Sun, Yuehua Hu, Jian Cao*, Zhiyong Gao* Synthesis of a Novel Collector Based on Selective Nitrogen Coordination for Improved Separation of Galena and Sphalerite against Pyrite. *Chem. Eng. Sci.* **2020**, 226, 115860.

[10]   Dan Liu, Meng-Jie Jiao, **Zhi-Tao Feng**, Xing-Zhi Wang, Guo-Qiang Xu, Peng-Fei, Xu*. The Design, Synthesis and Application of Highly Reducing Organic Visible-Light Photocatalysts. *Org. Lett.* **2018**, 20(18), 5700-5704

[11]   Hui Liang, Guo-Qiang Xu, **Zhi-Tao Feng**, Zhu-Yin Wang*, Peng-Fei Xu*. Dual Catalytic Switchable Divergent Synthesis: An Efficient Asymmetric Visible-light Photocatalytic Approach to Fluorine-containing γ-Keto Acid Frameworks. *J. Org. Chem.* **2019**, 84(1), 60-72.

[12]   Guo-Qiang Xu, Ji-Tao Xu, **Zhi-Tao Feng**, Hui Liang, Zhu-Yin Wang, Yong Qin, Peng-Fei Xu*. Dual C(sp^3)-H Bond Functionalization of N-heterocycles through Visible-Light Photocatalyzed Dehydrogenation/[2+2] Cycloaddition Sequential Reaction. *Angew. Chem. Int. Ed.* **2018**, 130(18), 5204-5208.

Chapter 1

Introduction to Computational Chemistry

Ever since Max Planck used the word 'quantum' in his black-body radiation paper, quantum mechanics have gone through hundreds of years rapid development and revolutionized multiple disciplines of science and technology.[1] In light of this, quantum chemistry has emerged as a powerful tool and an active field aiming at studying the properties of molecules and demystifying the mechanism behind chemical reactions.[2] The reactivity and selectivity models derived from computational chemistry experiments paves the way for rational design of new molecule and reactions that is essential to the human society.[3–9]

Wavefunction is at the core of the quantum mechanics. Any state of a microscopic system can be described by wavefunction and it carries all the information of the system. The evolution of a wavefunction is governed by the time-dependent Schrödinger equation (TDSE).

$$i\hbar \frac{\partial \Psi}{\partial t} = \widehat{H}\Psi$$

When the potential energy is time-independent, we can separate out time and obtain the time-independent Schrödinger equation (TISE):

$$\widehat{H}\Psi = E\Psi$$

where $\widehat{H}$ is the Hamiltonian operator of a system. The solutions of time-independent Schrödinger equation are the eigenstates of the Hamiltonian operator. For a system with N electrons and M nuclei, the Hamiltonian takes the form of:

$$\widehat{H} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{A=1}^{M}\sum_{i=1}^{N}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} - \frac{1}{2M_A}\sum_{A=1}^{M}\nabla_A^2 + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}}$$

Since nuclei are $10^3 \sim 10^5$ heavier than electrons, the motion of electrons and nuclei in most cases are decoupled. We can approximately view the electrons as moving in a field created by nuclei and the total electronic energy of the system depends only on the nuclei positions, which brings us the concept of potential energy surface. The Born-Oppenheimer (B-O) approximation

is at the center of computational chemistry as the nuclei essentially move on a potential energy surface governed by the electron. However, B-O approximation fails when the coupling between energy states is non-negligible. In such cases, the nuclei and the electrons are dynamically coupled and B-O approximation is not valid. This is common in some photophysical and photochemical processes.[10–14]

To solve the Schrödinger equation under the B-O approximation, we also need a set of basis functions to represent the wavefunction of the electrons called basis set. Linear combinations of these basis functions form molecular orbitals. Some commonly used basis sets include Gaussian-type basis set, Slater-type basis set and plane-wave basis set.

The Slater-type atomic orbital has the form of:

$$\phi(r) = \sqrt{\frac{\zeta^3}{\pi}} e^{-\zeta|r-R|}$$

The Gaussian-type orbital has the form of:

$$\phi(r) = \left(\frac{2\alpha}{\pi}\right)^{\frac{3}{4}} e^{-\alpha|r-R|^2}$$

ζ and α are Slater and Gaussian orbital exponent, respectively. Although Slater-type basis set is more physical, Gaussian-type basis sets are more widely used due to its better mathematical behavior for two-electron integrals.[15]

Schrödinger equation is a second-order multivariable partial differential equation and solving it necessitates complicated mathematical tools. The choice of how to solve the Schrödinger equation determines the method for the computation task. Hartree-Fock is the simplest one that uses mean field theory in treating electrons and it is the foundation of all *ab initio* methods. The HF method uses an iterative, self-consistent approach to solve for a

wavefunction by minimizing the energy. The wavefunction is written in the form of Slater determinants so as to satisfy the antisymmetric rule for electrons:

$$\Psi(1,2,\dots,n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \chi_1(x_1) & \cdots & \chi_n(x_1) \\ \vdots & \ddots & \vdots \\ \chi_1(x_n) & \cdots & \chi_n(x_n) \end{vmatrix}$$

χ_i is the spin orbital, which a spatial orbital (ψ) multiplied by a spin eigenfunction (α or β). The spatial orbital of a molecule, according to the linear combination of atomic orbitals (LCAO) approach, can be represented by a set of atomic orbitals:

$$\psi_i(\boldsymbol{r}) = \sum_{\mu=1}^{K} C_{\mu i} \phi_\mu(\boldsymbol{r})$$

The energy of the molecular system can be written as:

$$E = 2 \sum_{i=1}^{N/2} h_{ii} + \sum_{i=1}^{N/2} \sum_{j=1}^{N/2} (2J_{ij} - K_{ij})$$

In which h_{ii} represents the energy of one electron under the influence of all nuclei.

$$h_{ii} = \int \psi_i^*(r_1) \left(-\frac{1}{2}\nabla^2 - \sum_{A=1}^{M} \frac{Z_A}{r_{1A}} \right) \psi_i(r_1)\, dr_1$$

J_{ij} and K_{ij} represents the coulomb integral and the exchange integral, respectively.

$$J_{ij} = \int \psi_i^*(r_1)\psi_j^*(r_2) \left(\frac{1}{r_{12}} \right) \psi_i(r_1)\psi_j(r_2)\,dr_1 dr_2$$

$$K_{ij} = \int \psi_i^*(r_1)\psi_j^*(r_2) \left(\frac{1}{r_{12}} \right) \psi_i(r_2)\psi_j(r_1)\,dr_1 dr_2$$

The coulomb integral reflects the Coulombic repulsion between two electrons and the exchange integral reflects the potential energy arises from the Pauli repulsion of two electrons with the same spin.

The variational principle states that the expectation value of the energy must be greater than or equal to the true energy of the system. It means that if we were able to obtain a wavefunction with minimum energy, we can yield the accurate wavefunction of the system within the HF limit. The HF looks for one set of orbital coefficients that gives the lowest energy iteratively. It uses the wavefunction from the previous iteration to construct the pseudo one-electron Fock operator. Ideally this reaches convergence when the energy change between iterations is below a certain threshold. Such process is called the self-consistent field (SCF) method.

HF has its clear disadvantage of completely neglecting the electron correlation as when one electron changes its position, the mean electron field should change simultaneously. Many more advanced methods have been developed to circumvent this problem, and they are called post-HF methods, including configuration interaction (CI), Møller-Plesset (MP) perturbation theory and coupled-cluster (CC) theory. However, these methods are fairly expensive and are only practical for small systems. Density functional theory (DFT) is a way to account for the electron correlation in a cost-effective way and achieving an accuracy that is comparable, or even outperforms more expensive MP2 level of theory. There is no doubt that DFT have become the most widely utilized computational method and finds extensive applications in chemistry, biochemistry, drug design and material science.

The basic idea of DFT is that the ground state energy of a many-body system in a potential is defined uniquely the electron density. Instead of obtaining an accurate wavefunction, we try to get the correct electron density. According to the first Hohenberg-Kohn theorem, which states that the Hamiltonian of the system is exactly determined if we have the ground state electron density $\rho(r)$, all the properties of the multi-body system, including energy, is a functional of the electron density $\rho(r)$.[16] The second Hohenberg-Kohn theorem states that for

a given Hamiltonian we can define an energy functional $E[\rho(r)]$ such that if it is the ground electron density, the energy functional has the minimum energy.

To practically employ DFT to electronic structure calculations requires actual knowledge of the density functional, while Hohenberg-Kohn theorems only proved the existence of $E[\rho(r)]$. Kohn and Sham proposed that the main part of energy functional can be separated out first, namely the kinetic energy and Coulomb energy of non-interacting particles. The idea of orbitals is re-introduced to calculate the kinetic energy of the non-interacting electrons. The rest of the functional can be treated later with approximations. Thus, we seek for the solution for the Kohn-Sham equation shown below.[17]

$$\left[-\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{Z_A}{r_{1A}} + \int \frac{\rho(r_2)}{r_{12}}dr_2 + v_{XC} \right]\psi_i = \epsilon\psi_i$$

The v_{XC} term is called the exchange-correlation potential and it contains all the complicated aspects of electron exchange and correlation. The various forms of exchange-correlation are the key difference between the hundreds of DFTs. Based upon the form of the exchange-correlation term, functionals can be categorized into local density approximation (LDA), generalized gradient approximation (GGA) and meta-generalized gradient approximation (meta-GGA), hybrid generalized gradient approximation (hybrid-GGA) and hybrid meta-GGA functionals. LDA takes the simplest form the exchange-correlation functional in that the density is local. GGAs starts to incorporate the gradient of electron density into the exchange-correlation functional, thus they are non-local functionals. Meta-GGA functionals add second derivative of the electron density. Hybrid functionals adds partial exchange energy into the functional term since HF exchange energy of HF orbital is exact. Different hybrid functionals have distinct percentages of HF exchange energy.[18–20]

Hartree-Fock, MP2 and DFT methods are all single-reference methods, meaning that they are not suitable for systems with significant non-dynamical (or static) electron correlation. Non-dynamical electron correlation arises when other determinants have similar magnitude of coefficient in the total wavefunction shown below.

$$\Psi = c_0 \Psi_{HF} + c_1 \Psi_1 + c_2 \Psi_2 + \cdots + c_n \Psi_n$$

Static correlation is non-negligible when the HOMO and LUMO energy gap becomes small and substantial orbital mixing leads to singly occupied orbitals that are almost degenerate. This is common in systems with diradical character, metallic systems or homolytic bond breaking.

In these cases, we need a multiconfigurational wavefunction that consists of multiple configurations. For instance, a two-configuration wavefunction has the form of:

$$\Psi = c_0 \psi_0 + c_1 \psi_1$$

Each ψ is a configuration, which can be either single determinantal, or contains multiple determinants (the case of complete active space SCF with restricted open shell HF).

The multiconfigurational SCF not only finds the optimum coefficients for the orbitals, but also the optimum coefficient for each configuration. In the case of full configuration interaction, one considers all possible configurations of a system with an intractable complexity of *O(n!)*. To specify configurations to be included, we need to specify an active space of a molecule. Only the orbitals within the set may be occupied by zero, one or two electrons. Other orbitals remain doubly occupied or empty. A complete active space contains all possible permutations of active electrons in active space and such calculation is called complete active space self-consistent field (CASSCF). In practice, a HF calculation is first performed to obtain reference orbitals for constructing configurations. The orbitals can be delocalized canonical

orbitals or localized orbitals. Then the multiconfigurational SCF (MCSCF) procedure updates the MO coefficients and the configuration coefficients.

Although CASSCF is able to capture static correlations, it still lacks sufficient description of dynamical correlation. We can use MCSCF as the reference for perturbation theory, for instance the commonly used second-order complete active space perturbation theory (CASPT2), or for configuration interaction method, such as multireference CI. Multiconfigurational methods are not only suitable for systems with apparent static correlation, but also necessary to excited-state problems, as excited state configurations are inherently included in the wavefunction.

Solving the Schrödinger equation yields the energy and the wavefunction of the system. B-O approximation states that adding the electronic energy and the nuclear repulsion energy, we obtain the total energy of the molecule for a specific geometry. In other words, the energy of a molecule is a function of nuclear coordinates, and by systematically varying the nuclear coordinates, the corresponding energy changes, depicting the potential energy surface (PES). PES is at the center of the mechanistic investigations in computational chemistry, as a chemical reaction is effectively an exploration of certain regions of the PES. For a nonlinear molecule with N atoms, the PES is a 3N-6 hypersurface. For most of the reactions we only pick one critical dimension to study, referred to as the reaction coordinate. A chemical reaction usually follows the minimal energy path (MEP) along the reaction coordinate so mechanistic investigations focus on the critical points on the PES, such as minima and saddle points. The transition state structure refers to the highest-energy point along the reaction path on the potential energy surface. It connects the structures of reactants and products through intrinsic reaction coordinate (IRC, mass-weighted MEP).[21–23]

Minima and TSs both have zero gradients but their curvatures are different. Since the energy of a molecule is a multivariate function of atom positions, the curvature is a 3N*3N matrix, namely the Hessian matrix. Diagonalization of the Hessian matrix yields three translational, three rotational (these frequencies are zero) and 3N-6 vibrational frequencies (square root of eigenvalue) and modes (eigenvectors), for non-linear molecules. We treat these vibrational modes as quantum harmonic oscillators and it is usually valid for not very low frequency modes. The vibrational modes of minima on the PES have all positive eigenvalues (ω^2, square of frequency), implying that perturbing the molecule along any vibrational mode and direction, the energy will increase. TS have exactly one negative eigenvalue and following this imaginary frequency the energy decreases instead. The imaginary frequency is highly related to the reaction coordinate and it depicts the picture that TS sits on the top of a hill. A TS is a maximum and bottleneck of a reactive flux that passes through reaction coordinate, while it is a minimum for the rest of the 3N-7 dimension.

The information obtained from frequency calculations is essential. The vibrational frequencies can be used to calculate the vibrational contribution to the partition function, thus yields the free energy correction to the electronic energy. Another usage of frequency analysis result is that, in MD simulations it is required to have an appropriate sampling of the initial velocities with zero-point energies under certain temperature. The energy of the lowest vibrational level is $\hbar\omega/2$ and the sum of all vibrational levels defines the zero-point energy. Moreover, under non-zero temperature conditions, the molecule might populate vibrational excited states. The initiation of an MD simulation will be discussed in detail later.

Transition state theory plays a pivotal role in computational chemistry. TST is a statistical theory that quantitatively rationalizes the rate of chemical transformation.[24] The rate of a chemical reaction can be calculated from the Eyring equation which is derived from evaluating

the reactive flux that passes through a dividing surface. Few assumptions were made during the derivation of the equation. Apparently, B-O approximation must apply as the reactive flux is on the PES. Secondly, the motion of the nuclei is treated classically in conventional TST so there should be no crossing between electronic states. Additionally, when the reactive flux reaches the dividing surface, it has to pass through the surface without recrossing. Furthermore, the derivation starts from the equilibrium constant between the reactant and the so-called 'activated complex', so the reactant and the 'activated complex' have to be in equilibrium. With the normal mode properties of the TS and statistical mechanics, Henry Eyring and Michael Polanyi derived the Eyring equation of absolute reaction rate that provides theoretical understanding of reaction kinetics.[25]

Modern theoretical organic chemistry relies on TST. Barrier height determines the rate of an organic reaction, offering us information of the feasibility of the transformation. Organic chemists are also interested in the regioselectivity or enantioselectivity of a reaction, which in statistical cases, is directly related to the TST energy difference of the rate determining steps (so called $\Delta\Delta G$) between different pathways, provided that TST can be applied.[26,27]

There are, of course, cases when TST cannot be applied that requires a dynamic or non-classical inspection. It is usually called non-statistical or dynamic effect. One of the most widely studied one is bifurcation that one TS can lead to two distinct products. It occurs when two TSs are directly connected without the presence of an intermediate. After passing through the first TS, the reaction travels along the ridge and eventually falls to one of the products before reaching the second TS that interconverts the two products. Bifurcations have been found in many synthetic relevant reactions and enzymatic systems. The controlling factor of a bifurcation is the topology of the PES, which is a two-dimensional problem. By modulating

the PES around the ambimodal TS, one can eventually control the outcome of the reaction.[26,28–32]

Other situations that TST fail can be summarized as the oversimplification of the model led by the assumptions for the transition state theory. The failure of the assumptions leads to distinct non-statistical dynamic effect that has attracted increasing interests from synthetic and computational chemists. TST is a classical theory that inherently neglects quantum effect, so modeling the rate of a chemical reaction that displays significant quantum mechanical nuclear tunneling effect requires semiclassical corrections on top of conventional TST. Practical models include small-curvature tunneling (SCT), large-curvature tunneling (LCT) and zero-curvature tunneling (ZCT).[33]

Non-equilibrium effect is a direct violation of the quasi-equilibrium approximation. It is usually the consequence of inefficient intramolecular vibration-energy redistribution (IVR). In exergonic reactions, it takes about 1ps to 10ps for the molecule to transfer the excess energy to the surrounding environment and to redistribute the energy to each normal mode. In cases when a low barrier step follows a highly exergonic process, the reaction will bypass the shallow intermediate before fully equilibrated, and the molecule keeps the memory of the previous step. The excess energy can thus affect the outcome of the subsequent reactions, and the reaction outcome matches with the reaction vector of the previous step. Such effect is usually called dynamic matching or memory effect. A wide range of reactions have been found to be controlled by dynamic matching, including carbocation reactions, Diels-Alder reaction[34], hydroboration of alkenes[35,36], α-cleavage of alkoxy radicals and deazetization reactions.[31,37–39]

Recrossing is another type of dynamic effect, but less studied. Significant amount of recrossing is anticipated when the TS resides on a fairly flat PES. Although employing VTST can in principle provide us information about the free energy maxima and mitigate recrossing,

11

in some cases, 'non-statistical recrossing' is inherent in a reaction, and it is able to influence the reaction outcome. Recrossing can serve as a probe for hidden dynamical bottlenecks and alter selectivity.

It is not a trivial task to model dynamic effects in a chemical reaction and sometimes multiple dynamic effects can be involved in one reaction. The most widely used tool to study dynamic effect is *ab initio* molecular dynamics (AIMD). MD simulations provide information about the system evolution over time. AIMD incorporate electronic structure information into MD simulations by solving Schrödinger equation and Newton's equation. In Ehrenfest MD, the two equations are solved simultaneously, and the nuclei and the electrons are coupled. Such approach is often used in non-adiabatic MD (NAMD) simulations that requires evaluation of electron-nuclei couplings. Car-Parrinello MD (CPMD) is another one that can reduce the computational cost of MD by include electrons as degrees of freedom.

Nonetheless, Born-Oppenheimer MD (BOMD) is the most widely used method in AIMD mechanistic studies. BOMD treats nuclei and electrons separately. *Ab initio* computation is employed first to obtain the energy and force of the molecule. Then the trajectory is propagated by a numerical integration algorithm, such as the velocity Verlet algorithm or the Leapfrog algorithm. Molecular dynamics simulations start from either TS (downhill MD) or reactant (uphill MD). For downhill MD the trajectories are propagated along both directions to make sure that they connect the reactant and the product. For uphill MD, an additional energy is usually added to the system along the reaction vector to accelerate the reaction and enhance the sampling procedure, as a chemical reaction is a rare event and it would take extremely long time for the trajectory to overcome the barrier. The initiation of a molecular dynamic trajectory requires special care. Commonly used sampling methods include Boltzmann sampling and Wigner sampling that are both based on vibrational modes. Boltzmann sampling is a classical

sampling method that the most probable normal mode coordinate is at the turning points. However, in quantum mechanical distribution, the probability peaks at the equilibrium. Wigner sampling is common way to use quantum distribution in phase space and it has been applied to the initiation of coordinates and momenta. MD simulations can also be done in different thermodynamic ensembles. If there is no additional potential added and the integration is accurate enough, the total energy is conserved and the MD is conducted under the NVE ensemble. For simulations in NVT ensemble, forces are added to adjust the kinetic energy of the system so as maintain a constant temperature (Langevin MD).

AIMD is a powerful tool to explore the dynamics of the system. By examining how one system property evolves with time, one can obtain a more detailed and intuitive understanding of complicated systems. However, as mentioned earlier, along the relaxation process of electronic excited state to ground state, it is possible that the scale of electron and nuclei movement becomes close and the coupling between them is non-negligible. In such case, the Bohn-Oppenheimer (or the adiabatic) approximation that separate the two degrees of freedom can lead to large error. A process that cannot be treated with adiabatic approximation is called a non-adiabatic process and the corresponding molecular dynamic simulations is called the non-adiabatic molecular dynamics (NAMD). There are multiple theoretical frameworks that can treat the non-adiabatic problem. For instance, we can treat the both the electrons and the nuclei quantum mechanically. However, this is extremely time-consuming and is only applicable to tiny systems. Mixed quantum chemical methods are more widely used instead as it considers nuclei as a classical subsystem that are coupled with electrons (QM subsystem).[40] The evolution of the QM subsystem is influenced by the classical subsystem and hopping between different electronic states is allowed. In the meantime, the classical subsystem will respond to the state hopping and the treatment of the response necessitates special care. To

address this, two different schemes, namely the Ehrenfest dynamics[41–43] and the surface hopping dynamics have been developed[40,44–46].

In Ehrenfest dynamics, the classical particle can be viewed as evolving on a potential energy surface that is the weighted sum of all electronic states. It generally gives reasonable results to the hopping probability and is usually employed in systems where there is no apparent difference between electronic states. However, the correlation between the QM subsystem and the classical subsystem is not correctly depicted and it would give unphysical result when the states differ greatly. On the other hand, the surface hopping scheme introduces the idea of random hopping to deal with the surface hopping problem. The equation of motion of the classical subsystem can be written as:

$$M\ddot{R} = -\nabla E_m(R)$$

Where E_m is the potential energy surface of the kth state. The classical subsystem evolves on a specific potential surface and the coupling between states is described by the non-adiabatic coupling matrix. The degree of the coupling determines the hopping probability. The most widely used algorithm is the Tully's fewest-switches surface hopping algorithm. The probability to hop from state m to state n is calculated as,

$$P_{m \to n} = -\frac{2}{|C_m|^2} Re(C_n C_m^* \sigma_{mn}) \, \delta t$$

C is the time-dependent coefficient of the time-dependent wavefunction in the form of linear combination of adiabatic states. σ_{mn} is the non-adiabatic coupling between m and n state.

$$\sigma_{mn} = \left\langle \psi_m \left| \frac{\partial}{\partial t} \psi_n \right. \right\rangle$$

However, ab initio MD simulations are particularly costly, as each point in a trajectory requires on-the-fly calculation of the energy and gradient and it prohibits the application to larger systems and longer timescale simulations. It would be ideal to perform AIMD simulation

14

under molecular mechanics cost while maintain its accuracy. In this regard, machine learning, a rapidly evolving field that has proved its applicability in chemistry and materials science have been employed in the aim of achieving such a goal.[47–53] The general idea of a machine-learning potential is to look for a mapping from the cartesian coordinates to the energy and the gradient of the molecule with neural networks. First, an initial dataset of molecules with their energies and gradients is generated via QM computations. The dataset consists of critical points on the PES (minima, conical intersections and TSs) and points along the reaction pathways. An initial model is trained with the initial dataset and it is subsequently used for the adaptive sampling procedure to capture underrepresented structures in order to improve model accuracy. Usually, two distinct models with different architecture and complexity (layers, number of neurons per layer, regularization) are trained and their predictions are compared. If the two models exhibit discrepancies on the prediction of the same structure, it implies the structure is not well learned by the models and it will be added to the dataset. Then the model is retrained with the updated dataset. This iterative process is carried out until the mean absolute error (MAE) is below threshold, or the length of ML trajectories reaches certain limit. The ML potential is then used for large scale MD simulations.[54,55]

Machine learning can achieve up to thousand-fold acceleration to MD simulations (see also chapter 4) and it has made MD investigation of large systems possible. For instance, Galib *et al* have utilized reactive ML potential to study the hydrolysis of N_2O_5 at the liquid-air interface which is critical to the uptake of N_2O_5 by atmospheric aerosol.[56]Ang *et al* have employed graph convolutional neural network (GCNN) potential to study the PTSB of several pericyclic reactions and they have achieved a quantum chemical accuracy.[50] ML is also helpful to NAMD, as it requires calculations of multiple electronic states using expensive multireference methods. The size of the system practical for NAMD investigation used to be very limited (<20 atoms, femtosecond timescale). However, with the help of ML, we can now

achieve picosecond scale NAMD on larger systems (see chapter 4 for more examples). These advanced simulation techniques pave the way for constructing more complicated molecular reaction dynamic theories that can not only help us understand the nature around us, but also enable rational design of novel reactions.

References

1. Kragh, H. Max Planck: the reluctant revolutionary. *Physics World* **13**, 31–36 (2000).

2. Tantillo, D. J. *Applied Theoretical Organic Chemistry*. (World Scientific Publishing Europe Ltd., 2018).

3. Iribarren, I., Garcia, M. R. & Trujillo, C. Catalyst design within asymmetric organocatalysis. *WIREs Computational Molecular Science* **12**, (2022).

4. Wheeler, S. E., Seguin, T. J., Guan, Y. & Doney, A. C. Noncovalent Interactions in Organocatalysis and the Prospect of Computational Catalyst Design. *Acc Chem Res* **49**, 1061–1069 (2016).

5. Poree, C. & Schoenebeck, F. A Holy Grail in Chemistry: Computational Catalyst Design: Feasible or Fiction? *Acc Chem Res* **50**, 605–608 (2017).

6. Sabe, V. T. *et al.* Current trends in computer aided drug design and a highlight of drugs discovered via computational techniques: A review. *Eur J Med Chem* **224**, 113705 (2021).

7. Tantillo, D. J., Jiangang, C. & Houk, K. N. Theozymes and compuzymes: theoretical models for biological catalysis. *Curr Opin Chem Biol* **2**, 743–750 (1998).

8. Chen, Z., Yu, Y., Gao, Y. & Zhu, Z. Rational Design Strategies for Nanozymes. *ACS Nano* (2023) doi:10.1021/acsnano.3c04378.

9. Ahn, S., Hong, M., Sundararajan, M., Ess, D. H. & Baik, M.-H. Design and Optimization of Catalysts Based on Mechanistic Insights Derived from Quantum Chemical Reaction Modeling. *Chem Rev* **119**, 6509–6560 (2019).

10. Ishida, T., Nanbu, S. & Nakamura, H. Clarification of nonadiabatic chemical dynamics by the Zhu-Nakamura theory of nonadiabatic transition: from tri-atomic systems to reactions in solutions. *Int Rev Phys Chem* **36**, 229–285 (2017).

11. Wang, L., Akimov, A. & Prezhdo, O. V. Recent Progress in Surface Hopping: 2011–2015. *J Phys Chem Lett* **7**, 2100–2112 (2016).

12. Tully, J. C. Molecular dynamics with electronic transitions. *J Chem Phys* **93**, 1061–1071 (1990).

13. Barbatti, M. Nonadiabatic dynamics with trajectory surface hopping method. *Wiley Interdiscip Rev Comput Mol Sci* **1**, 620–633 (2011).

14. Mai, S., Marquetand, P. & González, L. Nonadiabatic dynamics: The SHARC approach. *WIREs Computational Molecular Science* **8**, (2018).

15. Pritchard, B. P., Altarawy, D., Didier, B., Gibson, T. D. & Windus, T. L. New Basis Set Exchange: An Open, Up-to-Date Resource for the Molecular Sciences Community. *J Chem Inf Model* **59**, 4814–4820 (2019).

16. Hohenberg, P. & Kohn, W. Inhomogeneous Electron Gas. *Physical Review* **136**, B864–B871 (1964).

17. Kohn, W. & Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review* **140**, A1133–A1138 (1965).

18. Orio, M., Pantazis, D. A. & Neese, F. Density functional theory. *Photosynth Res* **102**, 443–453 (2009).

19. Cohen, A. J., Mori-Sánchez, P. & Yang, W. Challenges for Density Functional Theory. *Chem Rev* **112**, 289–320 (2012).

20. Verma, P. & Truhlar, D. G. Status and Challenges of Density Functional Theory. *Trends Chem* **2**, 302–318 (2020).

21. Maeda, S., Harabuchi, Y., Ono, Y., Taketsugu, T. & Morokuma, K. Intrinsic reaction coordinate: Calculation, bifurcation, and automated search. *Int J Quantum Chem* **115**, 258–269 (2015).

22. Fukui, K. The path of chemical reactions - the IRC approach. *Acc Chem Res* **14**, 363–368 (1981).

23. Gonzalez, Carlos. & Schlegel, H. Bernhard. Reaction path following in mass-weighted internal coordinates. *J Phys Chem* **94**, 5523–5527 (1990).

24. Eyring, H. The Activated Complex in Chemical Reactions. *J Chem Phys* **3**, 107–115 (1935).

25. Hänggi, P., Talkner, P. & Borkovec, M. Reaction-rate theory: fifty years after Kramers. *Rev Mod Phys* **62**, 251–341 (1990).

26. Rehbein, J. & Carpenter, B. K. Do we fully understand what controls chemical selectivity? *Physical Chemistry Chemical Physics* **13**, 20906 (2011).

27. Ess, D. H. *et al.* Bifurcations on Potential Energy Surfaces of Organic Reactions. *Angewandte Chemie International Edition* **47**, 7592–7601 (2008).

28. Hare, S. R. & Tantillo, D. J. Post-transition state bifurcations gain momentum – current state of the field. *Pure and Applied Chemistry* **89**, 679–698 (2017).

29. Bai, M., Feng, Z., Li, J. & Tantillo, D. J. Bouncing off walls – widths of exit channels from shallow minima can dominate selectivity control. *Chem Sci* **11**, 9937–9944 (2020).

30. Hare, S. R., Pemberton, R. P. & Tantillo, D. J. Navigating Past a Fork in the Road: Carbocation−π Interactions Can Manipulate Dynamic Behavior of Reactions Facing Post-Transition-State Bifurcations. *J Am Chem Soc* **139**, 7485–7493 (2017).

31. Carpenter, B. K. Energy Disposition in Reactive Intermediates. *Chem Rev* **113**, 7265–7286 (2013).

32. Martin-Somer, A., Xue, X.-S., Jamieson, C. S., Zou, Y. & Houk, K. N. Computational Design of a Tetrapericyclic Cycloaddition and the Nature of Potential Energy Surfaces with Multiple Bifurcations. *J Am Chem Soc* (2023) doi:10.1021/jacs.2c12871.

33. Bao, J. L. & Truhlar, D. G. Variational transition state theory: theoretical framework and recent developments. *Chem Soc Rev* **46**, 7548–7596 (2017).

34. Wang, Z., Hirschi, J. S. & Singleton, D. A. Recrossing and Dynamic Matching Effects on Selectivity in a Diels-Alder Reaction. *Angewandte Chemie International Edition* **48**, 9156–9159 (2009).

35. Oyola, Y. & Singleton, D. A. Dynamics and the Failure of Transition State Theory in Alkene Hydroboration. *J Am Chem Soc* **131**, 3130–3131 (2009).

36. Bailey, J. O. & Singleton, D. A. Failure and Redemption of Statistical and Nonstatistical Rate Theories in the Hydroboration of Alkenes. *J Am Chem Soc* **139**, 15710–15723 (2017).

37. Engel, P. S., Nalepa, C. J., Soltero, L. R., Horsey, D. W. & Keys, D. E. Photolysis of reluctant azoalkanes. Effect of structure on photochemical loss of nitrogen from 2,3-diazabicyclo[2.2.2]oct-2-ene derivatives. *J Am Chem Soc* **105**, 7108–7114 (1983).

38. Engel, P. S., Hayes, R. A., Keifer, L., Szilagyi, S. & Timberlake, J. W. Extrusion of nitrogen from cyclic and bicyclic azo compounds. *J Am Chem Soc* **100**, 1876–1882 (1978).

39. Edmunds, A. J. F. & Samuel, C. J. Photochemical deazetation of 2,3-diazabicyclo[2.2.2]oct-2-ene: pseudorotation of the cyclohexanediyl biradical. *J Chem Soc Perkin 1* 1267 (1989) doi:10.1039/p19890001267.

40. Tully, J. C. Mixed quantum–classical dynamics. *Faraday Discuss* **110**, 407–419 (1998).

41. Ojanperä, A., Havu, V., Lehtovaara, L. & Puska, M. Nonadiabatic Ehrenfest molecular dynamics within the projector augmented-wave method. *J Chem Phys* **136**, (2012).

42. Cheng, S. C., Zhu, C., Liang, K. K., Lin, S. H. & Truhlar, D. G. Algorithmic decoherence time for decay-of-mixing non–Born–Oppenheimer dynamics. *J Chem Phys* **129**, (2008).

43. Nijjar, P., Jankowska, J. & Prezhdo, O. V. Ehrenfest and classical path dynamics with decoherence and detailed balance. *J Chem Phys* **150**, (2019).

44. Tully, J. C. Molecular dynamics with electronic transitions. *J Chem Phys* **93**, 1061–1071 (1990).

45. Wang, L., Trivedi, D. & Prezhdo, O. V. Global Flux Surface Hopping Approach for Mixed Quantum-Classical Dynamics. *J Chem Theory Comput* **10**, 3598–3605 (2014).

46. Schmidt, J. R., Parandekar, P. V. & Tully, J. C. Mixed quantum-classical equilibrium: Surface hopping. *J Chem Phys* **129**, 044104 (2008).

47. Westermayr, J. & Marquetand, P. Machine Learning for Electronically Excited States of Molecules. *Chem Rev* **121**, 9873–9926 (2021).

48. Gallegos, L. C., Luchini, G., St. John, P. C., Kim, S. & Paton, R. S. Importance of Engineered and Learned Molecular Representations in Predicting Organic Reactivity, Selectivity, and Chemical Properties. *Acc Chem Res* **54**, 827–836 (2021).

49. Wang, W. & Gómez-Bombarelli, R. Coarse-graining auto-encoders for molecular dynamics. *NPJ Comput Mater* **5**, 125 (2019).

50. Ang, S. J., Wang, W., Schwalbe-Koda, D., Axelrod, S. & Gómez-Bombarelli, R. Active learning accelerates ab initio molecular dynamics on reactive energy surfaces. *Chem* **7**, 738–751 (2021).

51. Zhang, J. *et al.* Rich Dynamics Underlying Solution Reactions Revealed by Sampling and Data Mining of Reactive Trajectories. *ACS Cent Sci* **3**, 407–414 (2017).

52. Li, C. & Voth, G. A. Using Machine Learning to Greatly Accelerate Path Integral *Ab Initio* Molecular Dynamics. *J Chem Theory Comput* **18**, 599–604 (2022).

53. Li, J. & Lopez, S. A. A Look Inside the Black Box of Machine Learning Photodynamics Simulations. *Acc Chem Res* **55**, 1972–1984 (2022).

54. Li, J. & Lopez, S. A. Excited-State Distortions Promote the Photochemical 4π-Electrocyclizations of Fluorobenzenes via Machine Learning Accelerated Photodynamics Simulations. *Chemistry – A European Journal* **28**, (2022).

55. Li, J., Stein, R., Adrion, D. M. & Lopez, S. A. Machine-Learning Photodynamics Simulations Uncover the Role of Substituent Effects on the Photochemical Formation of Cubanes. *J Am Chem Soc* **143**, 20166–20175 (2021).

56. Galib, M. & Limmer, D. T. Reactive uptake of N_2O_5 by atmospheric aerosol is dominated by interfacial processes. *Science (1979)* **371**, 921–925 (2021).

Part 1

Dynamic Effects in Chemical Reactions

2.1 Introduction

One could argue that the field of carbocation rearrangement reactions was birthed in 1859 when Fittig treated pinacol with sulfuric acid and discovered the pinacol rearrangement.[1] Since then, carbocation rearrangements have been shown to play critical roles in biological,[2] synthetic,[3] and physical organic chemistry,[4] both because of the efficiency by which such rearrangements can be used to construct complex polycyclic systems and because they have served as testing grounds for non-traditional bonding and mechanistic concepts such as three-center two-electron bonding,[5] flat potential energy surfaces (PESs) in which expected intermediates are not minima,[6] anchimeric assistance,[7] post-transition state bifurcations (PTSB),[4] and non-statistical dynamic effects (including memory effects).[8] In most cases, it is theoretically possible for more than one group to migrate to a given carbocation center during rearrangement. The tendency of a group to migrate, its "migratory aptitude", depends both on the inherent properties of the group and on context. The latter is related to, for example, what type of carbocation will be formed upon migration (e.g., tertiary vs. secondary) and the orientation of the migrating group relative to the formally empty p-orbital on the carbocation (carbenium ion) center (e.g., hyperconjugated or not). Here we focus on the former, especially the dynamic behavior of the migrating group, in pursuit of a new scale of "migratory aptitude that accounts for dynamic effects".

Much work has been done in the field of carbocation migratory aptitudes over the past few decades, from both experimental and theoretical perspectives. Experimental efforts focused on measuring rates for various simple rearrangements and comparing yields between different migrations.[9] For instance, Vogel and co-workers demonstrated the high migratory tendency of acyl groups in multiple synthetically relevant scenarios.[10] Frontier and co-workers examined the balance between electronic effects and steric hindrance in their examination of

sequential Narazov electrocyclization/Wagner-Meerwein rearrangement reactions.[11] These are but two examples that showcase the relevance of the migratory aptitude concept to synthetic organic chemistry.

Early theoretical work utilized density functional theory (DFT) to locate transition state structures (TSSs) and compare their energies to predict migratory aptitudes.[12] More recently, molecular dynamics (MD) simulations by Riveros, Longo and co-workers were used to probe non-statistical dynamic effects on migratory aptitudes associated with the absence of minima corresponding to secondary carbocations.[4,6,13] These two examples highlight the two common types of approaches — static and dynamic — employed in computational studies on carbocation rearrangement reactions.

While it is well-known that various factors contribute to the magnitudes of migratory aptitudes, a common rule of thumb for estimating migratory aptitude is: groups that better stabilize a carbocation will migrate preferentially. This guideline captures, to some degree, the fact that positive charge generally builds up on a group during migration to a carbocation center. But there are notable exceptions to this rule of thumb. For example, in some cases less substituted CR_2 groups have been shown to migrate preferentially over more substituted CR_2 groups, even when significant differences between the stability of resulting carbocations are not apparent.[14] Also, H atoms can display a high migratory tendency.[11,15] Recent work also has shown that the inherent properties of the migrating group can be overwhelmed by the properties of the framework over which migration occurs.[15b] Here we revisit the concept of migratory aptitude in pursuit of a theoretical framework that provides both fundamental understanding and principles that are useful for synthesis design.

Inspired by the pinacol rearrangement,[1,16] which has served as a useful platform for evaluating migratory aptitudes in the past and for which previous simulations indicate the

importance of non-statistical dynamic effects,[17,18] we designed the model system shown in Scheme 2.1. In this system, both R^1 and R^2 are oriented similarly with respect to the leaving group and carbocation center formed upon its departure, so that conformational issues are avoided. Migration of either R^1 or R^2 will lead to a benzylic cation so that product stability issues are minimized. The group that migrates moves towards the non-migrating group, which should minimize differences in steric congestion as one or the other group migrates. In short, this system was designed to minimize context effects so that inherent migratory aptitudes can be exposed. In addition, the leaving group chosen here was water, rather than an anionic leaving group, to minimize charge separation during reaction and associated complications.[7] This system was also designed such that initial leaving group loss would lead to a secondary carbocation expected not to be a minimum on the PES,[4,6,13] which would lead to a PTSB,[19] i.e., there would be a single ambimodal[20] TSS followed by pathways downhill in energy to the products of migration of R^1 and R^2, thereby forcing migratory preferences to be controlled by non-statistical dynamic effects. While our results bear out the expectation that the secondary cation is not a minimum, separate TSSs for concerted (but asynchronous) water loss and migration were found for R^1 and R^2; this scenario complicated our analysis but also introduced an opportunity to examine additional aspects of migratory behavior.

We applied several computational approaches and reactivity models to characterize behavior during migration. Our main approach involved uphill *ab initio* molecular dynamics (AIMD) simulations,[4c,15j,21] in which additional momenta were added to reactants to promote water loss. The validity of such an approach is supported by comparisons with experimental product distributions (*vide infra*). The relative abundance of trajectories for migration of R^1 and R^2 provides a measure of migratory ability. We compared these predictions to predictions arrived at by comparing free energies of TSSs on PESs[22] and by comparing free energies of variational transition states (VTSs).[23] Intrinsic reaction coordinates (IRCs)[24] and models of

25

PESs, as well as the results of downhill AIMD simulations (initiated from VTSs) were also explored; these pointed to potential roles of recrossing and pathway bifurcations. We arrived at a ranking of migratory aptitudes and a model in which this ordering is rationalized based on dynamic effects. Ultimately, we tested our methods on a real-world synthetic example and these not only revealed non-statistical dynamic effects, but also accurately predicted the product ratio observed in the experiment (where other methods and models did not).[25]
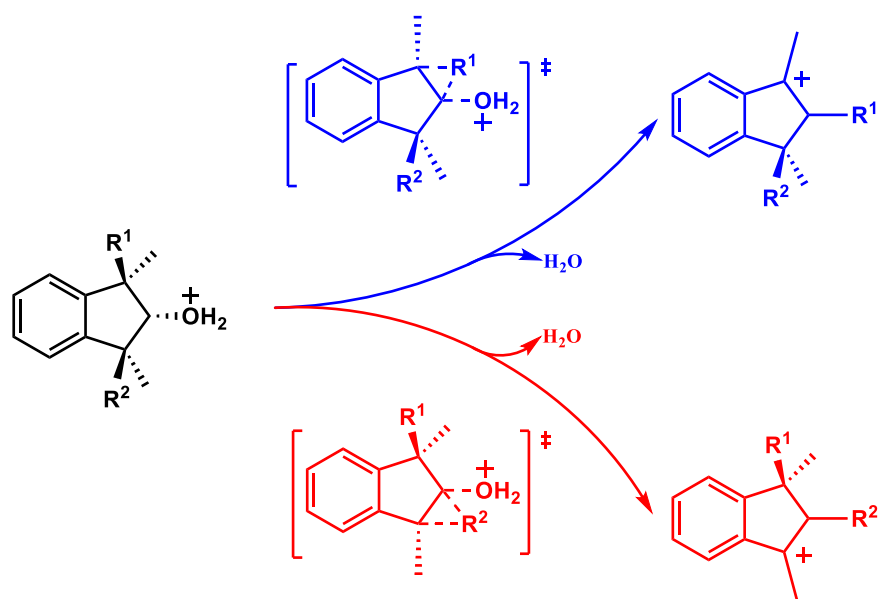
Scheme 2.1 Model reaction for comparing migratory aptitudes of R^1 and R^2.

2.2 Computational Methods

Geometry optimizations and frequency calculations were carried out with the *Gaussian16* C.01 package[26] at the M06-2X-D3/6-31G(d)[27] level of theory (see Supporting Information Table SA.1 for results of tests with more levels of theory). DLPNO-CCSD(T)[28] calculations were performed with ORCA 4.2.1.[29] PES minima and TSSs were identified by the number of imaginary frequencies obtained in frequency calculations, with 0 for minima and 1 for TSSs. VTSs and free energy reaction coordinates were calculated with *Gaussrate/Polyrate*.[29] Quasi-classical AIMD simulations from reactants and VTSs were performed using the *Progdyn* script

with a temperature of 298 K in the NVE ensemble.[31] The vibrational modes and forces were calculated quantum mechanically while the trajectories were propagated classically using the velocity Verlet algorithm with a timestep of 1 fs.[32] For uphill dynamics, breaking of the C–O bond was promoted by adding energy to the momentum vectors that have the direction of elongating the C–O bond, i.e., we only examine trajectories associated with water loss. Different amounts of excess energy were tested and compared with the experimental results. While the selectivity of migration varies (see Table SA.2 and section 3.6 in Appendix for details), the results indicate that the variation is not large (maximum deviation of less than 10%). Consequently, we used an excess energy of 25.0 ± 0.1 kcal/mol for all systems. The thresholds of C–C and C–H bond formation used to terminate trajectories were 1.6 Å and 1.1 Å, respectively. Due to the rigidity of the molecule and the distance between the migrating groups, the influence of conformational flexibility on uphill dynamic simulations is expected to be small and was not examined exhaustively (see Appendix for details). *CYLview* was used for visualization of molecular geometries.[33]

2.3 Results & Discussion

2.3.1 Shape of PES

To provide context for the results of our dynamics simulations, a 2D contour plot of the PES for the secondary carbocation formed for the R^1 = acyl, R^2 = *iso*-propyl system was calculated (Figure 2.1). The PES was derived by scanning the C^1–R^1 and C^1–R^2 bond lengths while relaxing all other degrees of freedom. The void space at the bottom left of the PES results from the two migrating groups being too close to each other and associated SCF convergence failures. Some regions of discontinuity in energy are also likely the result of the use of only two degrees of freedom. Despite these complications associated with reduced-dimension

(much less than 3N-6 dimensions corresponding to all degrees of freedom) PESs,[34] important features are revealed in this plot. The three valleys correspond to the reactant (top right) and two products (top left and bottom right). Points corresponding to the bond lengths in both TSSs (and VTSs, *vide infra*) are labeled. Again, the reduced dimensionality of this PES is partly responsible for these TSSs not residing at points that look like simple first-order saddle points. Uphill trajectories also have been projected onto the PES (curved lines), and these will be discussed below. One noteworthy feature of this PES is that the reactant minimum is surrounded by roughly oval contours that connect the pathways to each of the two products, i.e., the two competing pathways are not separated by a high energy feature in the region between transition states. As a result, leakage between these two pathways is possible, i.e., trajectories initially progressing towards one product may end up in the other by traversing this relatively flat region of the PES. Such a feature implies that the behavior after passing through the transition state regions cannot be neglected, a scenario for which transition state theory fails to account.[23a] This scenario mandates caution in attempting to make selectivity predictions based on energies of transition states alone.
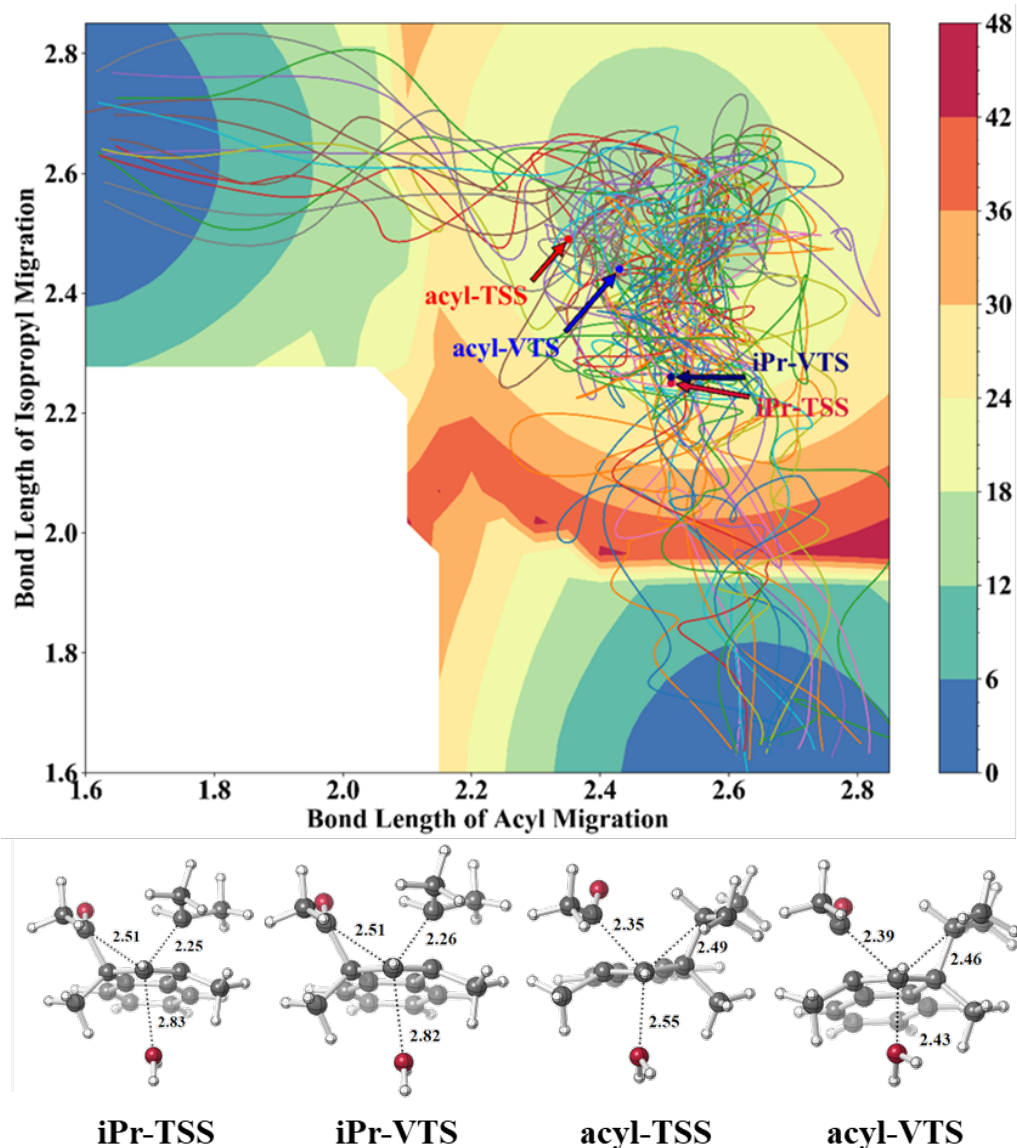
Figure 2.1 PES with representative trajectories for the R^1 = acyl, R^2 = *i*-Pr system at the M06-2X-D3/6-31G(d) level projected onto it. The axes correspond to the length (Å) of one or the other C–C bond formed upon migration of R^1 or R^2. The valley at the top right corresponds to the reactant region and the valleys at the top left and bottom right correspond to the two product regions. Points corresponding to the bond lengths in both TSSs and VTSs are labeled and the geometries of these structures are shown below the plot.

2.3.2 Uphill Dynamics

Uphill AIMD trajectories were initiated from the reactant for systems with a variety of R^1 and R^2 groups (reasonably representative of those of interest to synthetic chemists; Table 2.1). The percentage of trajectories leading to each product, along with barriers for formation of each product based on free energies of TSSs (and VTSs for selected systems; see below) are shown in Table 2.1. Since each system has a built-in competition between two migrating groups, we arrive at an overall, qualitative, ordering of migratory abilities, assuming that the migratory aptitude is transitive. The AIMD results indicate the following order of migratory aptitude: cyclobutyl > vinyl ≈ Ph > CH_2-cyclopropyl ≈ cyclopropyl ≈ *iso*-propyl ≈ acyl > allyl ≈ ethyl > methyl ≈ H > methoxymethyl. While a similar ordering is arrived at using free energies of TSSs, the magnitudes of some preferences vary significantly in AIMD-based and transition state free energy-based predictions, especially with H or acyl groups as migrators.

Table 2.1 Results of uphill dynamic simulations, conventional and variational free energy barriers and predicted product ratios (M06-2X-D3/6-31G*), energies in kcal/mol).

	R^1			R^2			trajectory-based product ratio	TSS-based product ratio	VTS-based product ratio	Total trajectories
	%	$\Delta G^{\ddagger}_{TS}$	$\Delta G^{\ddagger}_{VTS}$	%	$\Delta G^{\ddagger}_{TS}$	$\Delta G^{\ddagger}_{VTS}$				
1	H—R 32%	9.5	9.7	H₃C—R 45%	9.7	9.8	42:58	58:42	54:46	88
2	H—R 44%	9.4		[structure] R 55%	9.1		44:56	38:62		102
3	H—R 29%	9.2	9.3	[structure]—R 69%	6.6	7.0	30:70	1:99	2:98	116
4	H—R 26%	8.8	9.1	[structure] R 53%	4.8	4.9	27:73	0:100	0:100	204
5	H—R 9%	9.0	9.1	Ph—R 83%	4.6	4.8	10:90	0:100	0:100	82
6	H₃C—R 79%	9.6	10.2	[structure] 13%	9.7	10.3	86:14	54:46	54:46	176
7	H₃C—R 27%	8.3		[structure] R 54%	7.8		33:67	29:71		77
8	[structure] R 42%	7.5		[structure] R 36%	7.8		55:45	63:37		123
9	[structure] R 75%	7.8		H₃C—R 11%	8.3		87:13	71:29		119
10	[structure] R 49%	7.0		[structure] R 36%	7.1		58:42	54:46		105
11	[structure] R 59%	5.1	6.7	[structure] R 29%	7.9	7.9	67:33	99:1	89:11	76
12	[structure] R 50%	6.4	7.0	[structure]—R 47%	7.1	7.4	53:46	77:23	66:34	72
13	[structure] R 50%	6.2	6.2	[structure] R 50%	5.2	5.5	50:50	16:84	24:76	137
14	[structure] R 39%	6.5	6.5	[structure] R 58%	4.7	5.6	40:60	4:96	17:83	128
15	H₃C—R 26%	7.0	7.0	[structure] R 74%	5.1	5.5	26:74	2:98	8:92	100
16	H₃C—R 24%	7.4		[structure] R 69%	6.5		26:74	19:81		114
17	[structure] R 33%	8.2	8.3	[structure] R 58%	6.5	6.8	36:64	5:95	8:92	143
18	[structure] R 33%	7.4	7.5	[structure] R 50%	5.6	5.8	40:60	4:96	6:94	76
19	[structure]—R 34%	6.3		[structure]—R 64%	6.4		35:65	54:46		117

20	(structure) R	33%	6.7	(structure) —R	59%	6.7	36:64	50:50	123
21	(structure) R	48%	6.7	(structure) —R	52%	7.0	48:52	62:38	122
22	(structure) R	41%	4.5	(structure) —R	47%	5.7	45:55	88:12	116
23	(structure) R	33%	4.5	(structure) —R	67%	3.8	33:67	24:76	213

2.3.3 Variational Transition States

To make sure that differences between predictions based on AIMD and transition state energies were not the result of neglecting variational effects,[23a] canonical VTSs were located for a subset of systems. Compared with predicted product distributions based on conventional TSS free energies, some of those based on VTS free energies are closer to the results of uphill AIMD simulations (Table 2.1). Still, the VTS-based and AIMD-based predictions often differ substantially. Free energy-based reaction pathways were also computed. Representative pathways, for the R^1 = methyl, R^2 = methoxymethyl system, are shown in Figure 2.2. The black dashed line at the center indicates where the conventional TSSs are located; the VTSs are earlier than the TSSs, as expected for a fragmentation reaction based on assumptions about entropy. For this system, the geometries of the two VTSs are similar to each other, but skewed slightly toward the geometries of reactants to which they are connected. Several new intermediates (here, resulting from inclusion of ZPE; see Appendix for details),[35] are revealed along these pathways as well, although these are very shallow (barriers < 0.5 kcal/mol). While some differences between TSSs and VTSs and the pathways connected to them are present, these are not large enough to rationalize the differences in predicted product distributions derived from transition state energies and uphill AIMD simulations.
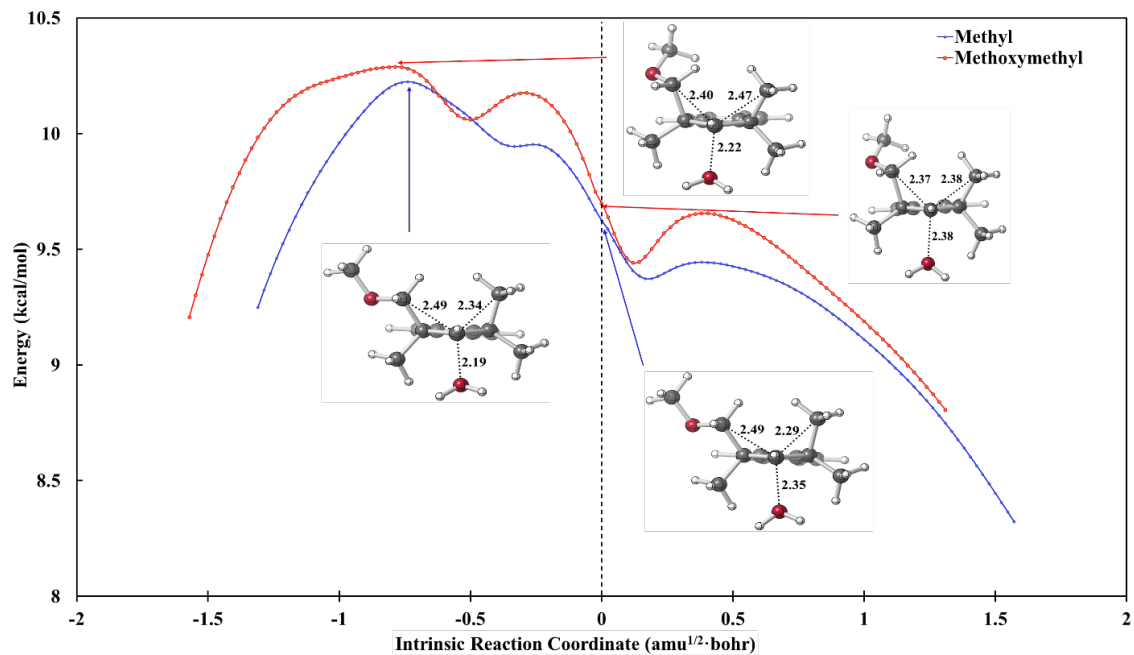
Figure 2.2 Free energy profiles for the R^1 = methyl, R^2 = methoxymethyl system. Geometries for the conventional TSSs correspond to the '0' point on the horizontal axis. Positions and geometries of VTSs are indicated. In this plot, reactants are on the left-hand side and products are on the right-hand side. See SI for plots with ZPE included but entropy excluded.
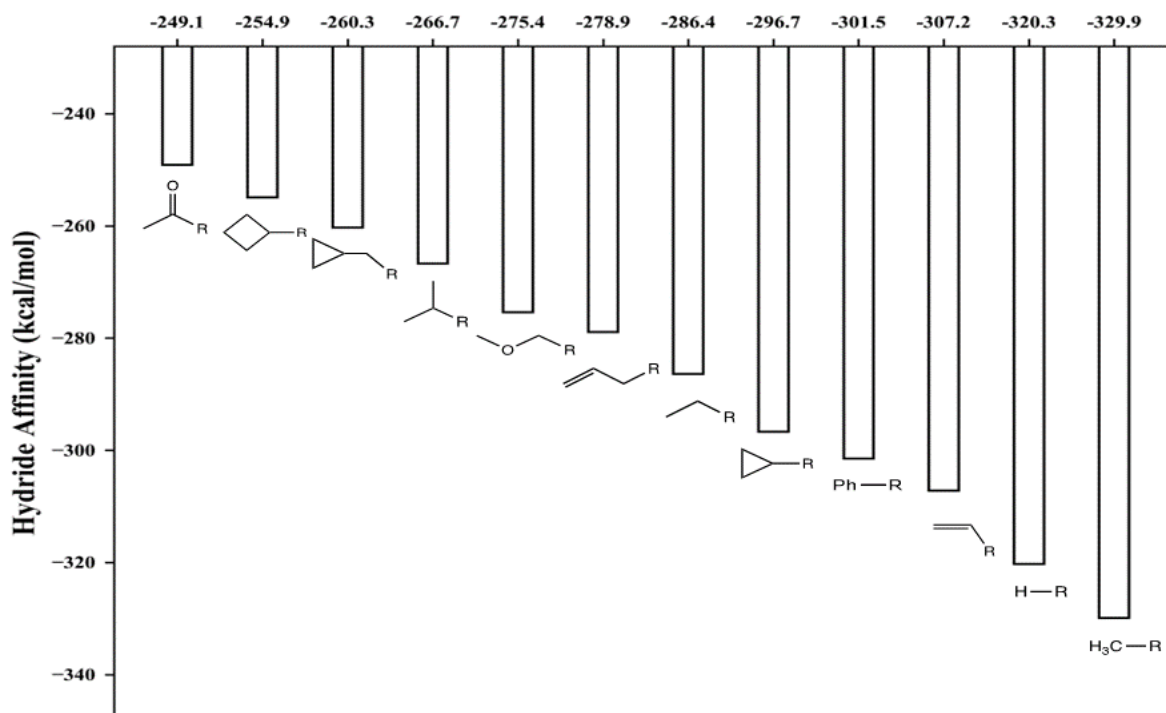
Figure 2.3 Hydride affinity of selected functional groups. The calculation is performed at M06-2X-D3/def2-TZVPP level of theory. The more negative the enthalpy is, the more thermodynamically unstable the corresponding cation is.

2.3.4 Hydride Affinities of Carbocations

Our predicted ordering of migratory aptitudes does not bear out the rule of thumb that better carbocation stabilizers migrate preferentially. One simple approach for quantifying carbocation stabilizing ability is to compare hydride affinities (HA) of cationic forms of the migrating groups—a reasonable approach in that migrating groups generally accumulate positive charge during migration (i.e., TSSs often resemble cations complexed to alkenes).[36] For HAs we use here enthalpies for the reaction of a cation R^+ with H^- to form neutral R–H. Using these values (Figure 2.3), we arrive at the following order of migratory aptitudes: acyl > cyclobutyl > CH_2-cyclopropyl > *iso*-propyl > methoxymethyl > allyl > ethyl > cyclopropyl > phenyl > vinyl > methyl. Some significant deviations are observed when comparing this ordering to those predicted based on uphill AIMD or transition state energies. For instance,

although acyl cation is the most thermodynamically stable cation among all the functional groups tested, it is predicted to be a worse migrator compared with cyclobutyl (Table 2.1). In addition, the migratory tendencies of the phenyl and cyclobutyl groups are greatly underestimated compared to our dynamics-based predictions (Table 2.1); while one might imagine that this is a result of stepwise pathways involving phenonium ions or ring-expansion, we find only concerted pathways for our systems. The discrepancies between hydride affinity, statistical kinetic models and dynamic simulations demonstrate the complexity of accurately predicting migratory aptitudes.

2.3.5 Downhill Dynamics and Bifurcations

In pursuit of additional insights into the dynamic behavior of our systems, we carried out downhill AIMD simulations for three systems initiated from VTSs (Table 2.2). Our results reveal several important behaviors. First, a small portion (~5%) of trajectories can access both products, indicating that the two transition states are not completely separated, consistent with the discussion of 2D PESs and similarities of VTS structures above. In other words, either of the two transition states can be viewed as an ambimodal transition state, although well-defined valley-ridge inflection points are here replaced by a flat portion of the PES, which allows the minor product to be accessed from the region of the VTS for formation of the major product and vice versa (this cross-over behavior also can be seen in the uphill trajectories mapped onto the surface in Figure 2.1). The failure of transition state theory to make meaningful predictions when two pathways share one dividing surface is known.[37]

Table 2.2 Downhill AIMD results with trajectories initiated from VTSs.

TS	R^1	R^2	Recrossing	VTS energy	R^1/R^2 ratio (VTS)[a]	R^1/R^2 ratio (recrossing)[b]	R^1/R^2 ratio (uphill)	Total trajectories

H	47%	5%	48%	9.3				233
					2:98	7:93	30:70	
cyclopropyl	2%	37%	61%	7.0				321
methyl	63%	3%	34%	10.2				108
					54:46	54:46	86:14	
methoxymethyl	5%	64%	31%	10.3				121
vinyl	84%	2%	14%	6.8				161
					8:92	7:93	36:64	
ethyl	47%	0%	53%	8.3				110

a. The R^1/R^2 ratio (VTS) is directly calculated from the Gibbs free energy difference ($\Delta\Delta G$) of the two VTSs.

b. Recrossing corrections are incorporated into the R^1/R^2 ratio (recrossing) based on the downhill dynamic results, where it is calculated as $\frac{R^1_{VTS1}P_{VTS1}+R^1_{VTS2}P_{VTS2}}{R^2_{VTS1}P_{VTS1}+R^2_{VTS2}P_{VTS2}}$. The subscripts VTS1 and VTS2 indicate the ratio of product 1 or 2 (indicated in superscripts) in the downhill dynamics initiated from the corresponding VTSs. The P_{VTS} indicates the ratio of each VTS based on Boltzmann distribution.

Second, we find a large amount of recrossing, even when initiating trajectories at VTSs. Although transition state theory occupies a central position in the realm of chemical kinetics because it is applicable to most chemical reactions,[23] it is subject to local-equilibrium and non-recrossing assumptions that may become invalid in particular cases.[35,37,39] The non-recrossing assumption states that a trajectory initiated from the reactant region will pass through the dividing surface and advance towards the product region without crossing back through the dividing surface, but deviations from this behavior have been reported. For example, Singleton and co-workers have shown how recrossing can enhance selectivity.[39] Singleton and co-workers also have pointed out that extensive recrossing can serve as a probe for hidden dynamical bottlenecks.[35] Variational transition state theory allows for some recrossing, but a VTS should be associated with minimal recrossing.[38] For some of the cases in Table 2.2, a large amount of recrossing is observed, up to 61%, despite initiating trajectories at VTSs. Such a

high proportion of recrossing trajectories is consistent with a fairly flat surface near the VTS.[8b] The recrossing behavior also is captured in uphill dynamics simulations.

These cross-over and recrossing behaviors are manifested in changes to distances between migrating groups and the carbon that bore the water leaving group during trajectories. In Figure 2.4 we plot these distances for representative trajectories for a representative reaction ($R^1 =$ cyclopropyl, $R^2 = H$); the horizontal line across each plot indicates the corresponding C–H distance in the H-migration VTS. In the top panel of Figure 2.4 are shown trajectories that ultimately form the H-migration product, and it is clear that the C–H distance sometimes dips below and then rises well above the C–H distance in the VTS before dropping to the distance in the product. In addition, the bottom panel of Figure 2.4, which shows cyclopropyl-migration trajectories, indicates that some of the trajectories sample C–H distances similar to or shorter than those in the region of the VTS for H-migration.

Accounting for cross-over and recrossing observed in downhill AIMD simulations, however, does not change the selectivity predicted using VTS very much. Since we suspect there is a flat surface connecting the two VTSs (in analogy to Figure 2.1), we expect the product ratio to be sensitive to the position of the starting point for downhill dynamics trajectories, i.e., slightly different C–C or C–H bond distances may lead to drastically different results. To verify this hypothesis, downhill dynamics simulations for **3** ($R^1 =$ cyclopropyl, $R^2 = H$) were initiated at positions different from the two VTSs (Figure 2.5). Even though deviations in initial geometries were not large, deviations to product ratios were observed. Specifically, more minor product and, in some cases, less recrossing is observed. Thus, it appears that trajectories can roam around the flat inter-transition state region before committing to formation of a particular product (see also Figure 2.1). Simple VTSs are unable to fully capture the topography of the energy surface.

37
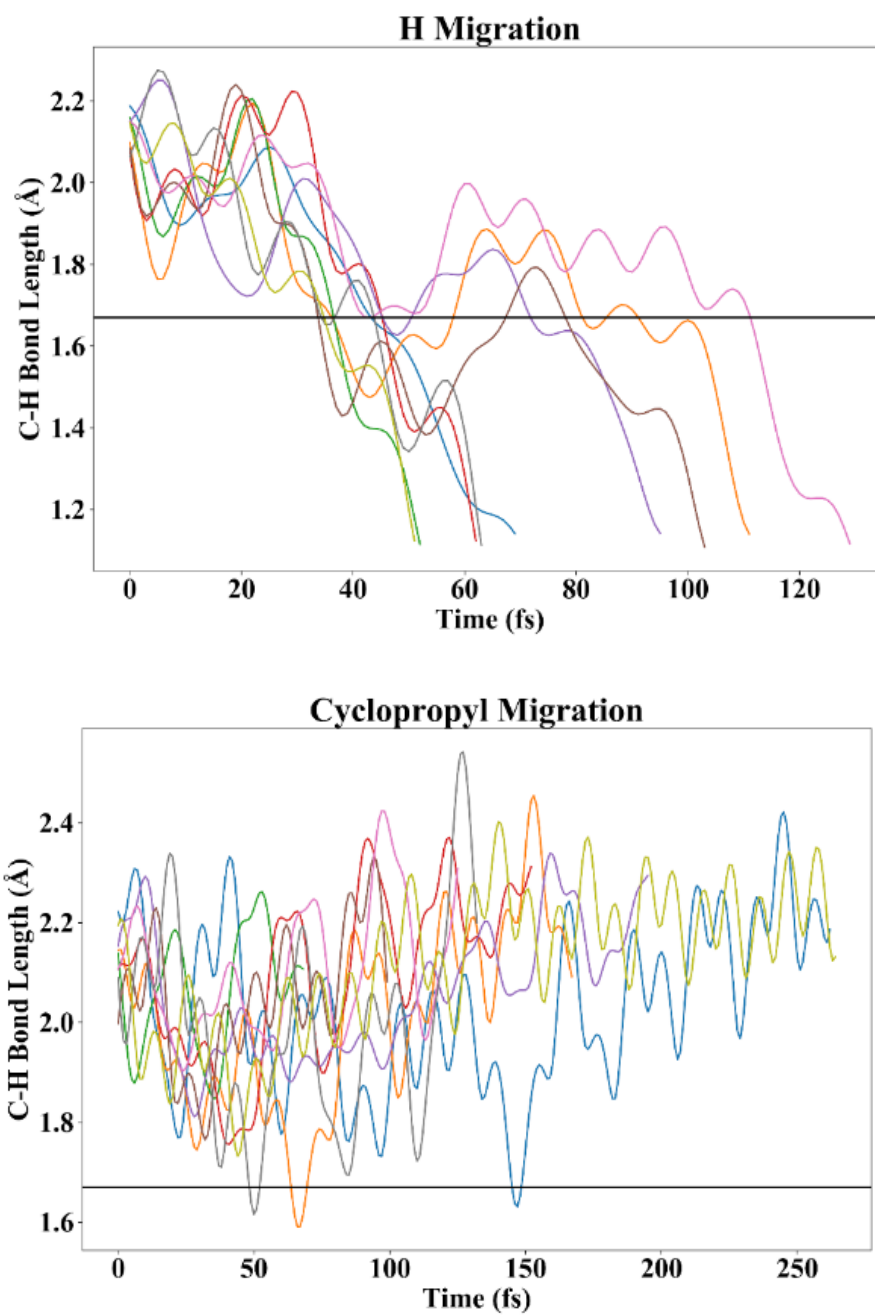
Figure 2.4 Variations in C–H distances over time for the R^1 = cyclopropyl R^2 = H system. The top panel shows a subset of trajectories that lead to H-migration and the bottom panel shows a subset of trajectories that lead to vinyl-migration. The horizontal black line indicates the C–H distance in the VTS for H-migration.
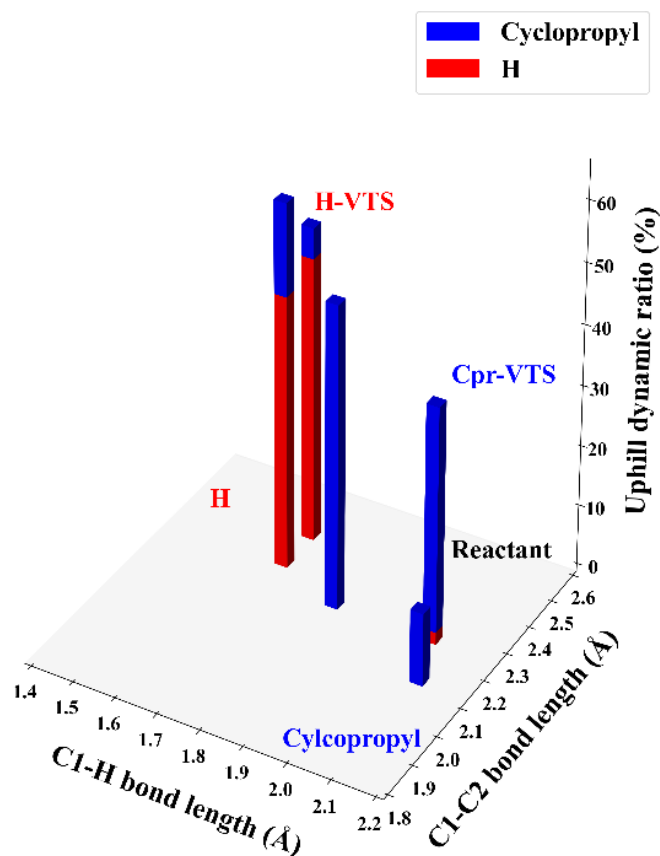
Figure 2.5 Downhill AIMD simulation results for system **3** (R^1 = cyclopropyl R^2 = H) initiated from different starting points. The colored bars indicate the ratios of products. Blue corresponds to cyclopropyl migration and red corresponds to hydrogen migration. The total height of each bar indicates the total ratio of productive trajectories versus recrossing trajectories. The regions of reactant, cyclopropyl and H migration products are also labeled. All bond lengths are in Å.

2.3.6 Synthetic Relevance

To probe how relevant the approach described above is to actual synthetic and biosynthetic reactions, we investigated a carbocation rearrangement used in Coates and co-workers' biomimetic synthesis of nopsane sesquiterpenes.[25] The synthetic approach involved Wagner-Meerwein rearrangement of a putative secondary carbocation and competition between methyl migration and ring expansion (Scheme 2.2; secondary carbocation not drawn). The secondary carbocation itself is not a minimum on the PES, and both migrations have their own TSSs –

this scenario matches that in our model systems. Here we assume that the experimentally determined product ratio reflects kinetic selectivity for formation of carbocations **24-A** and **24-B**. While we cannot definitively rule out thermodynamic equilibration, in that predicted reverse barriers are only ~12 kcal/mol, we note that the experimentally observed product ratio is not in line with the computed free energy difference for **24-A** and **24-B** and product formation in this case involves trapping by solvent, a process whose rate is difficult to predict accurately.

Firstly, we tested whether predictions using conventional transition state theory match the experimental selectivity. Several levels of theory were examined (see Appendix for details) and the highest, DLPNO-CCSD(T)/def2-QZVPP//M06-2X-D3/def2-TZVP),[40] predicts a Gibbs free energy difference between competing TSSs of 3.3 kcal/mol. This difference implies that no methyl migration product should be observed. However, some methyl migration is observed experimentally, with the observed ratio of products (88:12) corresponding (if transition state theory is employed) to a $\Delta\Delta G^{\ddagger}$ of 1.1 kcal/mol. The discrepancy between theory and experiment is not resolved by employing VTSs (negligible corrections; < 0.2 kcal/mol) or by incorporating implicit solvent ($\Delta\Delta G$ > 2.0 kcal/mol, see Appendix for details).

A : B Ratio
Conventional TST: **100:0**
Variational TST: **100:0**
Downhill dynamics from VTSs : **100**:**0**
Uphill dynamics: **93**:**7**
Uphill dynamics (additional momenta): **89**:**11**
Uphill dynamics (SMD solvent model): **89**:**11**
Experiment: **88**:**12**

Uphill Dynamics: **81 : 19**
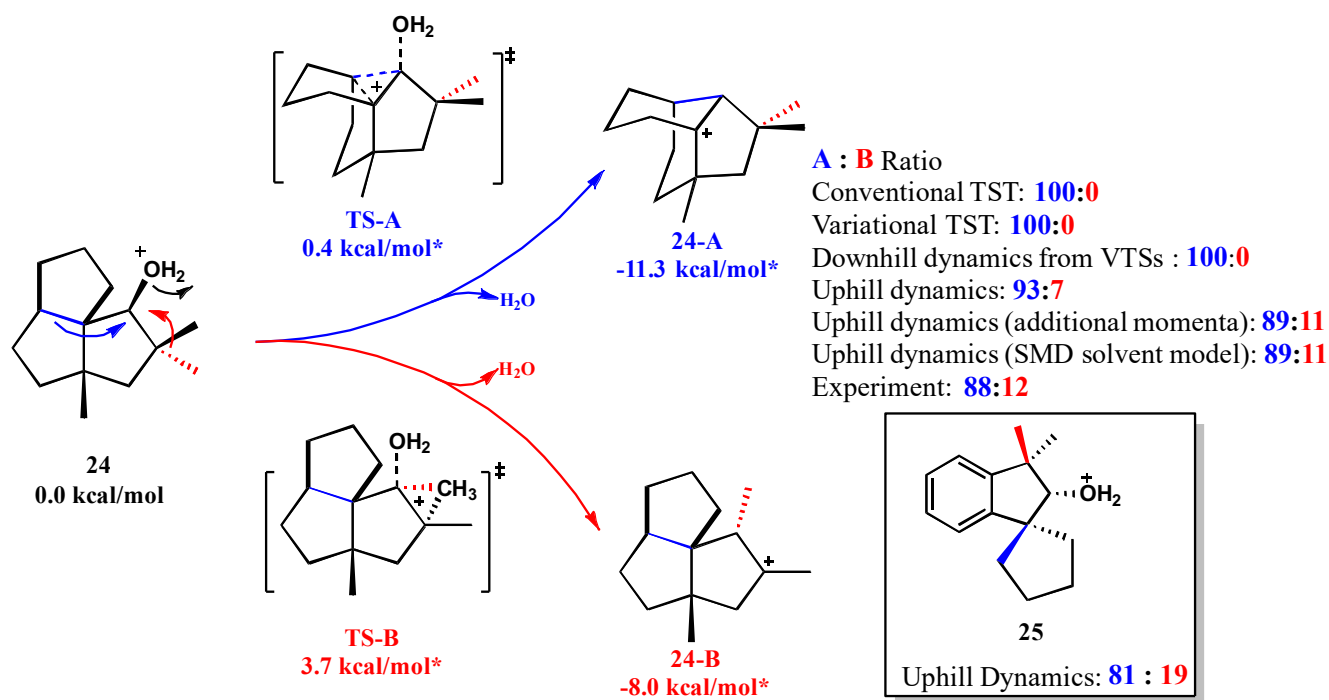
*Free energies at DLPNO-CCSD(T)/def2-QZVPP//M06-2X-D3/def2-TZVP level of theory using tight-PNO.

Scheme 2.2 Carbocation rearrangement in the synthesis of nopsane sesquiterpene.

Might non-statistical dynamic effects be at play here? To investigate this possibility, downhill AIMD simulations were initiated from the ring-expansion VTS. In the 152 trajectories obtained, no methyl migration was observed. Clearly the presence of a significant amount of the minor product cannot be rationalized by only taking into account non-statistical effects associated with downhill dynamics from the VTSs.

If a flat TS region of the PES spans the two competing transition states as described above, uphill trajectories will not necessarily come near to a VTS. Thus, we again carried out uphill AIMD simulations, here both with and without additional momenta associated with leaving group departure. Additional kinetic energy ranging from 0 to 30 kcal/mol was included in uphill dynamics simulations and a predicted ratio of ~90:10 ratio was found for all the energies (Scheme 2.2; 365 trajectories with additional momenta; 79 trajectories without additional

momenta), consistent with the experimentally observed ratio of 88:12. Additionally, we examined the effect of implicit solvent. Uphill dynamics simulations (118 trajectories) using the SMD model[41] led to a predicted product ratio of 89:11, consistent with the gas phase prediction (see Appendix for additional details). This result implies that solvent does not influence the dynamic effects observed for this system, although this contention should be tested with explicit solvent (a goal for our future work).

Could we have predicted this outcome using our model systems alone? Yes. As shown at the bottom right of Scheme 2.2, competing a cyclopentane ring expansion against a methyl shift in our framework leads to a predicted ratio from uphill AIMD simulations of 81:19, a less accurate but reasonable match.

In summary, the presence of the minor product observed in experiment was not predicted by comparing TSS or VTS free energies or by carrying out downhill AIMD simulations from either. Only uphill AIMD simulations – which allow consequences of a flat inter-transition state region of the PES to be expressed, led to reasonable predictions.

2.4 Conclusions

We set out to reexamine the concept of migratory aptitude of organic groups from a perspective that includes dynamic effects. While we present here a ranking of migratory aptitudes based on uphill dynamics simulations on model systems designed to capture inherent preferences, the behavior of these model systems turned out to be more complex than expected. This complexity arises in large part from using a molecular framework that allows two migrations to compete with each other. However, this complexity has allowed us to highlight several unusual and/or oft-overlooked behaviors that can influence selectivity. These include momentum effects associated with low barriers and recrossing and inter-transition state roaming associated with the sharing of dividing surfaces by competing transition states.[18,42]

42

While not all experimental systems will have competing transitions states with shared dividing surfaces, it is likely that many will, given the metastability of many carbocations. The examples described here highlight the relevance of such concepts to organic reactions of synthetic and biosynthetic relevance.[14,21a,43]

2.5 Acknowledgement

2.6 References

1. Fittig, R. Ueber einige Derivate des Acetons. *Justus Liebigs Ann. Chem*. **1860**, *114*, 54–63.

2. Tantillo, D. J. Biosynthesis via carbocations: Theoretical studies on terpene formation. *Nat. Prod. Rep.* **2011**, *28*, 1035-1053.

3. (a) Naredla, R. R.; Klumpp, D. A. Contemporary Carbocation Chemistry: Applications in Organic Synthesis. *Chem. Rev.* **2013**, *113*, 6905-6948 (b) Song, Z.; Fan, C.; Tu, Y. Semipinacol Rearrangement in Natural Product Synthesis. *Chem. Rev*. **2011**, *111*, 7523-7556.

4. (a) Tantillo, D. J. The carbocation continuum in terpene biosynthesis—where are the secondary cations? *Chem. Soc. Rev*. **2010**, *39*, 2847-2854. (b) Hong, Y. J.; Tantillo, D. J. A Maze of Dyotropic Rearrangements and Triple Shifts: Carbocation Rearrangements Connecting Stemarene, Stemodene, Betaerdene, Aphidicolene, and Scopadulanol. *J. Org. Chem*. **2018**, *83*, 3780-3793. (c) Hare, S. R.; Tantillo, D. J. Dynamic behavior of

rearranging carbocations–implications for terpene biosynthesis. *Beilstein J. Org. Chem.* **2016**, *12*, 377-390. (d) Bai, M.; Feng, Z.; Li, J.; Tantillo, D. J. Bouncing off walls–widths of exit channels from shallow minima can dominate selectivity control. *Chem. Sci.* **2020**, *11*, 9937-9944.

5. (a) Grob, C. A. Inductivity and bridging in carbocations. *Acc. Chem. Res.* **1983**, *16*, 426-431. (b) Brown, H. C. The energy of the transition states and the intermediate cation in the ionization of 2-norbornyl derivatives. Where is the nonclassical stabilization energy? *Acc. Chem. Res.* **1983**, *16*, 432-440. (c) Olah, G. A.; Prakash, G. K. S.; Saunders, M. Conclusion of the Classical-Nonclassical Ion Controversy Based on the Structural Study of the 2-Norbornyl Cation. *Acc. Chem. Res.* **1983**, *16*, 440-448. (d) Walling, C. An innocent bystander looks at the 2-norbornyl cation. *Acc. Chem. Res.* **1983**, *16*, 448-454. (e) Brown, H. C. (with comments by P. v. R. Schleyer) *The Nonclassical Ion Problem*; Plenum: New York, 1977.

6. (a) Tantillo, D. J. Importance of Inherent Substrate Reactivity in Enzyme-Promoted Carbocation Cyclization/Rearrangements. *Angew. Chem. Int. Ed.* **2017**, *56*, 10040-10045. (b) Ammal, S. C.; Yamataka, H.; Aida, M.; Dupius, M. Dynamics-driven reaction pathway in an intramolecular rearrangement. *Science* **2003**, *299*, 1555-1557. (c) Zhao, C. G.; Feng, Z. T.; Xu, G. Q.; Gao, A.; Chen, J. W.; Wang, Z. Y.; Xu, P. F. Highly Enantioselective Construction of Strained Spiro[2,3]hexanes through a Michael Addition/Ring Expansion/Cyclization Cascade. *Angew. Chem. Int. Ed.* **2020**, *59*, 3058-3062. (d) Carpenter, B. K. Dynamic matching: The cause of inversion of configuration in the [1, 3] sigmatropic migration? *J. Am. Chem. Soc.* **1995**, *117*, 6336-6344.

7. Schreiner, P. R.; Schleyer, P. V.; Schaefer, H. F. Why the Classical and Nonclassical Norbornyl Cations Do Not Resemble the 2-endo- and 2-exo-Norbornyl Solvolysis Transition States. *J. Org. Chem.* **1997**, *62*, 4216-4228.

8. (a) Ess, D. H.; Wheeler, S. E.; Iafe, R. G.; Xu, L.; Çelebi-Ö lçüm, N.; Houk, K. N. Bifurcations on potential energy surfaces of organic reactions. *Angew. Chem., Int. Ed.* **2008**, *47*, 7592−7601. (b) Reis, M. C.; López, C. S.; Faza, O. N.; Tantillo, D. J. Pushing the limits of concertedness. A waltz of wandering carbocations. *Chem. Sci.* **2019**, *10*, 2159-2170. (c) Ghigo, G; Maranzana, A; Tonachini, G. Memory Effects in Carbocation Rearrangements: Structural and Dynamic Study of the Norborn-2-en-7-ylmethyl-X Solvolysis Case. *J. Org. Chem.* **2013**, *78*, 9041-9050.

9. House, H. O.; Grubbs, E. J. The Semipinacolic Deamination of Certain 1-Alkyl-2-amino-1-phenylethanols. *J. Am. Chem. Soc.* **1959**, *81*, 4733–4737.

10. Le Drian, C.; Vogel, P. Facile. Facile migration of acyl groups in Wagner-Meerwein transpositions. Acid-catalyzed rearrangements of 5,6-exo-epoxy-7-oxa-2-bicyclo[2.2.1]heptanone and derivatives. *Tetrahedron Lett.* **1987**, *28*, 1523–1525.

11. Huang, J.; Lebœuf, D.; Frontier, A. J. Understanding the Fate of the Oxyallyl Cation following Nazarov Electrocyclization: Sequential Wagner−Meerwein Migrations and the Synthesis of Spirocyclic Cyclopentenones. *J. Am. Chem. Soc.* **2011**, *133*, 6307–6317.

12. Nakamura, K.; Osamura, Y. Theoretical study of the reaction mechanism and migratory aptitude of the pinacol rearrangement. *J. Am. Chem. Soc.* **1993**, *115*, 9112-9120.

13. de Souza, M. A. F.; Ventura, E.; Monte, S. A. do; Riveros, J. M.; Longo, R. L. Dynamic Effects Dictate the Mechanism and Selectivity of Dehydration–Rearrangement

Reactions of Protonated Alcohols [Me2(R)CCH(OH2)Me]+ (R=Me, Et, iPr) in the Gas Phase. *Chem. Eur. J.* **2014**, *20*, 13742–13754.

14.   Blumel, M.; Nagasawa, S.; Blackford, K.; Hare, S. R.; Tantillo, D. J.; Sarpong, R. Rearrangement of hydroxylated pinene derivatives to fenchone-type frameworks: computational evidence for dynamically-controlled selectivity. *J. Am. Chem. Soc*. **2018**, *140*, 9291-9298.

15.   (a) Wistuba, E.; Rüchardt, C. Intrinsic migration aptitudes of alkyl groups in a pinacol rearrangement. *Tetrahedron Lett.* **1981**, *22*, 4069–4072. (b)Stopka, T.; Schröder, S.; Maulide, N.; Niggemann, M. The Unusual Migratory Aptitude in a Case of α-Carbonyl Cation-Driven 1,2-Migration. *Tetrahedron,* **2020**, *76*, 131460.

16.   (a) Liang, T.; Zhang, Z.; Antilla, J. C. Chiral Brønsted acid catalyzed pinacol rearrangement. *Angew. Chem. Int. Ed*. **2010**, *49*, 9734-9736. (b) Snape, T. J. Recent advances in the semi-pinacol rearrangement of α-hydroxy epoxides and related compounds. *Chem. Soc. Rev.* **2006**, *36*, 1823-1842.

17.   (a) Hare, S. R.; Pemberton, R. P.; Tantillo, D. J. Navigating past a fork in the road: carbocation− π interactions can manipulate dynamic behavior of reactions facing post-transition-state bifurcations. *J. Am. Chem. Soc*. **2017**, *139*, 7485-7493. (b) Hare, S. R.; Tantillo, D. J. Post-transition state bifurcations gain momentum–current state of the field. *Pure Appl. Chem.* **2017**, *89*, 679-698. (c) Hare, S. R.; Tantillo, D. J. Post-transition state bifurcations induce dynamical detours in Pummerer-like reactions. *Chem. Sci.* **2017**, *8*, 1442-1449. (d) Siebert, M. R.; Manikandan, P.; Sun, R.; Tantillo, D. J.; Hase, W. L. Gas-phase chemical dynamics simulations on the bifurcating pathway of the pimaradienyl cation rearrangement: role of enzymatic steering in abietic acid biosynthesis. *J. Chem. Theory Comput.* **2012**, *8*, 1212−1222. (e) Pemberton, R. P.;

Hong, Y. J.; Tantillo, D. J. Inherent dynamical preferences in carbocation rearrangements leading to terpene natural products. *Pure Appl. Chem.* **2013**, *85*, 1949−1957. (f) Hong, Y. J.; Tantillo, D. J. Biosynthetic consequences of multiple sequential post-transition-state bifurcations. *Nat. Chem.* **2014**, *6*, 104−111. (g) Weitman, M.; Major, D. T. Challenges posed to bornyl diphosphate synthase: diverging reaction mechanisms in monoterpenes. *J. Am. Chem. Soc.* **2010**, *132*, 6349− 6360. (h) Lourderaj, U.; Park, K.; Hase, W. L. Classical trajectory simulations of post-transition state dynamics. *Int. Rev. Phys. Chem.* **2008**, *27*, 361–403.

18.   (a) Ma, X.; Hase, W. L. Perspective: chemical dynamics simulations of non-statistical reaction dynamics. *Phil. Trans. R. Soc. A* **2017**, *375*, 20160204 (b) Jayee, B.; Hase, W. L. Nonstatistical Reaction Dynamics. *Annu. Rev. Phys. Chem.* **2020**, *71*, 289-313. (c) Carpenter, B. K. Energy disposition in reactive intermediates. *Chem. Rev.* **2013**, *113*, 7265-7286. (d) Barnes, G. L.; Hase, W. L. Direct chemical dynamics simulations. *J. Am. Chem. Soc.* **2017**, *139*, 3570-3590.

19.   Campos, R. B.; Tantillo, D. J. Designing Reactions with Post-Transition-State Bifurcations: Asynchronous Nitrene Insertions into C–C σ Bonds. *Chem* **2019**, *5*, 227-236.

20.   (a) Pham, H. V.; Houk, K. N. Diels–Alder Reactions of Allene with Benzene and Butadiene: Concerted, Stepwise, and Ambimodal Transition States. *J. Org. Chem.* **2014**, *79*, 8968-8976. (b) Yu, P.; Patel, A.; Houk, K. N. Transannular [6+ 4] and ambimodal cycloaddition in the biosynthesis of heronamide A. *J. Am. Chem. Soc.* **2015**, *137*, 13518-13523.

21.   (a) Singleton, D. A.; Hang, C.; Szymanski, M. J.; Greenwald, E. E. A new form of kinetic isotope effect. dynamic effects on isotopic selectivity and regioselectivity. *J.*

Am. Chem. Soc. **2003**, *125*, 1176–1177. (b) Carpenter, B. K. Nonstatistical dynamics in thermal reactions of polyatomic molecules. *Annu. Rev. Phys. Chem.* **2005**, *56*, 57–89. (c) Lourderaj, U.; Hase, W. L. Theoretical and computational studies of non-RRKM unimolecular dynamics. *J. Phys. Chem. A*, **2009**, *113*, 2236–2253.

22. Peng, Q.; Paton, R. S. Catalytic control in cyclizations: From computational mechanistic understanding to selectivity prediction. *Acc. Chem. Res.* **2016**, *49*, 1042-1051.

23. (a) Bao, J. L.; Truhlar, D. G. Variational transition state theory: theoretical framework and recent developments. *Chem. Soc. Rev.* **2017**, *46*, 7548-7596. (b) Truhlar, D. G.; Garrett, B. C. Variational transition state theory. *Annu. Rev. Phys. Chem.* **1984**, *35*, 159-189. (c) Garrett, B. C.; Truhlar, D. G. Generalized transition state theory. Classical mechanical theory and applications to collinear reactions of hydrogen molecules. *J. Phy. Chem.* **1979**, *83*, 1052-1079. (d) Garrett, B. C.; Truhlar, D. G.; Grev, R. S.; Magnuson, A. W. Improved treatment of threshold contributions in variational transition-state theory. *J. Phy. Chem.* **1980**, *84*, 1730-1748.

24. (a) Gonzalez, C.; Schlegel, H. B. Reaction path following in mass-weighted internal coordinates. *J. Phys. Chem.* **1990**, *94*, 5523-5527. (b) Fukui, K. The path of chemical reactions-the IRC approach. *Acc. Chem. Res.* **1981**, *14*, 363-368. (c) Maeda, S.; Harabuchi, Y.; Ono, Y.; Taketsugu, T.; Morokuma, K. Intrinsic reaction coordinate: Calculation, bifurcation, and automated search. *Int. J. Quantum Chem.* **2015**, *115*, 258-269.

25. Davis, C. E.; Duffy, B. C.; Coates, R. M. Total synthesis of (±)-cameroonan-7α-ol and biomimetic rearrangements to related nopsane sesquiterpenes. *J. Org. Chem.* **2003**, *68*, 6935–6943.

26.   Gaussian 16, Revision C.01. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M.A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.;Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.;Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.;Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.;Liang,W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.;Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi,R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.;Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.;Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2016.

27.   Zhao, Y.; Truhlar, D. G. Density functionals with broad applicability in chemistry. *Acc. Chem. Res*. **2008**, *41*, 157-167.

28.   Riplinger, C.; Neese, F. An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys*. **2013**, *138*, 034106.

29.   (a) Neese, F. Software Update: The ORCA Program System,Version 4.0. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, No. e1327. (b) Neese, F. The ORCA Program System. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 73−78.

30.   (a) J. Zheng, J. L. Bao, S. Zhang, J. C. Corchado, R. Meana-Pañeda, Y.-Y. Chuang, E. L. Coitiño, B. A. Ellingson, and D. G. Truhlar, *Gaussrate 17*, University of Minnesota, Minneapolis, MN, 2017. (b) J. Zheng, J. L. Bao, R. Meana-Pañeda, S. Zhang, B. J.

Lynch, J. C. Corchado, Y.-Y. Chuang, P. L. Fast, W.-P. Hu, Y.-P. Liu, G. C. Lynch, K. A. Nguyen, C. F. Jackels, A. Fernandez Ramos, B. A. Ellingson, V. S. Melissas, J. Villà, I. Rossi, E. L. Coitiño, J. Pu, T. V. Albu, A. Ratkiewicz, R. Steckler, B. C. Garrett, A. D. Isaacson, and D. G. Truhlar, *POLYRATE17*, University of Minnesota, Minneapolis, 2017.

31.   Ussing, B. R.; Hang, C.; Singleton, D. A. Dynamic effects on the periselectivity, rate, isotope effects, and mechanism of cycloadditions of ketenes with cyclopentadiene. *J. Am. Chem. Soc*. **2006**, *128*, 7594−7607.

32.   Marx, D.; Hutter, J. *Ab initio molecular dynamics: Basic theory and advanced methods*; Cambridge University Press: New York, 2009.

33.   Legault, C.Y.; CYLview, 1.0b, Université de Sherbrooke, 2009 (http://www.cylview.org)

34.   (a) Chuang, H.-H.; Tantillo, D. J.; Hsu, C.-P. The Construction of Two-Dimensional Potential Energy Surfaces of Reactions with Post-Transition State Bifurcations. *J. Chem. Theory Comput.* **2020**, *16*, 4050-4060. (b) Carpenter, B. K. *Understanding Chemical Reaction Mechanisms, in Of Minds and Molecules, Eds. Nalini Bhushan and Stuart Rosenfeld.* Oxford University Press. New York, 2000, p. 223-224.

35.   Gonzalez-James, O. M.; Kwan, E. E.; Singleton, D. A. Entropic intermediates and hidden rate-limiting steps in seemingly concerted cycloadditions. Observation, prediction, and origin of an isotope effect on recrossing. *J. Am. Chem. Soc.* **2012**, *134*, 1914-1917.

36.   Cheng, J; Kishan, L. H.; Vernon D. P. Hydride affinities of carbenium ions in acetonitrile and dimethyl sulfoxide solution. *J. Am. Chem. Soc.* **1993**, *115*, 2655-2660.

37.   (a) Beno, B. R.; Houk, K. N.; Singleton, D. A. Synchronous or Asynchronous? An "Experimental" Transition State from a Direct

38.   Comparison of Experimental and Theoretical Kinetic Isotope Effects for a Diels−Alder Reaction. *J. Am. Chem. Soc.* **1996**, *118*, 9984-9985. (b) Nieves-Quinones, Y.; Singleton, D. A. Dynamics and the regiochemistry of nitration of toluene. *J. Am. Chem. Soc.* **2016**, *138*, 15167-15176. (c) Kurouchi, H.; Sanctis, I. L. A.-D.; Singleton, D. A. Controlling selectivity by controlling energy partitioning in a thermal reaction in solution. *J. Am. Chem. Soc.* **2016**, *138*, 14534-14537 (d) Hase, W. L.; Mondro, S. L.; Duchovic, R. J.; Hirst, D. M. Thermal rate constant for H+ CH3 CH4 recombination. 3. Comparison of experiment and canonical variational transition state theory. *J. Am. Chem. Soc.* **1987**, *109*, 2916-2922. (e) Pradhan, R.; Lourderaj, U. Can reactions follow non-traditional second-order saddle pathways avoiding transition states? *Phys. Chem. Chem. Phys.* **2019**, *21*, 12837-12842.

39.   Wang, Z.; Hirschi, J. S.; Singleton, D. A. Recrossing and dynamic matching effects on selectivity in a Diels–Alder reaction. *Angew. Chem. Int. Ed*. **2009**, *48*, 9156-9159.

40.   Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys*. **2005**, *7*, 3297-3305.

41.   Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B*. **2009,** *18*, 6378-6396.

42.   Townsend, D.; Lahankar, S. A.; Lee, S. K.; Chambreau, S. D.; Suits, A. G.; Zhang, X.; Rheinecker, J.; Harding, L. B.; Bowman, J. M. The roaming atom: straying from the reaction path in formaldehyde decomposition. *Science* **2004**, *306*, 1158-1161.

43. (a) Laconsay, C. J.; Tsui, K. Y.; Tantillo, D. J. Tipping the balance: theoretical interrogation of divergent extended heterolytic fragmentations. *Chem. Sci.* **2020**, *11*, 2231-2242. (b) Xu, L.; Doubleday, C. E.; Houk, K. N. Dynamics of carbene cycloadditions. *J. Am. Chem. Soc*. **2011**, *133*, 17848-17854. (c) Bailey, J. O.; Singleton, D. A. Failure and redemption of statistical and nonstatistical rate theories in the hydroboration of alkenes. *J. Am. Chem. Soc.* **2017**, *139*, 15710-15723.

Chapter 3

Bouncing Off Walls – Widths of Exit Channels from Shallow Minima Can Dominate Selectivity Control[*]

3.1 Introduction

Reactions between butadienes and allyl cations have been studied extensively from both mechanistic and synthetic perspectives. Reactions of these two types of components can lead to 4-, 5-, 6- or 7-membered rings via concerted or stepwise (formal) (2+2)/ [$_\pi2+_\pi2$], [1,2] (2+3)/[$_\pi2+_\pi2$],[3,4] (4+2)/[$_\pi4+_\pi2$][5-9] (the so-called "ionic Diels-Alder reaction") or (4+3)/[$_\pi4+_\pi2$][3,4,10-12] cycloaddition reactions, respectively (Scheme 3.1). Which products are observed can be influenced by which substituents are attached to each component and the environment in which the reaction is run. Here we reinvestigate the parent reaction of butadiene + allyl cation (in the gas phase and DMSO), adding results from density functional theory (DFT) calculations to Pascual-Teresa and Houk's previously reported HF and MP2 results,[13] and including an expanded view of the butadiene + allyl cation potential energy surface (PES). In addition, we have examined variational transition states (VTSs)[14] and subjected this reaction to *ab initio* molecular dynamics (AIMD) simulations, of both downhill and uphill varieties,[15] to assess whether or not non-statistical dynamic effects play a role in determining product selectivity.[16-19] Or results have led to the conclusion that the *widths of pathways to products can sometimes be more important than barrier heights*—an entropic effect that has often been overlooked but has potential utility in constructing useful models of selectivity and in designing new selective reactions.
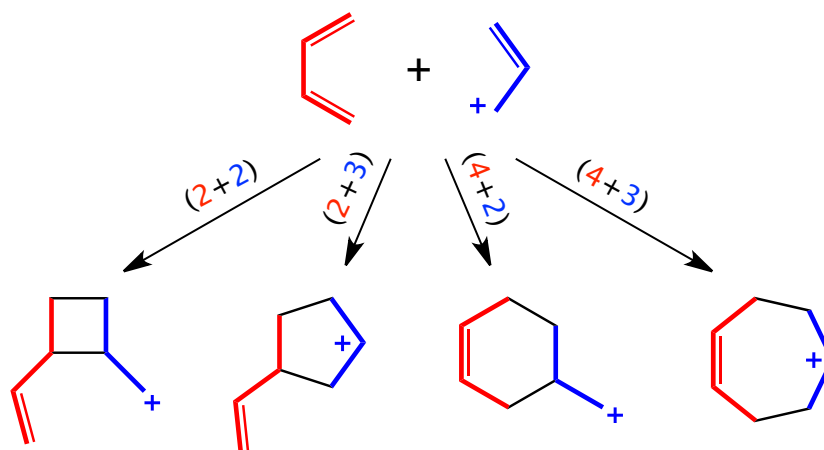
Scheme 3.1 Possible modes of cycloaddition for butadiene + allyl cation.

3.2 Computational Methods

The *Gaussian 09* software suite[20] was used for all DFT calculations. The geometries and frequencies of minima and transition-state structures (TSSs) were optimized at the B3LYP-D3/6-311+G(d,p),[21-25] M06-2X/6-311+G(d,p),[26-28] M06-2X/6-311G(d), and ωB97X-D/6-311G(d)[29] levels of theory. TSSs were defined as such by frequency calculations that predicted exactly one imaginary frequency,[30] while minima showed no imaginary frequencies. Molecular structures were visualized using *CYLview*.[31] Intrinsic reaction coordinate (IRC)[32-34] calculations were performed to confirm that minimum energy pathways (MEP) connected TSSs to the reactants and the products indicated. To explore the effect of solvation, calculations were also performed with M06-2X/6-311G(d) and the PCM solvation model for DMSO.[35] In addition, single-point energies were computed at the CCSD(T)-F12a/AVDZ[36] level using the *MOLPRO* package.[37] As shown in Figure 3.1, the relative energies of stationary points are similar among different levels. VTSs and associated reaction paths were determined using *Gaussrate*[38]/*Polyrate*[39] with application of the Reorientation Of the Dividing Surface (RODS) algorithm at the M06-2X/6-311G(d) level of theory.[14,40]
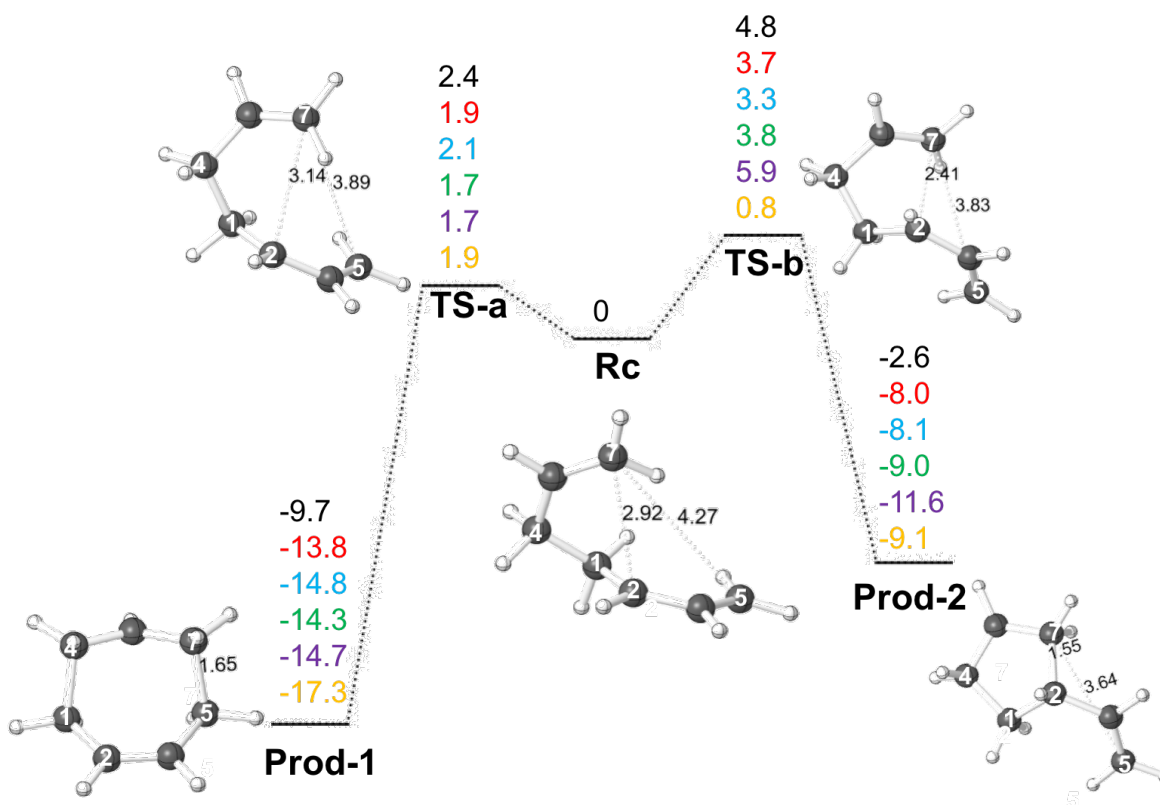
Figure 3.1 Relative free energies (single point energies for CCSD(T) results) of stationary points on the allyl cation + butadiene PES (in kcal/mol relative to **Rc**). The values shown are from top to bottom: B3LYP-D3/6-311+G (d,p) (black), M06-2X/6-311+G(d,p) (red), M06-2X/6-311G(d) (blue), ωB97X-D/6-311G(d) (green), PCM(DMSO)-M06-2X/6-311G(d) (purple), CCSD(T)-F12a/AVDZ// M06-2X/6-311G(d) (yellow). Distances shown (Å) are from M06-2X/6-311G(d). Structure labels are defined in the text below.

Both quasi-classical and classical AIMD simulations were performed using M06-2X/6-311G(d). For the former, zero-point energies are included in initial sampling while for the latter, they are not. Quasi-classical AIMD simulations also were performed using the ωB97X-D functional with the 6-311G(d) basis set. All trajectories were propagated using the *Progdyn*[41] script. To explore the effect of continuum solvent on dynamics, PCM(DMSO)-M06-2X/6-311G(d) was also used. Uphill (initiated in reactant, **Rc**, region) and downhill (initiated in

transition state region) dynamics simulations were both carried out. For downhill dynamics, starting geometries were generated from a Boltzmann distribution of vibrations at 298 K and trajectories were propagated in both product and reactant directions with the Verlet algorithm, using *Gaussian 09* to recalculate forces at each time step (the time step was set as 1 fs here).[42] For uphill dynamics simulations, starting geometries were generated from a Boltzmann sampling of vibrations at 298 K then propagated (1 fs time step) towards the product direction until one or the other of the products is formed. Starting points for uphill dynamics are overlaid in Figure 3.2. Note that our analysis here is for a "relaxed" **Rc**; i.e., momentum for its formation is not included. The following geometric criteria (atom numbers shown in Figure 3.1) were employed to halt trajectories: when the C1–C4 and C5–C7 distances both dropped below 1.70 Å, a trajectory was labeled as the (4+3) adduct (7-membered ring product, **Prod-1**); when the C1–C4 and C2–C7 distances both dropped below 1.60 Å, a trajectory was labeled as the (2+3) adduct (5-membered ring product, **Prod-2**).

(a)   **(b)**   **(c)**
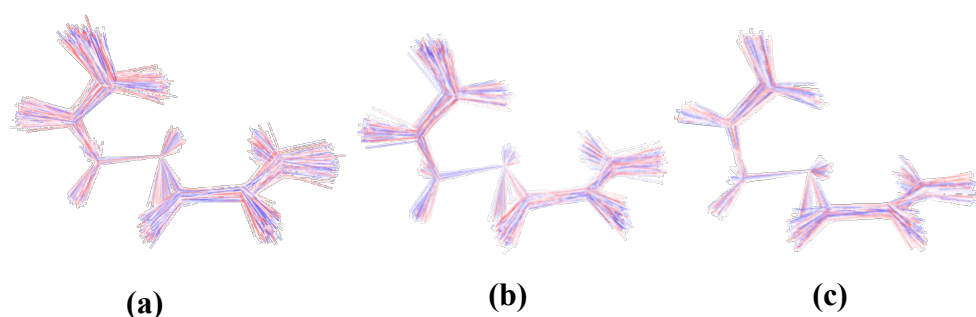
Figure 3.2 Overlays of all starting geometries consistent with a Boltzmann distribution at 298 K in the region of **Rc**: (a) M06-2X/6-311G(d), (b) ωB97X-D/6-311G(d), (c) PCM(DMSO)-M06-2X/6-311G(d).

3.3 Results and Discussion

3.3.1 Reaction Coordinates

Energy profiles from different levels of theory for the allyl cation + butadiene reaction are shown in Figure 3.1. No TSS for concerted cycloaddition or formation of **Rc**, the product of addition of one end of the butadiene π-system to one end of the allyl π-system, were found; it appears that adduct formation is barrierless, at least in the absence of explicit solvent (appropriate modeling of which is beyond our current capabilities). Cation **Rc** can then be converted to 7- or 5-membered ring-containing products via concerted processes involving **TS-a** and **TS-b**, respectively. Note that the C=C group involved in bond formation (that which originated in the allyl fragment) rotates in opposite directions to form the 7- and 5-membered ring-containing products (i.e., for an overall pathway connecting **Prod-1** to **Prod-2**, the migrating CH_2 group interacts antarafacially with the allyl group over which it migrates). While tempting to consider this observation as a result of "vestiges of orbital symmetry,"[43] that cannot be done with confidence without more detailed examination of the potential energy surface (PES) preceding **Rc**. At all DFT levels of theory (but not CCSD(T)), **TS-b** is predicted to be higher in energy than **TS-a** and **Prod-2** was predicted to be higher in energy than **Prod-1**. Namely, in terms of either kinetic or thermodynamic control, formation of the 7-membered ring is favored. However, **Rc** is a shallow minimum.
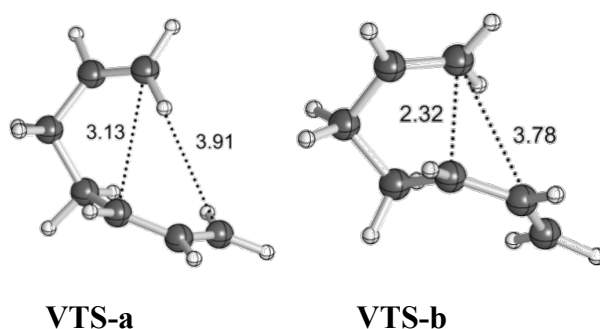
3.3.2 Variational Transition States

To address issues with using the harmonic approximation for computing entropies, reaction pathways derived using variational transition state theory (VTST) calculations (using M06-2X/6-311G(d)) were computed; these are shown in Figure 3.3. For both pathways, VTSs differ from the PES TSSs (compare the geometries of the VTSs in Figure 3.4 with those of the TSSs in Figure 3.1). The C5–C7 distance in **VTS-a** is 0.02 Å longer than that in **TS-a**; nevertheless, the C2–C7 distance in **VTS-a** is 0.01 Å shorter than that of in **TS-a**. The C5–C7 and C2–C7 distances in **VTS-b** are both shorter than those in **TS-b**, by 0.05 Å and 0.09 Å,

respectively. Both VTSs are higher in free energy than the corresponding TSSs. As shown in Figure 3.3, this effect is not simply an entropy effect, but also an effect of zero-point energy. Note, however, that the difference in energy between the peaks of the free energy and potential energy curves is small — tenths of a kcal/mol — consistent with flat surfaces, in terms of both potential energy and free energy, near the transition states. No evidence for "entropic intermediates" is found here.[44,45]
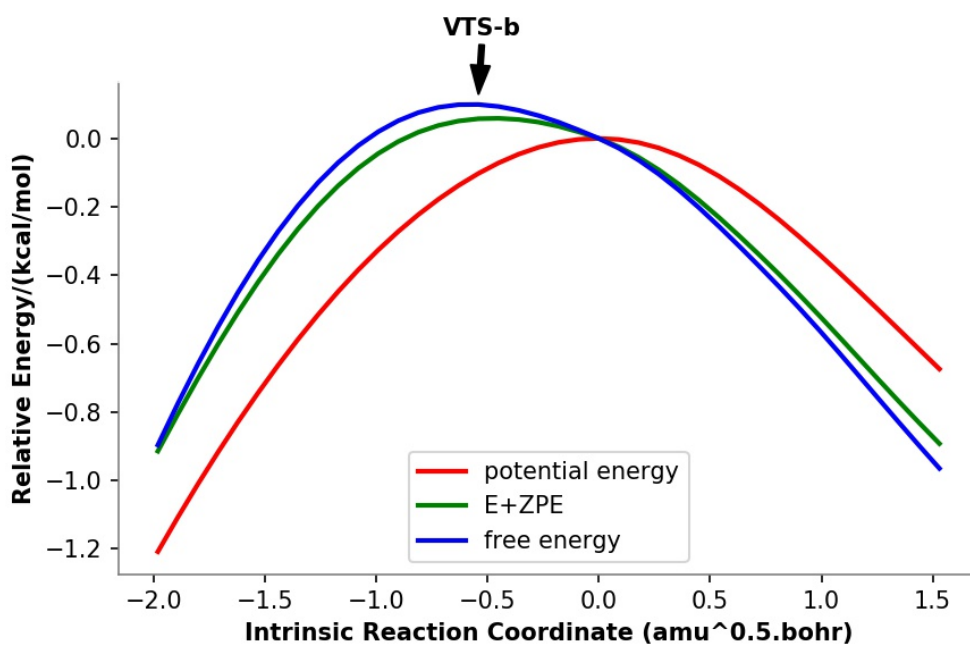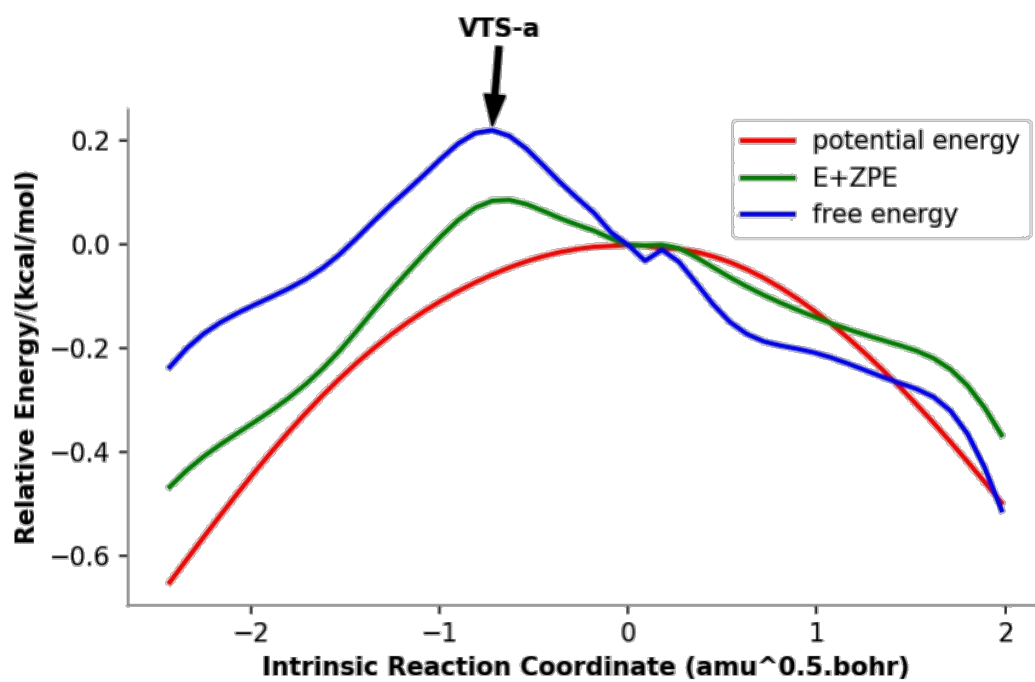
Figure 3.3 Reaction coordinate plots for pathways a (left) and b (right). The locations of**VTS-a** and

VTS-b are highlighted.

Figure 3.4 Geometries of variational transition states. Selected distances are shown in Å.

3.3.3 Direct Dynamics Trajectories

Given the flatness of the PES near **Rc** and the transition states for cyclization, AIMD trajectory simulations were employed to evaluate (1) whether either transition state can lead to both products and (2) the factors influencing product selectivity.

Downhill trajectories initiated from **TS-a** did in fact lead to both products, but just barely (Table 3.1), while trajectories initiated from **TS-b** did not. In both cases, however, a large amount of recrossing (here including trajectories connecting the reactant to the reactant, one product to the same product, or one product to the other product) was observed, consistent with a flat region of the PES.

Table 3.1 Results of downhill dynamics simulations initiated from **TS-a** and **TS-b**.

	Total trajectories	Trajectories that form **Prod-1**	Trajectories that form **Prod-2**	Trajectories that recross
TS-a	702	403	3	296
TS-b	699	0	451	248

Downhill trajectories were also initiated from **VTS-a** and **VTS-b**. Only **Prod-1** was formed from **VTS-a** and only **Prod-2** was formed from **VTS-b** (Table 3.2). While recrossing

61

trajectories were much reduced for **VTS-b**, compared to **TS-b**, a high percentage of recrossing was still observed for **VTS-a**.

Table 3.2 Results of downhill dynamics simulations initiated from VTSs.

	Total trajectories	Trajectories that form product	Trajectories that recross
VTS-a	139	64	75
VTS-b	103	93	10

Results for uphill AIMD trajectories initiated at **Rc** are shown in Table 3.3. At all levels of theory, formation of the 7-membered ring-containing product, **Prod-1**, is much more favorable. However, the predicted **Prod-1**/**Prod-2** ratio is lower than expected based on differences in TSS free energies (Figure 3.1). For example, at ωB97X-D/6-311G(d), the ratio from dynamics simulation is 90:10, while that from free energy differences for TSSs is 97:3. As described above, VTS free energies differ only slightly from TSS free energies. With PCM(DMSO)-M06-2X/6-311G(d), free energy differences predict essentially no **Prod-2**, but our dynamics simulations predict that 16.5% of trajectories lead to this product. What is the origin of this discrepancy? Is it technical or chemical in nature?

Table 3.3 Results of uphill dynamics simulations initiated from **Rc**.

method	Total trajectories	Trajectories that form **Prod-1**	Trajectories that form **Prod-2**
M06-2X/6-311G(d)	206	171（83%）	35（17%）
ωB97X-D/6-311G(d)	154	138（90%）	16（10%）
M06-2X (IEFPCM, DMSO)/6-311G(d)	91	76（84%）	15（16%）

Since our trajectories often involved thousands of fs to form products (see Supporting Information), we were concerned about ZPE leakage (generally a worry past ~0.5 ps).[46-49] To address this issue, we also carried out classical AIMD simulations. Results (using M06-2X/6-311G(d)) for uphill AIMD trajectories initiated at **Rc** and for downhill AIMD trajectories initiated at **VTS-a** are shown in Tables 3.4 and 3.5, respectively. For uphill simulations, the predicted **Prod-1**/**Prod-2** ratio does not change significantly for classical vs. quasi-classical simulations (compare Tables 3.3 and 3.4). For downhill simulations, a significantly smaller proportion of trajectories recrossed for classical vs. quasi-classical dynamics (compare Tables 3.2 and 3.5), but a large proportion of recrossing was still observed.

**Table 3.4.** Results of uphill dynamics simulations initiated from **Rc**.

	Total trajectories	Trajectories that form **Prod-1**	Trajectories that form **Prod-2**
Rc	90	75 (83%)	15 (17%)

Table 3.5. Results of downhill dynamics simulations initiated from **VTS-a**.

	Total trajectories	Trajectories that form product	Trajectories that recross

VTS-a	130	91	39

3.3.4 Potential Energy Surfaces.

To explore how shape of the PES influences the selectivity of the reaction, PESs for gas phase (M06-2X/6-311G(d)) and DMSO (PCM-M06-2X/6-311G(d)) reactions were built (Figure 3.5) using relaxed scans along C5–C7 and C2–C7 distances (see Figure 3.1 for atom labels) as x- and y-axes and energy displayed as height and color (top surfaces) or color alone (bottom surfaces, projections of top surfaces into the x-y plane). While these are reduced-dimension PESs, the geometric parameters used correspond to those of the forming bonds.[50,51] Small C5–C7 distances correspond to the **Prod-1** region, while small C2–C7 distances correspond to the **Prod-2** region. The locations of the two TSSs are marked, as are the IRC paths. One observation is immediately apparent from these surfaces: the pathway toward **Prod-1** is wider than the pathway toward **Prod-2**.
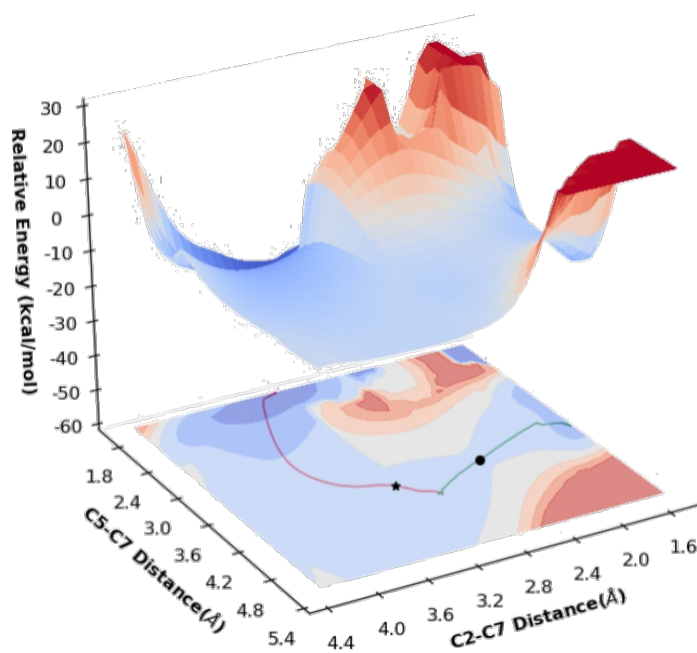

(a) gas phase
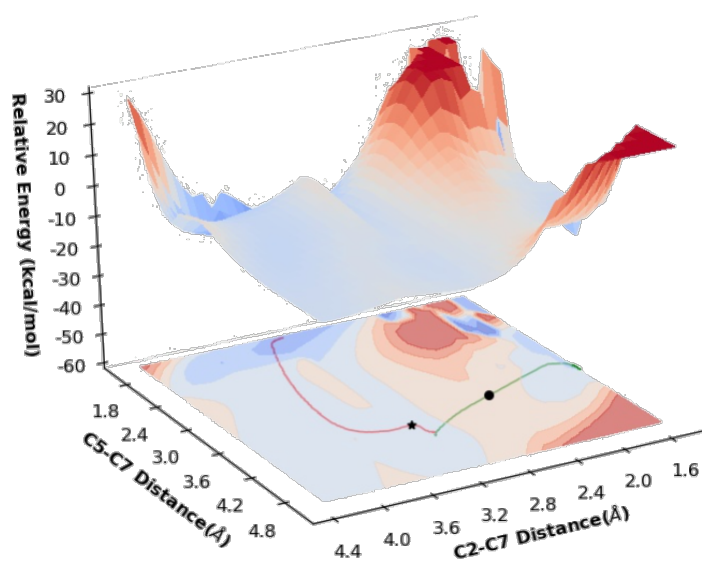
(b) DMSO

Figure 3.5. Potential energy surfaces: (a) M06-2X/6-311G(d), (b) PCM(DMSO)-M06-2X/6-311G(d). IRCs for **TS-a** are shown as red lines and IRCs for **TS-b** are shown as green lines. **TS-a** and **TS-b** are indicated as stars and circles, respectively.

Representative gas phase trajectories are plotted in Figure 3.6. As expected, long-time trajectories sample more of the **Rc** region on the PES before forming products. Importantly, some trajectories heading toward product "turn back" before crossing the dividing surface that separates reactant from products, i.e., they miss the exit channel and instead "bounce off the wall" flanking it. This is distinct from recrossing, where the dividing surface is passed before a trajectory turns back (at least for the dimensions we have plotted here).[16-19] This is also distinct from bouncing off the side walls of a valley en route to product(s), i.e., where deviations from an IRC still lead to product(s), not back to reactant(s).[16] The proportion of **Prod-1** forming trajectories that arose from "turning back" (62/171 or 36%) is larger than the proportion of **Prod-2** forming trajectories that arose from "turning back" (7/35 or 20%), i.e., more **Prod-1** is

65

formed from trajectories that initially miss the exit channel toward which they started, presumably because the exit channel toward **Prod-1** is wider. The product ratio predicted on the basis of "turn back" trajectories alone is 90:10 (62 vs. 7 trajectories). The ratio predicted on the basis of direct trajectories (here there are 137, making for a total of 206 as shown in Table 3.3) is 80:20, and the overall ratio accounting for all product-forming trajectories is 83:17 (Table 3.3). Thus, the ratio predicted from direct uphill trajectories does not reflect the predicted free energies of transition states, be they TSSs or VTSs, and the "turning back" phenomenon increases selectivity for the product formed via the wider pathway. A similar effect was recently discussed for a different carbocation reaction, but from the perspective of downhill trajectories and post-transition state bifurcations.[52]

A close-up view of the reactant region of the PES is shown in Figure 3.7, which also shows the locations of the VTSs. **TS-a** and **VTS-a** occupy similar positions in this plot, consistent with similar amounts of recrossing in downhill AIMD. **VTS-b** is noticeably later (closer to product in terms of the geometric parameters plotted here) than is **TS-b**, consistent with reduced recrossing. Being later, it is further into the least wide region of the path to products, consistent with it being even more difficult for uphill trajectories to reach **Prod-2** than if they only had to reach **TS-b**. This region of decreased width indicates that varying the C5–C7 distance happens to be particularly difficult near the transition state for formation of the C2–C7 bond. This phenomenon was unexpected, but is perhaps related to the fact that the forming σ-bond is much shorter in **TS-b** and **VTS-b** than in **TS-a** and **VTS-a** and conformational restrictions due to hyperconjugation with the forming σ-bond are expected, therefore, to be more severe.
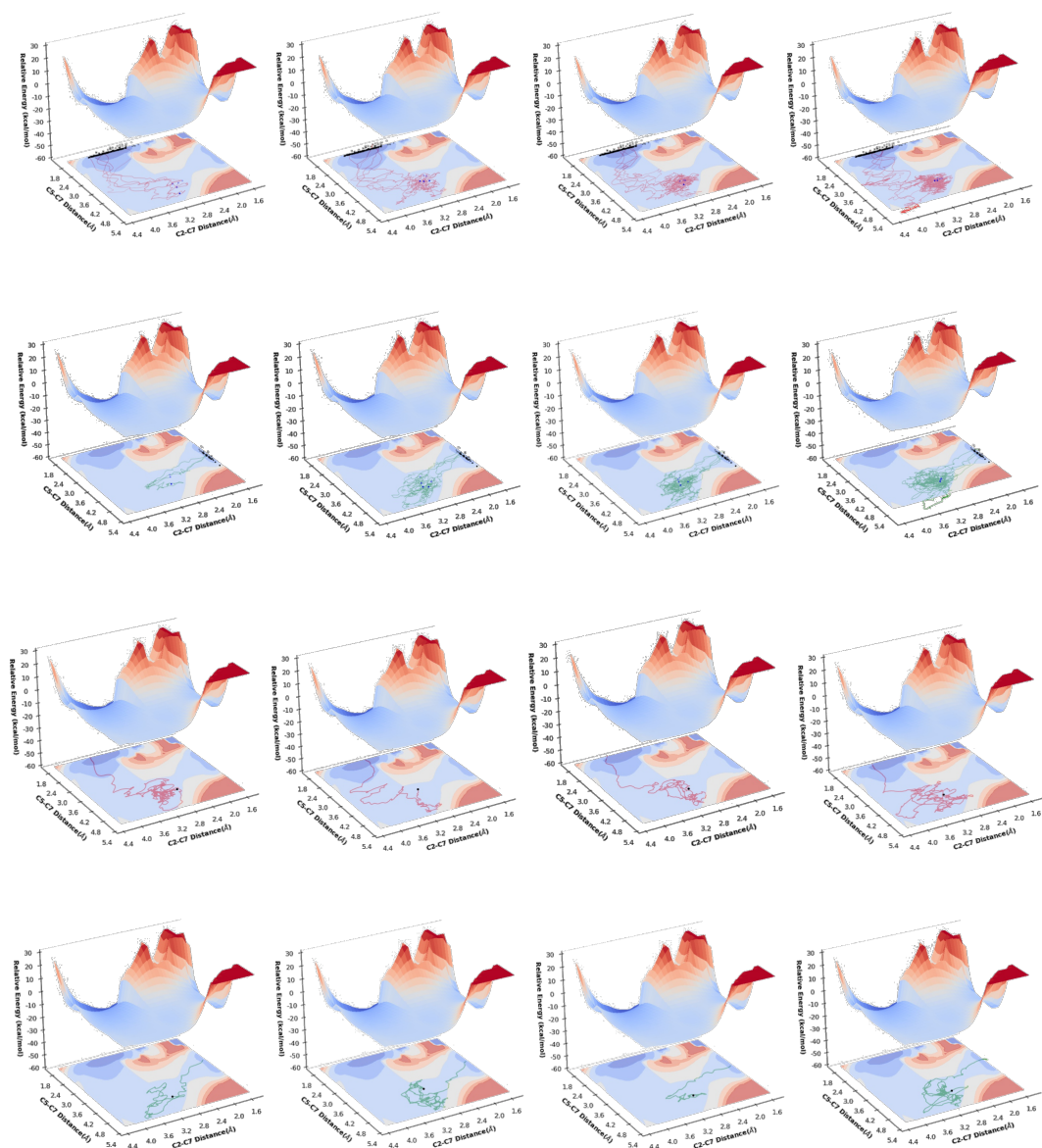
Figure 3.6. (1) Projections of trajectories onto the M06-2X/6-311G(d) PES, based on trajectory length (time). Left to right: 0-500 fs, 500-1000 fs, 1000-1500 fs, 1500-2000 fs. The first row corresponds to **Prod-1**-forming trajectories. The second row corresponds to **Prod-2**-forming trajectories. (2) Projections of representative "turn back" trajectories onto the M06-2X/6-311G(d) PES. The third row shows representative **Prod-1**-forming "turn back" trajectories. The bottom row shows representative **Prod-2**-forming "turn back" trajectories.
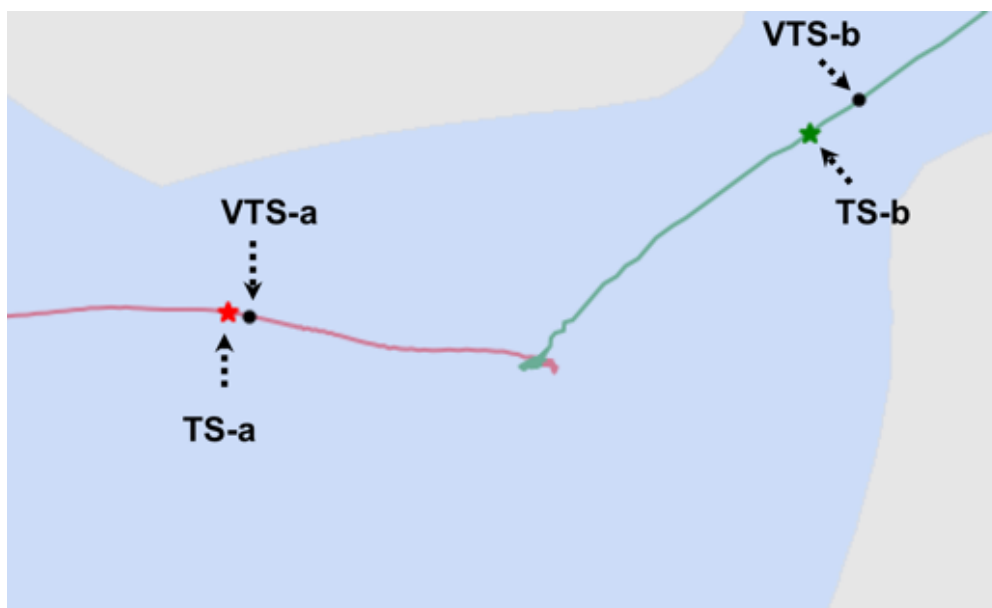
Figure 3.7. Close-up view of the reactant region of the surface from Figure 5a.

This pathway width effect is clearly connected to entropy, i.e., a wider path is related to more flexibility and more possible ways to cross through a dividing surface.[40,45,53-55] This effect appears not to be captured fully in VTS free energies, however, at least for the system examined here, which involves a flat energy surface, although we cannot definitively rule out the possibility that the apparent discrepancy is due to the particulars of running each type of calculation. Nonetheless, interpreting results of this type in terms of pathway widths allows one to better tie specific structural features to entropy changes along reaction coordinates.

3.4 Conclusions

We have explored the mechanism of the reaction between butadiene and allyl cation using quantum chemical computations on PESs and direct AIMD simulations (quasi-classical and classical). Consistent with results from previous calculations at other levels of theory,[13] concerted cycloaddition pathways are not found. Instead, stepwise pathways leading to products of formal (4+3) and (2+3) cycloaddition are present, but these involve a shallow

intermediate with low barrier exit channels. As for other reactions with shallow intermediates—twixtyls, calderas, mesas, para-intermediates[19,56-58]—non-traditional effects appear to play a key role in product formation. We propose that the product selectivity in the system examined here—and likely others[52]—is determined by dynamic effects related to the *width* of pathways to products. In particular, there is an entropy effect—transition states with wider energy wells orthogonal to the reaction coordinate have more possible paths through them—which is modulated, enhanced in this case, by an unequal probability of trajectories turning back before reaching transition states. We are currently expanding on this work to see if models based on pathway width can rationalize product selectivity for cases with existing experimental evidence for entropy-controlled selectivity where the origins of that entropy control are not known.

3.5 Acknowledgement

3.6 References

1.   Hong, Y. J.; Tantillo, D. J. How Cyclobutanes are Assembled in Nature - Insights from Quantum Chemistry. *Chem. Soc. Rev.* **2014**, *43*, 5042-5050.

2.   Gassman, P. G.; Lottes, A. C. Cyclobutane Formation in the $2\pi+2\pi$ Cycloaddition of Allyl and Related Cations to Unactivated Olefins. Evidence for the 2nd Step in the Proposed Mechanism of the Ionic Diels-Alder Reaction. *Tetrahedron Lett.* **1992**, *33*, 157-160.

3.   Hoffmann, H. M. R. The Cycloaddition of Allyl Cations to 1,3-Dienes -

General-Method for the Synthesis of 7-Membered Carbocycles. *Angew. Chem. Int. Edit.* **1984**, *23*, 1-19.

4.   Hoffmann, H. M. Syntheses of 7-Membered and 5-Membered Rings from Allyl Cations. *Angew. Chem. Int. Edit.* **1973**, *12*, 819-835.

5.   Haynes, R. K.; King, G. R.; Vonwiller, S. C. Preparation of a Bicyclic Analog of Qinghao (Artemisinic) Acid via a Lewis-Acid Catalyzed Ionic Diels-Alder Reaction Involving a Hydroxy Diene and Cyclic Enone and Facile Conversion into (+/-)-6,9-Desdimethylqinghaosu. *J. Org. Chem.* **1994**, *59*, 4743-4748.

6.   Gassman, P. G.; Gorman, D. B. Stepwise Mechanism for the Formation of $2\pi+4\pi$ Cycloadducts in the Ionic Diels-Alder Reaction. *J. Am. Chem. Soc.* **1990**, *112*, 8624-8626.

7.   Gassman, P. G.; Singleton, D. A. Acid-Catalyzed Intramolecular Diels-Alder Reactions - the Cycloaddition of Allyl Cations to 1,3-Dienes. *J. Am. Chem. Soc.* **1984**, *106*, 6085-6086.

8.   Gassman, P. G.; Singleton, D. A.; Wilwerding, J. J.; Chavan, S. P. Acrolein Acetals as Allyl Cation Precursors in the Ionic Diels-Alder Reaction. *J. Am. Chem. Soc.* **1987**, *109*, 2182-2184.

9.   Ko, Y. J.; Shim, S. B.; Shin, J. H. A Mechanistic Study on the Intramolecular Ionic Diels-Alder Reaction of 2-Methyl-3,9,11-Tridecatriene-2-ol and 2,11-Dimethyl-1,3,9,11-Dodecatetraene. *Tetrahedron Lett.* **2007**, *48*, 863-867.

10.   Woodward R B, H. R. The Conservation of Orbital Symmetry. *Elsevier* **2013**.

11.   Harmata, M. The (4+3)-Cycloaddition Reaction: Simple Allylic Cations as Dienophiles. *Chem. Commun.* **2010**, *46*, 8886-8903.

12.   Harmata, M. Intramolecular Cycloaddition Reactions of Allylic Cations. *Tetrahedron* **1997**, *53*, 6235-6280.

13.   dePascualTeresa, B.; Houk, K. N. The Ionic Diels-Alder Reaction of the Allyl

Cation and Butadiene: Theoretical Investigation of the Mechanism. *Tetrahedron Lett.* **1996**, *37*, 1759-1762.

14.   Bao, J. L.; Truhlar, D. G. Variational Transition State Theory: Theoretical Framework and Recent Developments. *Chem. Soc. Rev.* **2017**, *46*, 7548-7596.

15.   Pradhan, R.; Lourderaj, U. Can Reactions Follow Non-Traditional Second-Order Saddle Pathways Avoiding Transition States? *Phys. Chem. Chem. Phys.* **2019**, *21*, 12837-12842.

16.   Ma, X. Y.; Hase, W. L. Perspective: Chemical Dynamics Simulations of Non-Statistical Reaction Dynamics. *Philos. Trans. R. Soc. A* **2017**, *375*, 20160204.

17.   Lourderaj, U.; Park, K.; Hase, W. L. Classical Trajectory Simulations of Post-Transition State Dynamics. *Int. Rev. Phys. Chem.* **2008**, *27*, 361-403.

18.   Carpenter, B. K. Intramolecular Dynamics for the Organic Chemist. *Acc. Chem. Res.* **1992**, *25*, 520-528.

19.   Carpenter, B. K.: Nonstatistical Dynamics in Thermal Reactions of Polyatomic Molecules. In *Ann. Rev. Phys. Chem.*; Annual Review of Physical Chemistry; Annual Reviews: Palo Alto, 2005; Vol. 56; pp 57-89.

20.   Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli,

C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09. *Gaussian Inc.: Wallingford, CT* **2009**.

21.   Becke, A. D. A New Mixing of Hartree-Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98*, 1372-1377.

22.   Becke, A. D. Density-Functional Thermochemistry .3. The Role of exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648-5652.

23.   Lee, C. T.; Yang, W. T.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron-Density. *Phys. Rev. B* **1988**, *37*, 785-789.

24.   Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab-Initio Calculation of Vibrational Absorption and Circular-Dichroism Spectra Using Density-Functional Force-Fields. *J. Phys. Chem.* **1994**, *98*, 11623-11627.

25.   Tirado-Rives, J.; Jorgensen, W. L. Performance of B3LYP Density Functional Methods for a Large Set of Organic Molecules. *J. Chem. Theory Comput.* **2008**, *4*, 297-306.

26.   Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215-241.

27.   Zhao, Y.; Truhlar, D. G. Density Functionals with Broad Applicability in Chemistry. *Acc. Chem. Res.* **2008**, *41*, 157-167.

28.   Mardirossian, N.; Head-Gordon, M. How Accurate Are the Minnesota Density Functionals for Noncovalent Interactions, Isomerization Energies, Thermochemistry, and Barrier Heights Involving Molecules Composed of Main-Group Elements? *J. Chem. Theory Comput.* **2016**, *12*, 4303-4325.

29.   Sengupta, A.; Raghavachari, K. Solving the Density Functional Conundrum: Elimination of Systematic Errors To Derive Accurate Reaction Enthalpies of Complex Organic Reactions. *Org. Lett.* **2017**, *19*, 2576-2579.

30.   Tantillo, D. J. Biosynthesis via Carbocations: Theoretical Studies on Terpene Formation. *Nat. Prod. Rep.* **2011**, *28*, 1035-1053.

31.   Legault, C. Y. CYLview, 1.0b. *Université de Sherbrooke* **2009**.

32.   Gonzalez, C.; Schlegel, H. B. Reaction-Path Following in Mass-Weighted Internal Coordinates. *J. Phys. Chem.* **1990**, *94*, 5523-5527.

33.   Fukui, K. The Path of Chemical-Reactions - the IRC Approach. *Acc. Chem. Res.* **1981**, *14*, 363-368.

34.   Maeda, S.; Harabuchi, Y.; Ono, Y.; Taketsugu, T.; Morokuma, K. Intrinsic Reaction Coordinate: Calculation, Bifurcation, and Automated Search. *Int. J. Quantum Chem.* **2015**, *115*, 258-269.

35.   Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378-6396.

36.   Kong, L. G.; Bischoff, F. A.; Valeev, E. F. Explicitly Correlated R12/F12 Methods for Electronic Structure. *Chem. Rev.* **2012**, *112*, 75-107.

37.   Werner, H. J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schutz, M. Molpro: a General-Purpose Quantum Chemistry Program Package. *WIREs. Comput. Mol. Sci.* **2012**, *2*, 242-253.

38.   Zheng, J.; Zhang, S.; Corchado, J. C.; Chuang, Y.-Y.; Coitino, E. L.; Ellingson, B. A.; Truhlar, D. G. GAUSSRATE, version 2009-A; University of Minnesota: Minneapolis, MN. **2010**.

39.   Zheng, J.; al., e. POLYRATE-version 2010; University of Minnesota:

Minneapolis, MN. **2010**.

40.   Gonzalez-Lafont, A.; Villa, J.; Lluch, J. M.; Bertran, J.; Steckler, R.; Truhlar, D. G. Variational Transition State Theory and Tunneling Calculations with Reorientation of the Generalized Transition States for Methyl Cation Transfer. *J. Phys. Chem. A* **1998**, *102*, 3420-3428.

41.   Ussing, B. R.; Hang, C.; Singleton, D. A. Dynamic Effects on the Periselectivity, Rate, Isotope Effects, and Mechanism of Cycloadditions of Ketenes with Cyclopentadiene. *J. Am. Chem.Soc.* **2006**, *128*, 7594-7607.

42.   Marx, D., & Hutter, J. Ab initio molecular dynamics: basic theory and advanced methods. *Cambridge University Press* **2009**.

43.   Leach, A. G.; Catak, S.; Houk, K. N. Mechanism of the Forbidden 3s,5s - Sigmatropic Shift: Orbital Symmetry Influences Stepwise Mechanisms Involving Diradical Intermediates. *Chem-Eur. J.* **2002**, *8*, 1290-1299.

44.   Yang, Z. Y.; Jamieson, C. S.; Xue, X. S.; Garcia-Borras, M.; Benton, T.; Dong, X. F.; Liu, F.; Houk, K. N. Mechanisms and Dynamics of Reactions Involving Entropic Intermediates. *Trends Chem.* **2019**, *1*, 22-34.

45.   Gonzalez-James, O. M.; Kwan, E. E.; Singleton, D. A. Entropic Intermediates and Hidden Rate-Limiting Steps in Seemingly Concerted Cycloadditions. Observation, Prediction, and Origin of an Isotope Effect on Recrossing. *J. Am. Chem. Soc.* **2012**, *134*, 1914-1917.

46.   Bonnet, L.; Rayez, J. C. Quasiclassical Trajectory Method for Molecular Scattering Processes: Necessity of a Weighted Binning Approach. *Chem. Phys. Lett.* **1997**, *277*, 183-190.

47.   Doubleday, C.; Boguslav, M.; Howell, C.; Korotkin, S. D.; Shaked, D. Trajectory Calculations for Bergman Cyclization Predict H/D Kinetic Isotope Effects Due to

Nonstatistical Dynamics in the Product. *J. Am. Chem. Soc.* **2016**, *138*, 7476-7479.

48.   Bennun, M.; Levine, R. D. Conservation of Zero-point Energy in Classical Trajectory Computations by a Simple Semiclassical Correspondence. *J. Chem. Phys.* **1994**, *101*, 8768-8783.

49.   Reis, M. C.; Lopez, C. S.; Faza, O. N.; Tantillo, D. J. Pushing the Limits of Concertedness. A Waltz of Wandering Carbocations. *Chem. Sci.* **2019**, *10*, 2159-2170.

50.   Chuang, H.-H.; Tantillo, D. J.; Hsu, C.-P. Construction of Two-Dimensional Potential Energy Surfaces of Reactions with Post-Transition-State Bifurcations. *J. Chem. Theory Comput.* **2020**, 10.1021/acs.jctc.1020c00172.

51.   Carpenter, B. K. Understanding Chemical Reaction Mechanisms, in *Of Minds and Molecules*, Eds. Nalini Bhushan and Stuart Rosenfeld. *Oxford University Press* **2000**, New York, p.

52.   Hare, S. R.; Pemberton, R. P.; Tantillo, D. J. Navigating Past a Fork in the Road: Carbocation-π Interactions Can Manipulate Dynamic Behavior of Reactions Facing Post-Transition-State Bifurcations. *J. Am. Chem. Soc.* **2017**, *139*, 7485-7493.

53.   Anslyn, E. V.; Dougherty, D. A. *Modern Physical Organic Chemistry*. *University Science Books, Sausalito, p.* **2006**, *369*, 412-413.

54.   Wang, Z. H.; Hirschi, J. S.; Singleton, D. A. Recrossing and Dynamic Matching Effects on Selectivity in a Diels-Alder Reaction. *Angew. Chem. Int. Edit.* **2009**, *48*, 9156-9159.

55.   Harvey, J. N. Understanding the Kinetics of Spin-forbidden Chemical Reactions. *Phys. Chem. Chem. Phys.* **2007**, *9*, 331-343.

56.   Hoffmann, R.; Swaminat.S; Odell, B. G.; Gleiter, R. Potential Surface for a Nonconcerted Reaction-Tetramethylene. *J. Am. Chem. Soc.* **1970**, *92*, 7091-7097.

57.   Doering, W. V.; Ekmanis, J. L.; Belfield, K. D.; Klarner, F. G.; Krawczyk, B. Thermal Reactions of Anti- and Syn-Dispiro 5.0.5.2 Tetradeca-1,8-Dienes: Stereomutation and

Fragmentation to 3-Methylenecyclohexenes. Entropy-Dictated Product Ratios from Diradical Intermediates? *J. Am. Chem. Soc.* **2001**, *123*, 5532-5541.

58.　　Northrop, B. H.; O'Malley, D. P.; Zografos, A. L.; Baran, P. S.; Houk, K. N. Mechanism of the Vinylcyclobutane Rearrangement of Sceptrin to Ageliferin and Nagelamide E. *Angew. Chem. Int. Edit.* **2006**, *45*, 4126-4130.

Chapter 4

Analogies between Photochemical Reactions and Ground State Reactions with Post-Transition State Bifurcations – Prospects for Rational Design

4.1 Introduction

How does one achieve kinetic selectivity? When tackling this problem, one is usually considering competing reactions with separate transition states (Figure 4.1a) whose difference in free energy can be translated directly to an expected ratio of products, i.e., one applies transition state theory (TST).[3] In some cases, however, TST does not apply. For example, a large number of reactions have been found to proceed via ambimodal transition states, i.e., transition states followed by steepest-descent pathways on potential energy surfaces (PESs) that bifurcate before reaching a secondary transition state that interconverts products (Figure 4.1b).[4–10] In these reactions with post-transition state bifurcations (PTSBs), a single transition state leads to competing products, so product ratios are controlled by non-statistical dynamic effects.[11–15] Here we examine a third related scenario involving initial photoexcitation (Figure 4.1c).
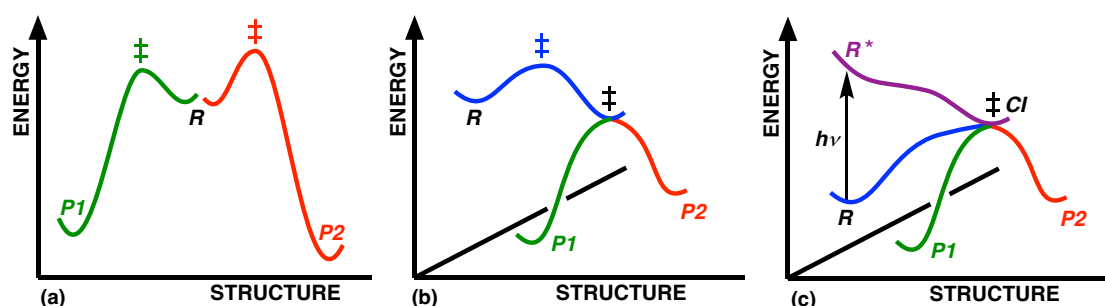
Figure 4.1 Three routes to kinetic selectivity: (a) transition state control, (b) dynamic control via a PTSB, (c) dynamic control initiated by photochemical excitation.

Photochemistry is often employed when a ground state reaction faces an unsurmountable barrier (under a particular set of reaction conditions and/or in comparison to that for a competing undesired reaction).[16–19] Excitation of a reactant (described here as a single molecule, but really an ensemble) to a species with the same geometry as the ground state species (i.e., vertical or Franck-Condon excitation) generally lands one at a point that is not a minimum

energy structure on the electronic excited state surface. Consequently, the molecule in question will rapidly proceed downhill in energy, ultimately crossing back to the ground state surface (perhaps crossing to one or more excited state surfaces along the way). If the hop to the ground state surface occurs near to a ground state transition state that interconverts the two competing products of interest (Figure 4.1c) – not simply the reactant and one product, as is common – then selectivity is controlled by dynamic effects on the excited state surface and the nature of surface hopping.

This third scenario is similar to that of a reaction with a PTSB in that both scenarios involve approaching a transition state that interconverts competing products from "above". For a reaction with a PTSB, above is from a higher energy transition state in the ground state. For the photo-promoted reaction just described, above is from an excited state. In both cases, dynamic matching is at play,[15,20–23] i.e., the part of the energy released along the downhill pathway that is not lost to solvent generates vibrationally excited 'hot' reactive species.[24–27] The momentum associated with the downhill process is "remembered" and plays a role in determining the reaction outcome. Noting the analogy between reactions with PTSBs and photochemical reactions, we set out to design a reaction with energy surfaces akin to those in Figure 4.1c. Our initial design plan involved reactions of the sort shown in Figure 4.2. Knowing that transition structures for Cope rearrangements can have significant diradical character,[28,29] we examined a variety of systems that could produce cyclohexane diyls through loss of different X groups. Ultimately, we arrived at a classic system whose chemistry was in need of explanation: photochemical deazetization of 2,3-diazabicyclo[2.2.2]oct-2-ene (DBO) (Figure 4.3). Modern modeling methods,[30–36] coupled with our PTSB/non-statistical dynamic perspective, allowed us to answer some lingering questions for this archetypal system, providing a jumping off point for future design.
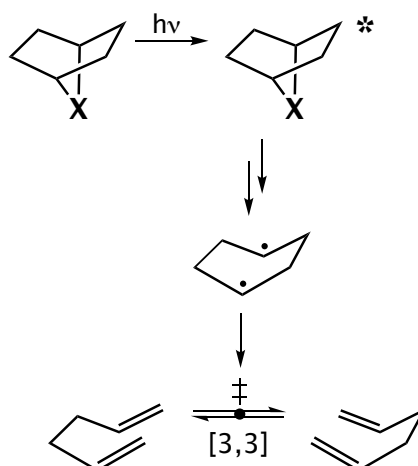
Figure 4.2 Initial design plan: photochemical removal of a group (X) to produce a diradical on an excited state surface that resembles the transition structure for a [3,3] sigmatropic shift.
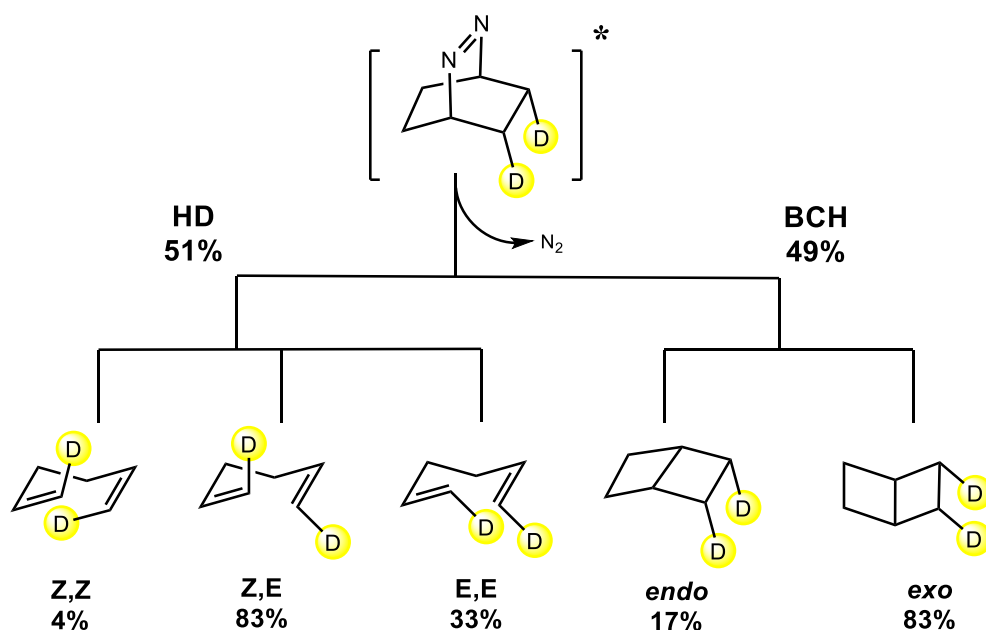
Figure 4.3 Experimentally studied system modeled here with results from references[37–39].

The photochemical reactivity of DBO has been studied under different reaction conditions (Figure 4.3). Grissom *et al.* investigated direct photolysis and reported a bicyclohexane (BCH; also called [2]-ladderane[40–43]):hexadiene (HD) ratio of 49:51,[38,43] which is consistent with results from other experiments.[37,44,45] Samuel and co-workers performed the photolysis of

deuterium-labeled DBO,[39] observing an excess (66%) of *exo*-BCH and determining the relative abundance of the HD isotopomers. Direct and sensitized DBO also have been shown to exhibit different reactivity, with more HD products observed and the *exo:endo* ratio being reduced to almost 1:1 for the latter.[39]

Multiple attempts have been made to rationalize these product distributions. Li *et al*. thoroughly investigated the first C–N bond dissociation process using high-precision multireference calculations that revealed a 12.2 kcal/mol barrier for dissociation and a subsequent downhill process leading to a four-fold (S_0-S_1-T_1-T_2) crossing point.[46] Their work successfully rationalized spectroscopic observations. Anderson *et al*. investigated the solvent, temperature, and magnetic field dependence of the product distribution.[38,43] They rationalized the product ratio based on the assumption of rapid intersystem crossing (ISC) and a statistical 25:75 distribution of singlet-triplet states. However, if ISC is rapid, we would expect a similar outcome for direct and sensitized reactions. Roberson *et al*. explored this problem from the perspective of excess energy and conformations.[45] They pointed out that the singlet-triplet gap becomes smaller along the reaction pathway and singlet DBO behaves like triplet DBO at the singlet-triplet crossing point. However, they also noticed that unlike heavy-atom-containing solvents, magnetic field perturbation has no effect on the product ratio, leading them to question the effect of singlet-triplet interconversion.[38,43] Heavy-atom tunneling have also been found in chemical reactions involving diradical coupling and [3,3]-sigmatropic shifts.[47–51]

Questions about DBO photochemistry still remain, however, that could be resolved with modern dynamics simulations. These types of simulations have only recently become feasible for tackling problems of the scope described here. In pioneering work, Martinez and co-workers used Born-Oppenheimer molecular dynamics (MD) simulations to investigate the photolysis of thioformaldehyde S-oxide sulfine, examining in detail how motion on the excited

state surface leads to ground state products and pointing out that the complexity of photochemical reactions can arise from a single conical intersection.[52] Lopez and co-workers have used related methods to study a variety of photochemical organic reactions, including denitrogenation reactions of smaller frameworks.[30]

However, conventional quantum chemistry-based NAMD (QC-NAMD) is very expensive and not practical for large systems like DBO. Recently, Lopez and co-workers have developed a machine-learning based workflow for NAMD simulations (*PyRAI^{2}MD*). The workflow aims at constructing a ML potential through adaptive sampling and conduct NAMD simulations.

This approach has been used to study a variety of photochemical organic reactions and is able to reproduce. For example, Li *et al.* have studied the formation of substituted cubanes via photochemical [2+2] cycloaddition reaction. They have pointed out that steric repulsions hamper the electrocyclic ring-opening and the noncovalent dispersions enhance the cycloaddition pathway. ML-NAMD trajectories successfully provide insights to the substituent effect of cubane formation. They have also studied the photochemical 4π-electrocyclization of different fluorobenzenes and have figured out that the electrocyclization is promoted by the momentum effect.

Here we use a combination of modern theoretical tools, i.e., the machine-learning assisted non-adiabatic MD (ML-NAMD) methods of Lopez[31,33–35] and the multiconfiguration pair-density functional theory (MC-PDFT)[53–56] of Truhlar and Gagliardi to interrogate DBO photochemistry at a new level of depth, allowing us to solve a long-standing reactivity/selectivity puzzle, thereby opening the door for forward design.

4.2 Reaction Pathway on the S$_1$ Surface

Photoexcitation of DBO involves a typical n-π* transition. revPBE(8,8)/ANO-S-VDZP calculations reveal a vertical S$_0$-S$_1$ excitation energy of 87.2 kcal/mol (3.78 eV), in reasonable

agreement with experiment[37,38] (76.4 kcal/mol (3.31 eV); 80.5 kcal/mol (3.49 eV) with CASPT2(8,8)/ANO-S-VDZP//revPBE(8,8)/ANO-S-VDZP).[57–61] Relaxation of DBO in the Frank-Condon region involves twisting the C–N=N–C bridge to form a minimum on the S_1 surface that is 3.6 kcal/mol lower in energy[46] (CASPT2//MC-PDFT). Unlike many photochemical reactions, the photodissociation of DBO appears to involve a transition state for dissociation on the excited state surface as a result of this facile relaxation process. The C–N bond breaking transition structure (S_1-TS) is associated with a barrier of 9.4 kcal/mol (0.41 eV) on the S_1 surface, which is in reasonable agreement with the 8.6 kcal/mol (0.37 eV) activation energy derived from kinetic experiments, as well as the low quantum yield and the observation of fluorescence.[37]

The minimum energy pathway (MEP) downhill from S_1-TS involves the elongation of one C–N bond. Proceeding through the geometries corresponding to the S_1 MEP, T_2 energies also decrease while S_0 and T_1 energies increase (Figure 4.4). Consequently, these four electronic states meet at the diazinyl diradical structure. Besides fast internal conversion (IC), ISC also can, in principle, occur along this pathway.[62,63] However, we have employed NAMD with MC-PDFT developed by Truhlar *et al*. using *SHARC* (See Appendix for more details)[64–67] and although a small amount of ISC is observed, it is predicted to be inefficient and negligible. Given the massive additional computational resources required to accurately calculate spin-orbit coupling (SOC) and the difficulty of predicting SOC with ML,[35,67] here, we focus on the direct photochemistry of DBO and the IC between singlet states. As described below, we are able to rationalize the experimental outcome for direct photolysis by considering the fate of the S1 state alone, suggesting that contributions from other states may be minor.
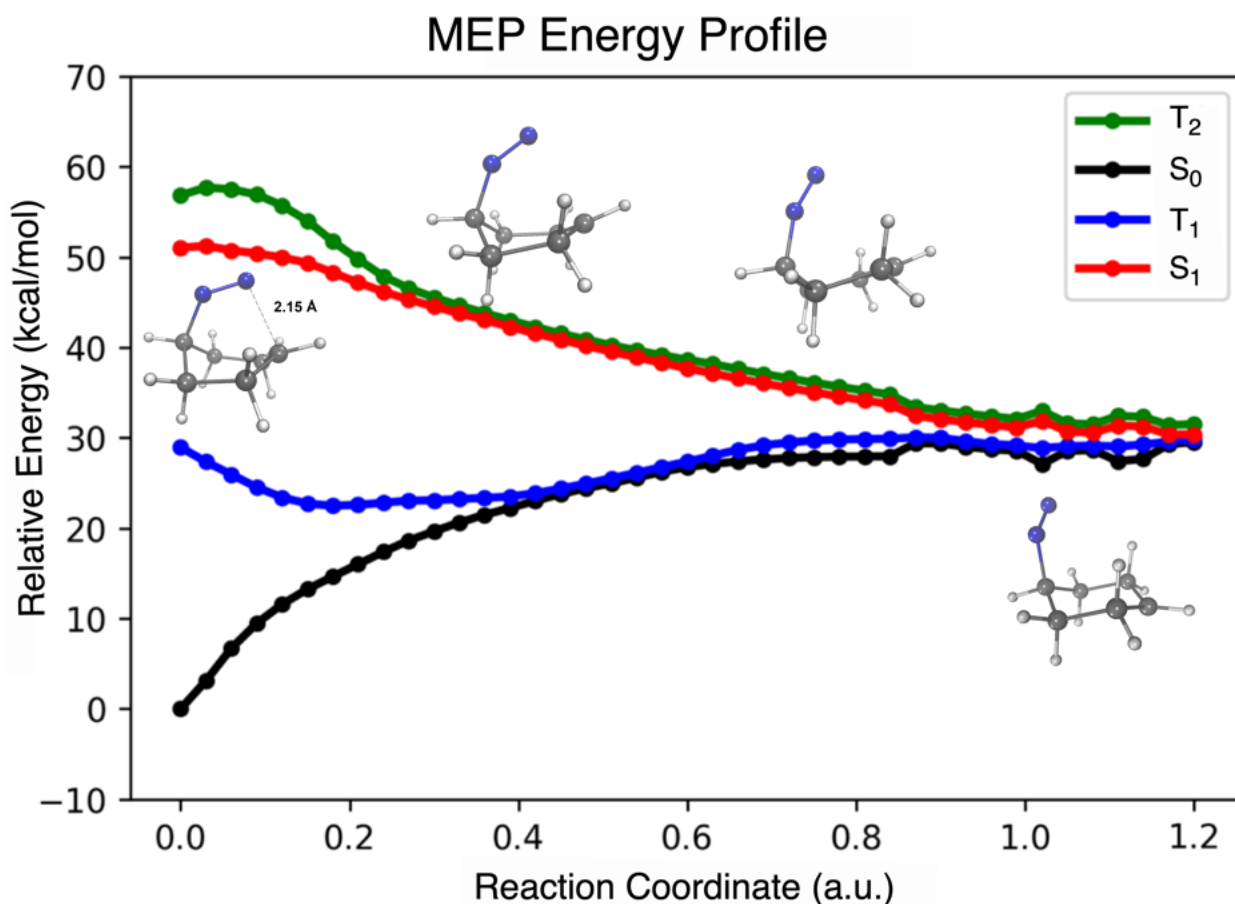
Figure 4.4. Minimal energy pathway (MEP) starting from the TS of C-N bond dissociation on S_1 surface computed with CASPT2(8,8)/ANO-S-VDZP//revPBE(8,8)/ANO-S-VDZP level of theory. '0' point of the MEP is defined as the S_0 energy of the corresponding TS. Relative energies are in kcal/mol.

Although the diazinyl diradical is a minimum on the T_1 surface, we failed to locate a corresponding minimum on the S_1 or S_0 surface with MC-PDFT or CASPT2, likely, we suspect, a result of the flat PES region around the diazinyl diradical. A NEVPT2(6,6)/def2-TZVPP//SS-CASSCF(6,6)/def2-SVP bond scan for S_0 shows that breaking of the second C–N bond, to form a 1,4-cyclohexyl diradical, is nearly barrierless (see Fig C7 in Appendix).

4.3 Machine-learning Assisted NAMD - Justification and Implementation

84

In that hopping to the S_1 surface is expected to occur for the diazinyl diradical (i.e., in the 4-fold crossing region, Figure 4.4), we find little to no barrier for loss of N_2 from the diazinyl diradical to form the 1,4-cyclohexyl diradical, and our MEP calculation (Figure 4.4) indicates that the diradical forms via a very exothermic pathway, we expect that intramolecular vibrational energy redistribution (IVR) is likely slower than subsequent reaction,[13,24–26] i.e., dynamic matching can play a major role in determining product distributions.[3,15,20,21,24–26,68] Consequently, we pursued fewest-switches surface hopping non-adiabatic molecular dynamics (FSSH-NAMD) simulations to determine the influence of non-statistical dynamic effects on the fate of the 1,4-cyclohexyl diradical resulting from the paths leading to it.[31,35,69–72]

Thus, our primary goal became identification of the reactivity of the 1,4-cyclohexyl diradical. In previous studies, 1,4-cyclohexyl diradicals were proposed to be key intermediates along photochemical reaction pathways,[38,47] and Stoltz and coworkers pointed out in a recent paper that a similar diradical resides on a flat energy surface with tiny barriers separating it from cycloaddition products.[73] We attempted to optimize the diradical at the CASPT2(4,4)/ANO-S-VDZP level of theory, but, to our surprise, the optimization converged to a transition structure for a [3,3] sigmatropic shift (implying that such a diradical is not a minimum at this level). MC-PDFT level resulted in a similar conclusion. Broken-symmetry DFT and CASSCF optimizations, on the other hand, successfully located a 1,4-cyclohexyl diradical intermediate in both chair and boat conformations. The sensitivity of the PES around the 1,4-cyclohexyl diradical to the level of theory led us to examine the potential energy landscape of the diradical in more detail. As shown in Figure 4.5, we performed a 2-dimensional scan of the forming/breaking bond lengths corresponding to the [3,3] sigmatropic shift and a [2+2] cycloaddition connected to the BCH product. While CASSCF revealed a shallow minimum for the diradical (Figure C8 in Appendix), NEVPT2//CASSCF single point energies led to the minimum region vanishing (Figure 4.5) for both PESs. In both cases, the

diradical resides on a flat region of the PES that connects to the [3,3] sigmatropic shift transition

state structure and [2+2]-cycloaddition transition state structure.
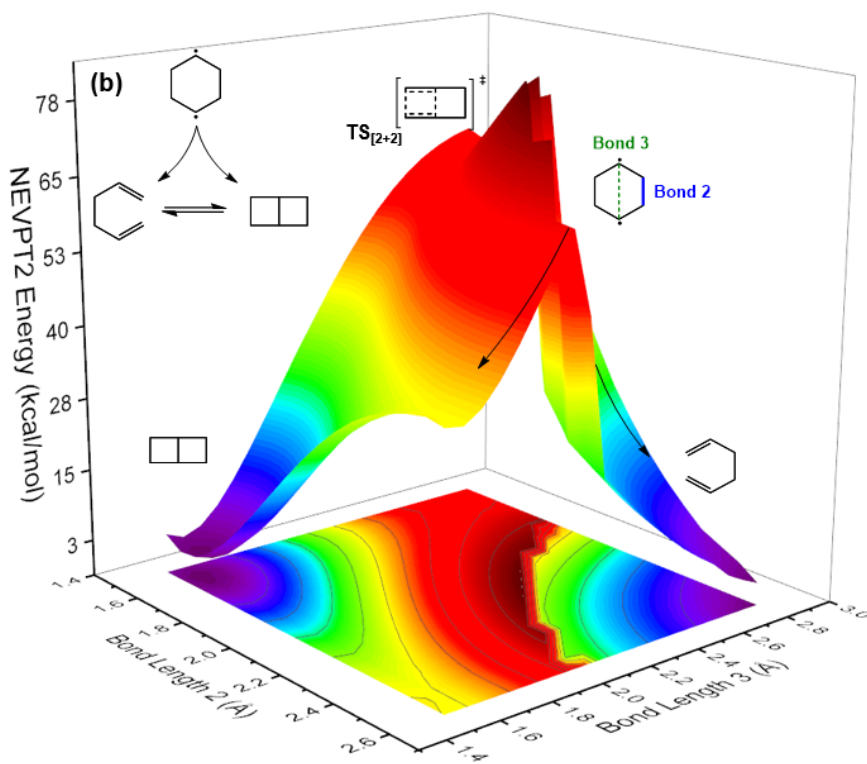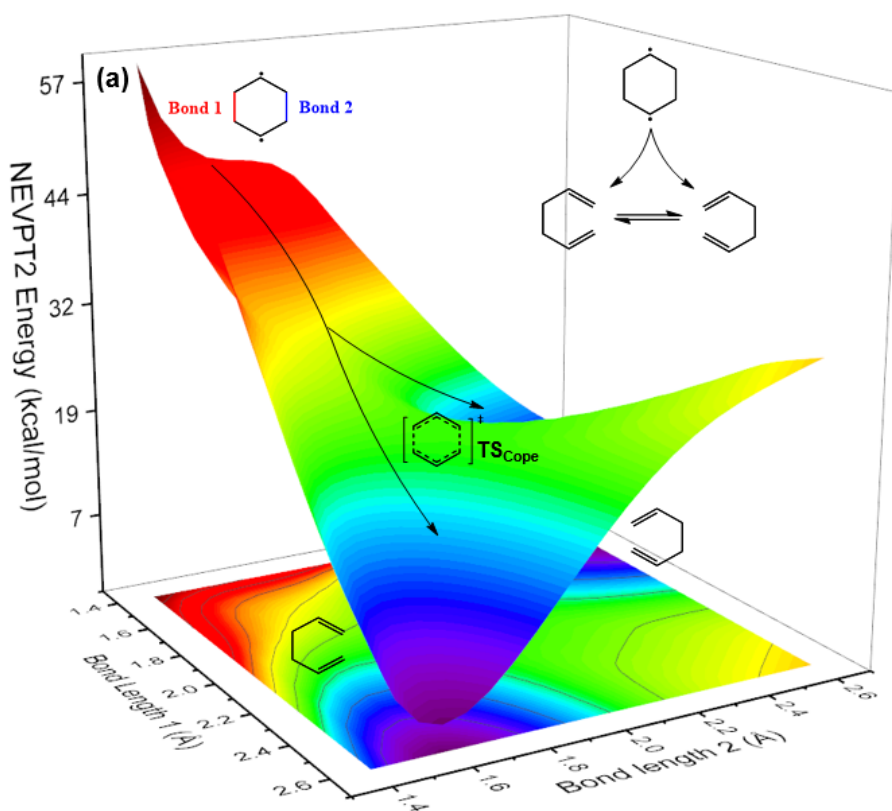
Figure 4.5. S_0-PES adjacent to the 1,4-cyclohexyl diradical under NEVPT2(6,6)/def2-TZVPP//SS-CASSCF(6,6)/def2-SVP level. (a) 1,4-cyclohexyl diradical and Cope-rearrangement TS (b) 1,4-cyclohexyl diradical and [2+2] cycloaddition TS.

Given the various products formed upon DBO photoexcitation, it seems likely that multiple reaction events occur after the first C–N bond breaking transition state and subsequent flat region of the PES. In that competing transition structures leading to different products do not appear to exist, the rationalization of product distribution is, of course, difficult. Dealing with this scenario is difficult for ground state surfaces,[3,4,8] but is even more difficult here when species must cross between excited and ground state surfaces. And, to make matters worse, according to our CASPT2//MC-PDFT computations and previous literature, the S_0-S_1 energy gap around the diazinyl diradical is small, indicating the feasibility of IC at this critical structure. Taken together, these factors strongly suggest that NAMD simulations are necessary for rationalization of product distributions in such a case.

Here we employed the ML-NAMD approach of Lopez (see 4.7 Methods section and Appendix for details),[31,33–35] initiated from different points along our computed MEP. 3600 trajectories were obtained for each NAMD task. The ML-NAMD results are summarized in Figure 4.6. For the ensemble of trajectories initiated from the C–N dissociation transition structure at S_1 state under 300K (Figure 4.6, left), 45% of the reactive trajectories form the HD product, while 55% form the BCH product, which is consistent with the product distribution observed in the experiment by Grissom *et al*. (Figure 4.6, right).[38,43] Among all the HD-forming trajectories, the *Z,E* isomer is predicted to dominate (76%), again consistent with experiment (83% based on isotopic labeling).[39] In addition, a significant excess of *exo*-product (96%) is observed, consistent with experiment again (83%).[38,43]
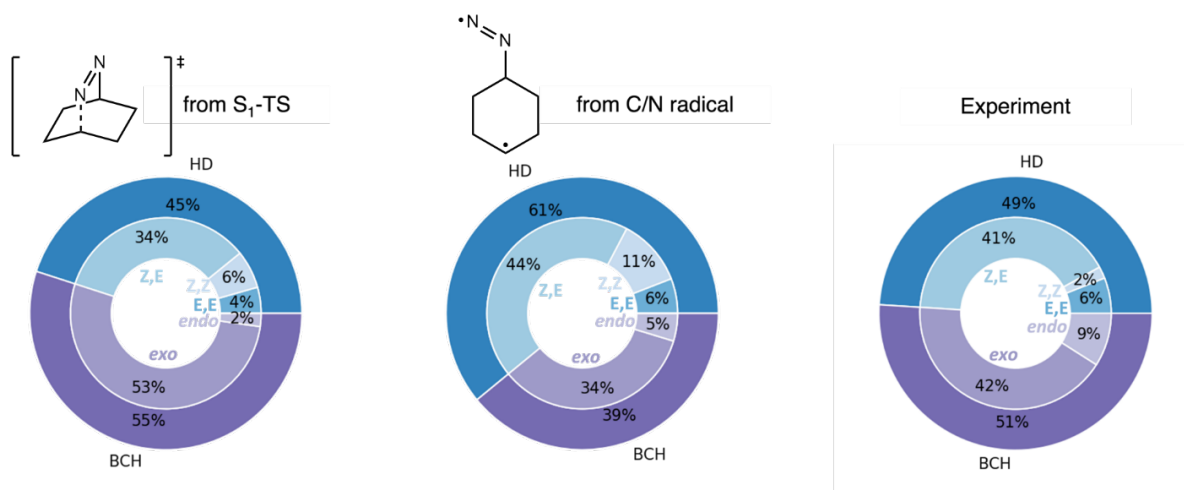
Figure 4.6 Comparison of ML-NAMD (left and center, initiated from different structures) and experimental (right) results.

To investigate how initial momentum and geometry affects the reaction outcome we also ran trajectories at different temperatures and with different initial geometries. The product distribution was found to be almost independent of the temperature and initial geometry on the MEP, except for trajectories initiated from the diazenyl diradical, which display 16% more HD product (Figure 4.6, center). Note that the excess of HD product is not consistent with experiment, implying that a non-statistical dynamic effect associated with approaching the diazenyl diradical specifically from the C–N dissociation transition structure may play an important role.

As expected, we observe that most of the NAMD trajectories hop back to S_0 in the region of the diazinyl diradical, i.e., at the 4-fold crossing region where S_0, S_1, T_1, T_2 states are almost degenerate (Figure 4.4 and Figure C5 in Appendix). Due to the efficient IC rate, subsequent reaction events occur on the ground-state S_0 surface. Thus, we now turn our attention to the issue of momentum effects on the 1,4-cyclohexyl diradical formed after surface hopping.

4.4 Dynamic Properties of the 1,4-cyclohexyl Diradical

Detailed analysis of the trajectories, which lead to an outcome consistent with the known experimental results, allows us to evaluate the means by which momentum influences the conformational dynamics of the 1,4-cyclohexyl diradical thereby determining which products are formed as this diradical is left behind. The fate of the diradical is generally thought to depend on its conformation and spin multiplicity. For example, Grissom and co-workers proposed previously that boat conformations eventually lead to a statistical distribution of singlet and triplet diradicals, which each lead to distinct products.[38] Short-lived boat conformations are thought to lead to the formation of BCH, while long-lived boat conformations undergo ISC to triplets, which then relax to energetically favored twist-boat conformations that lead to HD products. However, *both* HD and BCH products were observed in our NAMD simulations of pure singlets. Although we are not able to completely exclude the possibility of contributions from triplet reactions, the existing conformation-reactivity model appears to be in need of modification.

The conformations of the 1,4-cyclohexyl diradical can be described by two dihedral angles, as illustrated in Figure 4.7 (left). Each dihedral angle has three different operations, CW rotation (the dihedral angle rotates clockwise), CCW rotation (the dihedral angle rotates counterclockwise) and R (the dihedral retains its original value in the DBO reactant, which is close to 0). There are three extreme dihedral angles for these three operations, which we name *up* (U), *down* (D) and *middle* (M). In total, there are nine extreme dihedral angle pairs (Figure 4.7, right). The four structures on the two axes are half-chair conformations (UM,DM,MU,MD), with their two formal singly occupied p-orbitals approximately perpendicular to each other. The chair structures (UD and DU) are at the top right and bottom left corners, while the boat structures (UU and DD) are at the top left and bottom right. The transition state structure for the first C–N bond dissociation, which resides near the center of this plot, is MM.
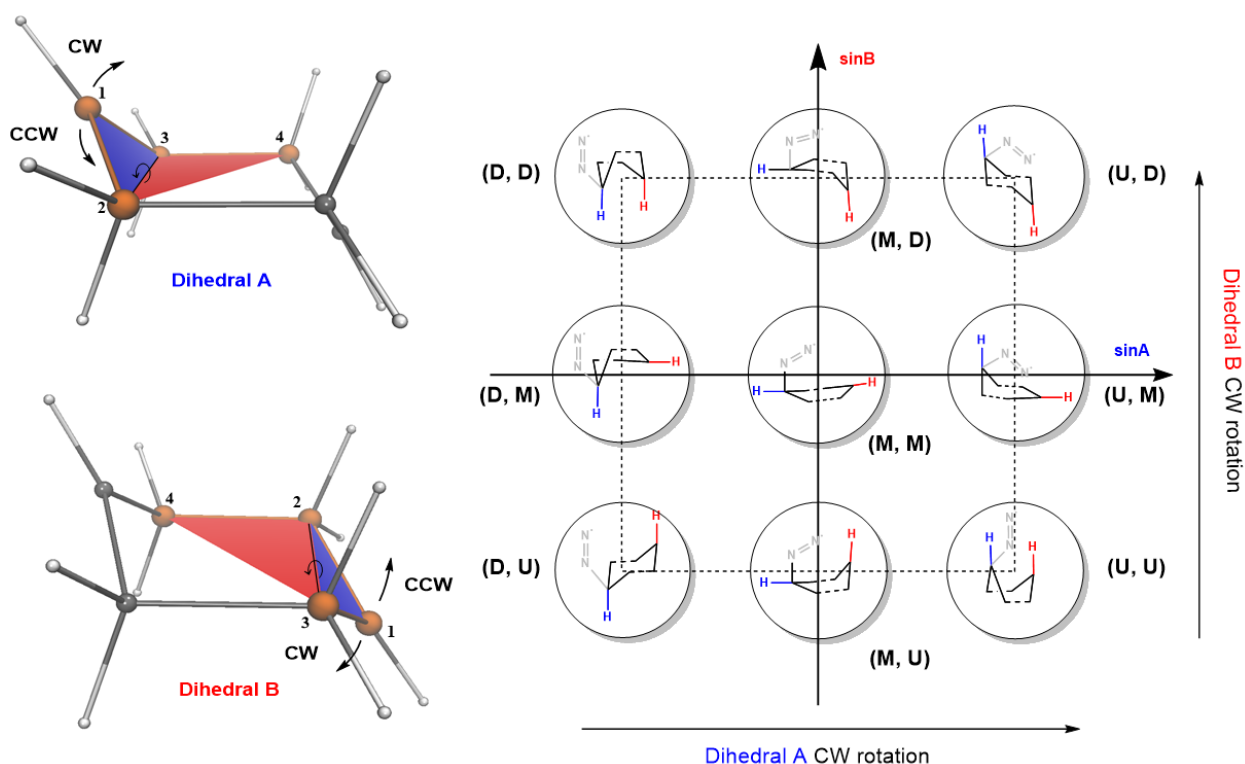
Figure 4.7 Graphical depiction of the dihedral angle space of the 1,4-cyclohexyl diradical. For a better demonstration of how the second C-N bond dissociation affects dihedral angle A, the position of the diaza group from the previous step is shown.

To address the relationship between the diradical conformation and the product distribution, we analyzed the dihedral angle evolution along trajectories initiated from the transition structure for breaking the first C–N bond. To distinguish the two reaction pathways (toward BCH and toward HD), we plotted the population difference of the dihedral angles for a given timeframe (Figure 4.8). The excess of the products is color-coded (red indicates an excess of HD and blue indicates an excess of BCH). As the first C–N bond cleaves, dihedral B acquires more freedom to spread out along the y-axis: depending on the degree of dissociation, dihedral B can do a CW rotation to approach the chair-like conformation (top right, UD), persist in a half-boat-like conformation (middle right, UM) or do a slight CCW rotation (bottom right, UD). As is evident from the separation of the blue and red regions in the first frame, the initial

90

geometry is correlated with the final product distribution. Along the propagation of the trajectories, starting from ~100 fs, the bifurcation of the reaction pathways leading to BCH and HD products becomes clear. Additionally, we have observed 'flips' of blue and red regions in the first 1800 a.u. (43 fs), which indicates "recoil" momentum arising from breaking the first C–N bond (more detailed illustrations with shorter time steps are displayed in SI). In the first 40 fs, the recoil momentum causes dihedral angle B to vibrate back and forth while dihedral angle A is restrained by the second C–N bond.
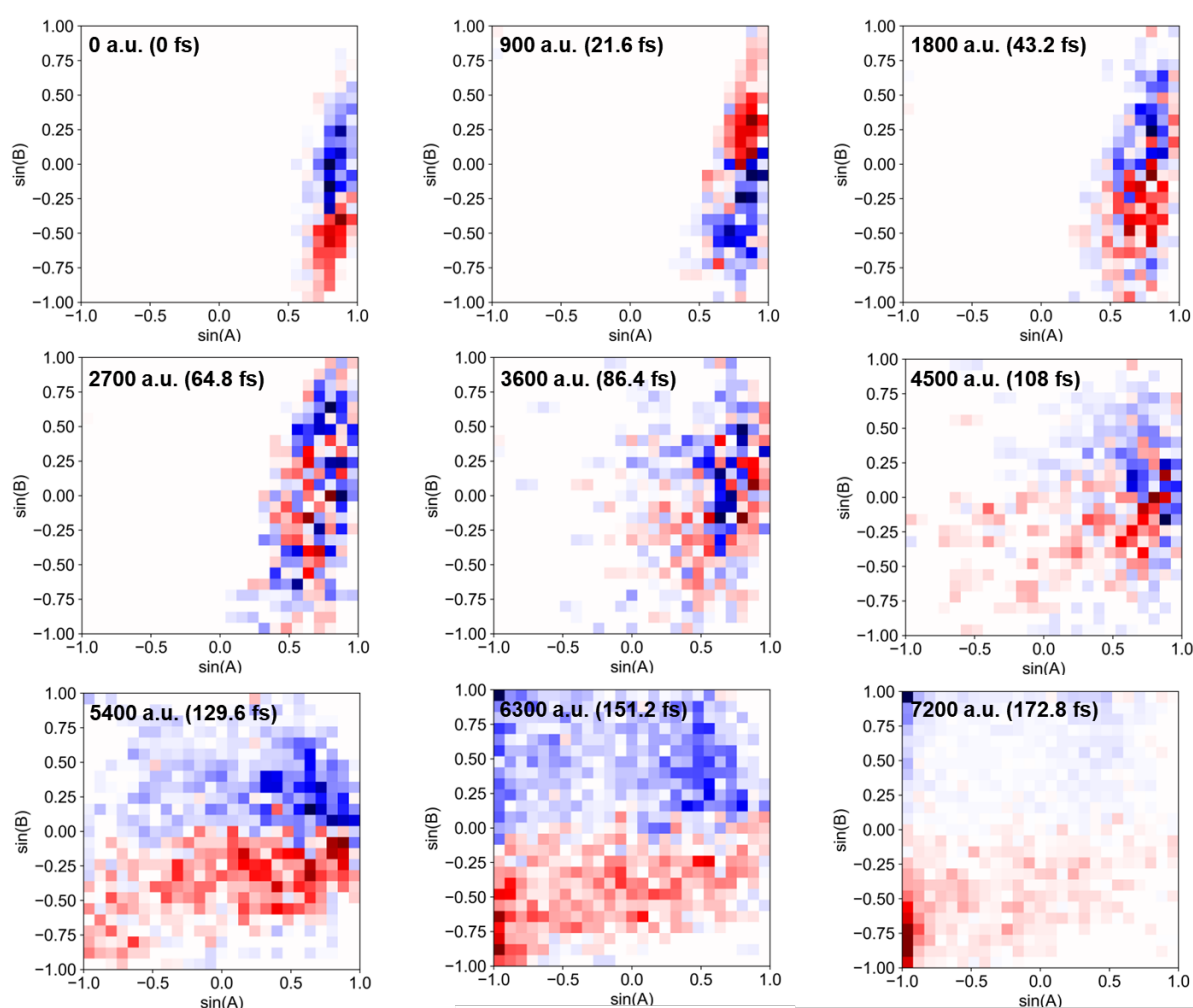
Figure 4.8 Population difference of the dihedral angle for the NAMD trajectories. To visualize, we define the population ρ as the number of trajectories passing through the dihedral angle grid at one certain timestep. The population difference is the subtraction of the number of BCH-forming trajectories from the HD-forming trajectories ($\Delta\rho = \rho_{BCH} - \rho_{HD}$) gives the population difference between the

91

two groups of trajectories. Red color indicates that the number of HD-forming trajectories exceeds that of BCH-forming trajectories and vice versa. The population difference is scaled to [-1,1] and a two-slope normalization is employed. White color indicates that either the number of BCH-forming and HD forming trajectories are equal, or no trajectories populate this grid. (Due to the fact that the min and max of the population difference is not is not symmetrically distributed on both sides of zero and in each picture, the min and max value are not the same, we need to scale the plots with different parameters to make sure that white color always indicates an equal amount of BCH and HD trajectories ($\Delta\rho = \rho_{BCH} - \rho_{HD} = 0$). To do so we map the population number difference $\Delta\rho_a$ on a grid to $2\frac{\Delta\rho_a - \Delta\rho_{min}}{\Delta\rho_{max} - \Delta\rho_{min}} - 1$. By doing so, all mapped data then fall into the range of [-1,1]. The equivalent point (equal amount of BCH and HD population) is mapped to $Eq = \frac{-2\Delta\rho_{min}}{\Delta\rho_{max} - \Delta\rho_{min}} - 1$. However, Eq is still not the middle point, because generally $|\Delta\rho_{max}| \neq |\Delta\rho_{min}|$. Thus, we employed a two-slope normalization of the data. The conceptual center, Eq, is placed at the center of the color bar and the two sides are normalized differently.

The second C–N bond dissociates on a much shorter timescale, as the diazinyl diradical sits on the top of an exothermic downhill pathway (Figure 4.5). The resulting recoil momentum from the ejection of the N_2 molecule leads to a CCW rotation of dihedral angle A. In the population difference plots, this manifests as an exploration of values along the -x direction. Due to the exothermicity and the absence of minima along the pathway, there is effectively no time for IVR. Consequently, for the majority of the trajectories, dihedral A undergoes a CCW rotation, producing *exo*-BCH and *Z,E*-HD (note that we overestimate the *exo:endo* selectivity. Figure 4.6). If dihedral B also initially undergoes a CCW rotation, a chair-like conformation is formed. The chair-like conformation of the 1,4-cyclohexyl diradical closely resembles the lower energy conformation of the transition structure for [3,3] sigmatropic shift,[28,29] allowing for formation of HD products by just small elongation of one of the C–C bonds. In contrast, if

92

dihedral B initially undergoes a CW rotation, and dihedral A undergoes a CCW rotation, a double-inverted boat-like conformation that quickly falls to *endo*-BCH is expected.

In short, during the breaking of the second C–N bond to break, the carbon that will bear the unpaired spin "flips away" from the nitrogen – this motion leads to the preferred stereoisomers of products. During that time, the distal carbon radical center can flip twice. If, at the transition state, the distal carbon was moving in the same direction as the N-bearing carbon, then the two radical carbons end up flipped away from each other as the 1,4-cyclohexyl diradical forms a chair-like structure, which tends to form HD products. Conversely, if the distal carbon was moving in the opposite direction as the N-bearing carbon, then the two radical carbons end up near to each other as the 1,4-cyclohexyl diradical forms a boat-like structure, allowing those radical centers to engage in bond formation, leading to BCH products.

4.5 Multiple Exit Pathways from the 1,4-cyclohexyl Diradical

The network of pathways from the 1,4-cyclohexyl diradical to BCH and HD products is complex. Two views of this network are shown in Figure 4.9a and b. Figure 4.9a shows a 3D scatter plot derived from the scanning the three critical forming/breaking C–C bonds that relate the BCH and HD products. The discontinuities in energy arise from dimensionality reduction, in particular, from changing dihedral angles. The three products – BCH and two equivalent HDs – are located in the three regions of low energy colored blue. Separating each pair of products, there is a ridge-like region of relatively high energy colored red. All of the diradical conformations discussed above reside at the center of the plot, as the bond length differences among these conformations are very small. Due to small energies required for conformational changes, this region is fairly flat. Another view of this network is shown in Figure 4.9b, where connectivity between key structures is emphasized. For example, the ridge-like regions in the 3D scatter plot correspond to the pathways 4 and 8 (for [2+2] cycloaddition, Figure 4.9d) and

pathway 9 (for [3,3] sigmatropic shift, Figure 4.9c) in the connectivity plot. As shown in the trajectories in Figure 4.9c and d, the diradicals move along one of the three ridges and fall to one of the products, depending on their initial geometries and momentum. Specifically for the [3,3] sigmatropic shift branch, the trajectories explore different sides of the ridge before committing to one of the products, similar to behavior observed for some reactions with ground state PTSBs[4–7,9]. Also of note is that some of the trajectories in Figure 4.9d explore the diagonal region, which corresponds to the dissociation of the 1,4 cyclohexyl diradical into two allylic radicals[28]. In our original design, there would have been a single ridge and two downhill paths to the two HD products (Figure 4.1c), but we find here paths to the BCH product as well.
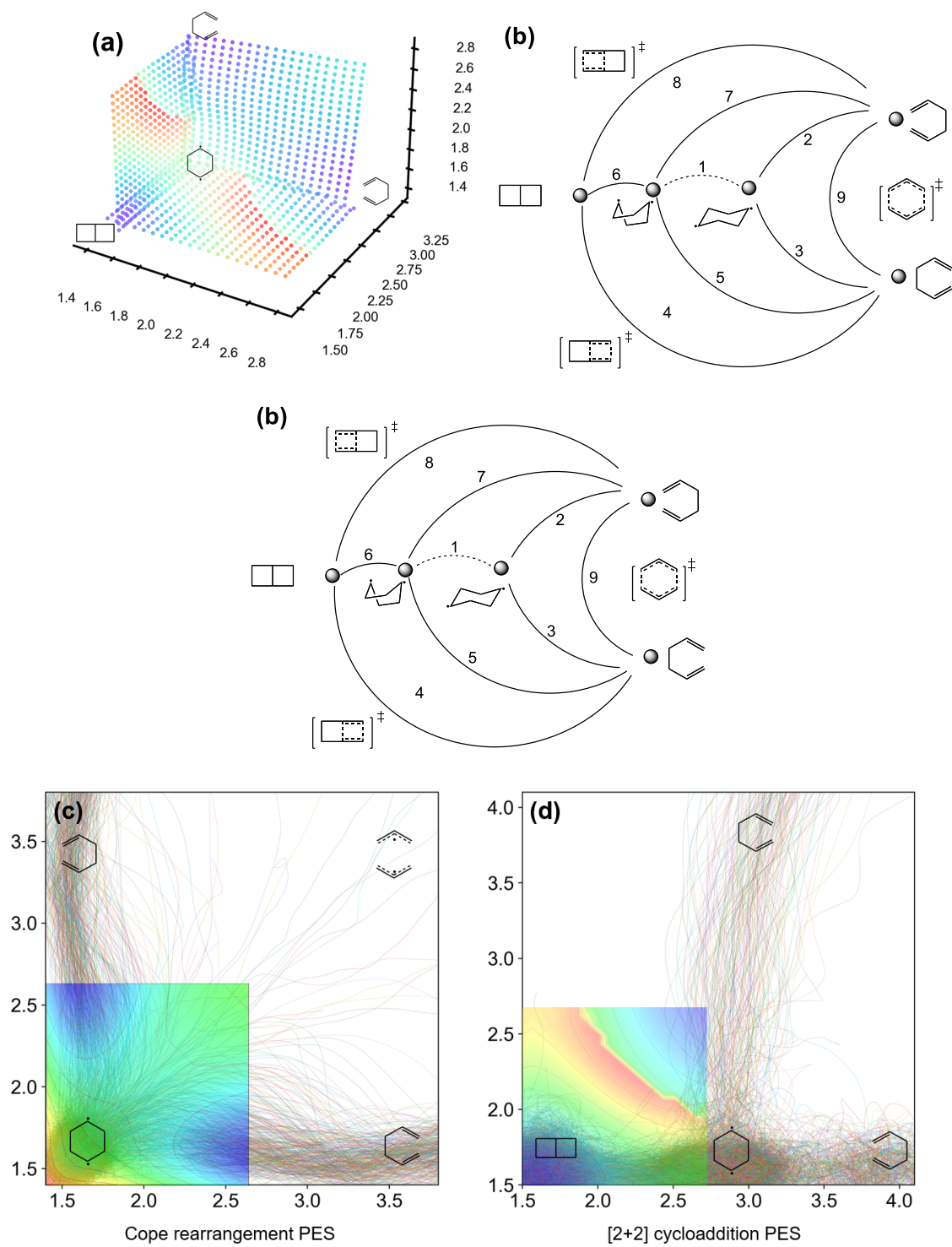
Figure 4.9 (a) 3D-PES adjacent to 1,4-cyclohexyl diradical with energies indicated by color. Energies are in kcal/mol. (b) Reaction network of cyclohexyl diradical (c) Cope-rearrangement branch of the PES with trajectories. (d) (2+2)-cycloaddition branch of the PES with trajecotories.

Previously we showed, for carbocation rearrangements, that ambimodal transition states can correlate with each other, being connected by relatively flat, low energy pathways that allow leakage of trajectories from one to the other.[11] The scenario just described and summarized in Figure 4.9 constitutes a three-dimensional example in terms of PES topography, although access to the flat inter-connecting zone comes from an excited state surface in this case (Figure 4.10).
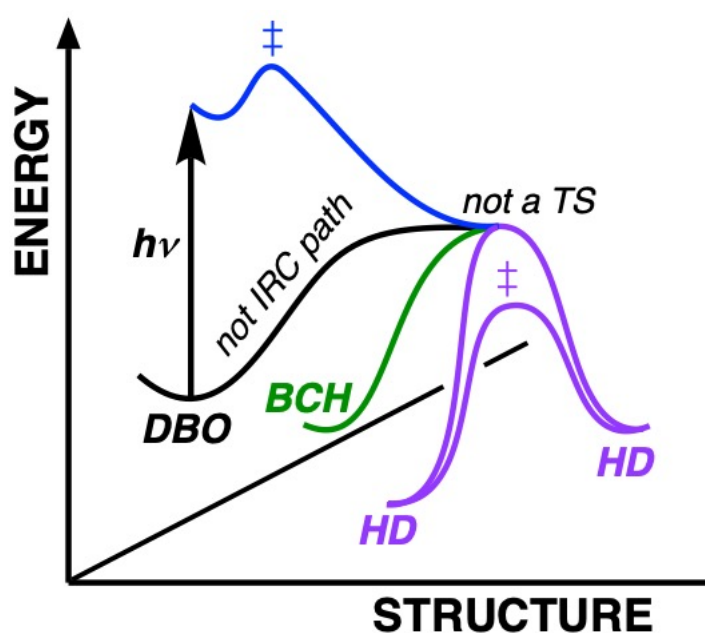
Figure 4.10. The scenario observed for DBO photolysis. The paths to at least two of the products must actually not be in the plane of the picture.

4.6 Conclusions

Using state-of-the art ML-NAMD simulations, we have answered several lingering questions about a classical photochemical reaction: deazetization of DBO. Perhaps most importantly, we propose a new detailed model that rationalizes the product distribution based on non-statistical dynamic (i.e., momentum, dynamic matching) effects on conformational behavior of 1,4-cyclohexyl diradicals formed via hopping from the S1 surface. In short, if the momentum as the 1,4-cyclohexyl diradical is reached is such that rotation of dihedral B leads

to UM or UD conformations (Figure 7), HD products dominate. Otherwise, BCH products form. The *exo:endo* ratio appears to be determined by momentum associated with dihedral A. We do not need to invoke triplets to rationalize the experimental product distribution.

This project was initiated as a means of testing the design principle outlined in Figure 4.1c: access to a transition state connecting two possible products from an excited state surface could allow control of product ratios by non-statistical dynamic effects, in analogy to ground state reactions involving PTSBs (Figure 4.1b). We demonstrate here that photochemical deazetization of 2,3-diazabicyclo[2.2.2]oct-2-ene does involve hopping from an excited state surface to a ground state surface near to a product-interconverting transition state, and the manner in which trajectories navigate from the excited state to the products is influenced by non-statistical dynamic effects – results that bode well for future design. The reaction studied here is even more complex, however, with exit routes to ladderane (BCH) products; one could view the presence of an additional competing product as a problem for design or an opportunity to tune for said product. We are excited about the latter possibility.

4.7 Methods

MC-PDFT, CASPT2 and MS-CASSCF calculations were carried out using OpenMolcas 21.[74,75] NEVPT2 and SS-CASSCF calculations were carried out with ORCA 5.0.3[76–81] Unless otherwise stated, all calculations were performed with MC-PDFT (revPBE)[55,56] and the ANO-S-VDZP basis set.[57–60] An active space of 8 orbitals and 8 electrons was chosen for $C_6H_{10}N_2$ species, and 4 electrons and 4 orbitals for 1,4-cyclohexyl diradicals. Multiwfn software was used for the visualization of the active orbitals.[82] For the ML-accelerated NAMD, we followed the general procedure developed by Lopez and coworkers for dataset generation. Model training NAMD was performed with *PyRAI²MD*.[31,33,35] The model was trained with dataset containing 9670 structures and it displays small error for energies

(MAE=0.049 eV, R^2= 0.9987) and gradients (MAE=0.15 eV/Å, R^2= 0.9923). For each simulation, 3600 trajectories were initiated from a Wigner sampling of the transition state and the trajectories were propagated with a step size of 20 a.u. (0.48 fs). The total energy is conserved in the S_1 state and the Nosé-Hoover thermostat is applied in the S_0 state (T=300K). The surface hopping probability was computed based on the Zhu-Nakamura theory[83] (see Appendix for details of the model training, grid searching, performance and NAMD simulation).

4.8 Acknowledgement

4.9 References

1. Rehbein, J. & Carpenter, B. K. Do we fully understand what controls chemical selectivity? *Physical Chemistry Chemical Physics* **13**, 20906 (2011).

2. Ess, D. H. *et al.* Bifurcations on Potential Energy Surfaces of Organic Reactions. *Angewandte Chemie International Edition* **47**, 7592–7601 (2008).

3. Martin-Somer, A., Xue, X.-S., Jamieson, C. S., Zou, Y. & Houk, K. N. Computational Design of a Tetrapericyclic Cycloaddition and the Nature of Potential Energy Surfaces with Multiple Bifurcations. *J Am Chem Soc* (2023) doi:10.1021/jacs.2c12871.

4. Campos, R. B. & Tantillo, D. J. Designing Reactions with Post-Transition-State Bifurcations: Asynchronous Nitrene Insertions into C–C σ Bonds. *Chem* **5**, 227–236 (2019).

5. Hare, S. R. & Tantillo, D. J. Post-transition state bifurcations gain momentum – current state of the field. *Pure and Applied Chemistry* **89**, 679–698 (2017).

6. Guo, W., Hare, S. R., Chen, S.-S., Saunders, C. M. & Tantillo, D. J. C–H Insertion in Dirhodium Tetracarboxylate-Catalyzed Reactions despite Dynamical Tendencies toward Fragmentation: Implications for Reaction Efficiency and Catalyst Design. *J Am Chem Soc* **144**, 17219–17231 (2022).

7. Hare, S. R. & Tantillo, D. J. Cryptic post-transition state bifurcations that reduce the efficiency of lactone-forming Rh-carbenoid C–H insertions. *Chem Sci* **8**, 1442–1449 (2017).

8. Hare, S. R. & Tantillo, D. J. Dynamic behavior of rearranging carbocations – implications for terpene biosynthesis. *Beilstein Journal of Organic Chemistry* **12**, 377–390 (2016).

9. Feng, Z. & Tantillo, D. J. Dynamic Effects on Migratory Aptitudes in Carbocation Reactions. *J Am Chem Soc* **143**, 1088–1097 (2021).

10. Bai, M., Feng, Z., Li, J. & Tantillo, D. J. Bouncing off walls – widths of exit channels from shallow minima can dominate selectivity control. *Chem Sci* **11**, 9937–9944 (2020).

11. Nieves-Quinones, Y. & Singleton, D. A. Dynamics and the Regiochemistry of Nitration of Toluene. *J Am Chem Soc* **138**, 15167–15176 (2016).

12. Roytman, V. A. & Singleton, D. A. Solvation Dynamics and the Nature of Reaction Barriers and Ion-Pair Intermediates in Carbocation Reactions. *J Am Chem Soc* **142**, 12865–12877 (2020).

13. Carpenter, B. K. Energy Disposition in Reactive Intermediates. *Chem Rev* **113**, 7265–7286 (2013).

14. Palmer, I. J., Ragazos, I. N., Bernardi, F., Olivucci, M. & Robb, M. A. An MC-SCF study of the S1 and S2 photochemical reactions of benzene. *J Am Chem Soc* **115**, 673–682 (1993).

15. Van der Lugt, W. Th. A. M. & Oosterhoff, L. J. Symmetry control and photoinduced reactions. *J Am Chem Soc* **91**, 6042–6049 (1969).

16. Bernardi, F., Olivucci, M. & Robb, M. A. Potential energy surface crossings in organic photochemistry. *Chem Soc Rev* **25**, 321 (1996).

17. Mai, S. & González, L. Molecular Photochemistry: Recent Developments in Theory. *Angewandte Chemie International Edition* **59**, 16832–16846 (2020).

18. Carpenter, B. K. Dynamic Behavior of Organic Reactive Intermediates. *Angewandte Chemie International Edition* **37**, 3340–3350 (1998).

19. Carpenter, B. K. Dynamic Matching: The Cause of Inversion of Configuration in the [1,3] Sigmatropic Migration? *J Am Chem Soc* **117**, 6336–6344 (1995).

20. Carpenter, B. K. Trajectories through an intermediate at a fourfold branch point. Implications for the stereochemistry of biradical reactions. *J Am Chem Soc* **107**, 5730–5732 (1985).

21. Reyes, M. B., Lobkovsky, E. B. & Carpenter, B. K. Interplay of Orbital Symmetry and Nonstatistical Dynamics in the Thermal Rearrangements of Bicyclo[*n* .1.0]polyenes. *J Am Chem Soc* **124**, 641–651 (2002).

22. Oyola, Y. & Singleton, D. A. Dynamics and the Failure of Transition State Theory in Alkene Hydroboration. *J Am Chem Soc* **131**, 3130–3131 (2009).

23. Bailey, J. O. & Singleton, D. A. Failure and Redemption of Statistical and Nonstatistical Rate Theories in the Hydroboration of Alkenes. *J Am Chem Soc* **139**, 15710–15723 (2017).

24. Quijano, L. M. M. & Singleton, D. A. Competition between Reaction and Intramolecular Energy Redistribution in Solution: Observation and Nature of Nonstatistical Dynamics in the Ozonolysis of Vinyl Ethers. *J Am Chem Soc* **133**, 13824–13827 (2011).

25. Kuan, K.-Y. & Singleton, D. A. Vibrationally Hot and Cold Triplets. Sensitizer-Dependent Dynamics and Localized Vibrational Promotion of a Di-π-methane Rearrangement. *J Am Chem Soc* **142**, 19885–19888 (2020).

26. Staroverov, V. N. & Davidson, E. R. Diradical Character of the Cope Rearrangement Transition State. *J Am Chem Soc* **122**, 186–187 (2000).

27. E. Doering, W. von & Roth, W. R. The overlap of two allyl radicals or a four-centered transition state in the cope rearrangement. *Tetrahedron* **18**, 67–74 (1962).

28. Li, J., Stein, R. & Lopez, S. A. A Theoretical Stereoselectivity Model of Photochemical Denitrogenations of Diazoalkanes Toward Strained 1,3-Dihalogenated Bicyclobutanes. *J Org Chem* **86**, 4061–4070 (2021).

29. Li, J. *et al.* Automatic discovery of photoisomerization mechanisms with nanosecond machine learning photodynamics simulations. *Chem Sci* **12**, 5302–5314 (2021).

30. Cox, J. M., Bain, M., Kellogg, M., Bradforth, S. E. & Lopez, S. A. Role of the Perfluoro Effect in the Selective Photochemical Isomerization of Hexafluorobenzene. *J Am Chem Soc* **143**, 7002–7012 (2021).

31. Li, J., Stein, R., Adrion, D. M. & Lopez, S. A. Machine-Learning Photodynamics Simulations Uncover the Role of Substituent Effects on the Photochemical Formation of Cubanes. *J Am Chem Soc* **143**, 20166–20175 (2021).

32. Li, J. & Lopez, S. A. Excited-State Distortions Promote the Photochemical 4π-Electrocyclizations of Fluorobenzenes via Machine Learning Accelerated Photodynamics Simulations. *Chemistry – A European Journal* **28**, (2022).

33. Li, J. & Lopez, S. A. A Look Inside the Black Box of Machine Learning Photodynamics Simulations. *Acc Chem Res* **55**, 1972–1984 (2022).

34. Schmidt, J. R., Parandekar, P. V. & Tully, J. C. Mixed quantum-classical equilibrium: Surface hopping. *J Chem Phys* **129**, 044104 (2008).

35. Engel, P. S., Nalepa, C. J., Soltero, L. R., Horsey, D. W. & Keys, D. E. Photolysis of reluctant azoalkanes. Effect of structure on photochemical loss of nitrogen from 2,3-diazabicyclo[2.2.2]oct-2-ene derivatives. *J Am Chem Soc* **105**, 7108–7114 (1983).

36. Anderson, M. A. & Grissom, C. B. Photolysis of 2,3-Diazabicyclo[2.2.2]oct-2-ene: Electronic Spin Determines the Distribution of Products. *J Am Chem Soc* **117**, 5041–5048 (1995).

37. Edmunds, A. J. F. & Samuel, C. J. Photochemical deazetation of 2,3-diazabicyclo[2.2.2]oct-2-ene: pseudorotation of the cyclohexanediyl biradical. *J Chem Soc Perkin 1* 1267 (1989) doi:10.1039/p19890001267.

38. Lodewyk, M. W. *et al.* The Correct Structure of Aquatolide—Experimental Validation of a Theoretically-Predicted Structural Revision. *J Am Chem Soc* **134**, 18550–18553 (2012).

39. Nouri, D. & Tantillo, D. They Came From the Deep: Syntheses, Applications, and Biology of Ladderanes. *Curr Org Chem* **10**, 2055–2074 (2006).

40. Epplin, R. C. *et al.* [2]-Ladderanes as isosteres for meta-substituted aromatic rings and rigidified cyclohexanes. *Nat Commun* **13**, 6056 (2022).

41. Anderson, M. A. & Grissom, C. B. Increasing the Heavy Atom Effect of Xenon by Adsorption to Zeolites: Photolysis of 2,3-Diazabicyclo[2.2.2]oct-2-ene. *J Am Chem Soc* **118**, 9552–9556 (1996).

42. Engel, P. S., Hayes, R. A., Keifer, L., Szilagyi, S. & Timberlake, J. W. Extrusion of nitrogen from cyclic and bicyclic azo compounds. *J Am Chem Soc* **100**, 1876–1882 (1978).

43. Engel, P. S. & Nalepa, C. J. Photochemical decomposition and isomerization of aliphatic azo compounds. *Pure and Applied Chemistry* **52**, 2621–2632 (1980).

44. Chen, H. & Li, S. Theoretical Study on the Photolysis Mechanism of 2,3-Diazabicyclo[2.2.2]oct-2-ene. *J Am Chem Soc* **127**, 13190–13199 (2005).

45. Roberson, M. J. & Simons, J. Ab Initio Study of the Mechanism of Photolytic Deazatization of 2,3-Diazabicyclo[2.2.2]oct-2-ene and 2,3-Diazabicyclo[2.2.1]hept-2-ene. *J Phys Chem A* **101**, 2379–2383 (1997).

46. Doubleday, C., Armas, R., Walker, D., Cosgriff, C. V. & Greer, E. M. Heavy-Atom Tunneling Calculations in Thirteen Organic Reactions: Tunneling Contributions are Substantial, and Bell's Formula Closely Approximates Multidimensional Tunneling at ≥250 K. *Angewandte Chemie International Edition* **56**, 13099–13102 (2017).

47. Chen, D., Xu, G., Zhou, Q., Chung, L. W. & Tang, W. Practical and Asymmetric Reductive Coupling of Isoquinolines Templated by Chiral Diborons. *J Am Chem Soc* **139**, 9767–9770 (2017).

48. Sponsler, M. B., Jain, R., Coms, F. D. & Dougherty, D. A. Matrix-isolation decay kinetics of triplet cyclobutanediyls. Observation of both Arrhenius behavior and heavy-atom tunneling in carbon-carbon bond-forming reactions. *J Am Chem Soc* **111**, 2240–2252 (1989).

49. Buchwalter, S. L. & Closs, G. L. Electron spin resonance and CIDNP studies on 1,3-cyclopentadiyls. A localized 1,3 carbon biradical system with a triplet ground state. Tunneling in carbon-carbon bond formation. *J Am Chem Soc* **101**, 4688–4694 (1979).

50. Mignolet, B., Curchod, B. F. E. & Martínez, T. J. Rich Athermal Ground-State Chemistry Triggered by Dynamics through a Conical Intersection. *Angewandte Chemie* **128**, 15217–15220 (2016).

51. Hoyer, C. E., Ghosh, S., Truhlar, D. G. & Gagliardi, L. Multiconfiguration Pair-Density Functional Theory Is as Accurate as CASPT2 for Electronic Excitation. *J Phys Chem Lett* **7**, 586–591 (2016).

52. Gagliardi, L. *et al.* Multiconfiguration Pair-Density Functional Theory: A New Way To Treat Strongly Correlated Systems. *Acc Chem Res* **50**, 66–73 (2017).

53. Li Manni, G. *et al.* Multiconfiguration Pair-Density Functional Theory. *J Chem Theory Comput* **10**, 3669–3680 (2014).

54. Sharma, P., Bao, J. J., Truhlar, D. G. & Gagliardi, L. Multiconfiguration Pair-Density Functional Theory. *Annu Rev Phys Chem* **72**, 541–564 (2021).

55. Widmark, P.-O., Persson, B. J. & Roos, B. O. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. *Theor Chim Acta* **79**, 419–432 (1991).

56. Pou-Amérigo, R., Merchán, M., Nebot-Gil, I., Widmark, P.-O. & Roos, B. O. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. *Theor Chim Acta* **92**, 149–181 (1995).

57. Widmark, P.-O., Malmqvist, P.-�ke & Roos, B. O. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. *Theor Chim Acta* **77**, 291–306 (1990).

58. Pierloot, K., Dumez, B., Widmark, P.-O. & Roos, B. O. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. *Theor Chim Acta* **90**, 87–114 (1995).

59. Finley, J., Malmqvist, P.-Å., Roos, B. O. & Serrano-Andrés, L. The multi-state CASPT2 method. *Chem Phys Lett* **288**, 299–306 (1998).

60. Penfold, T. J., Gindensperger, E., Daniel, C. & Marian, C. M. Spin-Vibronic Mechanism for Intersystem Crossing. *Chem Rev* **118**, 6975–7025 (2018).

61. Marian, C. M. Spin-orbit coupling and intersystem crossing in molecules. *Wiley Interdiscip Rev Comput Mol Sci* **2**, 187–203 (2012).

62. Mai, S., Marquetand, P. & González, L. A general method to describe intersystem crossing dynamics in trajectory surface hopping. *Int J Quantum Chem* **115**, 1215–1231 (2015).

63. Richter, M., Marquetand, P., González-Vázquez, J., Sola, I. & González, L. SHARC: *ab Initio* Molecular Dynamics with Surface Hopping in the Adiabatic Representation Including Arbitrary Couplings. *J Chem Theory Comput* **7**, 1253–1258 (2011).

64. Mitchell, E. C., Scott, T. R., Bao, J. J. & Truhlar, D. G. Application of Multiconfiguration Pair-Density Functional Theory to the Diels–Alder Reaction. *J Phys Chem A* **126**, 8834–8843 (2022).

65. Calio, P. B., Truhlar, D. G. & Gagliardi, L. Nonadiabatic Molecular Dynamics by Multiconfiguration Pair-Density Functional Theory. *J Chem Theory Comput* **18**, 614–622 (2022).

66. Chen, Z., Nieves-Quinones, Y., Waas, J. R. & Singleton, D. A. Isotope Effects, Dynamic Matching, and Solvent Dynamics in a Wittig Reaction. Betaines as Bypassed Intermediates. *J Am Chem Soc* **136**, 13122–13125 (2014).

67. Wang, L., Akimov, A. & Prezhdo, O. V. Recent Progress in Surface Hopping: 2011–2015. *J Phys Chem Lett* **7**, 2100–2112 (2016).

68. Barbatti, M. Nonadiabatic dynamics with trajectory surface hopping method. *Wiley Interdiscip Rev Comput Mol Sci* **1**, 620–633 (2011).

69. Tully, J. C. Molecular dynamics with electronic transitions. *J Chem Phys* **93**, 1061–1071 (1990).

70. Tully, J. C. Mixed quantum–classical dynamics. *Faraday Discuss* **110**, 407–419 (1998).

71. Zhang, T. *et al.* Investigations of an Unexpected [2+2] Photocycloaddition in the Synthesis of (−)-Scabrolide A from Quantum Mechanics Calculations. *J Org Chem* **87**, 14115–14124 (2022).

72. Aquilante, F. *et al.* Modern quantum chemistry with [Open]Molcas. *J Chem Phys* **152**, 214117 (2020).

73. Fdez. Galván, I. *et al.* OpenMolcas: From Source Code to Insight. *J Chem Theory Comput* **15**, 5925–5964 (2019).

74. Hellweg, A., Hättig, C., Höfener, S. & Klopper, W. Optimized accurate auxiliary basis sets for RI-MP2 and RI-CC2 calculations for the atoms Rb to Rn. *Theor Chem Acc* **117**, 587–597 (2007).

75. Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Physical Chemistry Chemical Physics* **8**, 1057 (2006).

76. Weigend, F. & Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Physical Chemistry Chemical Physics* **7**, 3297 (2005).

77. Neese, F., Wennmohs, F., Hansen, A. & Becker, U. Efficient, approximate and parallel Hartree–Fock and hybrid DFT calculations. A 'chain-of-spheres' algorithm for the Hartree–Fock exchange. *Chem Phys* **356**, 98–109 (2009).

78. Neese, F. Software update: the ORCA program system, version 4.0. *WIREs Computational Molecular Science* **8**, (2018).

79. Neese, F. The ORCA program system. *WIREs Computational Molecular Science* **2**, 73–78 (2012).

80. Lu, T. & Chen, F. Multiwfn: A multifunctional wavefunction analyzer. *J Comput Chem* **33**, 580–592 (2012).

81. Ishida, T., Nanbu, S. & Nakamura, H. Clarification of nonadiabatic chemical dynamics by the Zhu-Nakamura theory of nonadiabatic transition: from tri-atomic systems to reactions in solutions. *Int Rev Phys Chem* **36**, 229–285 (2017).

Part 3

Mechanistic Studies in collaboration with experiments

Chapter 5

Mechanistic Study of the Synthesis of Pupukeanane

Core through a Bio-Inspired Rearrangement.

5.1 Introduction

Recently, the Sarpong group at UC Berkeley reported a formal synthesis of 2-isocyanoallopupukeanane (**3**, Scheme 1A), a downstream secondary metabolite in the pupukeanane family, and demonstrated the utility of the allopupukeanane skeleton in accessing a related scaffold that is biosynthetically upstream.

a) Selected pupukeanane natural products

9-isocyano
pupukeanane (**1**)

9-epi-9-isocyano
pupukeanane (**2**)

2-isocyanoallo
pupukeanane (**3**)

b) Proposed biosynthetic pathway to the allopupukeanyl cation

amorphene (**4**)

5

proposed biosynthesis

6

this work 'contra-biosynthetic'
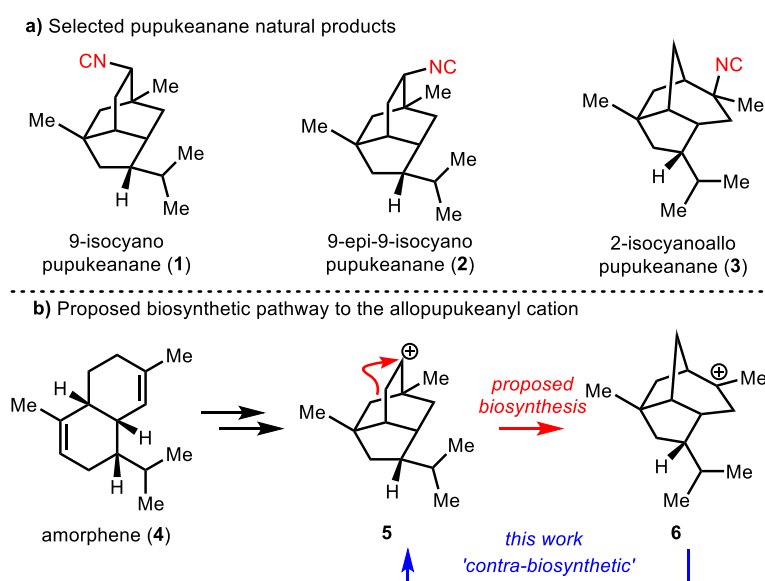
Scheme 5.1 (A) pupukeanane natural products; (B) proposed biosynthesis of the allopupukeanane skeleton (**6**)

The pupukeananes (selected members shown in Scheme 5.1A) are a family of structurally diverse marine-derived sesquiterpenes that possess caged tricyclic skeletons.[1] Many related natural products that possess diverse skeletons, as well as related functional groups,[2] have also been isolated. Notably, there is renewed interest in isocyanoterpenes such as the pupukeananes because of their emerging antimalarial activity which is associated with a novel mechanism of action.[3]
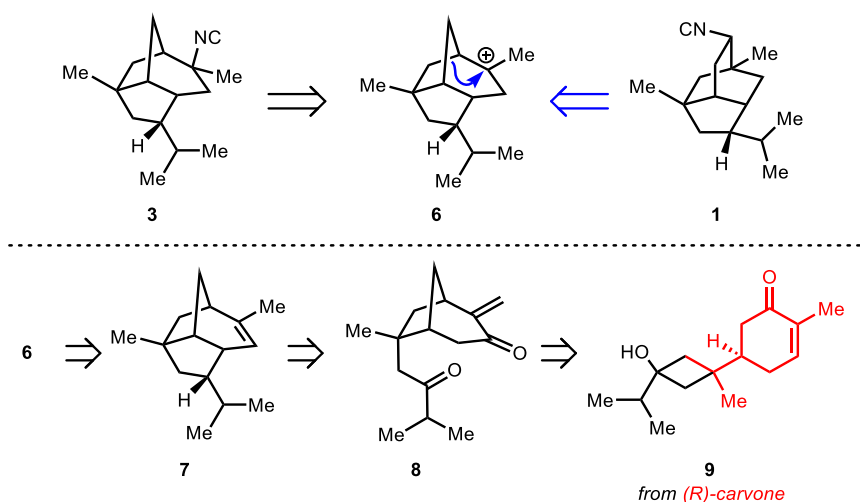
Scheme 5.2. Proposed retrosynthesis of 2-isocyanoallopupukeanane and the pupukeanane core.

Biosynthetically, the pupukeanane skeleton is proposed to arise from amorphene (**4**, Scheme 5.1B), which is converted to the pupukeanyl skeleton (**5**).[4] The 9-pupukeanyl cation (**5**) is proposed to undergo a Wagner–Meerwein type shift to give the 2-allopupukeanyl cation (**6**). Trapping of carbocation intermediates (**5** and **6**) with cyanide[5] would lead to the corresponding natural products 9-isocyanopupukeanane (**1**) and 2-isocyanoallopupukeanane (**3**) respectively. While many approaches to prepare natural products with the pupukeanane core have been reported,[6] only a single total synthesis, of (±)-2-isocyanoallopupukeanane (**3**), has been achieved to date.[7]

The Sarpong group proposed a short synthesis of **3** using strategic C–C bond disconnections that provide the opportunity to access the related pupukeanane natural products (Scheme 5.2). It is anticipated that carbocation **6** would lead to **3** through direct cyanide capture, whereas a late-stage 'contra-biosynthetic' rearrangement of **6** could yield **5** and ultimately 9-isocyanopupukeanane (**1**). The key intermediate, carbocation **6** could arise from corresponding alkene **7**. However, the formation of the pupukeanane core from an allopupukeanyl precursor proved particularly challenging. The Sarpong group observed only decomposition or capture of **6** by the exogenous nucleophile and when allopupukeanene **7** was subjected to H_2SO_4 and

glacial AcOH at elevated temperatures, the formation of small amounts of pupukeanyl sultone **15** is observed (Scheme 5.3). In this regard, computation is necessary to gain further mechanistic insights.
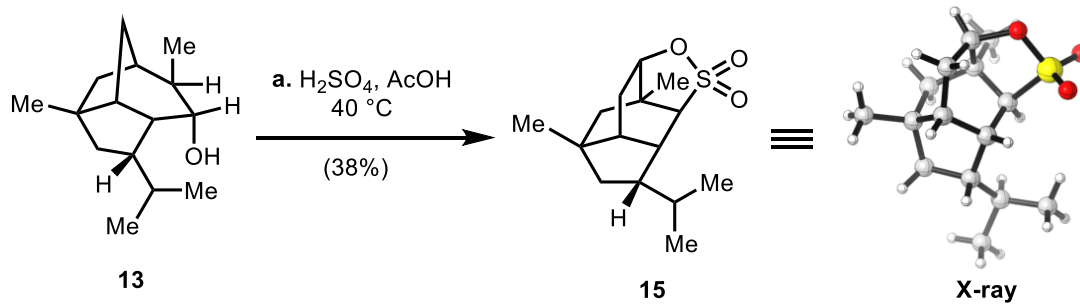
Scheme 5.3 Rearrangement to the pupukeanyl core.
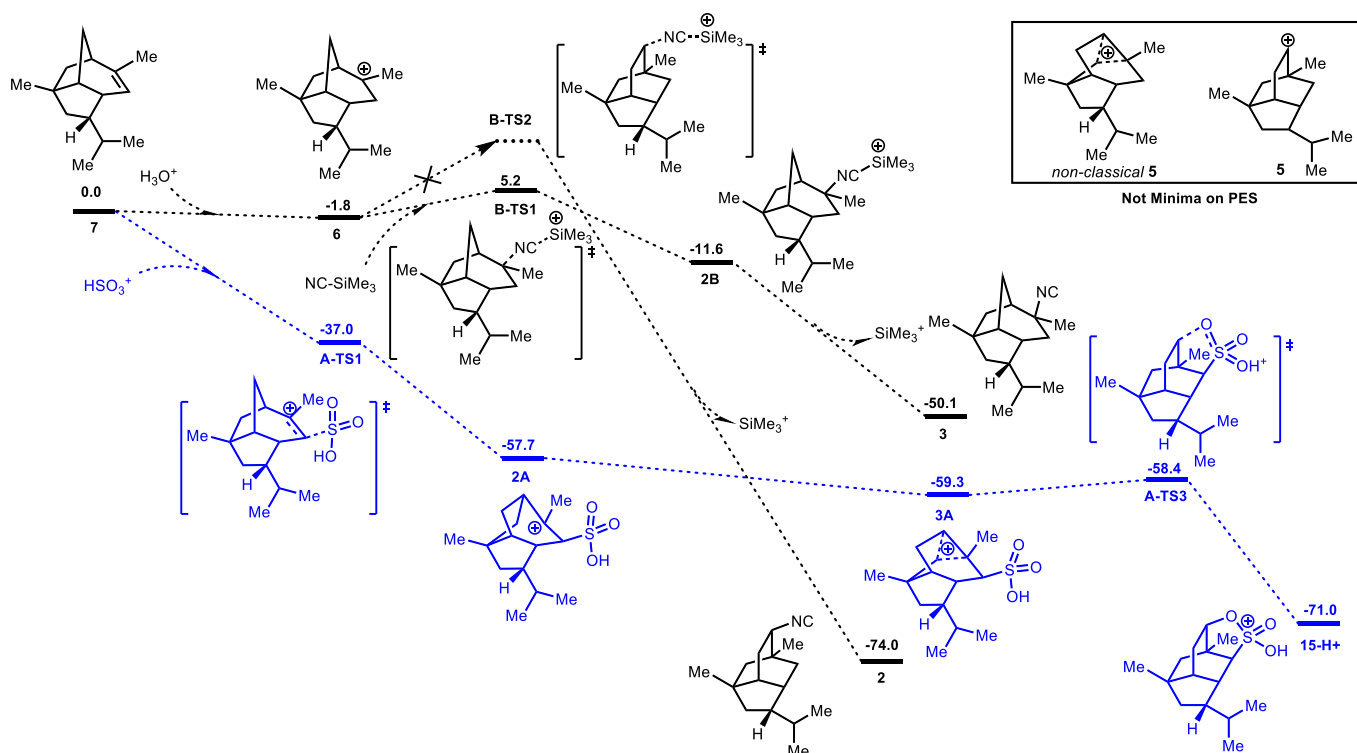
Scheme 5.4 Computational studies on the rearrangement of the allopupukeanane core

5.2 Method

Geometry optimizations were carried out with the Gaussian 16 software and the ωB97X-D(SMD, solvent=acetic acid)/def2-SVP level of theory. Potential energy surface (PES) minima and transition state structures (TSS) were identified by the number of imaginary frequencies obtained in frequency calculations conducted at the same level, with 0 for minima and 1 for TSS.

5.3 Results and Discussion

Calculations were conducted on the capture of **6** with a nucleophilic cyanide source, trimethylsilyl cyanide (Scheme 5.4). The corresponding transition state structure for TMSCN addition was located (**B-TS1**) and, as expected, the direct capture of **6** to give 2-isocyanoallopupukeanane (**3**) was calculated to be both kinetically and thermodynamically feasible. Additionally, the 9-pupukeanyl cation (**5**) was not observed to be a minimum on the potential energy surface. An alternative transition state structure for concerted attack of TMSCN and alkyl shift (see **B-TS2**) could not be located, despite the fact that 9-*epi*-9-isocyanopupukeanane (**2**) was predicted to be thermodynamically more stable than the 2-isocyanoallopupukeanane congener (**3**).

We have also modeled the reaction coordinate for the sulfonation of **7** with an activated sulfur electrophile to form intermediate **2A**. From this intermediate, the rearrangement is nearly barrierless, proceeding through a nonclassical carbocation (**3A**) on a flat region of the potential energy surface to reach protonated pupukeanyl sultone **15-H$^+$** as an energy minimum. Thus, it would appear that while a direct capture of the 9-pupukeanyl cation (**5**) is not feasible, by coupling the generation of **5** with a sulfonation process, it is possible to trap the allopupukeanane skeleton in the form of pupukeanyl sultone **15**. This transformation effectively reverses the proposed biosynthetic pathway to access the pupukeanane core from a downstream biosynthetic intermediate.

5.4 References

1. a) B. J. Burreson, P. J. Scheuer, J. Finer, J. Clardy, *J. Org. Chem.* **1975**, *97*, 4763–4764; b) M. R. Hagadone, B. J. Burreson, P. J. Scheuer, J. S. Finer, J. Clardy, *Helv Chim Acta* **1979**, *62*, 2484–2494; c) N. Fusetani, H. J. Wolstenholme, S. Matsunaga, H. Hirota, *Tetrahedron Lett.* **1991**, *32*, 7291–7294; d) F. Le Bideau, M. Kousara, L. Chen, L. Wei, F. Dumas, *Chem. Rev.* **2017**, *117*, 6110–6159.

2. a) H. Y. He, J. Salva, R. F. Catalos, D. J. Faulkner, *J. Org. Chem.* **1992**, *57*, 3191–3194; b) A. H. Marcus, T. F. Molinski, E. Fahy, D. J. Faulkner, C. Xu, J. Clardy, *J. Org. Chem.* **1989**, *54*, 5184–5186; c) J. S. Simpson, M. J. Garson, J. N. A. Hooper, E. I. Cline, C. K. Angerhofer, *Aust. J. Chem.* **1997**, *50*, 1123–1128; d) Y. Yasman, R. A. Edrada, V. Wray, P. Proksch, *J. Nat. Prod.* **2003**, *66*, 1512–1514.

3. a) H. H. Lu, S. V. Pronin, Y. Antonova-Koch, S. Meister, E. A. Winzeler, R. A. Shenvi, *J. Am. Chem. Soc.* **2016**, *138*, 7268–7271; b) S. V. Pronin, R. A. Shenvi, *J. Am. Chem. Soc.* **2012**, *134*, 19604–19606; c) C. A. Reiher, R. A. Shenvi, *J. Am. Chem. Soc.* **2017**, *139*, 3647–3650; d) M. E. Daub, P. C. Roosen, C. D. Vanderwal, *J. Org. Chem.* **2017**, *82*, 4533–4541; e) M. J. Schnermann, R. A. Shenvi, *Nat. Prod. Rep.* **2015**, *32*, 543–577; f) A. M. White, K. Dao, D. Vrubliauskas, Z. A. Konst, G. K. Pierens, A. Mandi, K. T. Andrews, T. S. Skinner-Adams, M. E. Clarke, P. T. Narbutas, D. C. Sim, K. L. Cheney, T. Kurtan, M. J. Garson, C. D. Vanderwal, *J. Org. Chem.* **2017**, *82*, 13313–13323; g) P. C. Roosen, C. D. Vanderwal, *Angew. Chem. Int. Ed.* **2016**, *55*, 7180–7183; h) A. S. Karns, B. D. Ellis, P. C. Roosen, Z. Chahine, K. G. Le Roch, C. D. Vanderwal, *Angew. Chem. Int. Ed.* **2019**, *58*, 13749–13752; i) M. E. Daub, J. Prudhomme, C. Ben Mamoun, K. G. Le Roch, C. D. Vanderwal, *ACS Med. Chem. Lett.* **2017**, *8*, 355–360; j) M. E. Daub, J. Prudhomme, K. Le Roch, C. D. Vanderwal, *J. Am. Chem. Soc.* **2015**, *137*, 4912–4915.

4. C. H. McCulley, D. J. Tantillo, *J. Phys. Chem. A* **2018**, *122*, 8058–8061.

5. M. J. Garson, *J. Chem. Soc., Chem. Commun.* **1986**, 35.

6. a) M. K. Brown, E. J. Corey, *Org. Lett.* **2010**, *12*, 172–175; b) E. J. Corey, M. Behforouz, M. Ishiguro, *J. Am. Chem. Soc.* **1979**, *101*, 1608–1609; c) H. Yamamoto, H. L. Sham, *J. Am. Chem. Soc.* **1979**, *101*, 1609–1611; d) A. Srikrishna, P. Ravi Kumar, *Tetrahedron Lett.* **2002**, *43*, 1109–1111; e) G. A. Schiehser, J. D. White, *J. Org. Chem.* **1980**, *45*, 1864–1868; f) E. Piers, M. Winter, *Liebigs Ann. Chem.* **1982**, *1982*, 973–984; g) S. L. Hsieh, C. T. Chiu, N. C. Chang, *J. Org. Chem.* **1989**, *54*, 3820–3823; h) N.-C. Chang, C.-K. Chang, *J. Org. Chem.* **1996**, *61*, 4967–4970; i) E. J. Corey, M. Ishiguro, *Tetrahedron Lett.* **1979**, *20*, 2745–2748; j) G. Fráter, J. Wenger, *Helv Chim Acta* **1984**, *67*, 1702–1706; k) K. Kaliappan, G. S. R. S. Rao, *Tetrahedron Lett.* **1997**, *38*, 2185–2186; l) A. J. Smaligo, M. Swain, J. C. Quintana, M. F. Tan, D. A. Kim, O. Kwon, *Science* **2019**, *364*, 681–685; m) A. Srikrishna, S. Danieldoss, *J. Org. Chem.* **1997**, *62*, 7863–7865; n) A. Srikrishna, P. Ravi Kumar, S. J. Gharpure, *Tetrahedron Lett.* **2001**, *42*, 3929–3931.

7. a) T. L. Ho, L. R. Kung, R. J. Chein, *J. Org. Chem.* **2000**, *65*, 5774–5779; b) A. Srikrishna, G. Satyanarayana, P. Ravi Kumar, *Tetrahedron Lett.* **2006**, *47*, 363–366.

Chapter 6

Difunctionalization of Nitrogen-containing Strained Carbocycles: DFT Reveals the Origin of Reactivity and the role of TMSCN.

6.1 Introduction

Distonic radical cation (DRC) is a class of chemical species that carry separate cation and radical sites in one molecule.[1] Traditionally, DRCs have been generated through the homolysis of charged molecules or via rearrangement of radical cations.[2] Since a DRC often entails a reactive radical site with an inert charge site, it has become an important tool for organic chemists to understand radical chemistry via online analysis of the charged species through mass spectrometry.[3] This approach is highly informative on neutral radical reactivities, such as reaction rate and product distribution, provided that the radical site on a DRC behaves similarly to neutral molecules.4

Ring-opening of cyclopropylamines to form DRC is one of the most studied systems due to its biological implication in enzyme inhibition.[4] The Zheng group at the University of Arkansas recently reported intermolecular [3+2] and [4+2] annulations[5] involving DRC under photoredox catalysis.[6] By comparing the DRC structures of *N*-aryl amine and *N*-alkyl amine, they hypothesized that the two types of amines would exhibit similar behaviour at the radical site unless the charged site does affect the reactivity of the radical site, despite being spatially separated from it.[3c, 3e, 7] The reaction is displayed in Scheme 6.1
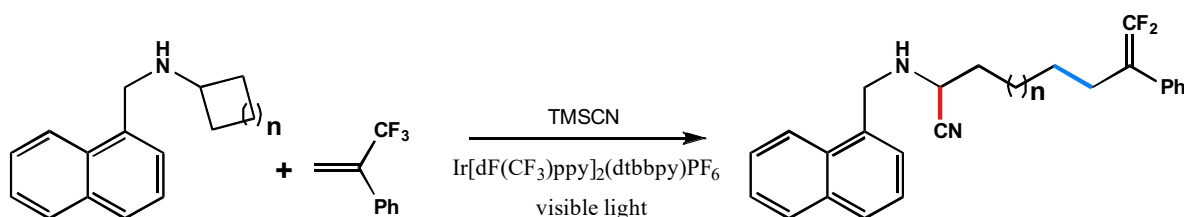
Scheme 6.1 Photoredox catalyzed annulation of *N*-alkyl amines.

Experiments on distinct *N*-alkyl amines and under different reaction conditions (whether TMSCN is added) demonstrated that this is a practical difunctionalization protocol (Scheme 6.2). Initially, TMSCN is added in the hope of quenching the charged site (iminium ion) of the

generated distonic radical cation to unveil a neutral radical for subsequent radical addition to olefins.

However, as can be seen in Scheme 6.2, TMSCN seems to have displayed additional roles more than charged site quencher and we were interested in whether the charged iminium ion site of DRC can affect the reactivity of terminal carbon radical. Control studies have shown that the presence of TMSCN is critical for radical addition to alkene. Without TMSCN, even with 20% of photocatalyst Ir[dF(CF$_3$)ppy]$_2$(dtbbpy)PF$_6$, no conversion of cyclobutylamine has been observed. Moreover, the reactivity is altered for spirocyclic cyclobutylamines when an amidinium ion or conjugation to an arene is introduced (Scheme 6.2) and the reaction occurs in decent yield without the presence of TMSCN, leading us to question the reactivity model and the role of TMSCN.
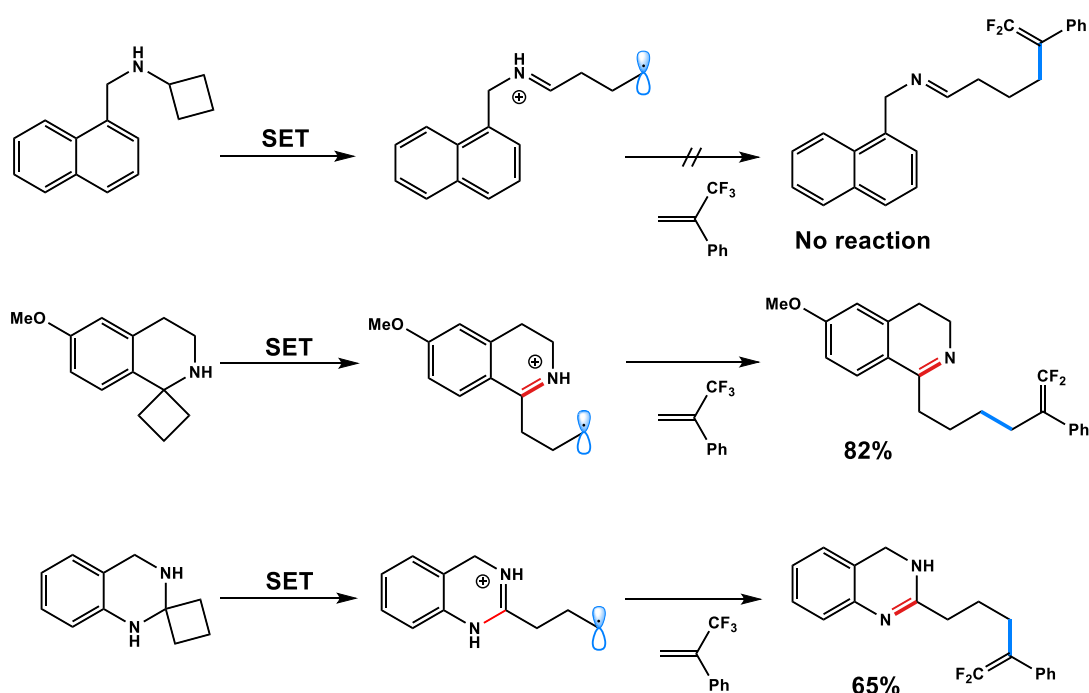
Scheme 6.2 reaction outcome of different *N*-alkyl amines.

The general reaction mechanism is depicted in Scheme 6.3. The excited-state Ir[dF(CF$_3$)ppy]$_2$(dtbbpy)PF$_6$ generated via absorption of visible light can undergo single electron transfer to oxidize the cyclobutylamine **1a** to its radical cation. The radical cation

induces adjacent C-C bond cleavage to form distonic radical cation (DRC). If TMSCN is present, it will quench the distonic radical cation via a Strecker reaction to give α-amino cyanide with a neutral radical. Subsequently the neutral radical undergoes addition reaction to olefin, followed by single electron transfer and elimination of fluoride to yield product **1**.
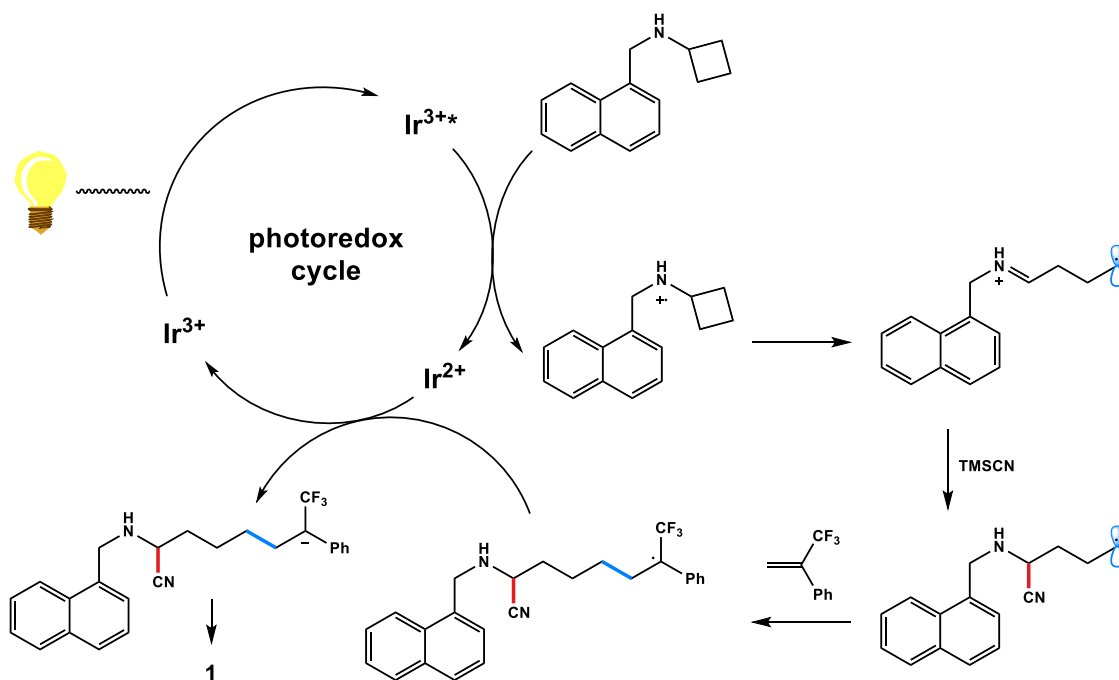
Scheme 6.3 Proposed catalytic cycle.

We initially speculate the unusual role of TMSCN and reactivity of spirocyclic cyclobutylamines lies in that an electrophilic iminium ion site interacts intramolecularly with nucleophilic carbon radical to reduce the reactivity of the radical. Hence, the addition of TMSCN serve as a quencher of the iminium ion to resume the reactivity of a neutral radical. The reduction of the electrophilicity of the iminium ion by attaching a π-electron donating group can weaken the interaction between the electrophilic charged site and nucleophilic radical site, thus increasing the reactivity of the radical site.

6.2 Computational Results and Discussion

To verify the proposed mechanism, we have calculated the reaction coordinate under M06-2X (SMD, solvent=DCM)/ma-def2-TZVP//def-TZVP level of theory using Gaussian 16 software, as shown below. On the doublet cationic radical surface, the cyclobutyl-ring opening and closing is kinetically reversable, with the DCR as a more thermodynamically stable radical. The following addition reaction is also kinetically feasible (barrier height of 15.6 kcal/mol). So one would expect an efficient addition to CF_3-styrene once DCR is formed, while it is not observed in the experiment. Other possibilities that could potentially hamper the addition reaction, including explicit solvent effect, counterion effect, defluorination and deprotonation of the iminium cation, are all ruled out. So what could be the factor that prevents the addition reaction from happening?

We first turned our attention to the cation radical DCR (A-2). There are two possible exit pathways, one is undergoing addition to CF_3-styrene or oxidize the Ir(II) photocatalyst. According to the experimental redox potential of the Ir-photocatalyst (Ir(III)/Ir(II) = -1.75 V vs SHE, -40 kcal/mol), cationic radical A-2 is able to oxidize the Ir(II) photocatalyst (-61.3 kcal/mol < -40 kcal/mol) to the Ir(III) photocatalyst. That means although the starting N-alkyl amine can be oxidized to a cationic diradical, its ring-opening product can take back the lost electron and return to the starting material. Since electron transfer is much faster than bimolecular reaction, the cationic radical is quenched to diradical before any addition reaction can happen. The diradical ring-closure is much lower in energy ($\Delta G = 5.7$ kcal/mol) comparing to the radical addition to styrene ($\Delta G = 14.5$ kcal/mol). Thus, no product is observed and the radical cannot be trapped by any radical acceptors. On the other hand, when TMSCN is present, the cation radical is quenched to a radical that can undergo a nucleophilic addition to the styrene (Scheme 6.3).
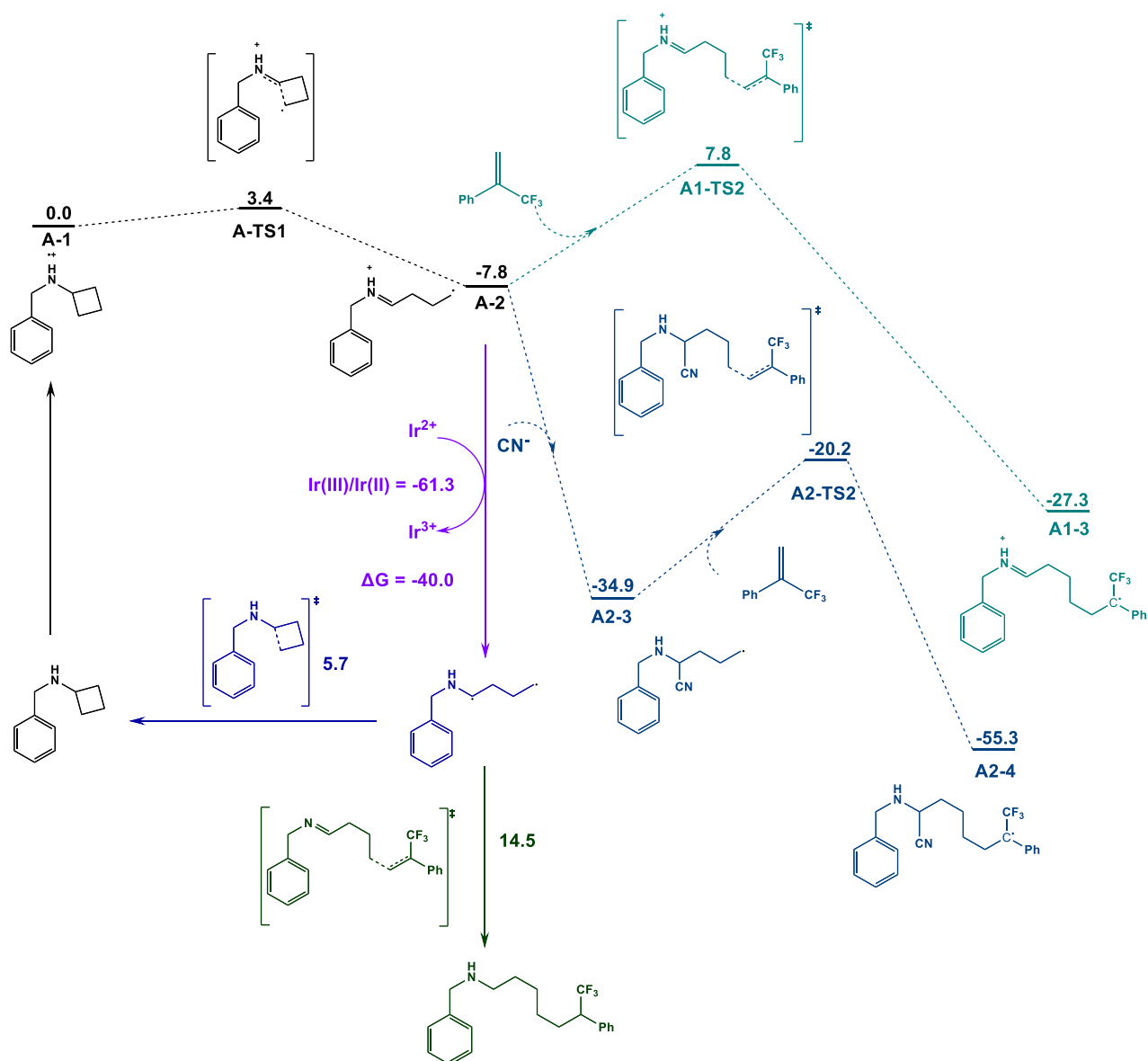
Scheme 6.3 Reaction coordinate of non-spirocyclic N-benzylcyclobutanamine with CF_3-styrene. Energies are in kcal/mol.

However, for spirocyclic cyclobutylamines that can react without TMSCN, the cationic radical is unable to be reduced by the photocatalyst (-25.6 kcal/mol > -40.0 kcal/mol), making it possible for DRC to undergo the subsequent addition reaction with CF_3-styrene (Scheme 6.4).
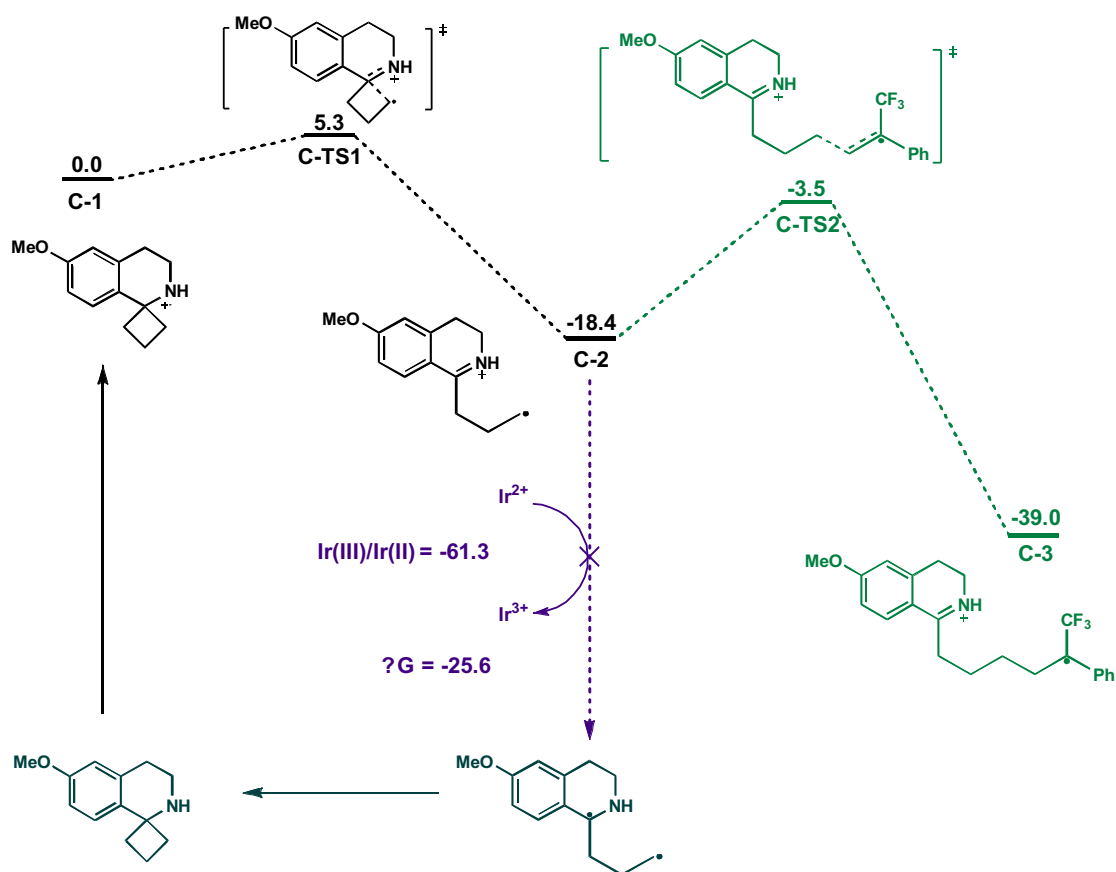
Scheme 6.4 Reaction coordinate of spirocyclic cyclobutylamines with CF_3-styrene. Energies are in kcal/mol.

6.3 Conclusion

In conclusion, we have modelled the photoredox catalyzed difunctionalization of N-alkyl cyclobutyl arenes. Our calculation successfully revealed that without the presence of TMSCN, DRC would be easily reduced by the Ir(II) photocatalyst and no reaction occur. Adding TMSCN can open up the pathway of anionic nucleophilic addition. Additionally, DRCs from spirocyclic cyclobutylamines exhibit higher reduction potential and the reaction can happen without TMSCN.

6.4 References

1. Yates, B. F.; Bouma, W. J.; Radom, L., Detection of the prototype phosphonium (CH_2PH_3), sulfonium (CH_2SH_2) and chloronium (CH_2ClH) ylides by neutralization-

reionization mass spectrometry: a theoretical prediction. *J. Am. Chem. Soc.* **1984,** *106* (20), 5805-5808.

2.   Stirk, K. M.; Kiminkinen, L. K. M.; Kenttamaa, H. I., Ion-molecule reactions of distonic radical cations. *Chem. Rev.* **1992,** *92* (7), 1649-1665.

3.   (a) Smith, R. L.; Chyall, L. J.; Stirk, K. M.; Kenttämaa, H. I., Radical-type reactivity of the methylenedimethylsulfonium ion, $(CH_3)_2S+-CH_2$˙. *Org. Mass Spectrom.* **1993,** *28* (12), 1623-1631; (b) Chyall, L. J.; Byrd, M. H. C.; Kenttamaa, H. I., Reactions of the Charged Radical $(CH_3)_2S+-CH_2$˙ bul. with Cyclic Alkenes. *J. Am. Chem. Soc.* **1994,** *116* (23), 10767-10772; (c) Osburn, S.; Chan, B.; Ryzhov, V.; Radom, L.; O'Hair, R. A. J., Role of Hydrogen Bonding on the Reactivity of Thiyl Radicals: A Mass Spectrometric and Computational Study Using the Distonic Radical Ion Approach. *The Journal of Physical Chemistry A* **2016,** *120* (41), 8184-8189; (d) Tureček, F.; Julian, R. R., Peptide Radicals and Cation Radicals in the Gas Phase. *Chem. Rev.* **2013,** *113* (8), 6691-6733; (e) Petzold, C. J.; Nelson, E. D.; Lardin, H. A.; Kenttämaa, H. I., Charge-Site Effects on the Radical Reactivity of Distonic Ions. *The Journal of Physical Chemistry A* **2002,** *106* (42), 9767-9775.

4.   (a) Jurva, U.; Bissel, P.; Isin, E. M.; Igarashi, K.; Kuttab, S.; Castagnoli, N., Model Electrochemical-Mass Spectrometric Studies of the Cytochrome P450-Catalyzed Oxidations of Cyclic Tertiary Allylamines. *J. Am. Chem. Soc.* **2005,** *127* (35), 12368-12377; (b) Sun, Q.; Zhu, R.; Foss, F. W.; Macdonald, T. L., In Vitro Metabolism of a Model Cyclopropylamine to Reactive Intermediate: Insights into Trovafloxacin-Induced Hepatotoxicity. *Chem. Res. Toxicol.* **2008,** *21* (3), 711-719; (c) Cerny, M. A.; Hanzlik, R. P., Cytochrome P450-Catalyzed Oxidation of N-Benzyl-N-cyclopropylamine Generates Both Cyclopropanone Hydrate and 3-Hydroxypropionaldehyde via Hydrogen Abstraction, Not Single Electron Transfer. *J. Am. Chem. Soc.* **2006,** *128* (10), 3346-3354; (d) Zhao, Z.;

Mabic, S.; Kuttab, S.; Franot, C.; Castagnoli, K.; Castagnoli, N., Rat liver microsomal enzyme catalyzed oxidation of 1-cyclopropyl-4-phenyl-1,2,3,6-tetrahydropyridine. *Biorg. Med. Chem.* **1998,** *6* (12), 2531-2539; (e) Grimm, M. L.; Allen, W. J.; Finn, M.; Castagnoli, N.; Tanko, J. M., Reaction of benzophenone triplet with aliphatic amines. What a potent neurotoxin can tell us about the reaction mechanism. *Biorg. Med. Chem.* **2011,** *19* (4), 1458-1463; (f) Macdonald, T. L.; Zirvi, K.; Burka, L. T.; Peyman, P.; Guengerich, F. P., Mechanism of cytochrome P-450 inhibition by cyclopropylamines. *J. Am. Chem. Soc.* **1982,** *104* (7), 2050-2052; (g) Shaffer, C. L.; Morton, M. D.; Hanzlik, R. P., N-Dealkylation of an N-Cyclopropylamine by Horseradish Peroxidase. Fate of the Cyclopropyl Group. *J. Am. Chem. Soc.* **2001,** *123* (35), 8502-8508; (h) Shaffer, C. L.; Harriman, S.; Koen, Y. M.; Hanzlik, R. P., Formation of Cyclopropanone during Cytochrome P450-Catalyzed N-Dealkylation of a Cyclopropylamine. *J. Am. Chem. Soc.* **2002,** *124* (28), 8268-8274.

5.   (a) Maity, S.; Zhu, M.; Shinabery, R. S.; Zheng, N., Intermolecular [3+2] Cycloaddition of Cyclopropylamines with Olefins by Visible-Light Photocatalysis. *Angew. Chem. Int. Ed.* **2012,** *51* (1), 222-226; (b) Jiang Wang, T. H. N., Nan Zheng, Photoredox-catalyzed [4+2] annulation of cyclobutylanilines with alkenes, alkynes, and diynes in continuous flow. *Sci China Chem* **2016,** *59* (2), 180-183; (c) Wang, J.; Zheng, N., The Cleavage of a C-C Bond in Cyclobutylanilines by Visible-Light Photoredox Catalysis: Development of a [4+2] Annulation Method. *Angew. Chem. Int. Ed. Engl.* **2015,** *54* (39), 11424-7; (d) Wang, Q.; Zheng, N., A Photocatalyzed Synthesis of Naphthalenes by Using Aniline as a Traceless Directing Group in [4 + 2] Annulation of Amino-benzocyclobutenes with Alkynes. *ACS Catalysis* **2017,** *7* (6), 4197-4201.

6.   Marshall, D. L.; Gryn'ova, G.; Poad, B. L. J.; Bottle, S. E.; Trevitt, A. J.; Coote, M. L.; Blanksby, S. J., Experimental evidence for long-range stabilizing and destabilizing

interactions between charge and radical sites in distonic ions. *Int. J. Mass spectrom.* **2019,** *435*, 195-203.

Chapter 7

Origin of Regioselectivity in the electrophilic aromatic

substitution of nitronium ions.

7.1 Introduction

Electrophilic aromatic substitution (EAS) constitutes one of the most important classes of transformations in the entirety of chemical synthesis.[1,2] To the present day, synthetic chemists rely heavily on traditional methods of attaching activating (or deactivating) groups directly to aromatic rings to control the rates and selectivities of substitutions. Imagine instead a functional group containing a lone pair of electrons suspended over the π-cloud of an aromatic ring. Theoretically, such an arrangement should activate the aromatic ring to approach of an electrophile. To test this hypothesis, the Lectka group at Johns Hopkins University has employed rigid bicyclic structures based on fused 9,10-dihydroanthracene-bicycloheptane scaffolds to lock select heteroatoms (e.g., fluorine, oxygen) in close spatial proximity to aromatic rings.[3, 4] To their delight the probe ring was functionalized almost exclusively, and intermolecular experiments indicated several hundred-fold rate enhancements (Scheme 7.1).
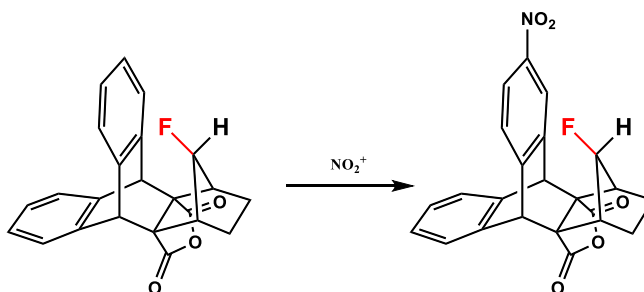
Scheme 7.1 Original model system of electrophilic aromatic substitution.

Conventionally, it is assumed that a *direct* activation of the probe ring by the proximate group would affect the electron density of the ring and lower the transition state energy (Scheme 7.2, I). This process is modeled and rationalized previously through the calculation of the corresponding Wheland intermediates, or σ-complexes.[5] However, there is no further evidence supporting this thermodynamic-controlled selectivity.

126
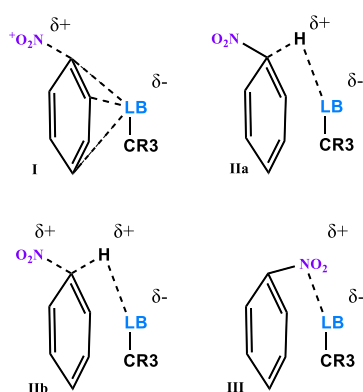
Scheme 7.2 Possible modes of mechanisms; direct activation through -complex stabilization (Types I, II) vs. relay mechanism (Type III).

Other modes of activation are also possible, including rate-determining deprotonation (Scheme 7.2, IIa)[6,7] and stabilization of the transition state through hydrogen bonding (Scheme 7.2, IIb). In this work, we show through calculation that the precomplexation of the electrophile to the LB serves as a relay to the aromatic ring (Scheme **7.2**, III). The Lectka group have examined multiple scaffolds and several more basic functional groups that would appear to favor precomplexation.

Singleton and co-workers have showed that modeling the regioselectivity of toluene nitration required careful consideration of the solvent complex coupled to molecular dynamics simulations, as a result of the flat potential (and free) energy surfaces associated with electrophilic addition.[8] Our computational results also point to fairly flat potential energy surfaces around the transition states through which activator–NO_2^+ π-complexes transform into σ-complexes that also make it difficult to predict the regioselectivity. Although a complete picture of regioselectivity would require exhaustive ab initio molecular dynamic simulations our current resources, our computation is able to rationalize these reactions and provide a practical model for rational design.

7.2 Experimental Results

The Lectka group have designed three different scaffolds (Scheme 7.3). Scaffold I has the virtue of rigidly positioning the Lewis basic site very close to the probe ring. Unfortunately, only analogues containing fluoro and hydroxy groups can be synthesized in. Scaffold II affords more choices, although the trade-off is that the Lewis base is positioned further from the probe ring. Scaffold III allows the placement of biologically relevant Lewis bases such as amides and esters.
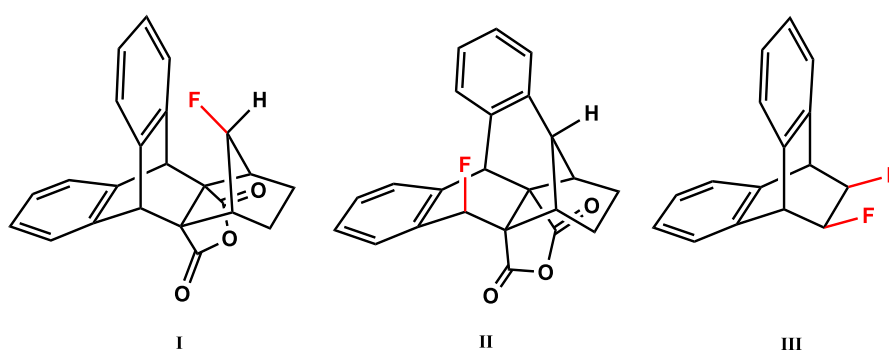
Figure 7.3 Representative model systems I, II, and III.

The original model based on scaffold I located heteroatoms in spatial proximity to the *top* aromatic ring while the *bottom* ring served as a control.[3,4] Interestingly, when was subjected to nitrating conditions, the alcoholic oxygen formed a nitrate ester before substitution occurred on the probe ring, thereby providing us the first hint of a possible relay mechanism rather than direct activation (Scheme 7.4).[3] Nevertheless, the nitrated alcohol still directs electrophilic nitration on the probe ring, in the presence of additional nitronium ions.
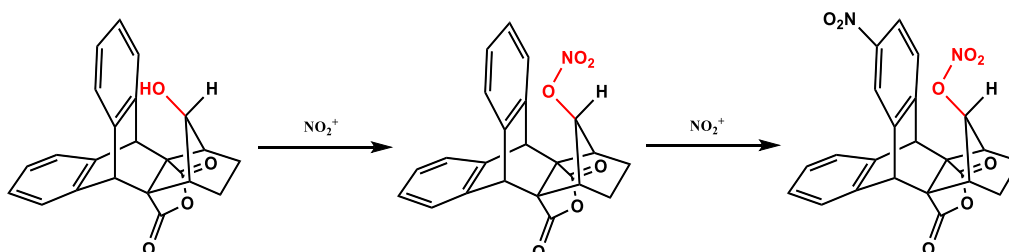
Scheme 7.4 Formation of nitrate ester **3** and its subsequent nitration.

128

The Lectka group attempted to generate secondary carbocation **4a₁** (Figure **6**), stabilized by a cation-π interaction, as a characterizable species inspired by a serendipitous discovery.[9-14] The distance between the carbon bearing positive charge and the adjacent aromatic ring decreases from 3.06 Å (in compound **5**) to 1.65 Å, congruent with expectations.

When subjected to nitrating conditions,[15] **6** predominantly nitrates the *"top"* ring at the positions indicated in Scheme 7.5 (~1.7:1 for **7a:7b**).
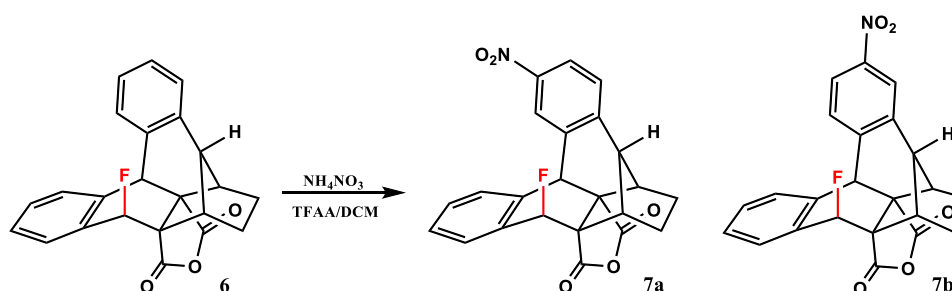
Scheme 7.5 Nitration of compound **6.**

We first performed energetic analysis of σ-complexes that predicts those derived from the *"top"* ring to be 8.6 kcal/mol lower in energy than those anchored on the *bottom* ring. At the same level, however, σ-complexes at the *top* ring were predicted to be 5.2 kcal/mol lower in energy in **1**,[4a] despite the shorter distance of the fluorine atom from the positively charged carbon atom of the probe ring. Quite possibly, the *bottom* ring in **6** is being affected by the inductive effects of fluorine.

Alternatively, the Lectka group envisioned that replacing the hydrogen atom in **6** (H$_A$, Scheme 7.5) with a phenyl group can provide an additional control ring free of steric as well as inductive effects. This molecule serves as another candidate for our investigation as it provides two control rings **B** and **C**, wherein **C** is free from ring strain and inductive effects, although its reactivity is relatively diminished by mono-substitution, as opposed to di-substitution.

Subjecting **10** to nitrating conditions resulted in two major products **11a** and **11b** (~1.5:1) with nitration predominately on the probe ring **A** (Scheme **7.6**). The exact positions of the nitro groups were confirmed by X-ray crystallographic analysis. ^{1}H NMR analysis of the reaction mixture, however, also revealed some unreacted **10** and very small amounts of seemingly di-nitration products. They correspond to subsequent nitration of control ring **C**, after initial nitration of the probe ring (*vide infra*). DFT calculations predict the σ-complexes of the probe ring to be 8.0 kcal/mol more favorable than those of both control rings. If the selectivity is thermodynamically controlled, we should not have observed formation of dinitrated products before **10** is completely nitrated, thereby providing hint at the insufficiency of evaluating energetics of σ-complexes.
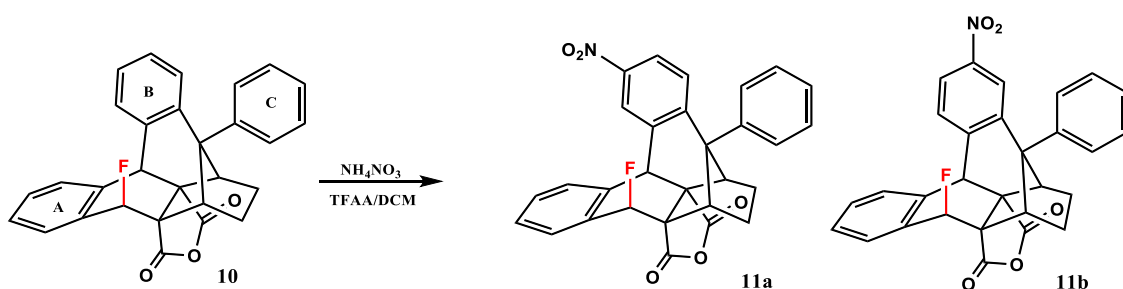
Scheme 7.6 Nitration of compound **10**

The synthetic exploitation of the structural rearrangement of secondary carbocation **4b** provides an attractive means to incorporate larger moieties within scaffold II. Tetrazoles, for instance, constitute an important class of organic molecules in medicinal chemistry, stemming from their isosterism with carboxylic acids and their relative metabolic stabilities.[15-22] While literature on their biological activities as well as physical properties is bountiful, their effect on aromatic reactivity is not as well-explored. When subjected to nitrating conditions, tetrazole **18** underwent exclusive nitration of the probe ring, predominantly on positions 2 and 3, as confirmed by X-ray analysis of the isolated products, but also to a lesser extent on positions 1 and 4 (Scheme 7.7).
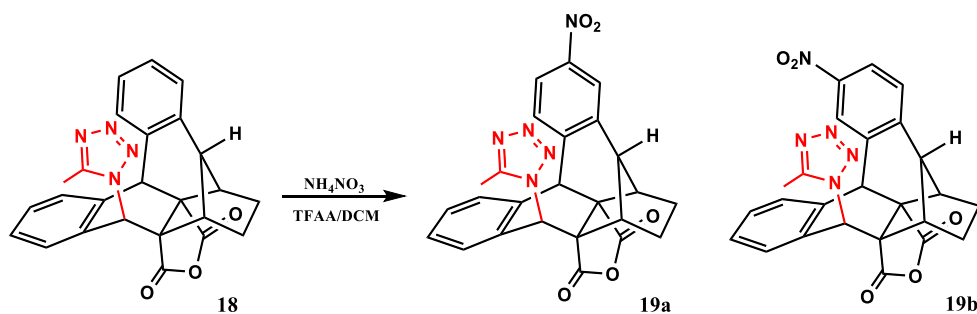
Scheme 7.7 Nitration of tetrazole **18.**

For Scaffold III, the Lectka group initiated the investigation of carboxyalkyl group mediated arene activation with attempted nitration of cis-di-ester **21.**[23] They observed exclusive nitration of the probe ring at the "ortho-meta" position (Scheme 7.8).
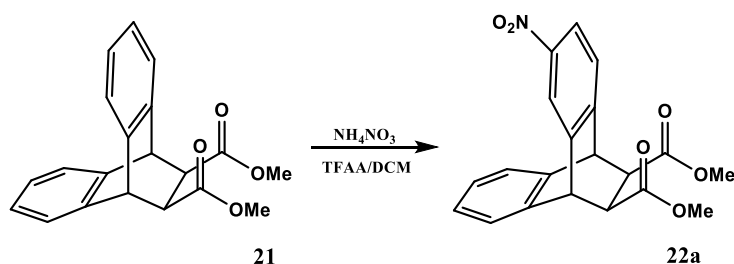
Scheme 7.8 Nitration of cis-diester **21.**
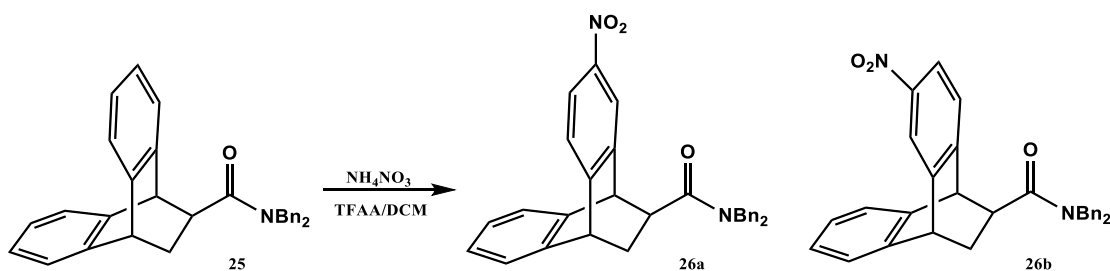
Scheme 7.9 Synthesis and nitration of compounds **25.**

Additionally, compound **25** is synthesized by functionalization of mono-ester**,** which in turn yielded exclusive nitration of the probe ring at the positions indicated in Scheme 7.9 (~1:1 ratio for **26a:26b**) as confirmed by the [^{1}H,^{1}H]-NOESY spectrum of the isolated products, wherein no aromatic proton on the substituted ring engages in through-space coupling with the

amide α-proton. It represents a clear failure of the Wheland intermediate model. By this juncture, we have already developed some skepticism that simply calculating energies of σ-complexes could provide a good energetic picture of this reaction and concluded that a more realistic approach would be called for.

7.3 Computational Results and Discussion

We initially assumed that in most cases the directing group stabilizes the forming arenium ion as the electrophile approaches from the opposite face of the aromatic ring (Scheme 7.2, Type I). However, our prior calculation (see Fig F1, F2, F3, F4 in Appendix) and a critical experimental observation (the nitration of the OH group, Scheme 7.4) prompted us to think something else may be going on. Combine these reasons with the failure of σ-complex analysis, a more rigorous approach to modeling was mandated. To start, we calculated the free energy profiles up to the formation of the putative σ-complexes to understand the roles of various through-space interactions in promoting selectivity. As mentioned above, Singleton has pointed out that the regioselectivity of nitration in toluene can be complicated to rationalize and predict, requiring MD simulations. In our case, we were initially interested in the problem of selectivity between different aromatic rings, which, we believe, can be controlled by precomplexation and a stabilizing relay between the directing group and the substrate aromatic ring.

We started our theoretical investigation with the trickiest case of diester. Experimentally a 1:1 mixture of compounds **21** and **23** is treated with 1 eq. of nitrating reagent. Integration of relevant peaks in the ^{1}H NMR of the reaction mixture revealed that cis-diester **21** reacts four times as fast as trans-di-ester **23**.

Our computation indicated that, after being generated, the electrophilic nitronium ion can form two different complexes with compound **21**. A typical π-complex involving the

132

unperturbed ring (**22b-PC** wherein the nitronium ion is perpendicular to the ring), is 7.1 kcal/mol endergonic comparing with isolated reactants. On the other hand, electron-rich carbonyl groups can interact with the nitronium ion to form complex **22a-PC** and it releases 9.5 kcal/mol free energy. The subsequent barrier that connects the precomplexation and their corresponding σ-complexes are very low, indicating a very shallow potential energy surface (PES) and a rapid transition from π-complex (or lone-pair) to σ-complex (Fig 7.1). Clearly from the potential energy surface, the selectivity-determining step is the precomplexation step and the formation of **22a-PC**, being exothermic, is preferred and it leads to the exclusive nitration of the adjacent probe ring. The precomplex then transforms to its corresponding σ-complex passing through transition state **22a-TS**. The vibration mode of the imaginary frequency displays migration of NO_2^+ from the carbonyl group to the benzene ring.
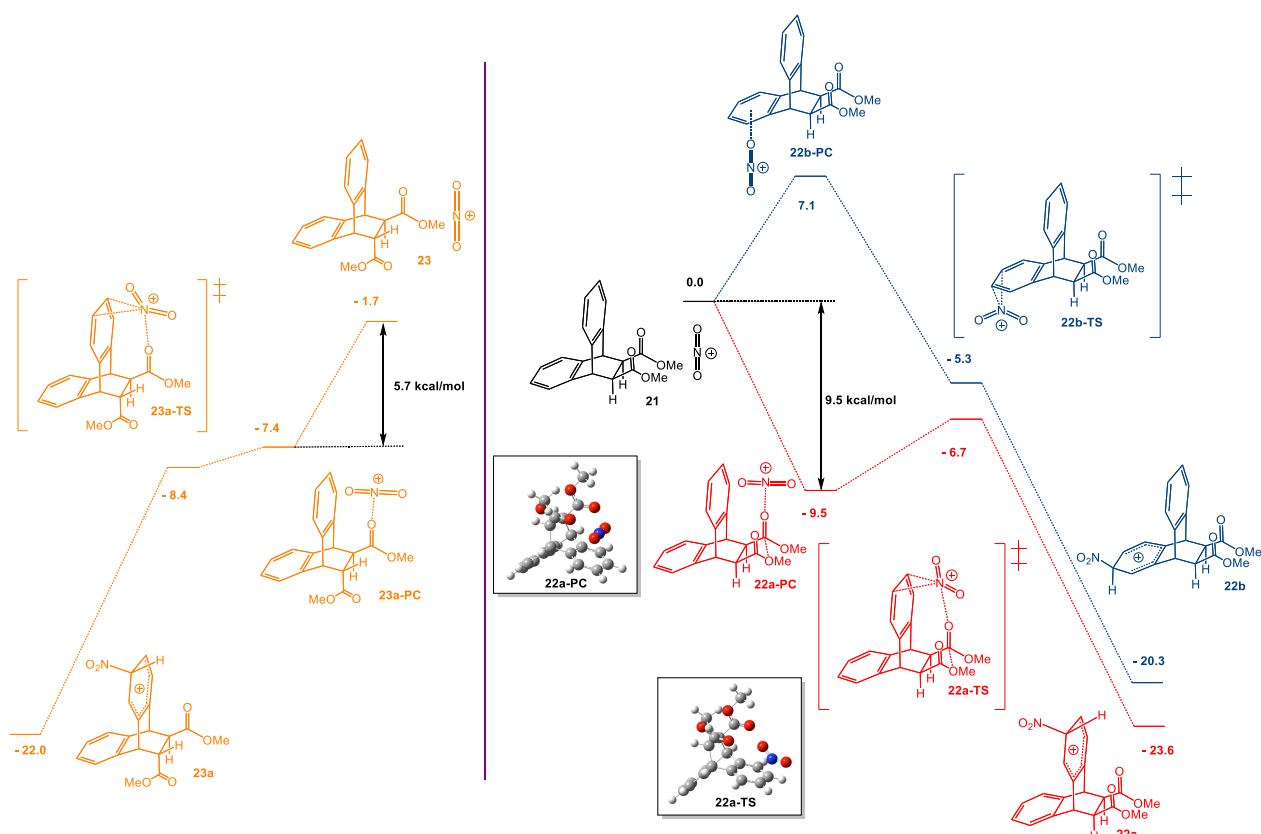
Figure 7.1 Free energy profile up to the formation of s complexes in cis- and trans-diesters **21** and **23** respectively involving precomplexes and transition states.

The precomplexation model is also capable of explaining the faster rate of **21** nitration. Although the *trans*-di-ester **23** is intrinsically 1.7 kcal/mol more stable than the *cis*-diester **21**, the *cis*-π-complex **22a-PC** is much lower in energy than the *trans*-π-complex **23a-PC**. Furthermore, the binding energy ΔG_{bind} is 3.8 kcal/mol lower for the formation of **22a-PC**. This implies that in solution, the concentration of the *cis*-π-complex is higher than the *trans*-π-complex, leading to a faster reaction rate. To validate that the *cis*-π-complex is indeed more stable, we calculated the binding energies with different functionals and a much larger basis set (def2-TZVPP), as accurately calculating energies of weakly-bound systems can be challenging. The results, shown in Table **1**, reveal that nitronium-carbonyl binding is consistent across all levels and the ΔG for the formation of *cis*-π-complex is roughly 2-3 kcal/mol lower than the corresponding *trans*-π-complex. In all, the observed 4:1 *cis/trans* nitration can be rationalized with the thermodynamic stability of the π-complexes.

Table 7.1 Binding energies for the formation of cis/trans π-complexes, calculated at different functionals with def2-TZVPP basis set ($\Delta\Delta G = \Delta G_{trans} - \Delta G_{cis}$)

Functional	ΔG_{cis}	ΔG_{trans}	$\Delta\Delta G$
M06-2X	-4.8	-1.5	3.3
ωB97X-D	-2.8	-0.7	2.1
MN15	-5.0	-1.4	3.5

Similar computational results are revealed on amides, in which exothermic formation of the carbonyl–NO_2^+ π-complex (**29a-PC**) overcomes its endothermic complexation with the

unperturbed ring (**29b-PC**), thereby directing it to the probe ring selectively in amide **x**. The complex (**29a-PC**) then passes through **29a-TS** to form the corresponding σ-complex.

Finally, to confirm that precomplexation is really the selectivity controlling factor, we aimed at turning off this selectivity enhancement by synthesizing control imide **30**[24] whose structural restraints lock both carbonyl oxygens farther and orthogonally from the probe ring. Our model predicts that the position and the orientation of the carbonyl is not able to form reactive complex.   A complex mixture of unreacted, mono-nitrated and dinitrated products is observed when **9** is treated with 1 eq. of nitrating agent (Scheme 7.10).
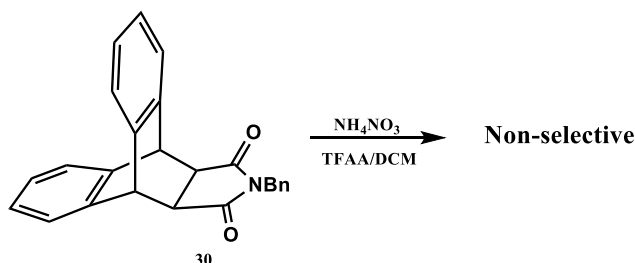
Scheme 7.10 Nitration of compound **30**.

7.4 References

1.   L. N. Ferguson, *Chem. Rev.* **1952,** *50*, 47-67.

2.   B. Galabov, D. Nalbantova, P. V. Schleyer, H. F. Schaefer, *Acc. Chem. Res.* **2016,** *49*, 1191-1199.

3.   a) L. Guan, M. G. Holl, C. R. Pitts, M. D. Struble, M. A. Siegler, T. Lectka, *J. Am. Chem. Soc.* **2017,** *139*, 14913-14916; b) M. Kazim, L. Guan, A. Chopra, R. Sun, M. A. Siegler, T. Lectka, *J. Org. Chem.* **2020**, *85*, 9801-9807.

4.   a) M. G. Holl, M. D. Struble, P. Singal, M. A. Siegler, T. Lectka, *Angew. Chem. Int. Ed.* **2016,** *55*, 8266-8269; b) For a study of lone-pair–p interactions, see; J. Novotny, S. Bazzi, R. Marek, J. Kozelka, *Phys. Chem. Chem. Phys.* **2016**, *18*, 19472-19481.

5.   G. W. Wheland, *J. Am. Chem. Soc.* **1942,** *64*, 900-908.

6.   J. Y. Wang, K. Choi, S. J. Zuend, K. Borate, H. Shinde, R. Goetz, J. F. Hartwig, *Angew. Chem. Int. Ed.* **2021,** *60*, 399-408.

7.   A. C. Knipe, J. Loundkeast, N. Sridhar, *J. Chem. Soc. Chem. Comm.* **1976**, 765-766.

8.   Y. N. Quinones, D. A. Singleton, *J. Am. Chem. Soc.* **2016,** *138*, 15167-15176.

9.   M. Kazim, M. A. Siegler, T. Lectka, T. *Org. Lett.* **2019,** *21*, 2326-2329.

10. R. Bhatia, K. Kadyan, M. Duhan, M. Devi, R. Singh, R. C. Kamboj, P. Kumar, *ChemistrySelect* **2019,** *4*, 10417-10424.

11. S. Z. Zard, *Chem. Commun.* **2002**, 1555-1563.

12. R. N. Grimes, *J. Organomet. Chem.* **2013,** *747*, 4-15.

13. H. E. Ho, A. Pagano, J. R. Ashton, J. R. Donald, R. G. Epton, J. C. Churchill, M. J. James, P. O'Brien, R. J. K. Taylor, W. P. Unsworth, *Chem. Sci.* **2020,** *11*, 1353-1360.

14. M. Kazim, M. A. Siegler, T. Lectka, *J. Org. Chem.* **2019,** *84*, 15765-15765.

15. J. V. Crivello, *J. Org. Chem.* **1981,** *46*, 3056-3060.

16. R. J. Herr, *Bioorg. Med. Chem.* **2002,** *10*, 3379-3393.

17. P. Kushwaha, S. Fatima, A. Upadhyay, S. Gupta, S. Bhagwati, T. Baghel, M. I. Siddiqi, A. Nazir, K. V. Sashidhara, *Bioorg. Med. Chem. Lett.* **2019,** *29*, 66-72.

18. Y. J. Li, K. K. Pasunooti, R. J. Li, W. K. Liu, S. A. Head, W. Q. Shi, J. O. Liu, *J. Med. Chem.* **2018,** *61*, 11158-11168.

19. J. Y. Zhang, S. Wang, Y. Y. Ba, Z. Xu, *Eur. J. Med. Chem.* **2019,** *178*, 341-351.

20. F. Gao, J. Q. Xiao, G. Huang, *Eur. J. Med. Chem.* **2019,** *184*, 111744.

21. S. Q. Wang, Y. F. Wang, Z. Xu, *Eur. J. Med. Chem.* **2019,** *170*, 225-234.

22. P. Lassalas, G. Gay, C. Lasfargeas, M. J. James, V. Tran, K. G. Vijayendran, K. R. Brunden, M. C. Kozlowski, C. J. Thomas, A. B. Smith, D. M. Huryn, C. Ballatore, *J. Med. Chem.* **2016,** *59*, 3183-3203.

23. J. Roh, G. Karabanovich, H. Vlckova, A. Carazo, J. Nemecek, P. Sychra, L. Valaskova, O. Pavlis, J. Stolarikova, V. Klimesova, K. Vavrova, P. Pavek, A. Hrabalek, *Bioorg. Med. Chem.* **2017,** *25*, 5468-5476.

24. H. G. Oddy, J. Am. Chem. Soc. 1923, 45, 2156-2160.

Appendix

A. Supporting Information for chapter 2

Method, Conformation, Excess Energy Exploration and additional details of Molecular Dynamics

Influence from level of theory

Table A1. Uphill AIMD simulations of system **4** (R^1 = H, R^2 = Vinyl) with different level of theories.

Level of theory	H%	Vinyl%	Total trajectory number
M06-2X-D3/6-31G(d)	27	73	75%
ωB97X-D/6-31G(d)	14	86	79%
M06-2X-D3/6-311G(d,p)	25	75	83%

Influence from conformation

The structure of the skeleton of the carbocation is overall rigid and we would expect small influence from conformational flexibility. Nonetheless, to confirm that the product distribution is invariant with the conformations, we have initiated uphill dynamic simulation of system **17** with different orientations of the vinyl group. The ratio of ethyl:vinyl migration is shown in **Fig A1**. The 3% difference is small and negligible.

33:67   **36:64**
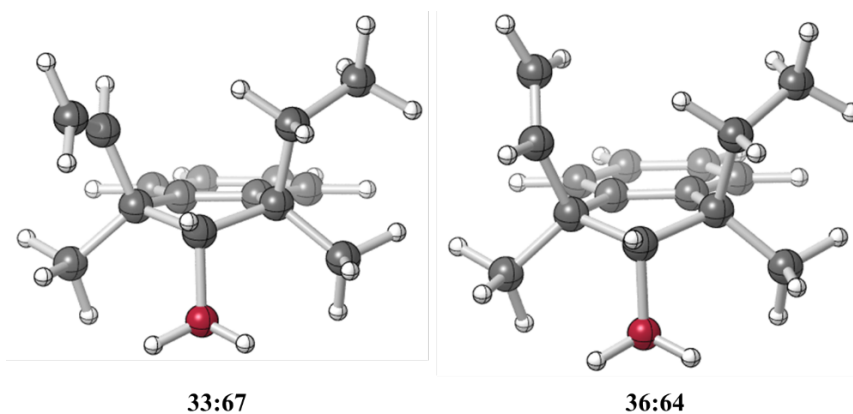
Figure A1. Uphill AIMD simulations of system **17** (R^1 = Ethyl, R^2 = Vinyl) with different conformations as shown. The difference in the overall product ratio is small (3%)

Influence from excess energy

Table A2. Uphill AIMD simulations of the R^1 = Me, R^2 = acyl system with different excess energies (kcal/mol).

Excess Energy	# of Me migration trajectories	# of acyl migration trajectories	% of acyl migration
22.0	25	76	75%
25.0	26	74	74%
28.0	22	84	79%
31.0	17	81	83%

To verify that choosing different excess energy will not make significant influence on the dynamic outcome, we compared the results with the experimental data shown below in Table A3. It is clear that with excess energy ranging from 0 to 30 kcal/mol, the results of AIMD simulation are within ± 5% range of the experimental outcome. Thus the amount of excess energy accessible in the MD simulation will not affect our conclusions.

Table A3. Uphill AIMD simulations of the R^1 = Me, R^2 = acyl system with different excess energies (kcal/mol).

Excess Energy	Total number of trajectories	% of ring expansion
0.0	79	93%
15.0	84	94%
20.0	78	93%
25.0	365	89%
30.0	96	86%
Experiment	-	88%

Detailed derivation of R^1/R^2 ratio incorporating recrossing in Table 1.2.

Since significant amount of recrossings are observed in the downhill dynamic simulations starting from VTSs, it is vital to reflect this in our predicted ratio. We can divide all trajectories from the same VTS-1 into three parts, the R represents the ratio of such type of trajectory:

$$R^1_{VTS1} + R^2_{VTS1} + R^{recrossing}_{VTS1} = 100\%$$

In which the superscripts correspond to each product. Assume that the possibility of arriving at any of the VTSs follows the Boltzmann distribution, the product contributions from VTS-1 and VTS-2 are:

$$P_{VTS1} = \frac{1}{1 + e^{2\frac{\Delta\Delta G}{k_B T}}}$$

$$P_{VTS2} = \frac{1}{1 + e^{-2\frac{\Delta\Delta G}{k_B T}}}$$

Since both of the two VTSs are ambimodal VTSs, meaning that we need to add the contributions from the two VTSs together.

$$R^1 = R^1_{VTS1}P_{VTS1} + R^1_{VTS2}P_{VTS2} = \frac{R^1_{VTS1}}{1 + e^{2\frac{\Delta\Delta G}{k_B T}}} + \frac{R^1_{VTS2}}{1 + e^{-2\frac{\Delta\Delta G}{k_B T}}}$$

And similarly,

$$R^2 = R^2_{VTS1}P_{VTS1} + R^2_{VTS2}P_{VTS2} = \frac{R^2_{VTS1}}{1 + e^{2\frac{\Delta\Delta G}{k_B T}}} + \frac{R^2_{VTS2}}{1 + e^{-2\frac{\Delta\Delta G}{k_B T}}}$$

And the ratio can be written as:

$$\frac{R^1}{R^2} = \frac{\dfrac{R^1_{VTS1}}{1 + e^{2\frac{\Delta\Delta G}{k_B T}}} + \dfrac{R^1_{VTS2}}{1 + e^{-2\frac{\Delta\Delta G}{k_B T}}}}{\dfrac{R^2_{VTS1}}{1 + e^{2\frac{\Delta\Delta G}{k_B T}}} + \dfrac{R^2_{VTS2}}{1 + e^{-2\frac{\Delta\Delta G}{k_B T}}}} = \frac{R^1_{VTS1}P_{VTS1} + R^1_{VTS2}P_{VTS2}}{R^2_{VTS1}P_{VTS1} + R^2_{VTS2}P_{VTS2}}$$

Dynamic simulation of secondary carbocations without water

Table A4. Dynamic simulation results of secondary carbocations without water, which serves as a putative model of 'equilibrated carbocations'. The results are compared with the uphill dynamic simulations with water and predictions from TST, VTS and VTS with corrections on recrossing.

	R^1	%	R^2	%	R^1/R^2 ratio (w/ water)	R^1/R^2 ratio (w/o water)	VTS ratio	R1/R2 ratio (recrossing)	Total trajectories
3	H—R	29%	▷—R	69%	30:70	28:72	2:98	12:88	116
6	H₃C—R	79%	＼ₒ⌒R	13%	86:14	62:38	54:46	54:46	65
11	⅄R	59%	⌒⌒R	29%	67:33	62:38	89:11		63
17	⌒R	33%	▔＼R	58%	36:64	50:50	8:92	5:95	143

Variational TST Calculations with *Gaussrate/Polyrate*

General methods

All variational TSs calculations and related IRCs were obtained with gaussrate/polyrate. See

attached Gaussrate input file below for details of options we chose

*General	13 C
	14 H
TITLE	15 C
H-Me	16 H
END	17 H
	18 H
ATOMS	19 C
1 C	20 H
2 C	21 H
3 C	22 H
4 C	23 C
5 O	24 H
6 C	25 H
7 C	26 H
8 H	27 C
9 H	28 H
10 H	29 H
11 H	30 H
12 C	END

```
                                            3

       NOSUPERMOL                           4

                                            5

       INPUNIT AU                           6

                                            7

       MDMOVIE ON                           8

                                            9

       *OPTIMIZATION                        10

                                            11

         OPTMIN  OHOOK                       12

         OPTTS   OHOOK                       13

                                            14

       *SECOND                              15

         HESSCAL HHOOK                       16

                                            17

       *REACT1                              18

       INITGEO  HOOKS                        19

       SPECIES NONLINRP                      20

       STATUS 0                             21

       GEOM                                 22

        1                                   23

        2                                   24
```

25	9
26	10
27	11
28	12
29	13
30	14
END	15
# end of react1 section	16
	17
*PROD1	18
INITGEO HOOKS	19
SPECIES NONLINRP	20
STATUS 0	21
GEOM	22
1	23
2	24
3	25
4	26
5	27
6	28
7	29
8	30

END

end of prod1 section

```
*START

INITGEO  HOOKS

SPECIES NONLINRP

STATUS 0

GEOM

 1

 2

 3

 4

 5

 6

 7

 8

 9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30


END

# end of start section


*PATH

SCALEMASS  1.00
```

INTMU 3

SSTEP   0.005

INH   9

SRANGE

 SLP   3.6

 SLM -1.0

END

RPM pagem

SIGN REACTANT

IDIRECT 1

COORD CART

RODS

 PRPATH

 COORD   1 11

 INTERVAL 1

XMOL

END

*RATE

FORWARDK

SIGMAF 1

CVT

CUS 2

PRDELG ON

#PRGIGT ON

#ICVT ON

MUVT ON

MUVTOPT

prenergy 4

niter 30

END

TEMP

200.

298.15   END

Supplementary free energy IRC
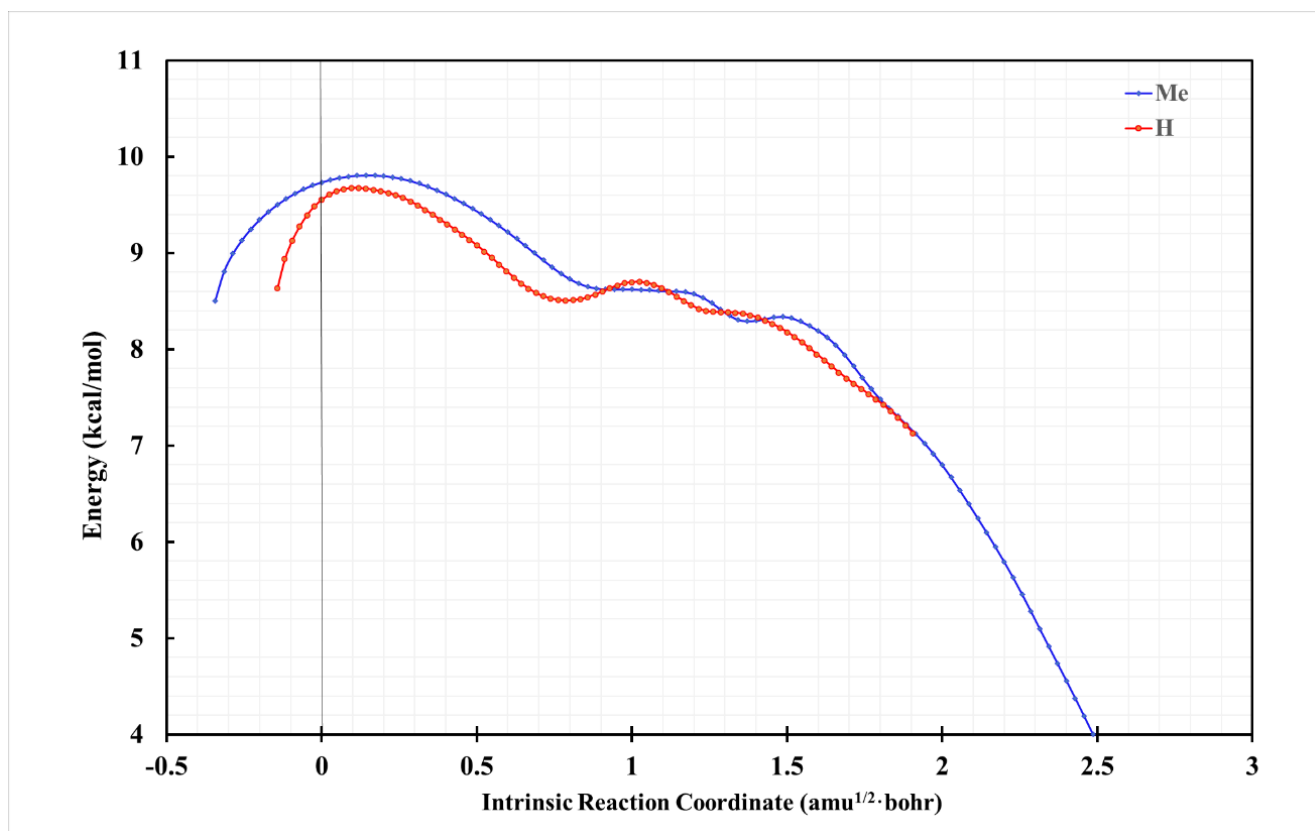
Figure A2. IRC of (R^1 = H, R^2 = Me)
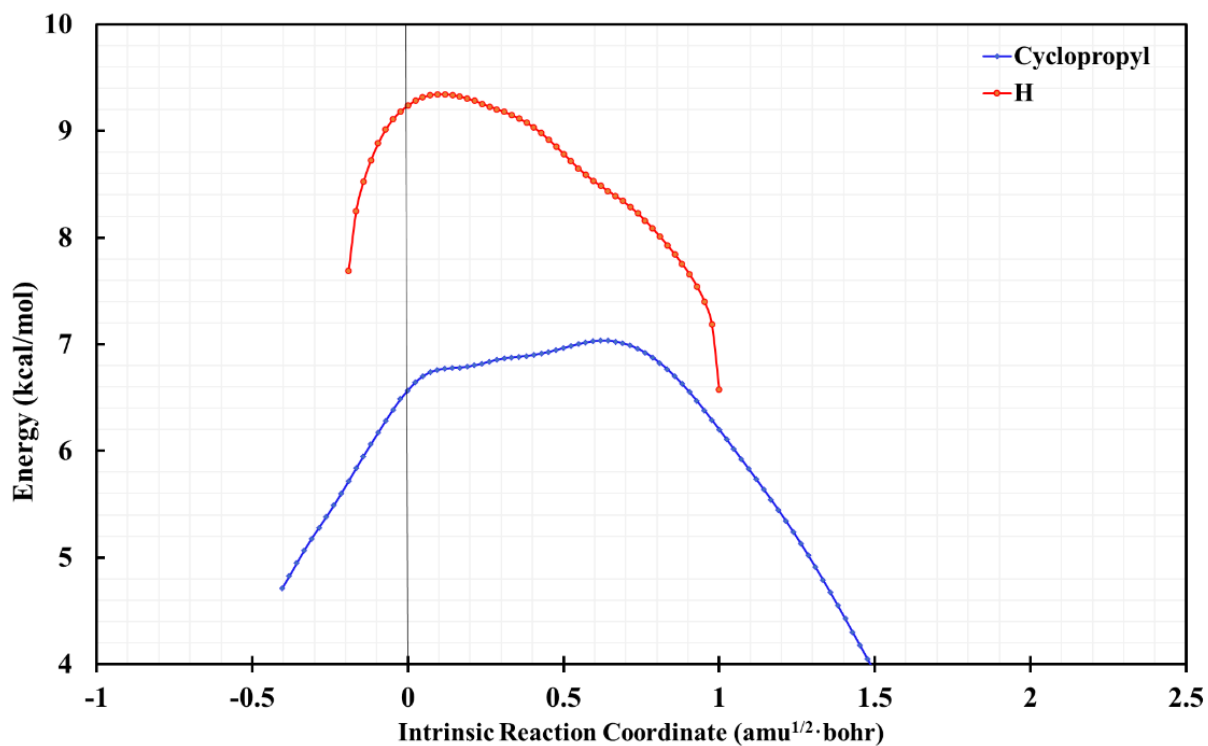
Figure A3. IRC of (R^1 = H, R^2 = Cyclopropyl)
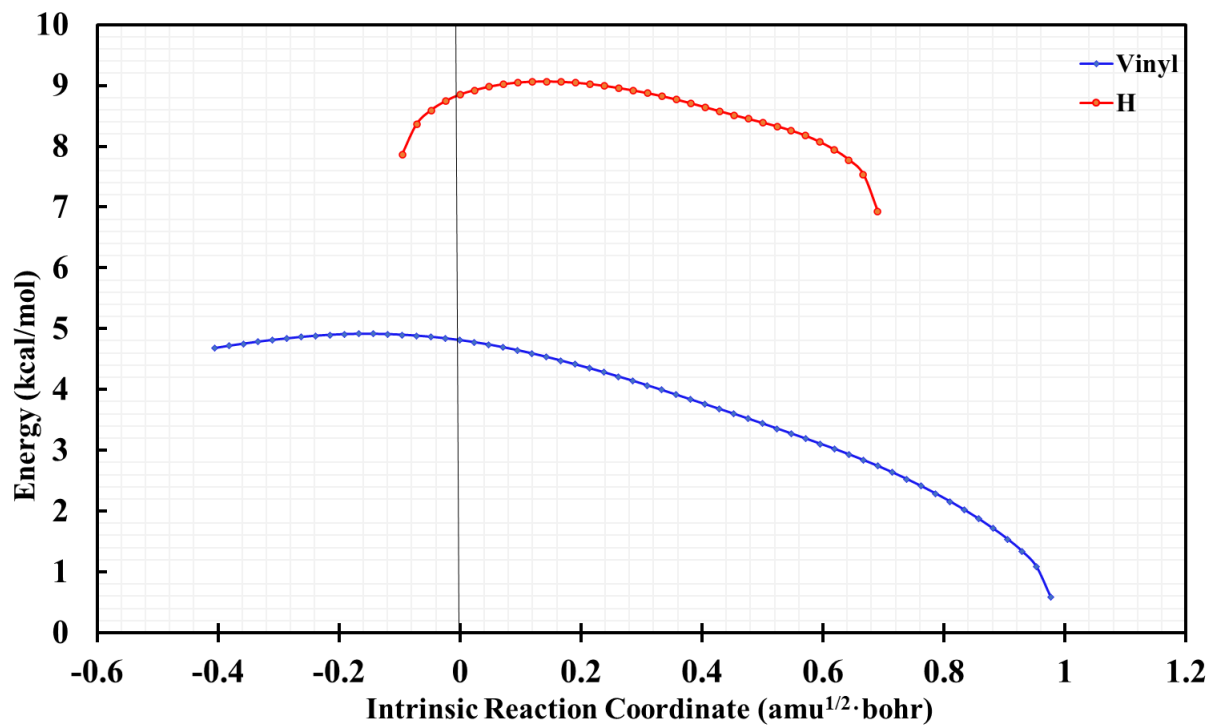
Figure A4. IRC of (R^1 = H, R^2 = Vinyl)
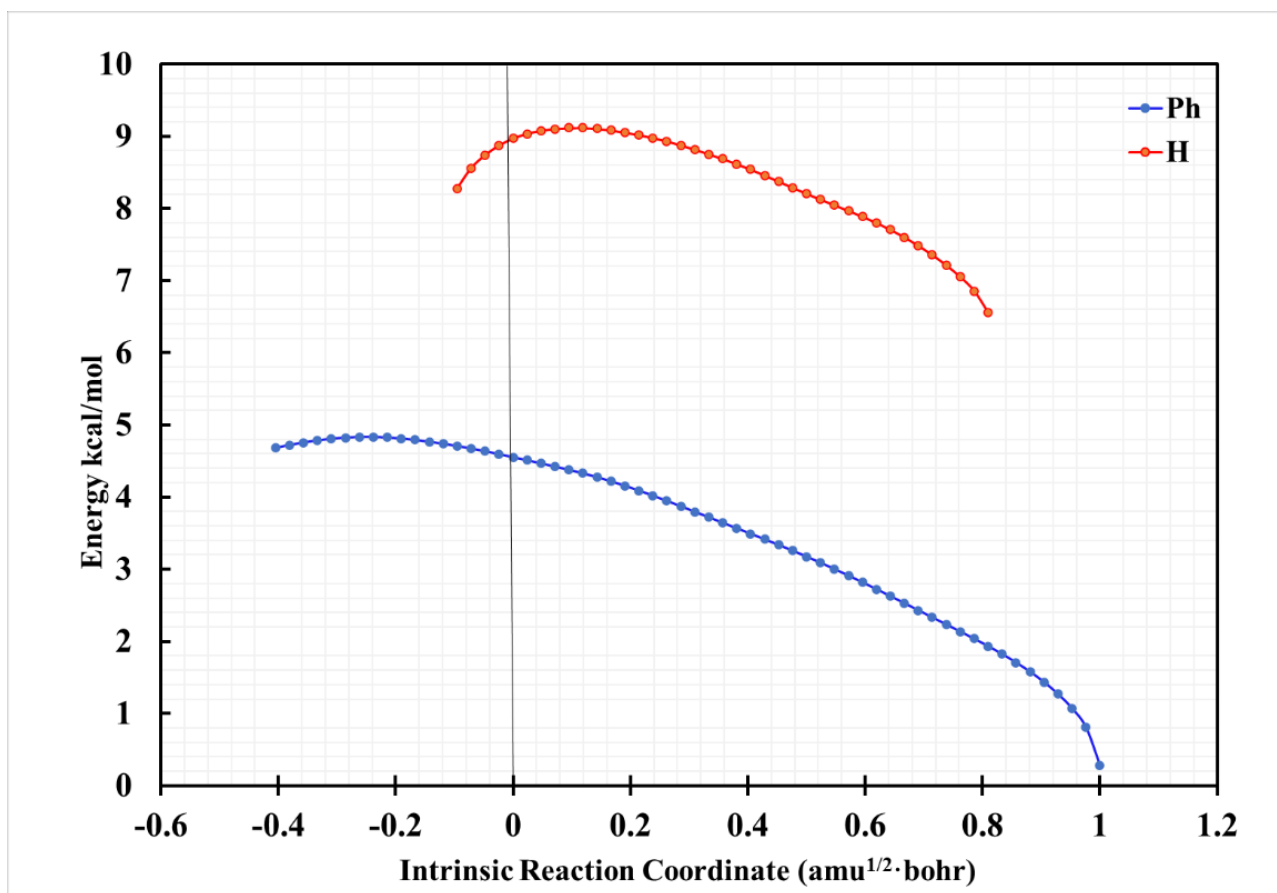
Figure A5. IRC of (R^1 = H, R^2 =Phenyl)
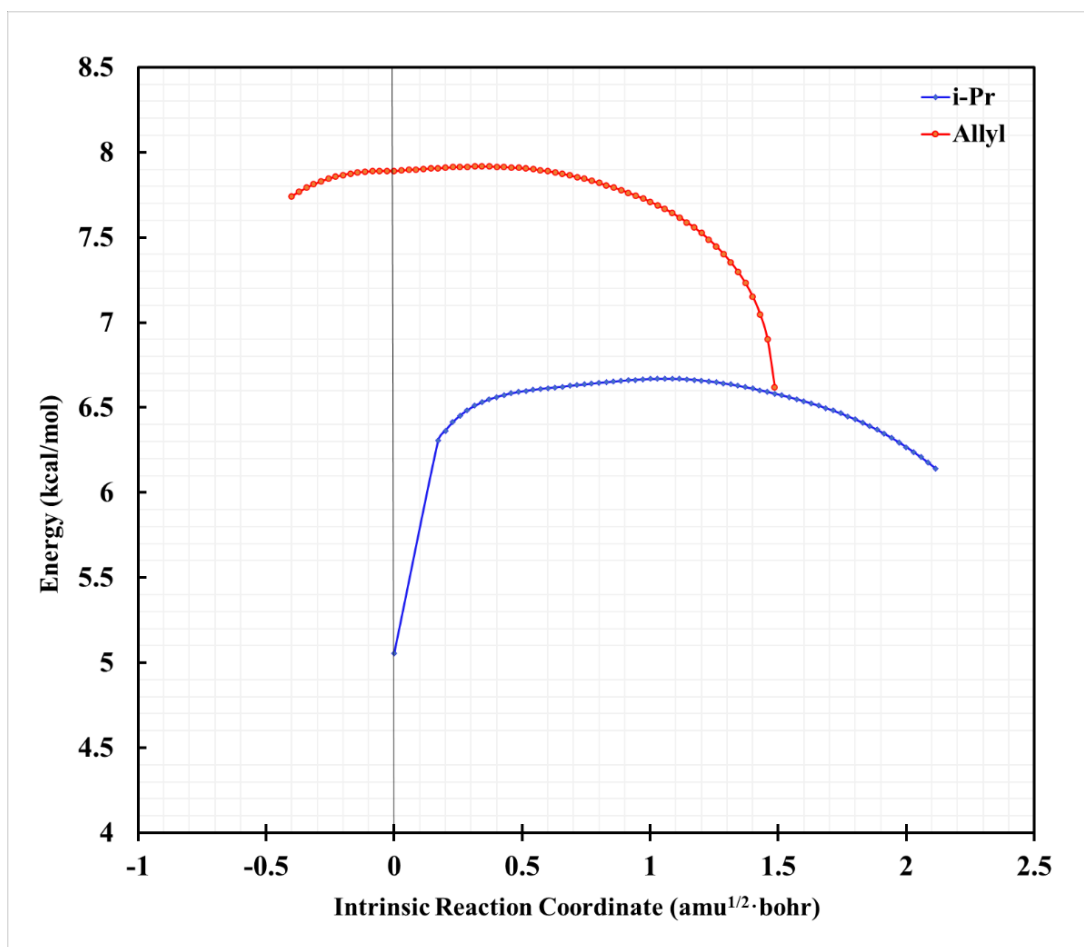
Figure A6. IRC of ($R^1 = i$-Pr, R^2 =Allyl)
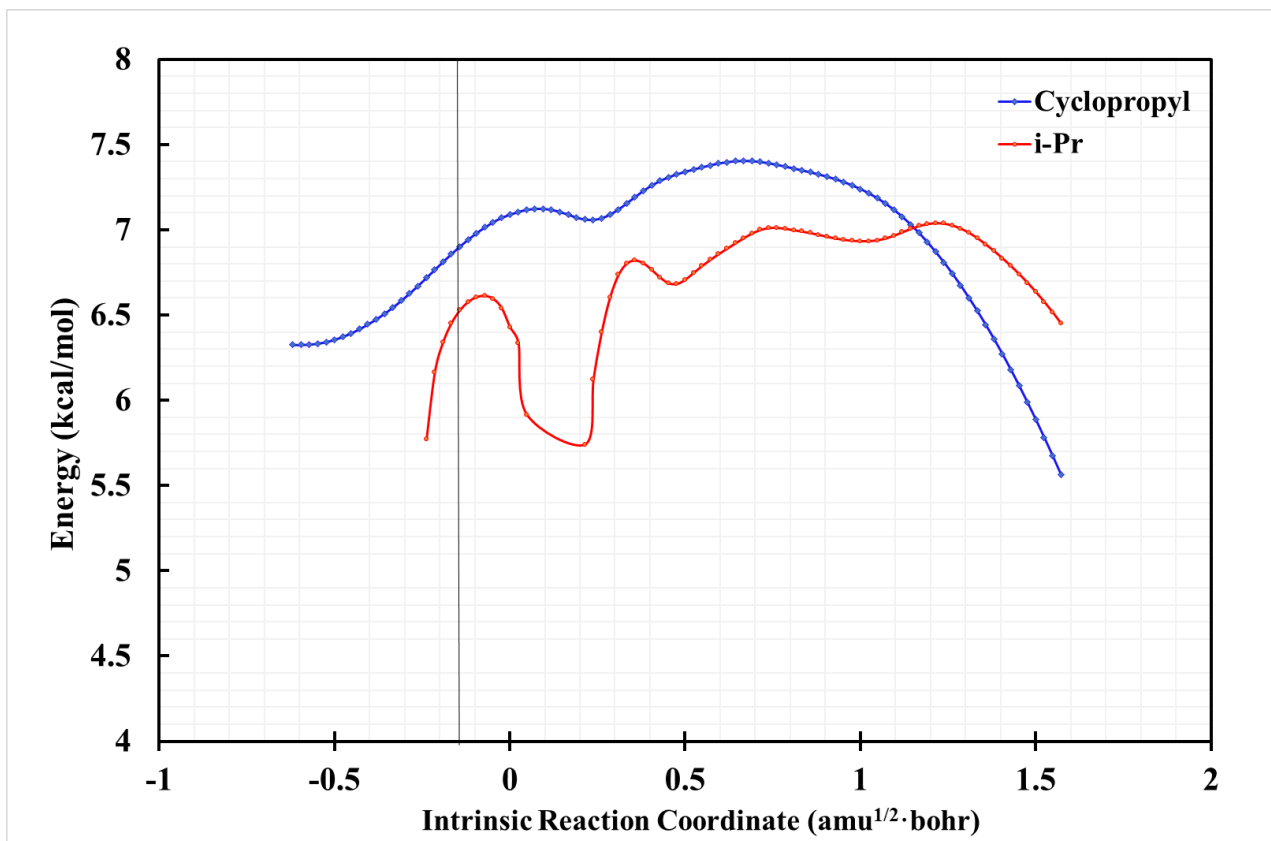
Figure A7. IRC of (R^1 = *i*-Pr, R^2 =Cyclopropyl)
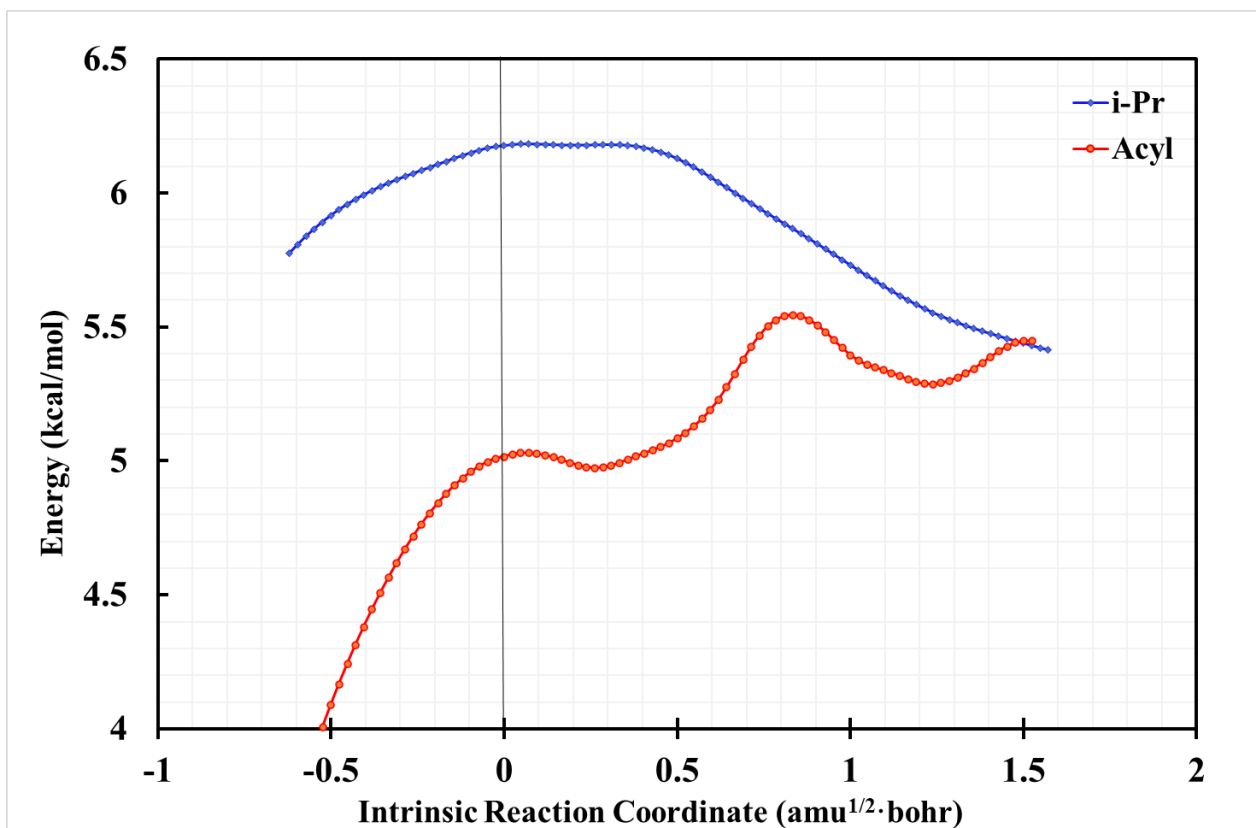
Figure A8. IRC of ($R^1 = i$-Pr, $R^2 =$ Cyclopropyl)
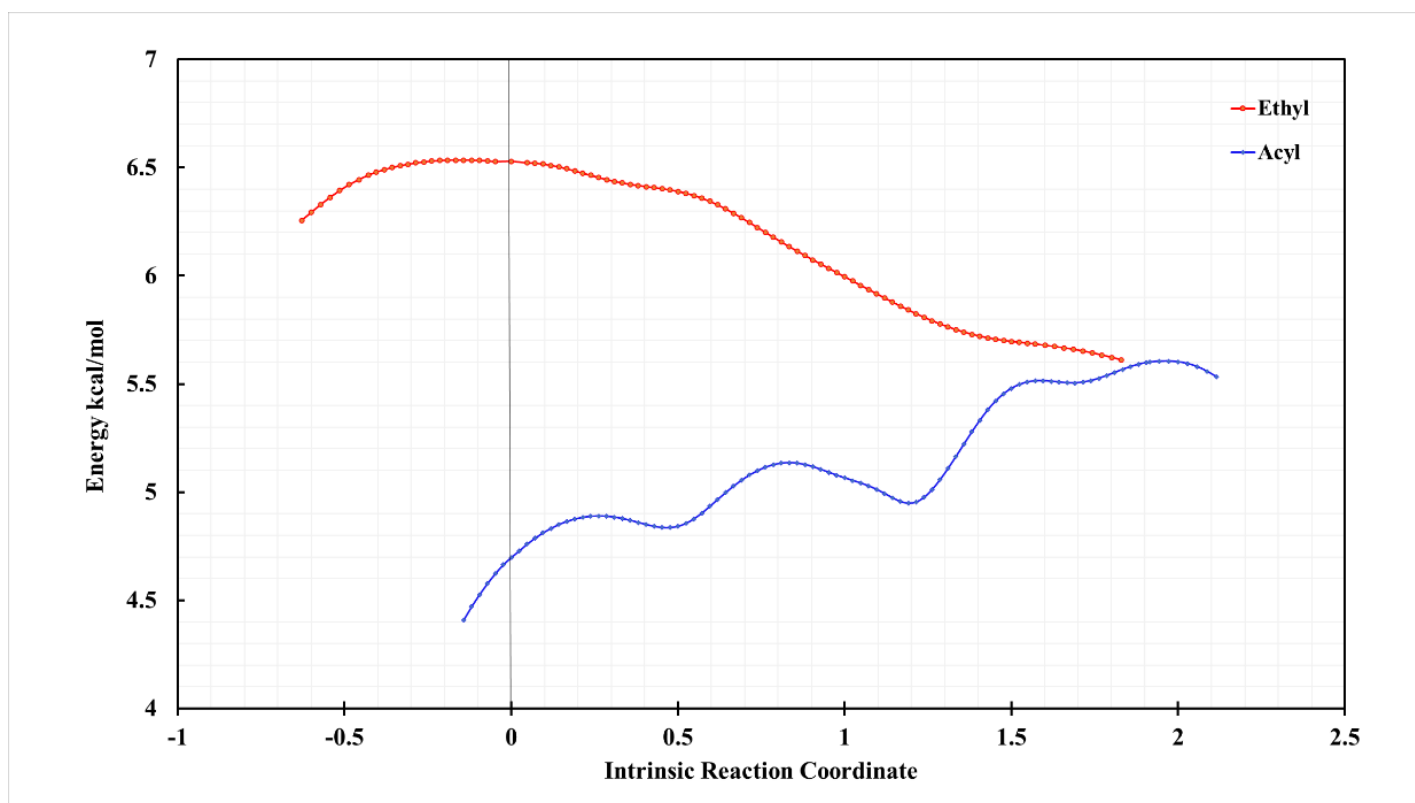
Figure A9. IRC of ($R^1 = $ Ethyl, $R^2 =$ Acyl)
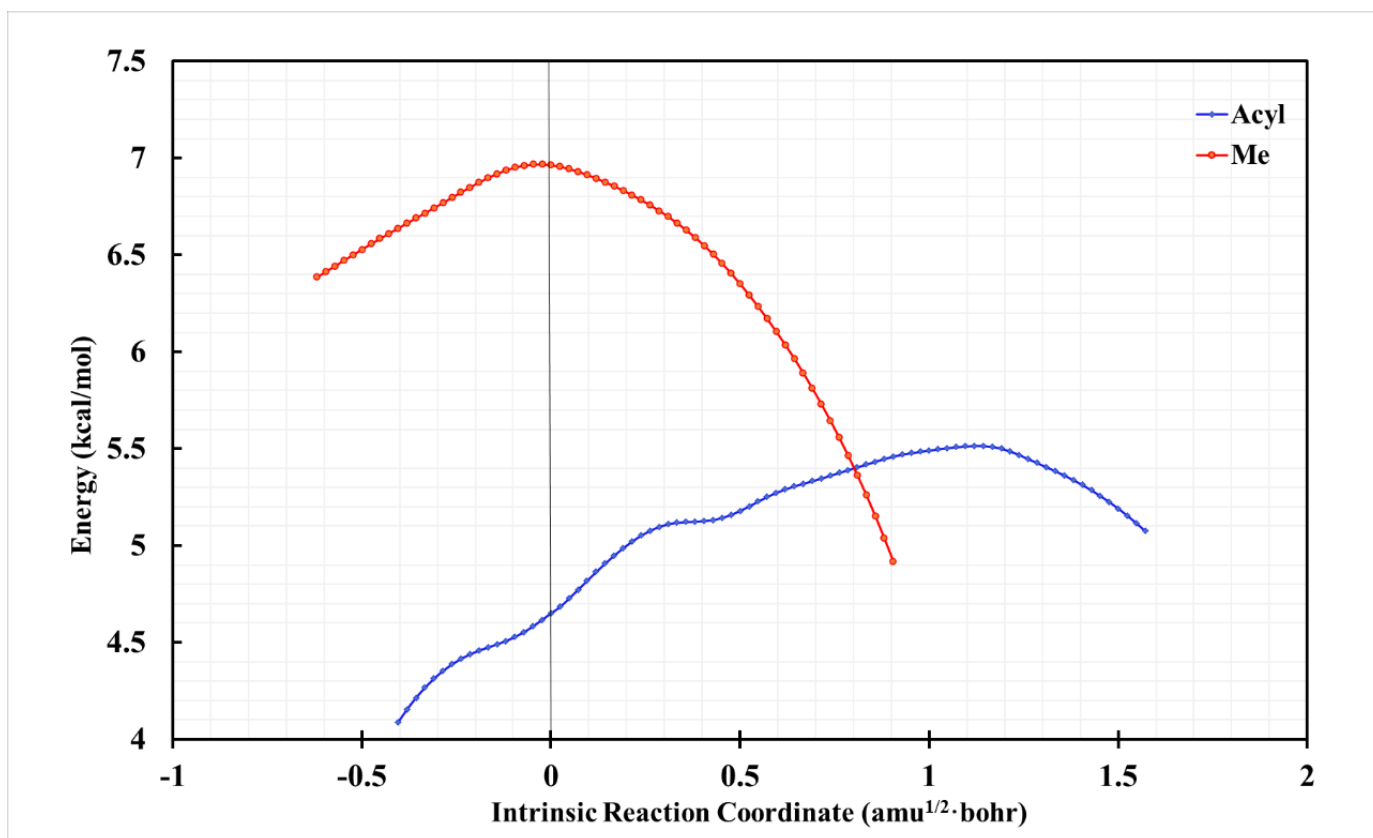
Figure A10. IRC of (R^1 = Methyl, R^2 =Acyl)
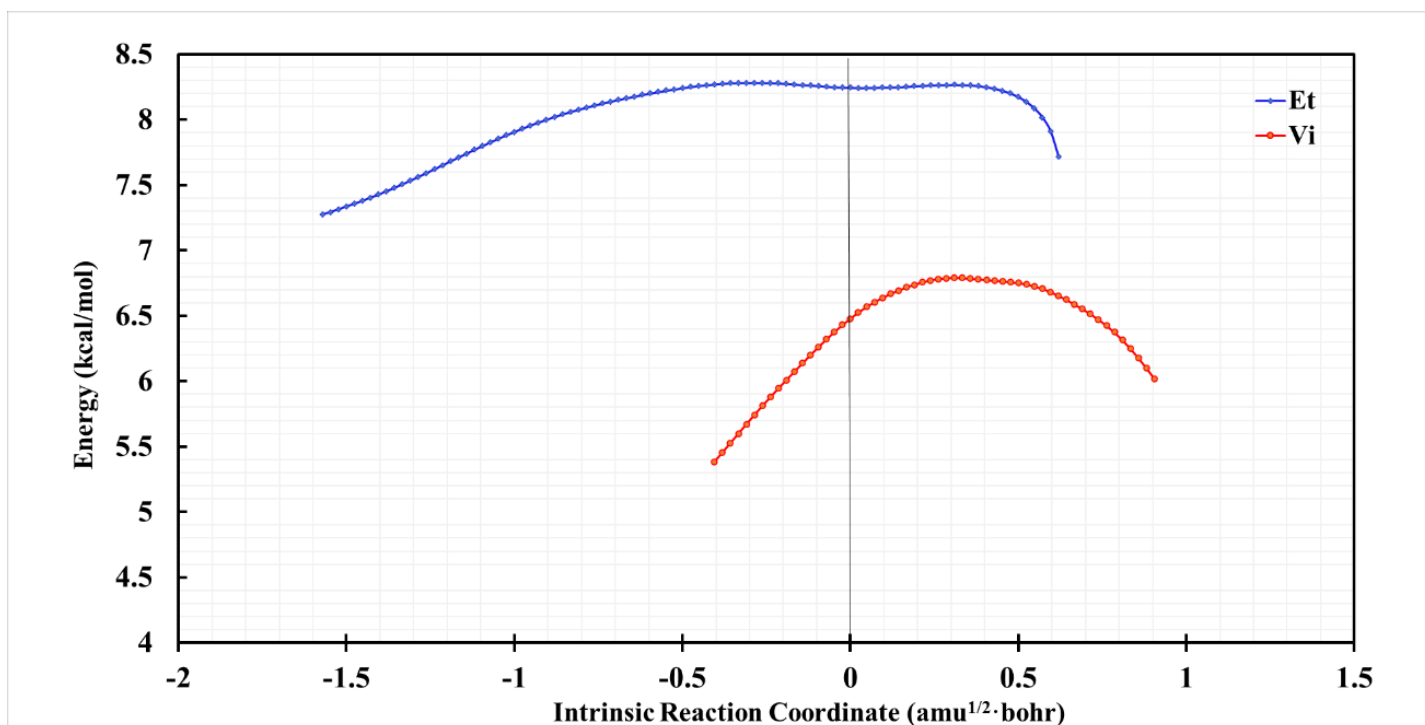
Figure A11. IRC of (R^1 = Ethyl, R^2 =Vinyl)

154
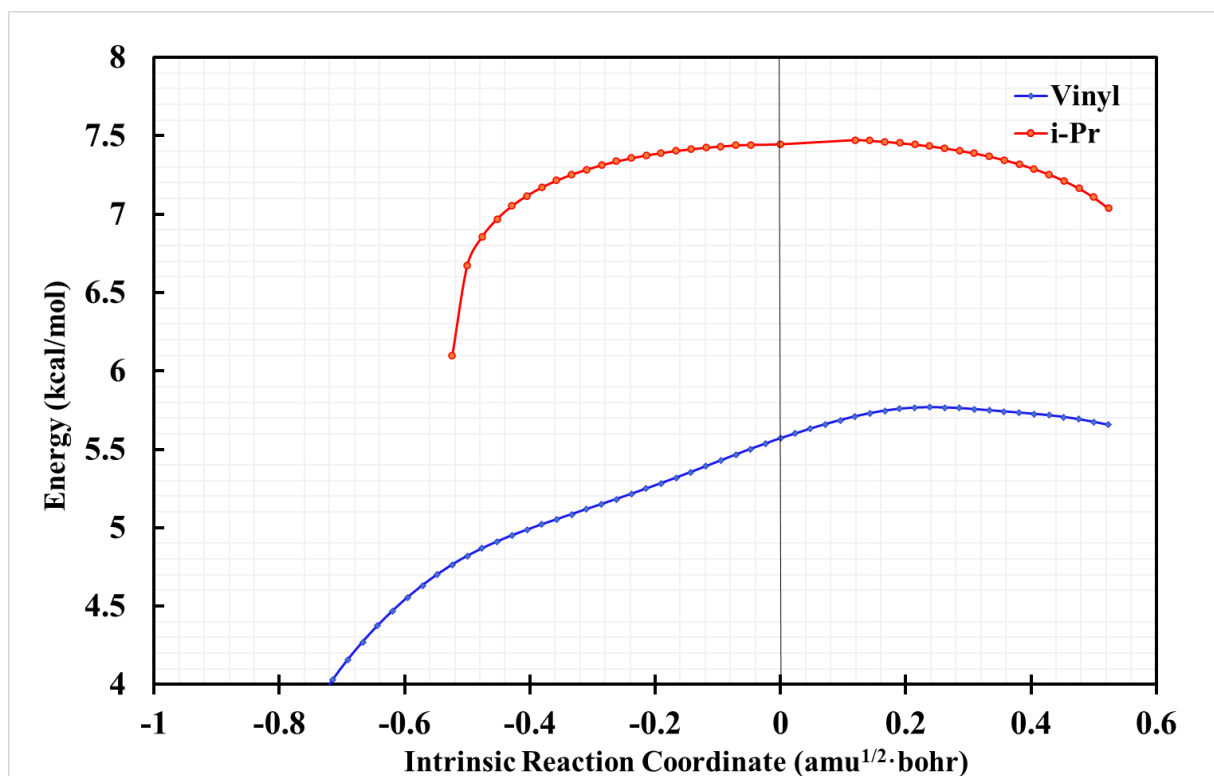
Figure A12. IRC of ($R^1 = i$-Pr, $R^2 =$Vinyl).
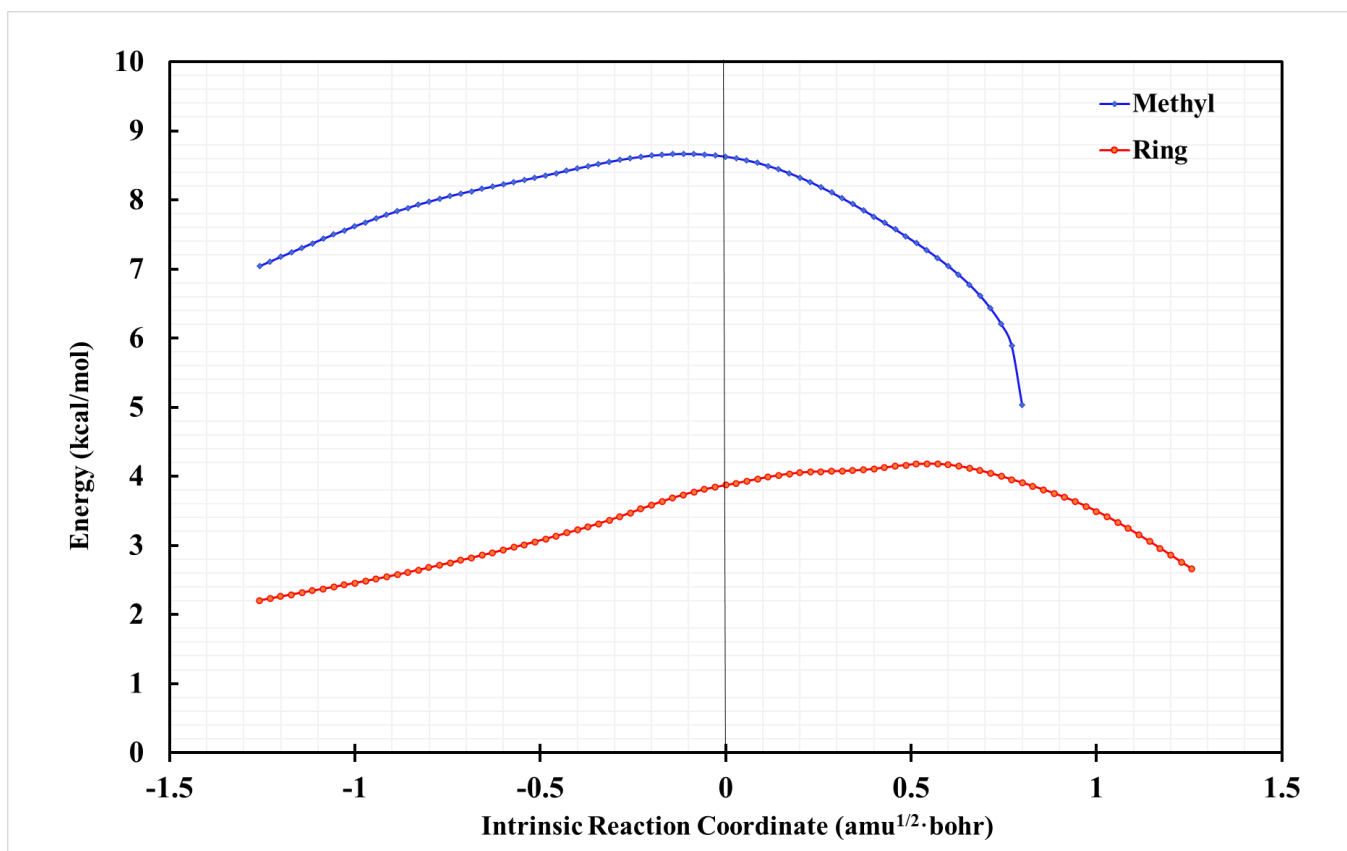
Figure A13. IRC of the Nopsane system.

Supplementary electronic potential energy (PE) and zero-point energy (ZPE)

6. (R^1 = Methyl, R^2 = Methoxymethyl)
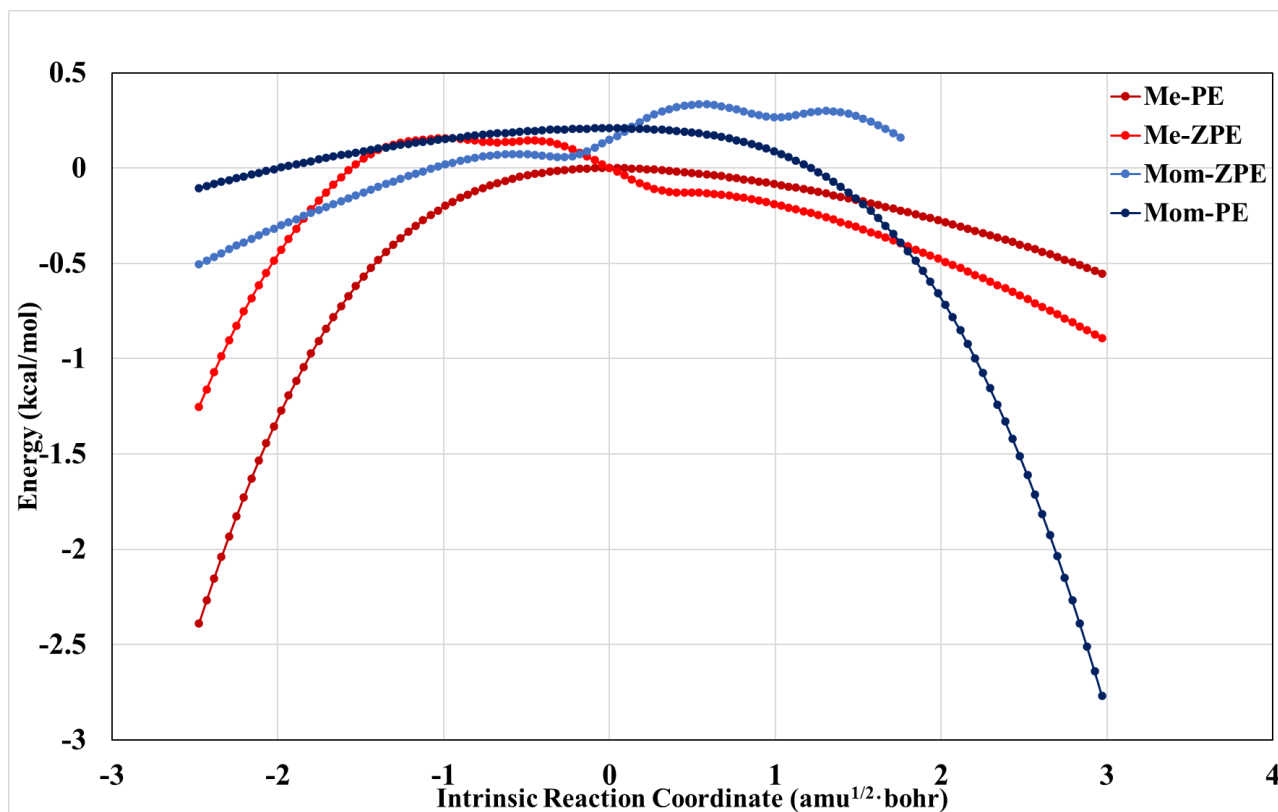
Figure A14. PE and ZPE along IRC for (R^1 = Methyl, R^2 = Methoxymethyl)
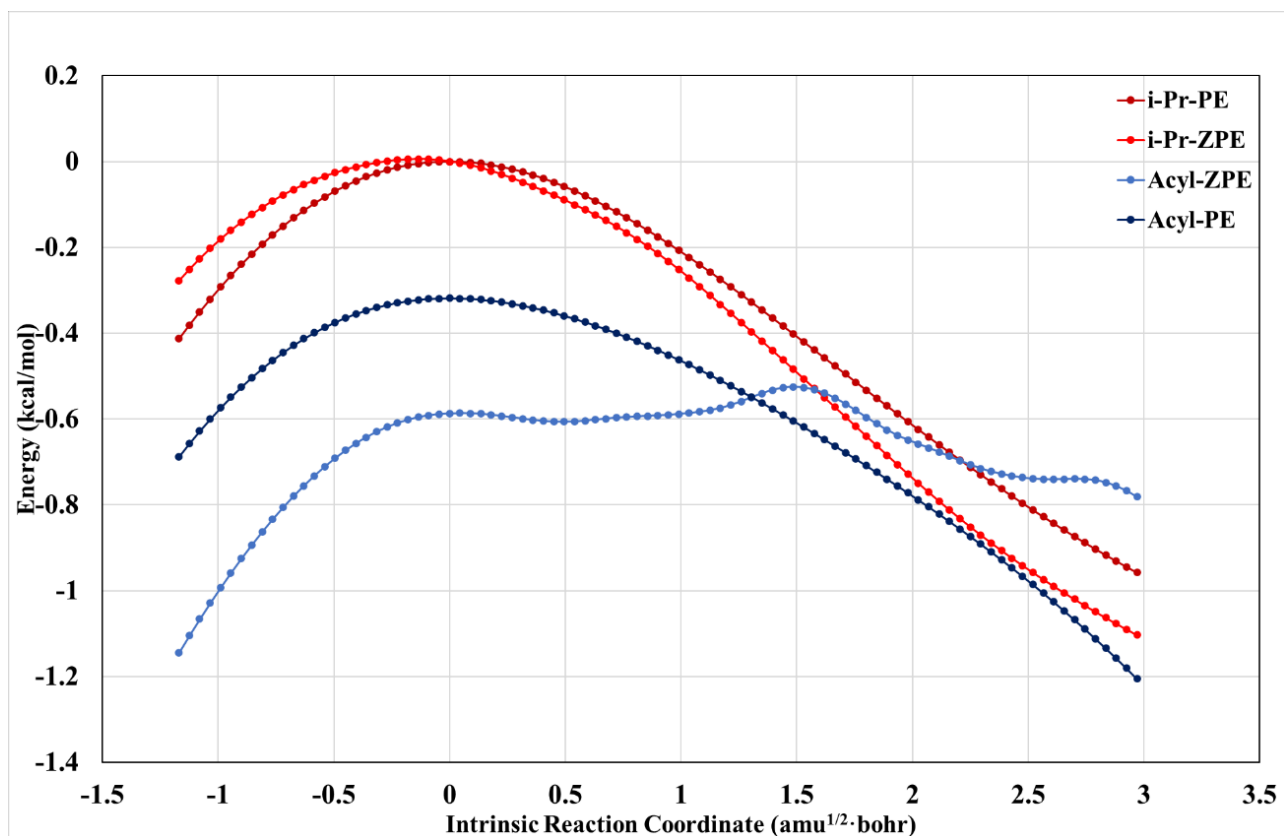
Figure A15. PE and ZPE along IRC for (R^1 = *i*-Pr, R^2 =Acyl)
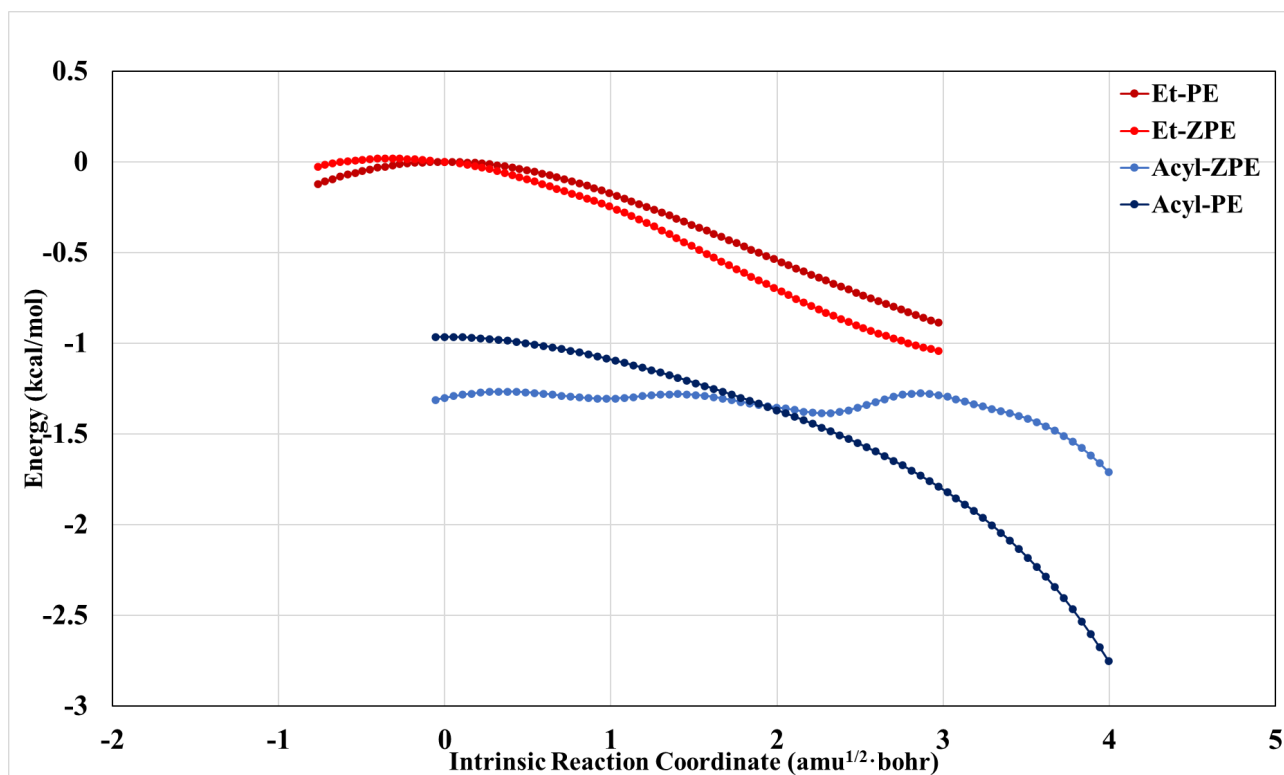
Figure A16. PE and ZPE along IRC for (R^1 = Ethyl, R^2 =Acyl)

Free Energy Barriers in Nopsane Synthesis

To verify how different levels of theory could affect the $\Delta\Delta G^{\ddagger}$ of the migrations, we have picked several methods, basis sets and solvation models to do geometry optimizations and single point energies. Tight-PNO is used for all DLPNO-CCSD calculations.

Table A5. Relative free energies of carbocation rearrangement in the synthesis of nopsane sesquiterpene **in vacuum.**

Level of theory	$\triangle$G(TS-A)	$\triangle$G (TS-B)	$\triangle\triangle$G‡
M06-2X-D3/6-31G*	3.9	8.6	4.8
M06-2X-D3/def2-SVP	5.0	9.3	4.3
M06-2X-D3/6-311G*	3.0	8.1	5.1
M06-2X-D3/def2-TZVP	2.0	4.4	2.4
B3LYP-D3(BJ)/def2-TZVP	-0.6	1.8	2.5
ωB97X-D/ def2-TZVP	0.8	2.7	2.0
PBE0-D3(BJ)/def2-TZVP	1.0	3.6	2.5
MN15/def2-TZVP	1.6	4.4	2.9
DLPNO-CCSD(T)/def2-TZVPP[1]	0.2	3.1	3.0
DLPNO-CCSD(T)/def2-QZVPP[1]	0.4	3.7	3.3

1. Optimization at M06-2X-D3/def2-TZVP level of theory

Gibbs free energy plot of hydride affinity
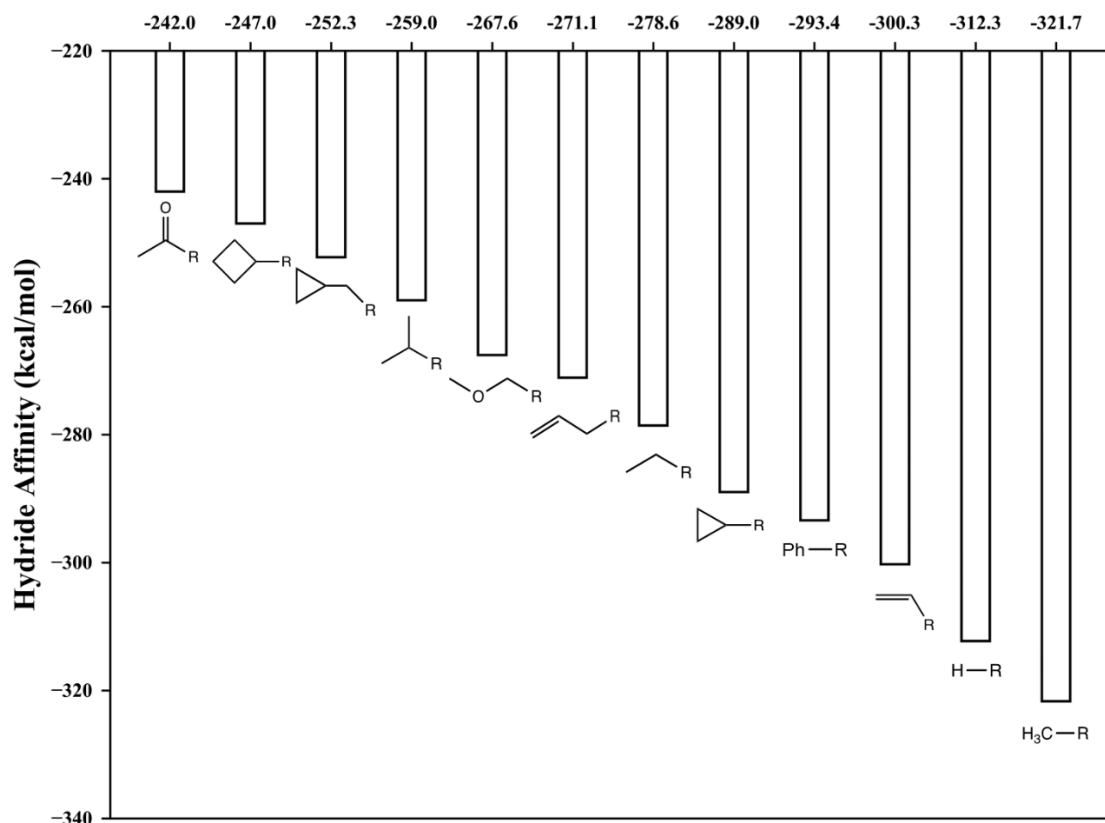
Fig A17. Hydride affinity of Gibbs free energy. The calculation is performed at M06-2X-D3/def2-TZVPP level of theory. The result is consistent with that of the enthalpy.

Further discussions on solvent effect.

Most of the results discussed in the main text are in the context of the gas-phase. Although implicitly or explicitly incorporating solvent, raises concerns and can be associated with technical difficulties, its potential significance should not be overlooked. To verify how solvent could affect our conclusions, we performed additional studies using an implicit solvent model.

First, we examined the influence on $\Delta\Delta G^{\ddagger}$ from solvent for the nopsane synthesis case. The SMD model was employed and the activation barriers were calculated at several levels. Although the absolute barrier heights are generally higher than those in the gas phase, all the

$\Delta\Delta G^{\ddagger}$ values are still significantly higher than 1.1 kcal/mol. Thus, taking solvent into account in this way does not improve the overall TST predictions.

Table A6. Relative free energies of carbocation rearrangement in the synthesis of nopsane sesquiterpene **with SMD solvation model (acetone).**

Level of theory	$\triangle$G(TS-A)	$\triangle$G (TS-B)	$\triangle\triangle$G‡
M06-2X-D3/6-31G*	8.3	12.7	4.4
M06-2X-D3/def2-TZVP	5.6	7.9	2.4
B3LYP-D3(BJ)/def2-TZVP	0.3	3.9	3.6
B2PLYP-D3(BJ)/ def2-QZVPP[1]	1.8	5.0	3.3
ωB97M-V/def2-QZVPP [1]	4.9	6.9	2.0

1. Optimization at M06-2X-D3(SMD)/def2-TZVP level of theory.

The solvent effect also can impact the shape of the PES or Free energy surface (FES). Consequently, we computed the free energy IRC and located the VTS using SMD, as shown in Fig A17. The $\Delta\Delta G^{\ddagger}$ based on this VTS is 4.0 kcal/mol, which again fails to rationalize the experimental selectivity.
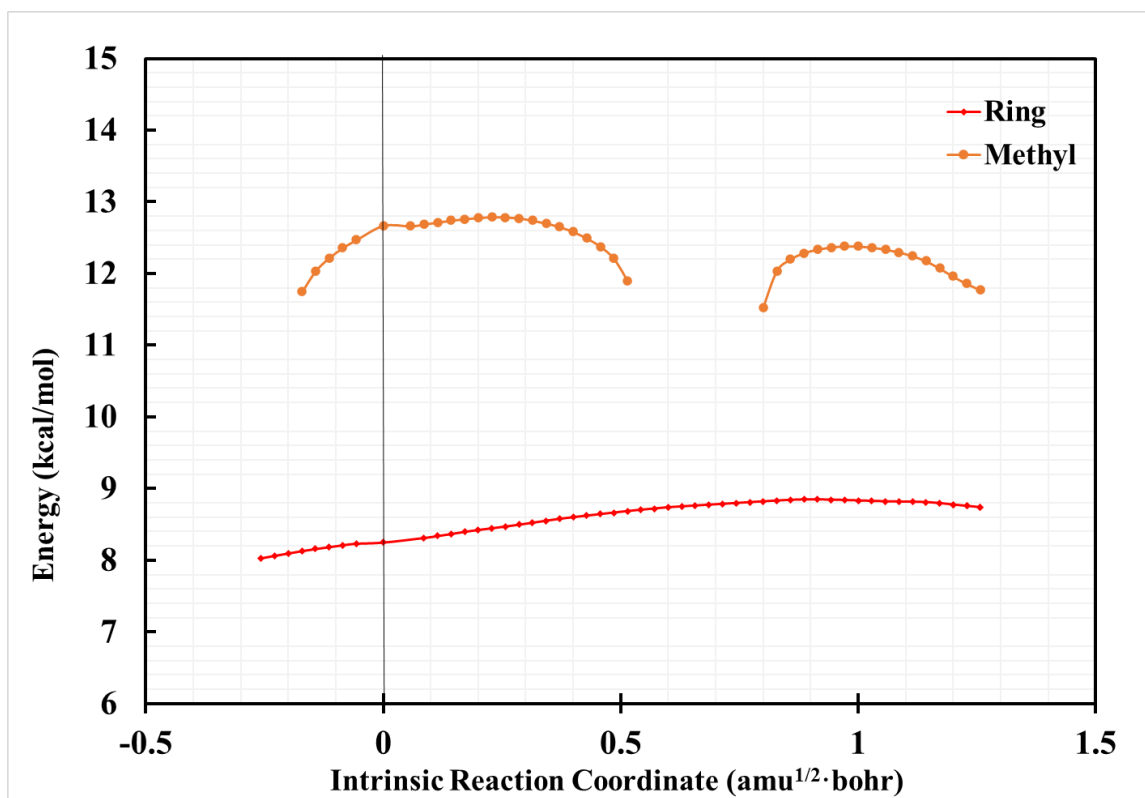
Figure A18. Free energy profiles for the nopsane system under M06-2X-D3(SMD)/6-31G(d) level of theory.

As is indicated in the main text, only by taking non-statistical effects into consideration were we able to reproduces the experimental selectivity. Thus, we incorporated implicit solvent into uphill dynamics simulations and observed essentially no change to the product ration from gas phase uphill AIMD (89:11, 118 trajectories).

We also examined the influence of implicit solvation for our model systems. First, the $\Delta\Delta G^{\ddagger}$ for various pairs of functional groups were calculated, and the results are shown in Table A7. In general, the barriers for migration are higher when solvent effect is taken into account, but $\Delta\Delta\Delta G^{\ddagger}$ values are much less affected.

Table A7. Results of conventional free energy barriers in gas-phase and solvent (M06-2X-D3/6-31G*), energies in kcal/mol).

	$\Delta G^{\ddagger}_{TSS(gas)}$	$\Delta G^{\ddagger}_{TSS(sol)}$	$\Delta G^{\ddagger}_{TSS(gas)}$	$\Delta G^{\ddagger}_{TSS(sol)}$	$\Delta\Delta G^{\ddagger}_{TSS(gas)}$	$\Delta\Delta G^{\ddagger}_{TSS(sol)}$

1 H—R	9.5	15.4	H₃C—R	9.7	16.5	0.2	1.1
3 H—R	9.2	14.9	▷—R	6.6	14.1	-2.6	-0.8
4 H—R	8.8	16.0	═\R	4.8	12.7	-4.0	-3.3
5 H—R	9.0	16.0	Ph—R	4.6	13.1	-4.4	-2.9
6 H₃C—R	9.6	15.9	⟍O⟋R	9.7	17.1	0.1	1.2
11 ⟍⟋R	5.1	14.2	⟋⟍R	7.9	15.6	2.8	1.4
13 ⟍⟋R	6.2	15.7	(acyl) R	5.2	14.4	-1.0	-0.7
14 ⟋⟍R	6.5	15.5	(acyl) R	4.7	13.9	-1.8	-1.6
15 H₃C—R	7.0	14.3	(acyl) R	5.1	15.3	-1.9	1.0

In addition, we ran uphill AIMD for a selection of systems: **4** (R^1 = H, R^2 = vinyl) and **15** (R^1 = methyl, R^2 = acyl). The results differ from the gas phase results. For system **4**, the ratio of vinyl:H migration is 54:46 (136 trajectories), indicating much worse selectivity. The selectivity of system **15** is reversed, with a methyl:acyl ratio of 67:33 (58 trajectories) comparing to a 26:74 ratio in the gas-phase. Future work will involve looking into the origins of such effects for this and other systems – we are not convinced that continuum solvation is sufficient for capturing the influence of solvent on non-statistical dynamic effects.**Error! Reference source not found.**

PROGDYN setup

This is the configuration file for PROGDYN. This file is read by progdynstarterHP and

the awk programs proggenHP, prog1stpoint, prog2ndpoint, and progdynb.

#The programs won't read anything past the first blank line,

#and this file must end with a blank line.

#The program has a number of default values but they are unlikely to be what you want.

#Do not delete lines - rather, comment out lines for unwanted options.

#The values here are read repeatedly and most can be changed in the middle of running jobs

#***The keywords are case sensitive. The following keywords should always be defined:***

#***method, charge, multiplicity, memory, processors, title

#*** method --The following word is copied exactly to the gaussian input file.

method m062x/6-311G(d,p)

#*** method2 --The options here are restricted, unrestricted, and read. restricted is the default

#If the method is U..., put unrestricted here and the .com files will have in them guess=mix.

#If you put read here, the .com files will contain guess=tcheck, which sometimes makes things faster, sometimes not.

#The use of read requires a specifically defined checkpoint file name using the keyword checkpoint.

method2 restricted

charge 1

multiplicity 1

processors 16

#*** memory --The following "word" is copied exactly to the gaussian input file after %mem=.

memory 48GB

#*** killcheck and checkpoint -- You can use a specifically defined checkpoint file name by putting

#the name after the keyword checkpoint. This is necessary if you use the read option with method2.

#Defined checkpoint names are an unnecessary modest hastle and if you do not want to bother, use killcheck 1

killcheck 1

#checkpoint dyn20.chk

#*** diagnostics -- 0 prints out nothing extra, 1 (default) prints out extra stuff to a

#file "diagnostics", 2 adds more stuff, 3 adds velocities to a file "vellist"

diagnostics 0

#*** title -- the title keyword must be followed by exactly four words

title a b c d

#*** initialdis -- 0 (default) turns off displacement of the normal modes, so that all trajectories start from the same place

and only the energies and signs of the motion in the modes are randomized

1 gives a flat distribution of displacements where all of the possible values are equally likely

2 (recommended) gives a QM-like gaussian distribution of displacements, so that displacements in the middle are more likely

that

those at the end by 1/e

initialdis 2

#*** timestep -- this is the time between points in the trajectory. Typical values would be 1E-15 or 0.5E-15 or 0.25E-15

timestep 1E-15

#*** scaling -- this lets you scale the gaussian frequencies by a constant

scaling 1.0

temperature 298.15

#*** method3, method4, method5, and method6 -- These keywords let you add extra lines to the gaussian input file.

#method3 and method4 add lines at the top of the input after the lines defining the method, and

#this is useful to implement things like the iop for mPW1k

#method5 and method6 add lines after the geometry, after a blank line of course

#only a single term with no spaces can be added, one per method line. Here are some examples to uncomment if needed

#method3 IOp(3/76=0572004280)

#method3 scrf=(pcm,Solvent=water)

#add the line below with big structures to get it to put out the distance matrix and the input orientation

method3 empiricaldispersion=gd3

#method4 int=ultrafine

#method5 radii=bondi

#method6

#*** methodfile -- This keyword lets you add more complicated endings to gaussian input files

#such as a gen basis set. Put after the keyword the number of lines in a file you create called

#methodfile that contains the test you want to add to the end of the gaussian input

methodfile 0

#*** numimag --This tells the program the number of imaginary frequencies in the starting structure.

#if 0, treats as ground state and direction of all modes is random

#if 1, motion along the reaction coordinate will start out in the direction defined by searchdir

#if 2, only lowest freq will go direction of searchdir and other imag mode will go in random direction

numimag 1

#*** searchdir -- This keyword says what direction to follow the mode associated with the imaginary frequency.

#The choices are "negative" and "positive". Positive moves in the direction defined in the gaussian frequency calculation

#for the imaginary frequency, while negative moves in the opposite direction. The correct choice can be made either

#by a careful inspection of the normal modes and standard orientation geometry, or by trial and error.

searchdir positive

#*** classical -- for quassiclassical dynamics, the default, use 0. for classical dynamics, use 1

#if there are no normal modes and the velocities are to be generated from scratch, use classical 2

classical 0

#*** DRP, saddlepoint, and maxAtomMove --to run a DRP use 'DRP 1' in the line below, otherwise leave it at 0 or comment it out

#the treatment of starting saddlepoints is not yet implemented so use saddlepoint no

#if DRP shows oscillations then decrease maxAtomMove

#DRP 1

#saddlepoint no

#maxAtomMove 0.01

#*** cannonball -- The program can "fire" a trajectory from a starting position toward a particular target, such as toward

#a ts. To use this, make a file cannontraj with numAtom lines and three numbers per line that defines the vector

#for firing the trajectory, relative to the starting geometry's standard orientation. The number following cannonball sets

#the extra energy being put into the structure in kcal/mol

cannonball 25

#*** keepevery --This tells the program how often to write the gaussian output file to file dyn, after the first two points.

#Use 1 for most dynamics to start with, but use a higher number to save on disk space or molden loading time.

keepevery 1

#*** highlevel --For ONIOM jobs, the following line states the number of highlevel atoms,

#which must come before the medium level atoms. Use some high value such as 999 if not using ONIOM

highlevel 999

#*** fixedatom1, fixedatom2, fixedatom3, and fixedatom4 - These fix atoms in space.

#Fixing one atom serves no useful purpose and messes things up, while fixing two atoms

#fixes one distance and fixing three has the effect of fixing three distances, not just two

#in current form fixed atoms only are meant to work with no displacements, that is, initialdis=0

#fixedatom1 2

#fixedatom2 3

#fixedatom3 19

#*** boxon and boxsize - With boxon 1, a cubic box is set such that atoms that reach the edge

#are reflected back toward the middle. Useful for dynamics with solvent molecules. This is a crude

#implementation that is ok for a few thousand femtoseconds but will not conserve energy long term.

#Set the box size so as to fit the entire initial molecule but not have too much extra room.

#The dimensions of the box are two times the boxsize, e.g. boxsize 7.5 leads to a box that is 15 x 15 x 15 angstroms

boxon 0

boxsize 7.5

#*** displacements -- This keyword lets you set the initialdis of particular modes by using a series of lines of the format

displacements NumberOfMode InitialDisForThatMode, as in the example below. You should be able to do as many of these as you like

you might consider this for rotations where a straight-line displacement goes wrong at large displacements

The choices for InitialDisForThatMode are 0, 1, 2, and 10, where 10 does the same thing as 0 but is maintained for now because

a previous version of the program had a bug that made 0 not work.

#displacements 1 0

#*** etolerance --This sets the allowable difference between the desired energy in a trajectory and the actual

#energy, known after point 1 from the potential energy + the kinetic energy in the initial velocities.

#The unit is kcal/mol and 1 is a normal value for mid-sized organic systems. For very large and floppy molecules, a larger value

#may be needed, but the value must stay way below the average thermal energy in the molecule (not counting zpe).

#If initialdis is not 0 and few trajectories are being rejected, decrease the value.

etolerance 999999

#*** controlphase --It is sometimes useful to set the phase of particular modes in the initialization of trajectories.

#The format is controlphase numberOfModeToControl positive or controlphase numberOfModeToControl negative.

#controlphase 2 positive

#*** damping -- The damping keyword lets you add or subtract energy from the system at each point, by multiplying the velocities

#by the damping factor. A damping of 1 has no effect, and since you mostly want to change the energy slowly, normal values range

#from 0.95 to 1.05. The use of damping lets one do simulated annealing - you add energy until the structure is moving enough

#to sample the kinds of possibilities you are interested in, then you take away the energy slowly.

damping 1

#*** reversetraj --This keyword sets the trajectories so that both directions from a transition state are explored.

reversetraj false

B. Supporting Information for chapter 3:

Intrinsic Reaction Coordinates
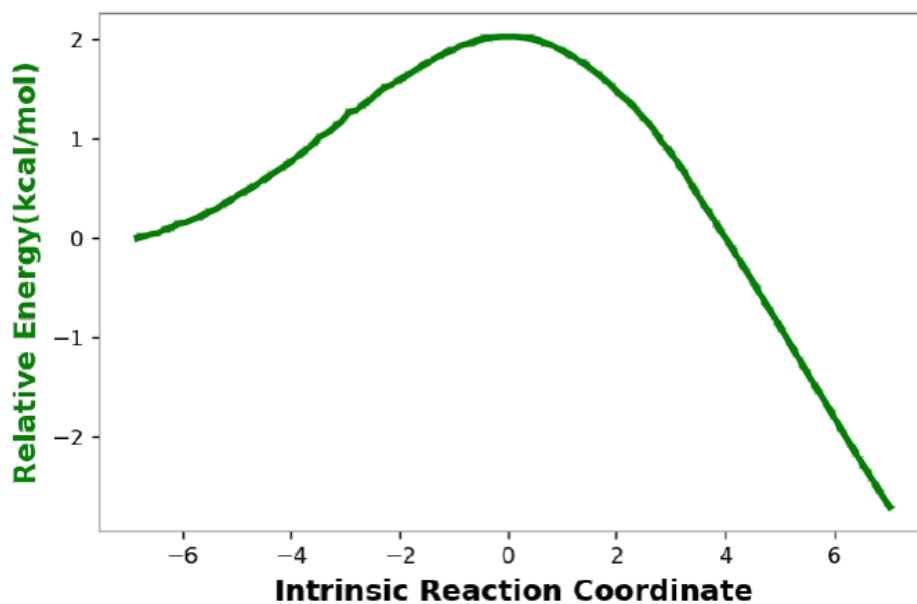
TS-a
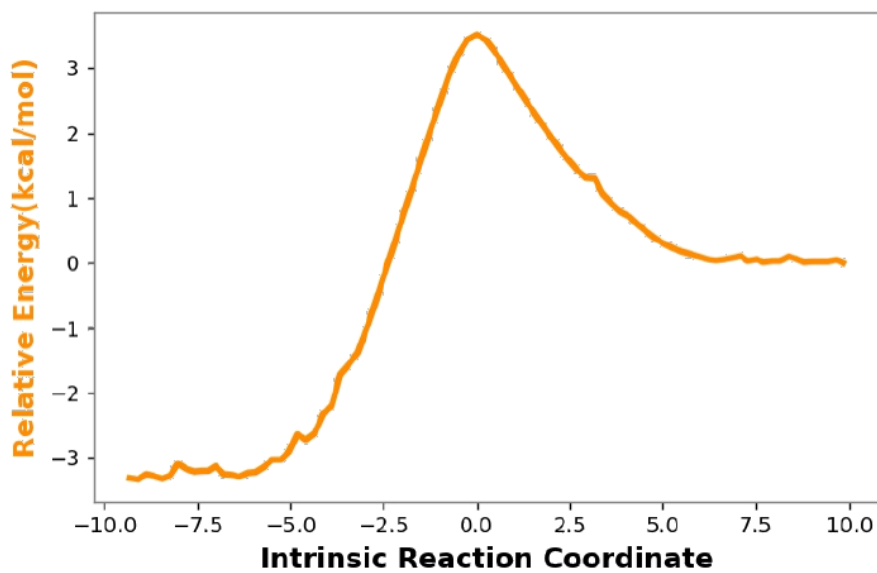
TS-b

Figure B1. IRC under B3LYP-D3/6-311+G(d,p) level of theory.
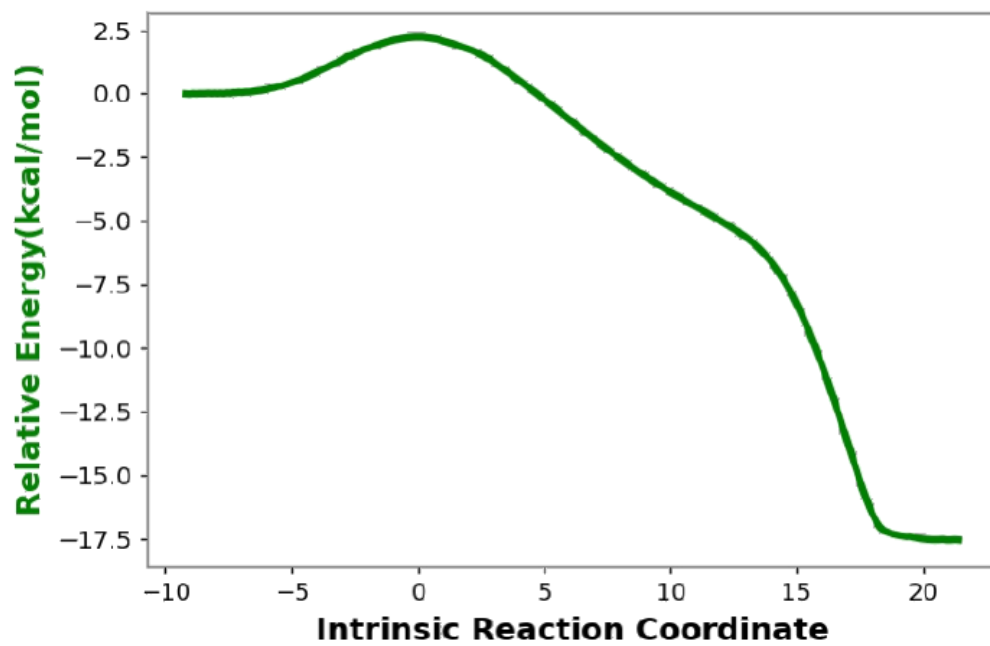
TS-a
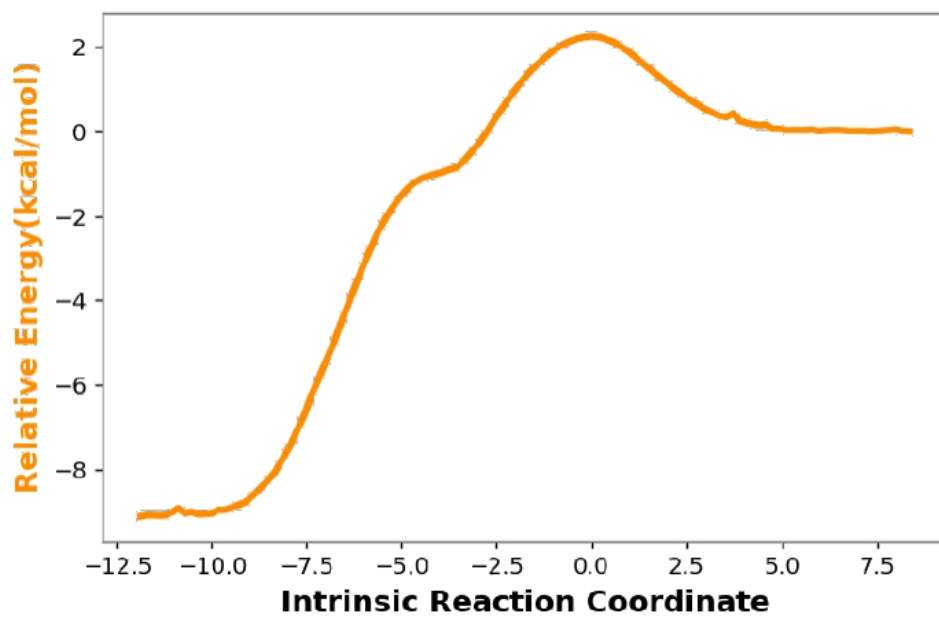
TS-b

Figure B2. IRC under M06-2X/6-311+G(d,p) level of theory.
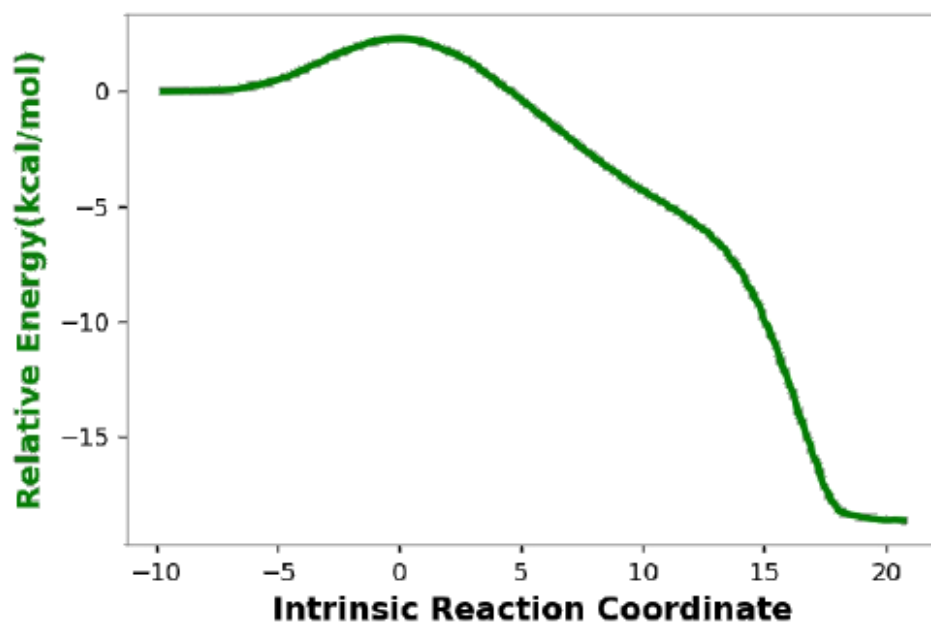
TS-a
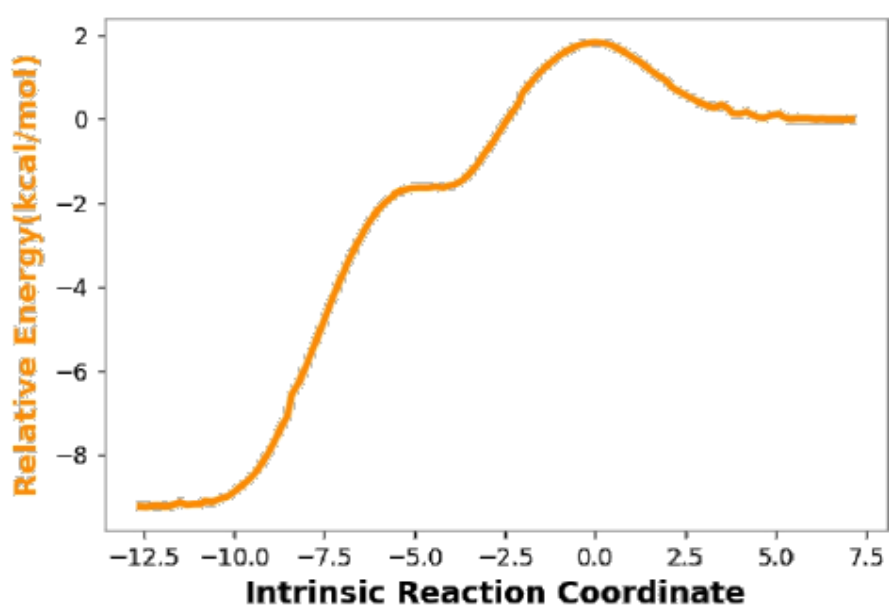
TS-b

Figure B3. IRC under M06-2X/6-311G(d) level of theory.
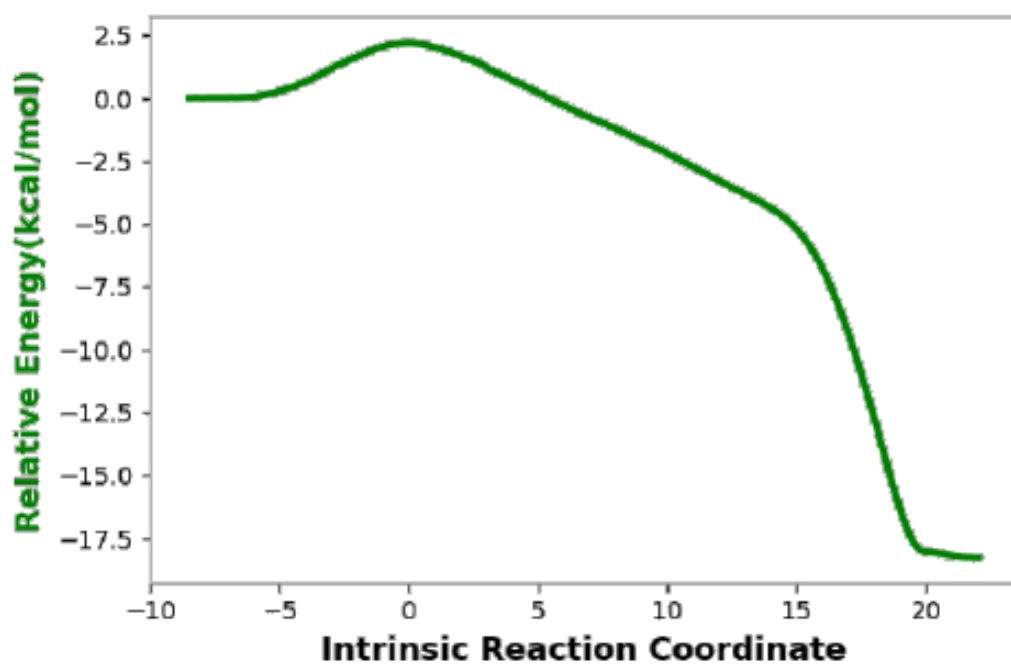
TS-a
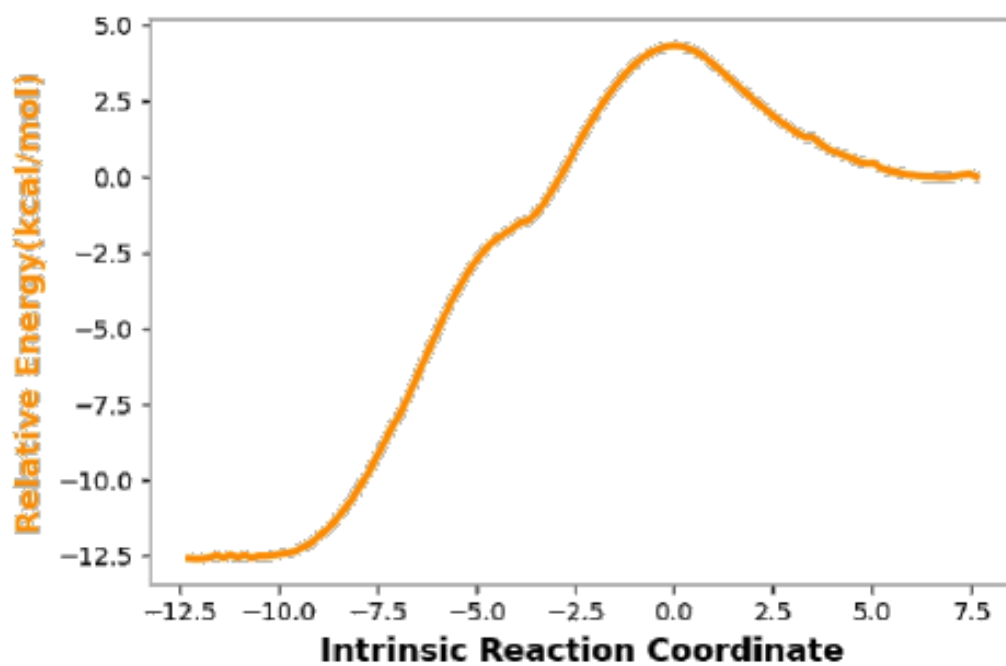
TS-b

Figure B4. IRC under PCM(DMSO)-M06-2X/6-311G(d) level of theory.
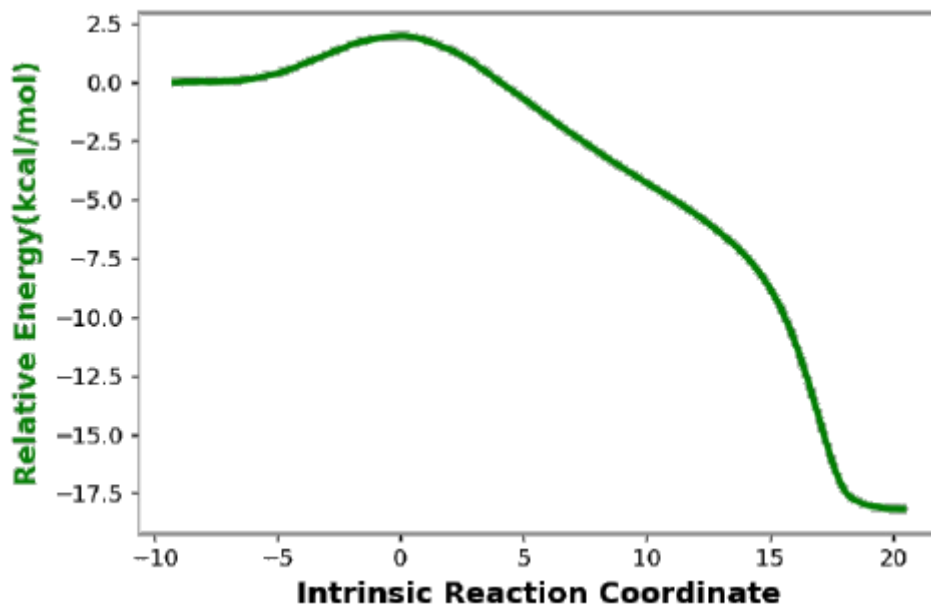
TS-a
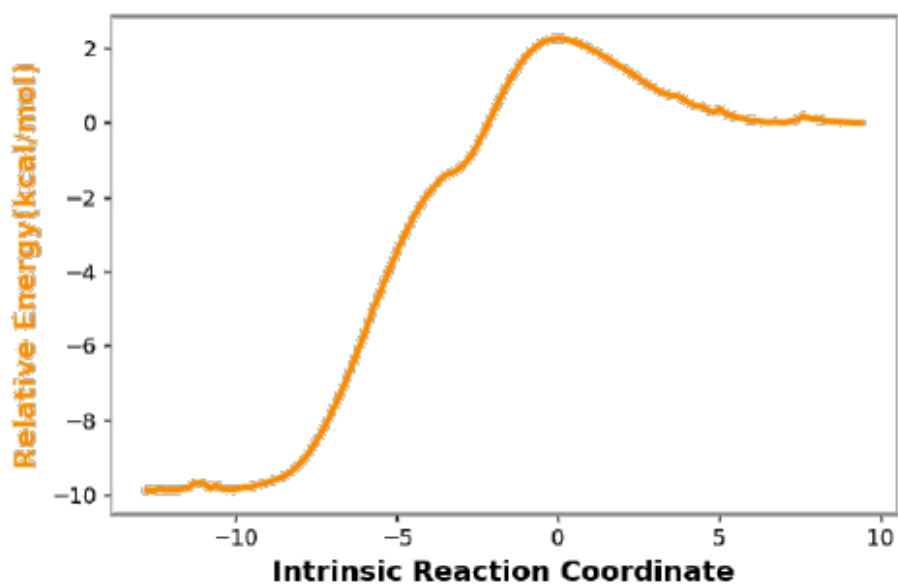
TS-b

Figure B5. IRC under wB97X-D/6-311G(d) level of theory.

Projections of trajectories onto the PCM(DMSO)-M06-2X/6-311G(d) PES.
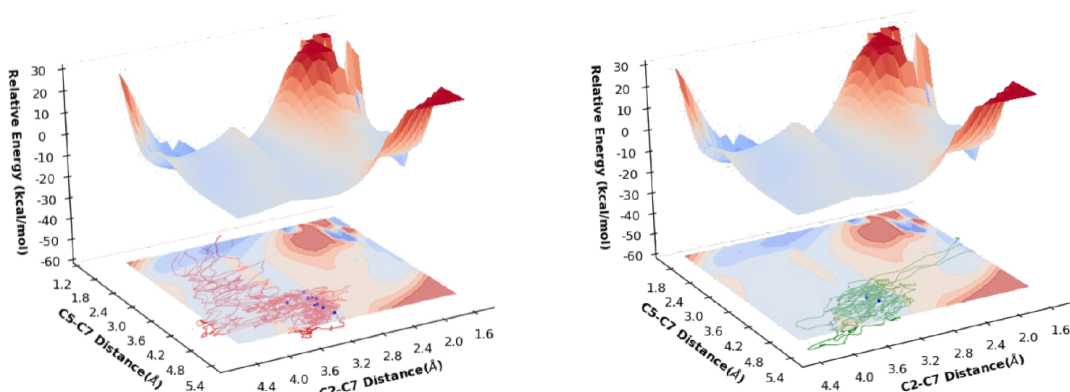
Figure B6. Projections of trajectories onto the potential energy surface at PCM(DMSO)-M06-2X/6-311G(d) level of theory. The left and right plots demonstrate (4+3) and (2+3) addition trajectories, respectively.

Timing of reactions
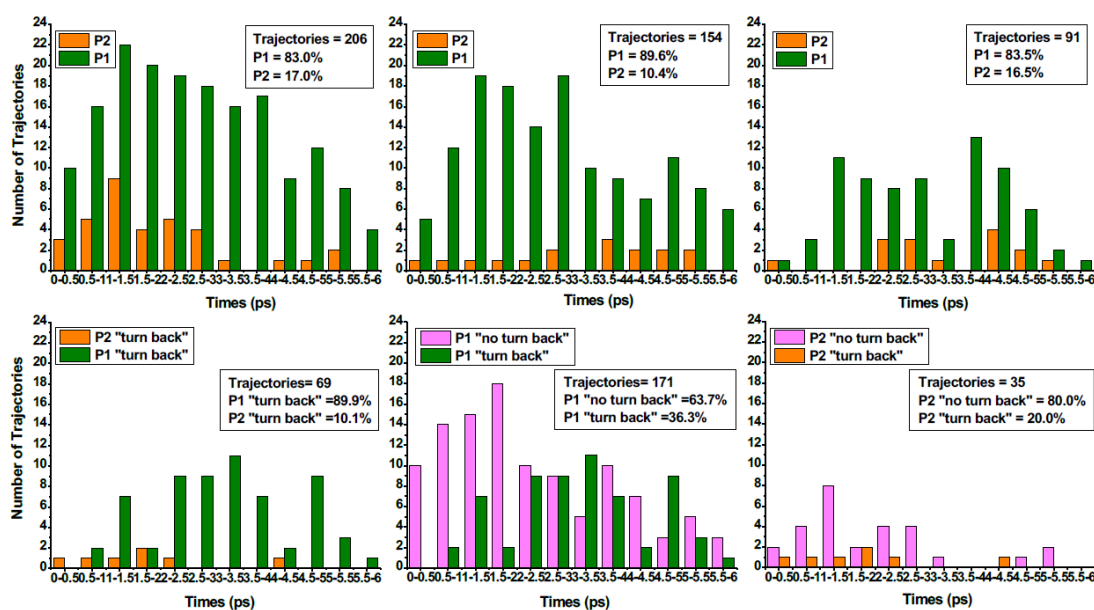
Figure B7. Histograms that bin (top) uphill trajectories (bin widths of 0.5 ps) based on the timing of formation for **Prod-1** and **Prod-2** calculated at three different levels of theory. (left to right: M06-2X/6-311G(d), ωB97X-D/6-311G(d) and PCM(DMSO)-M06-2X/6-311G(d)). Green bins correspond to **Prod-1** (P1) formation (i.e., when the C5−C7 distance drops below 1.7 Å) while yellow bins correspond to **Prod-2** (P2) formation (i.e., when the C2−C7 distance drops below 1.6 Å). (2) Histograms that bin (bottom) uphill trajectories based on the timing of product-forming "turn back"

172

trajectories discussed at M06-2X/6- 311G(d) level of theory. Green bins correspond to **Prod-1**-forming "turn back" trajectories and yellow bins correspond to **Prod-2**-forming "turn back" trajectories. Pink bins correspond to "no turn back" trajectories.

Results of "turn back" trajectories in uphill dynamics simulations.

Table B1. Results of "turn back" trajectories in uphill dynamics simulations in M06-2X/6-311G(d) level of theory.

	Total	Turn back toward Prod-1	Turn back toward Prod-2
Turn back	69	62 (89.9%)	7 (10.1%)
Prod-1	171	62 (36.3%)	
Prod-2	35		7 (20.0%)

C. Supporting Information for Chapter 4:

NN model and validation.

We follow the general procedure developed by Lopez and coworkers in dataset generation and training the neural network. We employed a modified version of Lopez's PyRAI^{2}MD software to accommodate for MC-PDFT energy and gradient. To capture all relevant structures in the photochemistry of DBO, the initial dataset consists of Wigner sampling of the critical stationary points (minima and TSs) and the interpolation of the reaction pathway connecting them ([2+2] cycloadditions, [3,3] sigmatropic shift and C-N bond dissociations). Each reaction pathway (a, b, c, d and e shown in Fig C1) consists of 10 interpolated intermediate structures and we sample 64 points around each structure, making in total 3200 datapoints. We have put an emphasis on sampling the S_1-TS (**2**) and the four-fold crossing point (**3**) with additional 1024 datapoints and the rest of the data points are Wigner sampling of structure **1**, **4** and **5**.
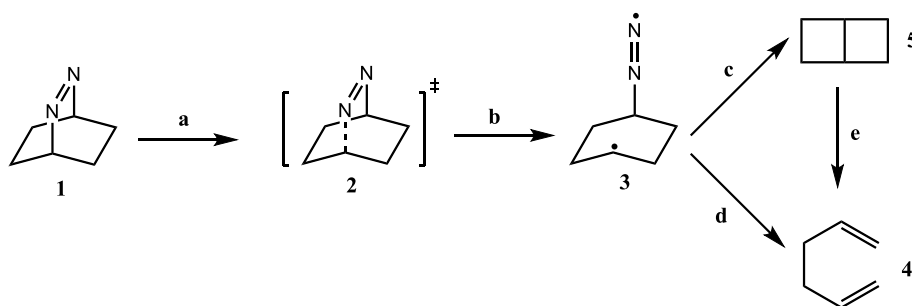
Figure C1. Pathways and structures sampled for the initial dataset.

In this way, the initial dataset is able to adequately describe the equilibrium geometries while capturing the geometrical changes during the photochemistry process.

With the initial dataset containing 5112 structures, the NN was trained with the inverse distance matrix as the input and a loss function of 1:1 combination of energies and sum of gradients. The model predicts the energies and atomic forces on S_0 and S_1 states. We used L2

174

regularization to prevent overfitting. We utilized a stepwise learning rate schedular that reduces the learning rate (1e-3, 5e-4, 1e-4, 1e-5) after (500,1200,250,50) epochs.

$$L = \frac{1}{N}\left(\hat{E} - E\right)^2 + \frac{1}{3N}\sum_{i=1}^{3N}\left(\frac{\partial\hat{E}}{\partial x_i} - \frac{\partial E}{\partial x_i}\right) + \alpha\,\|\omega\|_2^2$$

Since 2,3-diazabicyclo[2.2.2]octhept-2-ene (DBO) belongs to C_{2v} point group, we take the advantage of symmetry and expand the dataset by adding symmetrically equivalent structures. Utilizing the permutation map method embedded in PyRAI^{2}MD software, we achieved 3-fold expansion of the dataset. The permutation is illustrated in Fig C2.
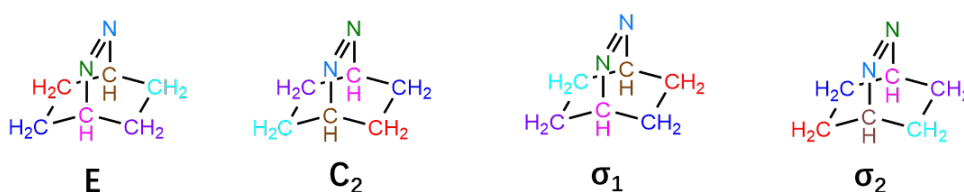
Figure C2. Demonstration of the permutation map. Atom indices are color-coded.

We conducted grid search to determine the best set of hyperparameters. The initial dataset is split to 90% of training set and 10% of test set. The results are displayed in Table C1.

Table C1. Grid search result

Layers	Nodes	L2	Energy1	Gradient1
3	200	1.00E-09	0.08974	0.25699
3	200	1.00E-08	0.093516	0.256423
3	200	1.00E-07	0.08652	0.257686
3	200	1.00E-06	0.091318	0.256556
3	200	1.00E-05	0.09363	0.255197
3	300	1.00E-09	0.069295	0.211828

3	300	1.00E-08	0.069304	0.207927
3	300	1.00E-07	0.06445	0.207525
3	300	1.00E-06	0.070471	0.211158
3	300	1.00E-05	0.078633	0.22538
3	400	1.00E-09	0.057357	0.183499
3	400	1.00E-08	0.055999	0.179409
3	400	1.00E-07	0.057685	0.186341
3	400	1.00E-06	0.054209	0.175639
3	400	1.00E-05	0.072187	0.208564
3	500	1.00E-09	0.042728	0.147208
3	500	1.00E-08	0.04622	0.16068
3	500	1.00E-07	0.048593	0.163302
3	500	1.00E-06	0.047531	0.162337
3	500	1.00E-05	0.065958	0.195072
3	600	1.00E-09	0.036269	0.134254
3	600	1.00E-08	0.041173	0.146372
3	600	1.00E-07	0.036821	0.135466
3	600	1.00E-06	0.04089	0.143994
3	600	1.00E-05	0.066458	0.193743
4	200	1.00E-09	0.085228	0.242334
4	200	1.00E-08	0.081547	0.236145
4	200	1.00E-07	0.088395	0.248581

4	200	1.00E-06	0.084593	0.240169
4	200	1.00E-05	0.095532	0.257752
4	300	1.00E-09	0.060514	0.194287
4	300	1.00E-08	0.060281	0.189274
4	300	1.00E-07	0.063824	0.199396
4	300	1.00E-06	0.063753	0.198708
4	300	1.00E-05	0.078635	0.225855
4	400	1.00E-09	0.048711	0.16704
4	400	1.00E-08	0.047235	0.162927
4	400	1.00E-07	0.049191	0.164794
4	400	1.00E-06	0.048919	0.162164
4	400	1.00E-05	0.068269	0.201342
4	500	1.00E-09	0.037381	0.134861
4	500	1.00E-08	0.037738	0.138001
4	500	1.00E-07	0.042422	0.147838
4	500	1.00E-06	0.043304	0.150101
4	500	1.00E-05	0.064382	0.195853
4	600	1.00E-09	0.028384	0.10976
4	600	1.00E-08	0.032324	0.123292
4	600	1.00E-07	0.035247	0.133335
4	600	1.00E-06	0.038218	0.137657
4	600	1.00E-05	0.063454	0.189839

5	50	1.00E-08	0.177547	0.382242
5	100	1.00E-08	0.128301	0.317618
5	200	1.00E-08	0.080571	0.232537
5	300	1.00E-08	0.0608	0.194685
5	400	1.00E-08	0.05235	0.172645
5	500	1.00E-08	0.043062	0.148911
5	600	1.00E-08	0.032441	0.12232
5	700	1.00E-08	0.034867	0.128552
5	800	1.00E-08	0.03592	0.13573
6	50	1.00E-08	0.179693	0.385689
6	100	1.00E-08	0.118656	0.300128
6	200	1.00E-08	0.082048	0.236378
6	300	1.00E-08	0.067041	0.203851
6	400	1.00E-08	0.054893	0.178979
6	500	1.00E-08	0.047234	0.161144
6	600	1.00E-08	0.049803	0.163498
6	700	1.00E-08	0.078247	0.226325
6	800	1.00E-08	0.056601	0.184724
7	50	1.00E-08	0.176471	0.381513
7	100	1.00E-08	0.125942	0.314563
7	200	1.00E-08	0.08536	0.240333
7	300	1.00E-08	0.081234	0.233579

7	400	1.00E-08	0.054779	0.17502
7	500	1.00E-08	0.065086	0.202756
7	600	1.00E-08	0.090721	0.261724
7	700	1.00E-08	0.09244	0.269161
7	800	1.00E-08	0.142156	0.348524
8	50	1.00E-08	0.17323	0.386057
8	100	1.00E-08	0.122711	0.305378
8	200	1.00E-08	0.095223	0.262289
8	300	1.00E-08	0.06873	0.208273
8	400	1.00E-08	0.091078	0.256849
8	500	1.00E-08	0.081801	0.240414
8	600	1.00E-08	0.093068	0.265962
8	700	1.00E-08	0.110138	0.287005
8	800	1.00E-08	0.1976	0.460217
9	50	1.00E-08	0.175029	0.386673
9	100	1.00E-08	0.132743	0.323126
9	200	1.00E-08	0.098073	0.267416
9	300	1.00E-08	0.101187	0.277633
9	400	1.00E-08	0.080541	0.244347
9	500	1.00E-08	0.091577	0.257502
9	600	1.00E-08	0.141488	0.349767
9	700	1.00E-08	0.106032	0.279388

Table C2. Search of batch size

Layers	Nodes	Batch Size	Energy1	Gradient1
4	400	32	0.075246	0.224985
4	400	64	0.061949	0.19048
4	400	128	0.049104	0.161013
4	400	512	0.046369	0.161518
4	400	1024	0.072855	0.222251

Table C3. Hyperparameters chosen for the two working models.

Model	Layer	Neurons/Layer	L2 alpha	Batch size	Activation alpha
1	4	400	1e-8	512	0.03
2	5	300	1e-8	512	0.03

We eventually picked the following hyperparameters shown in Table C3, as they achieve a reasonable balance between accuracy and model complexity.

MD simulations often explore much broader conformational space than the interpolated initial dataset. An adaptive sampling procedure is followed after obtaining the initial NN model to include the underrepresented structures. To evaluate the uncertainty of the predictions, two distinct NNs were trained with different hyperparameters (Table C3). Trajectories are propagated until the prediction uncertainty exceeds the threshold (MAE (energy) = 0.03 eV, MAE (gradient) = 0.15 eV/Å). The last uncertain structure is added to the dataset, and a new model is trained based on the new dataset. This iterative procedure is carried out until the model converges, i.e., all the trajectories either reach the 4000fs time limit, or form one of the three products.

The adaptive sampling takes 20 iterations and 240 trajectories were run for each iteration. In total 4800 trajectories were sampled and 4558 points were added into the dataset. The final dataset contains 9670 (38680 including symmetry) structures.

Fig C3 displays the geometrical distribution of the initial dataset and the following adaptive sampling. The adaptive sampling procedure clearly explores a broader range of the geometry space.

Fig C4 displays how the model is adapted and refined by evaluating the average trajectory length. Longer trajectory length indicates higher confidence of on the energy and gradient prediction.
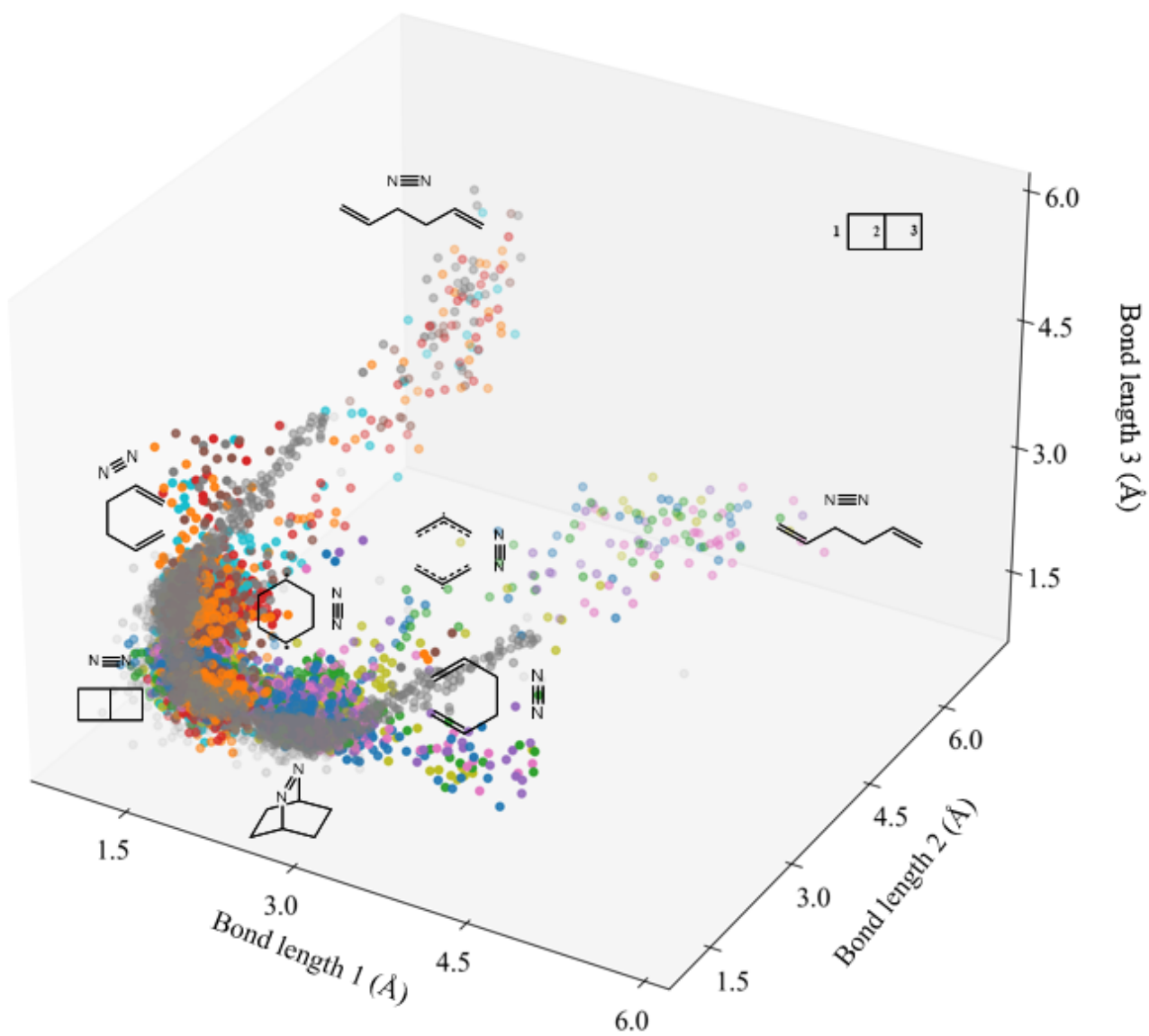
Figure C3. Distribution of the initial dataset (grey) and the data points included from adaptive sampling (colored)
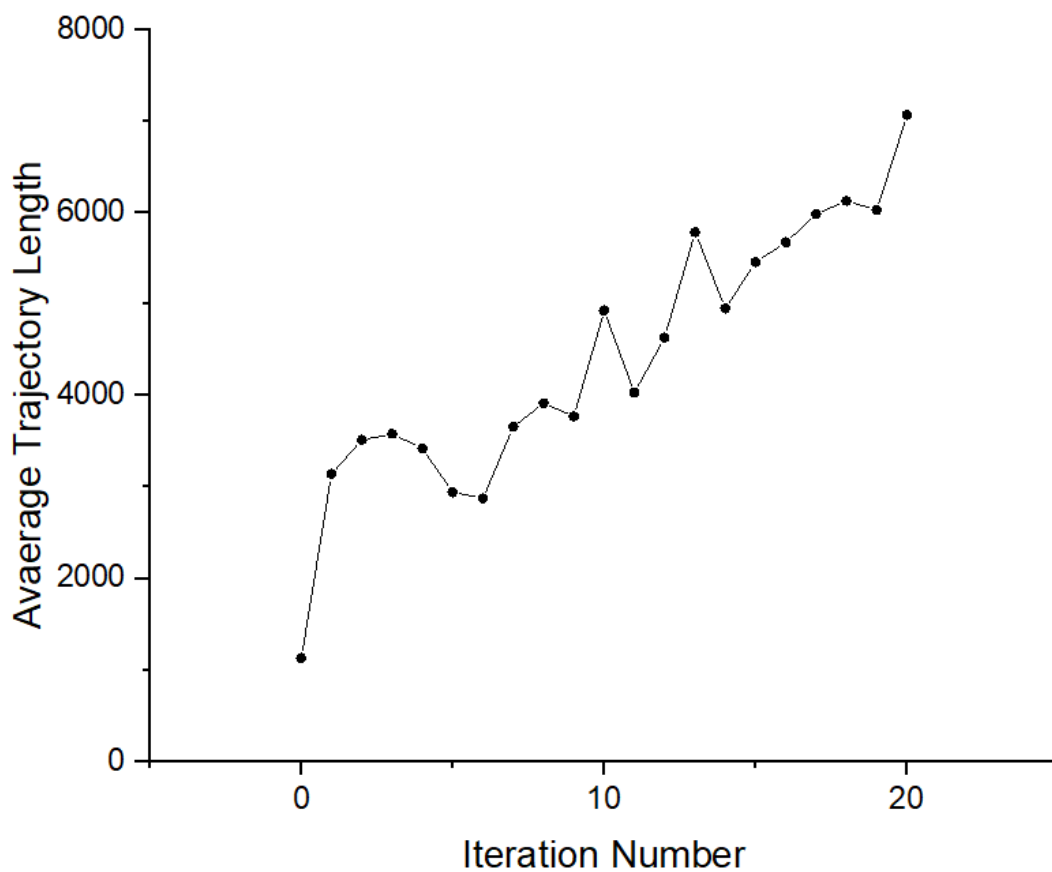
Figure C4. Average trajectory length of each iteration in adaptive sampling process.

MCPDFT-NAMD using SHARC
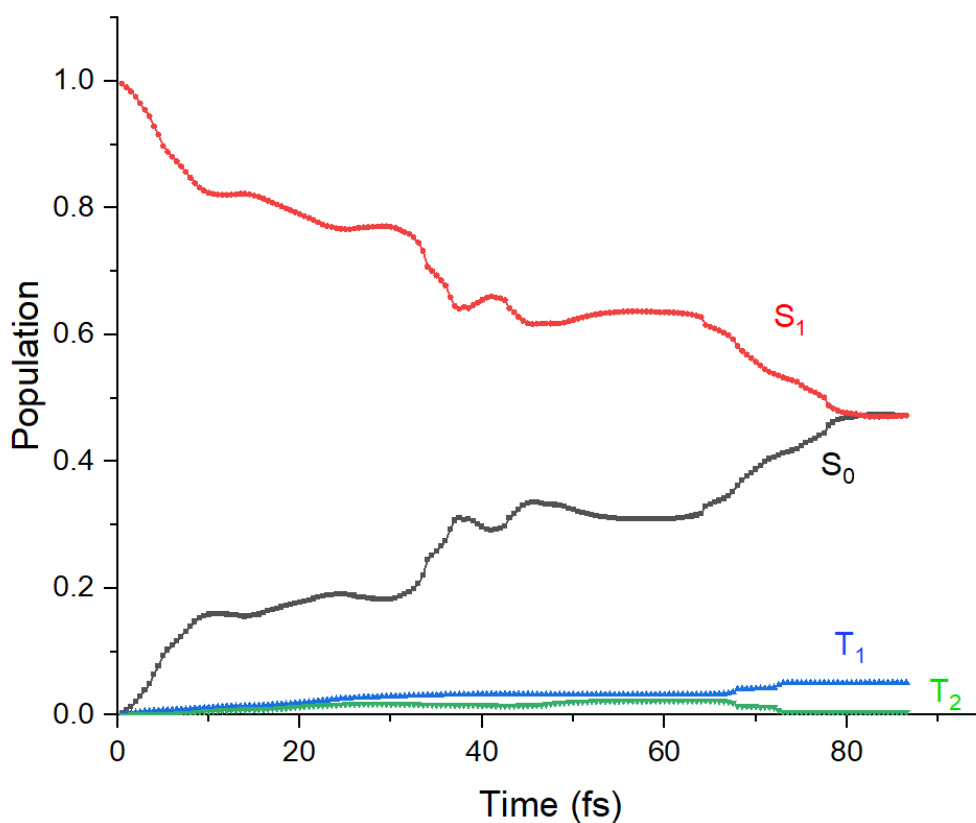
Figure C5. Population evolution of electronic states in SHARC-NAMD trajectories.

We performed full-QM based NAMD with MC-PDFT method with NACs and SOCs calculated on-the-fly during each step utilizing SHARC software. Thus singlet-singlet internal conversion and singlet-triplet intersystem crossings are both considered in order to compare their rate. Two singlets and three triplet states were included in state-average calculation and S_0, S_1, T_1, T_2 states are set active in NAMD simulation. The timestep of the simulation is 0.5fs. We have employed an EDC decoherence scheme with a factor of 0.1 and kinetic correction to maintain energy conservation. The trajectories were initiated from TSS optimized under trevPBE(8,8)/ANO-S-VDZP level of theory.

From the full-QM NAMD simulation, we can see a fast S_1-S_0 IC within the 100fs range. This is due to the reduction of S_1-S_0 energy gap along the first C-N bond dissociation, in consistent with MEP calculation. When either of the product is formed, it will stay on the

ground-state surface. At the same time, the population of the two triplets slightly increases, implying an increase in the SOC along the pathway. However, the triplet population is less than 5% and is negligible comparing with S_1-S_0 transition. One possible source of error is that the SOC calculated by MC-PDFT level utilizes CASSCF wavefunction. Nonetheless we believe that IC is much faster than ISC and it is reasonable to neglect ISC in modelling the direct photochemistry of DBO.

We have also calculated the magnitude of SOC between S_1 and T_2 state, displayed in Fig C6.
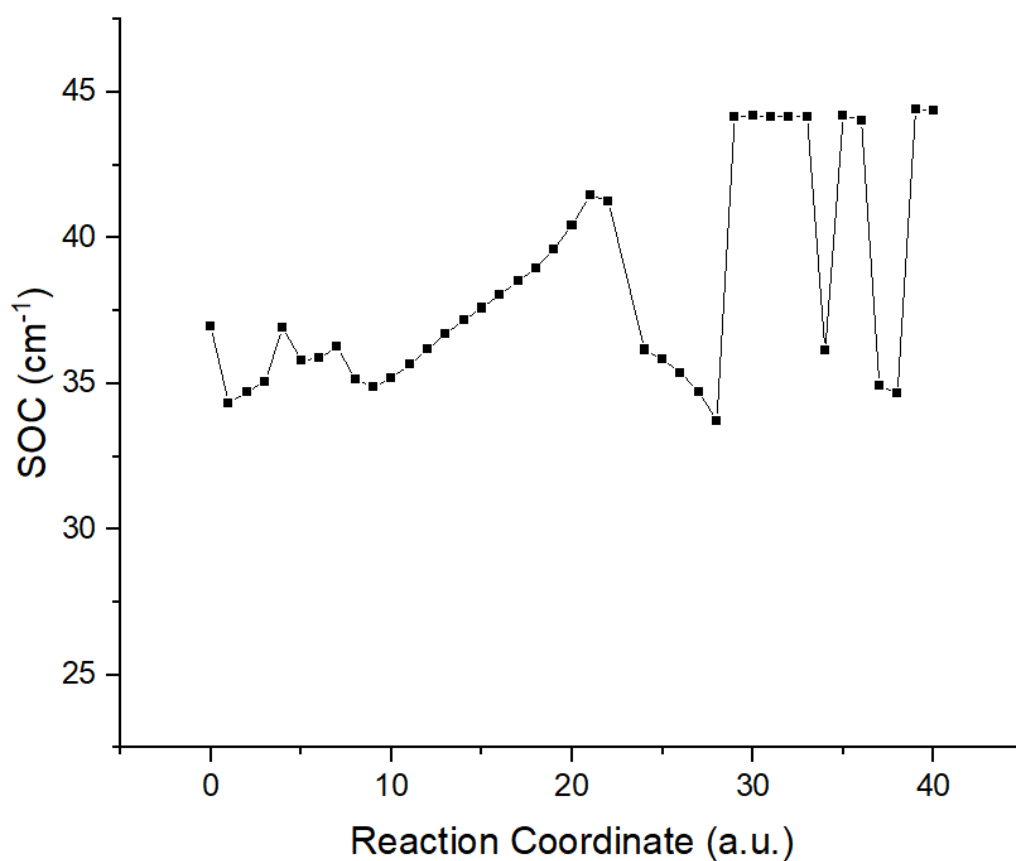
Figure C6. SOC magnitude between S_1 and T_2 state
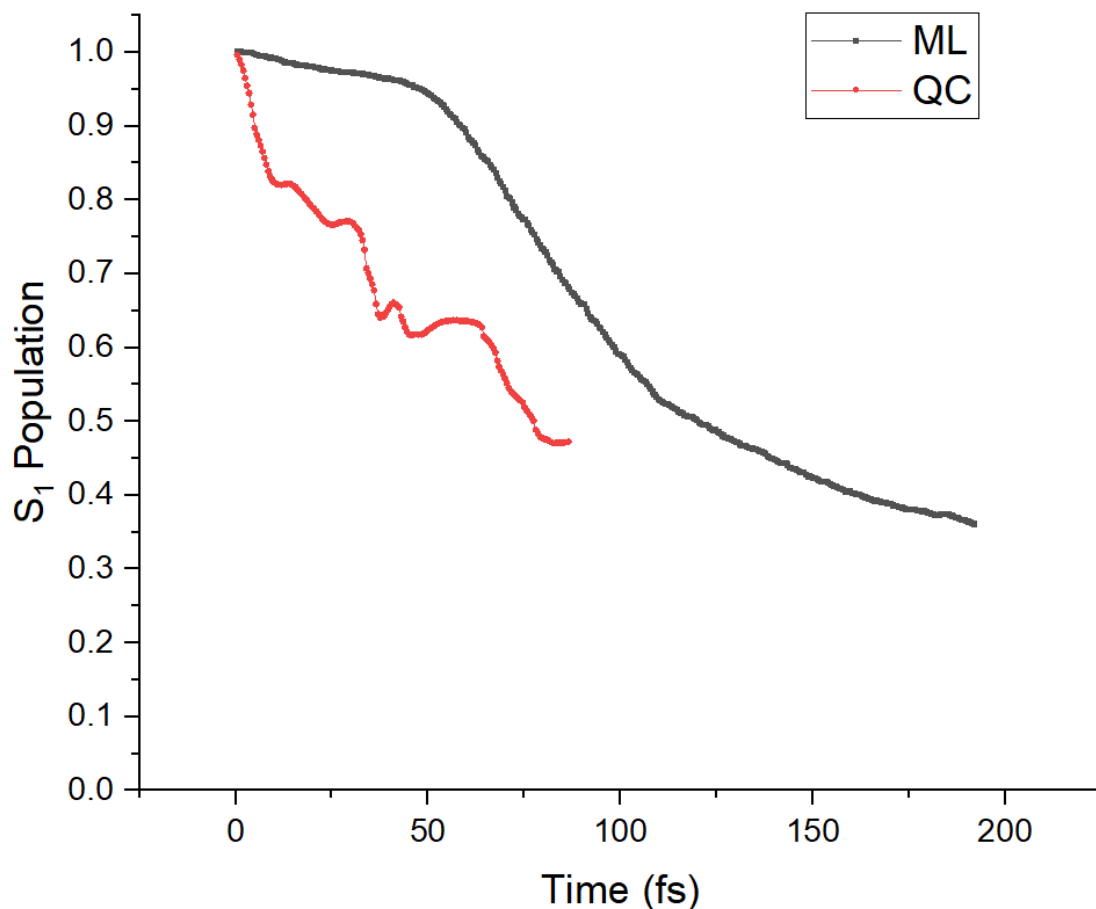
Figure C7. Comparison of ML-NAMD and SHARC QC-NAMD of S_1 population evolution.

We have compared the S_1 population derived from ML-NAMD and QC-NAMD trajectories (Fig C7). Both of them indicate short-live S_1 state and a fast decay to ground state within 200fs.

CASSCF Potential Energy Surfaces

Potential energy surfaces adjacent to Cope rearrangement and [2+2] cycloaddition TSS under SS-CASSCF(6,6)/def2-SVP level of theory are shown in Fig S3 and S4. The most notable difference in comparison with NEVPT2 PES is the presence of a 1,4-cyclohexyl diradical minima. Higher-precision CASPT2 calculations have eventually proved that such minimum is a flaw of CASSCF level.

186
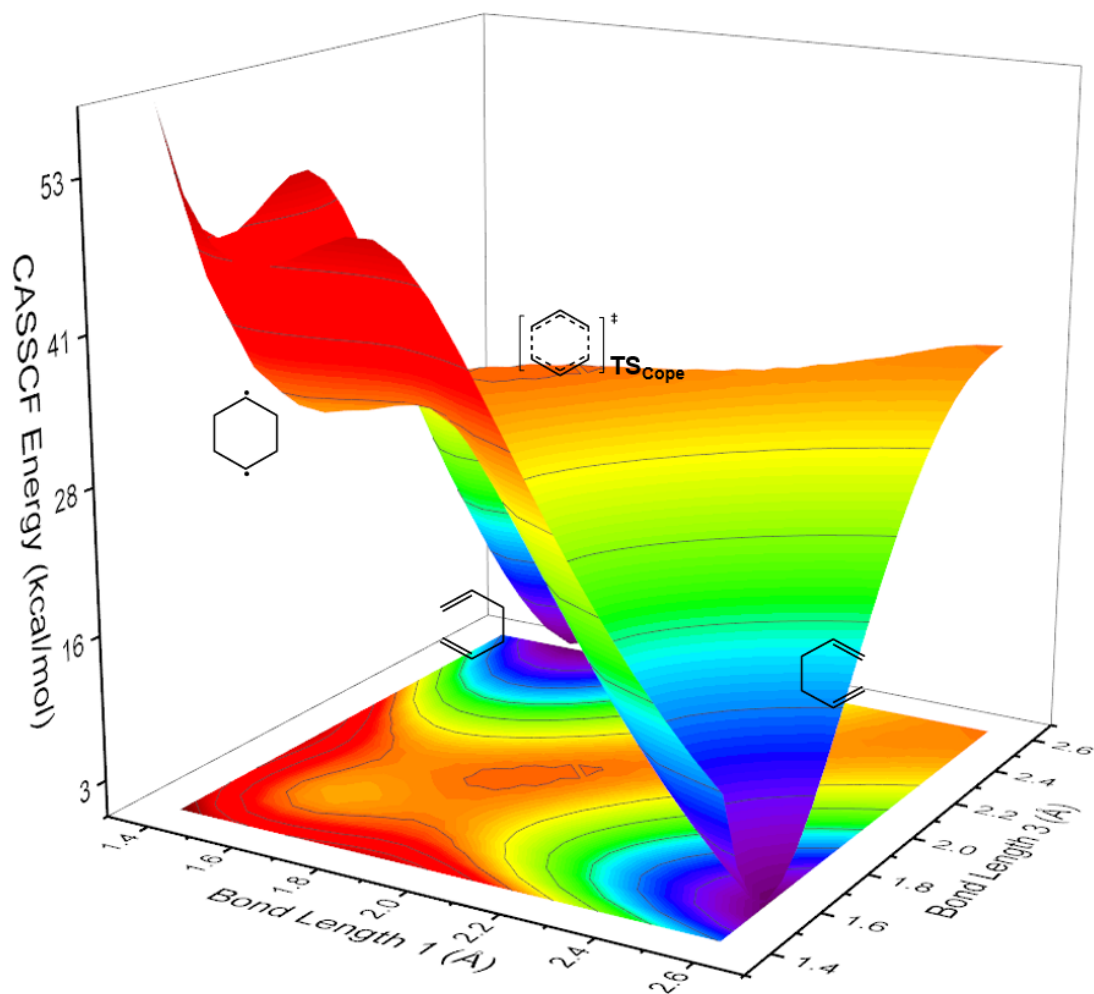
Figure C8. S_0-PES adjacent to the 1,4-cyclohexyl diradical and Cope rearrangement TS under SS-CASSCF(6,6)/def2-SVP level. ((b) 1,4-cyclohexyl diradical and [2+2] cycloaddition TS.
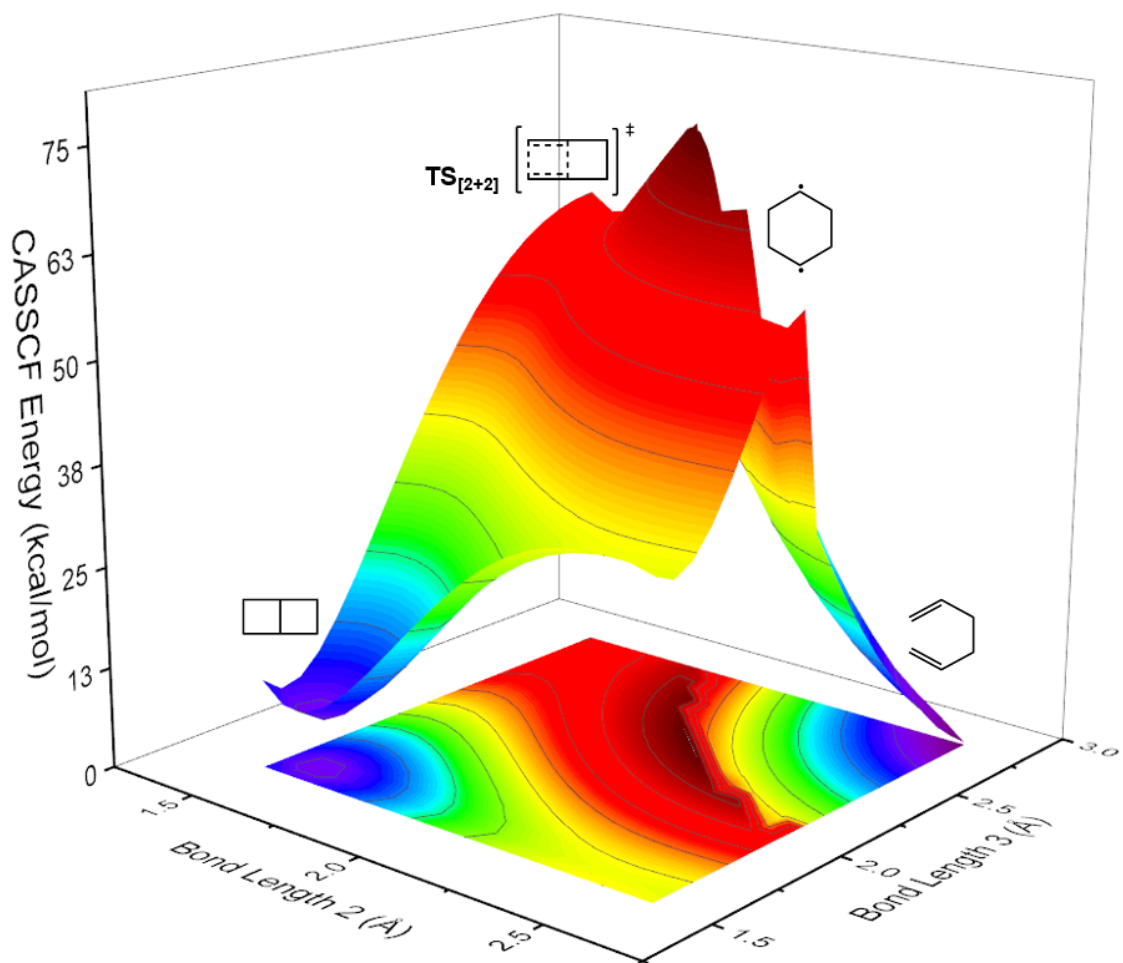
Figure C9. S_0-PES adjacent to the 1,4-cyclohexyl diradical and [2+2] cycloaddition TS under SS-CASSCF(6,6)/def2-SVP level.

MEP under MC-PDFT and CASSCF Level of Theory

We have calculated single-point energies of each point on the MC-PDFT based MEP. MC-PDFT shows remarkable agreement with CASPT2 level of theory in terms of shape and the 4-fold crossing region. CASSCF, on the contrary, significantly overestimates the S_1-S_0 gap around the diazinyl region, implying the importance of incorporating dynamic correlation in our simulations.
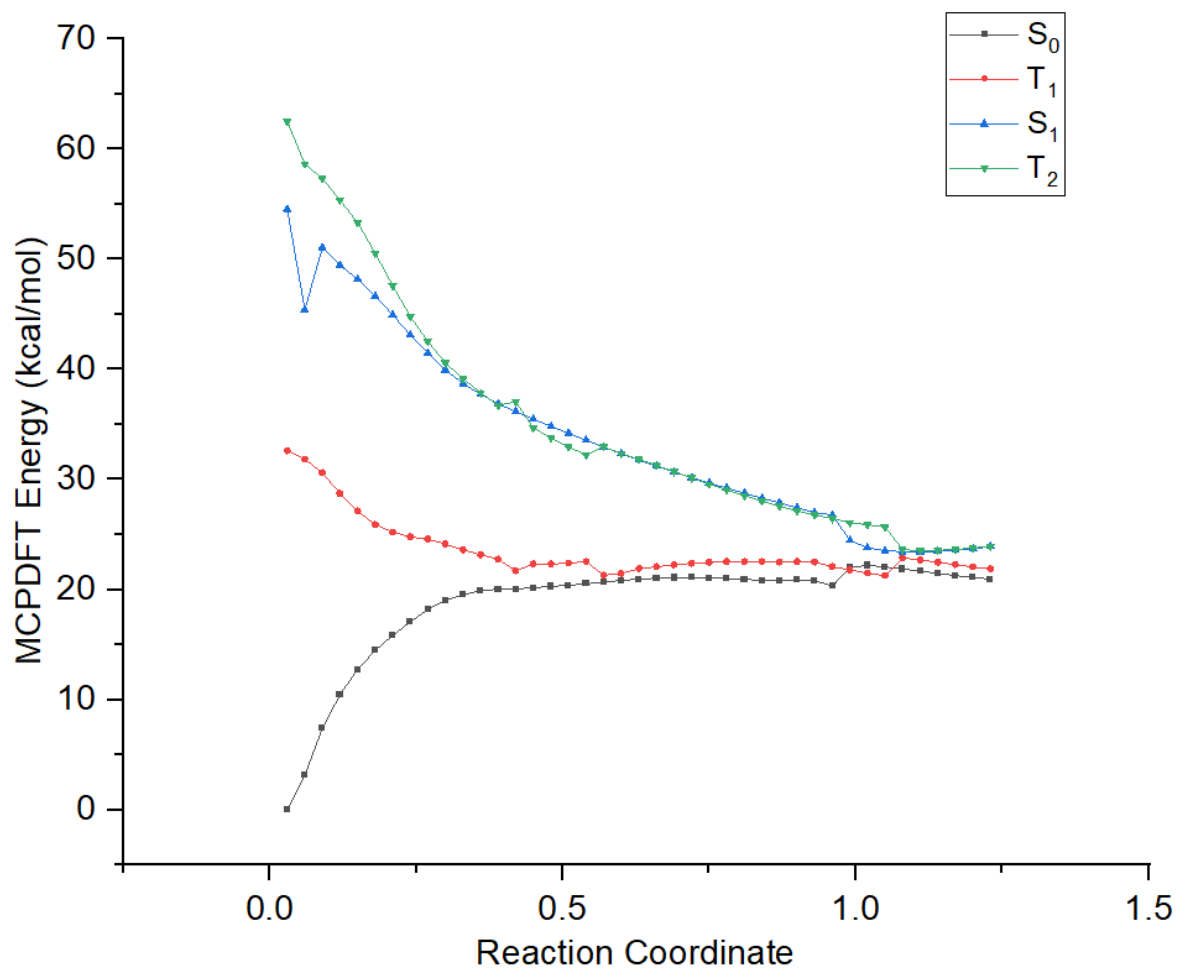
Fig C10. Minimal energy pathway (MEP) starting from the S_1-TS of first C-N bond dissociation on computed with revPBE(8,8)/ANO-S-VDZP level of theory. '0' point of the MEP is defined as the S_0 energy of the corresponding TS. Relative energies are in kcal/mol.
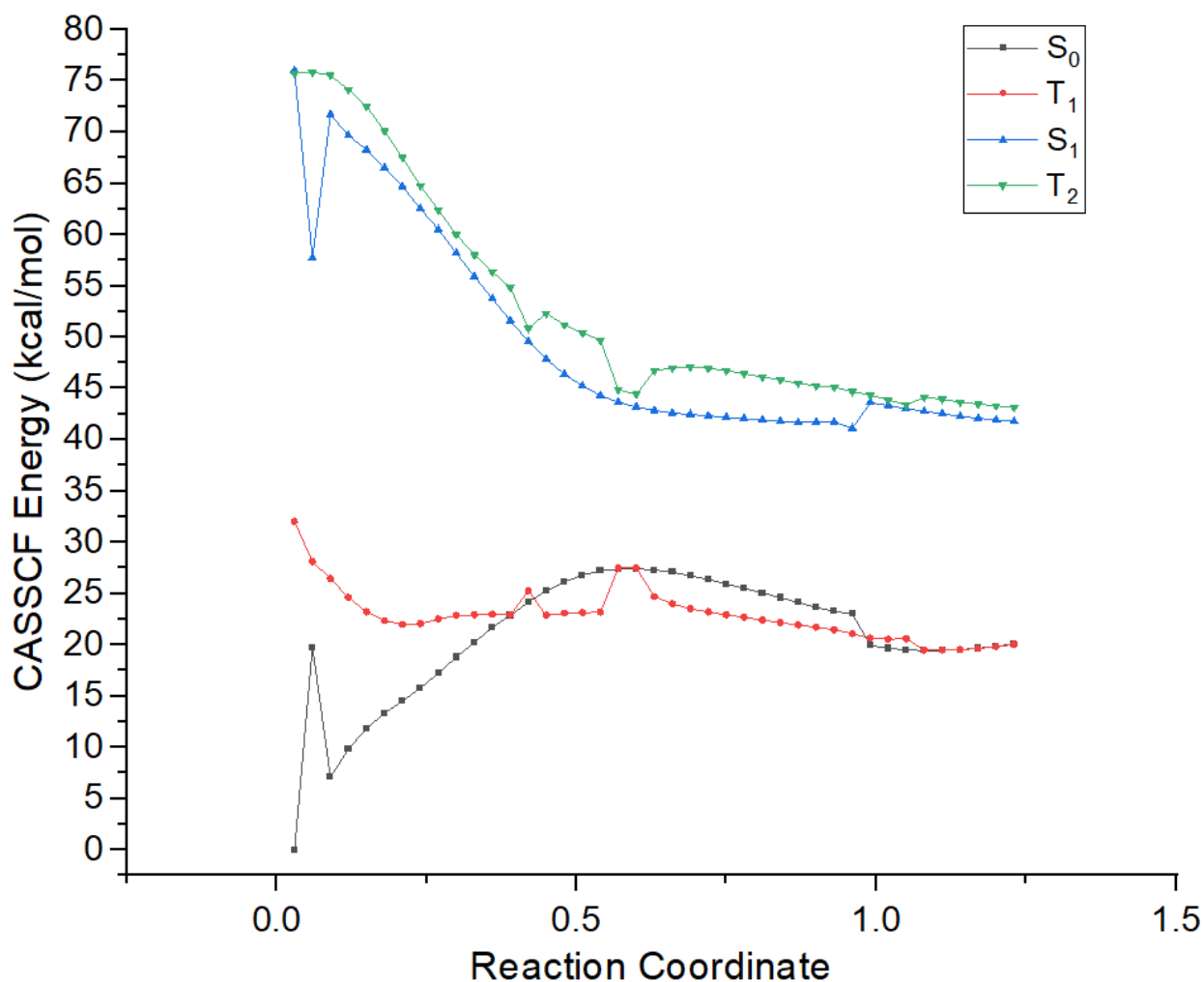
Figure C11. Minimal energy pathway (MEP) starting from the S_1-TS of first C-N bond dissociation on computed with SA(2)-CASSCF(8,8)/ANO-S-VDZP//revPBE(8,8)/ANO-S-VDZP level of theory. '0' point of the MEP is defined as the S_0 energy of the corresponding TS. Relative energies are in kcal/mol.

Second C-N Bond Dissociation

To prove that the second bond dissociation is a barrierless transition, we scanned the C-N bond starting from the diazinyl structure under SS-CASSCF(8,8)/def2-SVP level of theory and calculated single point energies under NEVPT2/def2-TZVPP level using ORCA 5.0.3 (Fig C12). Both levels of theory display a flat surface around the diazinyl diradical. Although there

190

is a small barrier of 1.8 kcal/mol under level of theory, we believe that it is trivial comparing

with the massive energy released from the first C-N bond dissociation.
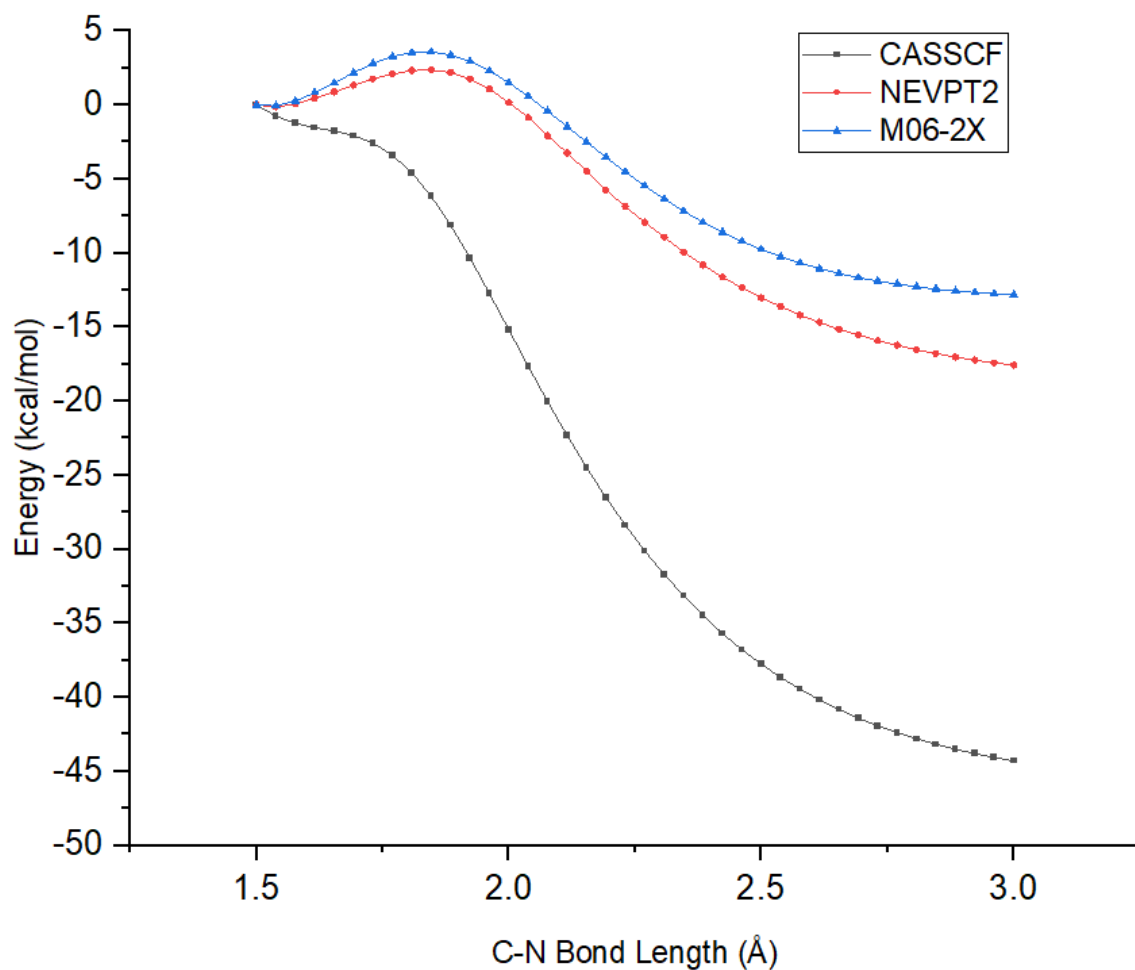
Figure C12. Energy profile of the C-N bond scan. (1). SS-CASSCF(6,6)/def2-SVP (2).

NEVPT2(6,6)// SS-CASSCF(6,6)/def2-SVP (3). M06-2X/def2-SVP. Energies are in kcal/mol.

Active orbitals
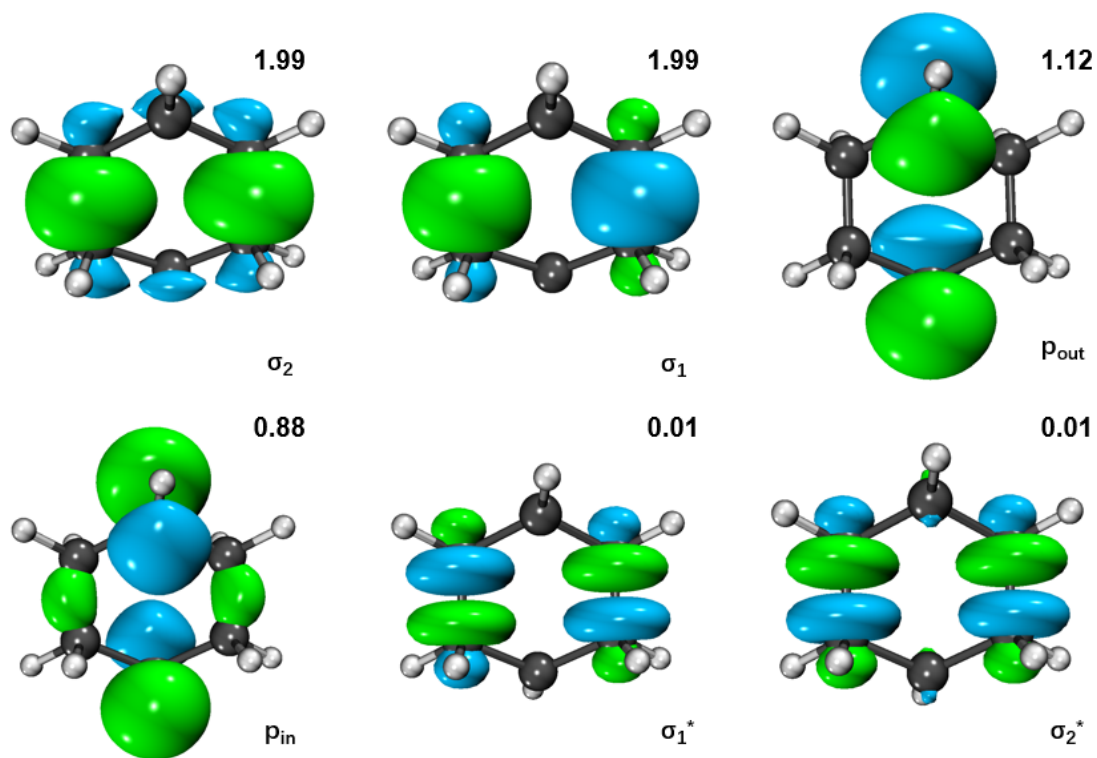
Fig C13. Active orbitals for 1,4-cyclohexyl diradical
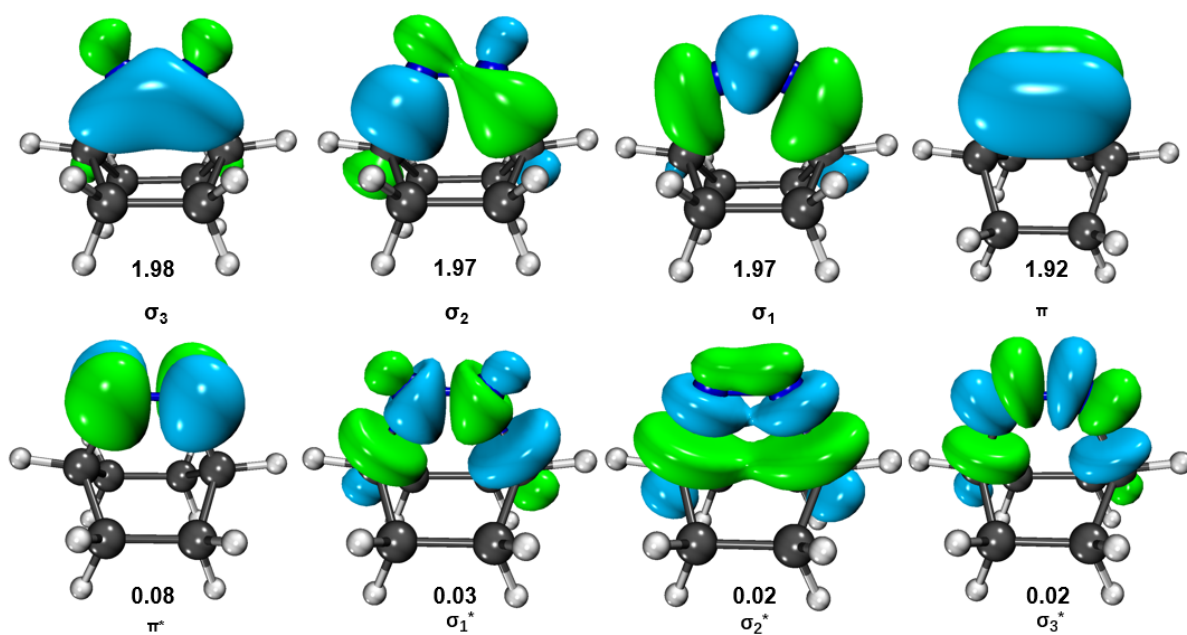
Fig C14. Active orbitals for $C_6H_{10}N_2$ species

Additional information of 1,4-cyclohexyl diradical.

1,4-cyclohexyl diradical structure displays unique sensitivity on computational level. To prove this, we employed CASPT2 level for the optimization of the diradical (Fig C15). When started from chair-conformation, the optimization eventually lands on the Cope rearrangement TSS, while the boat conformation lands on a fairly twisted structure that is 7.8 kcal/mol in energy than the Cope rearrangement TSS.

Chair   **Boat**
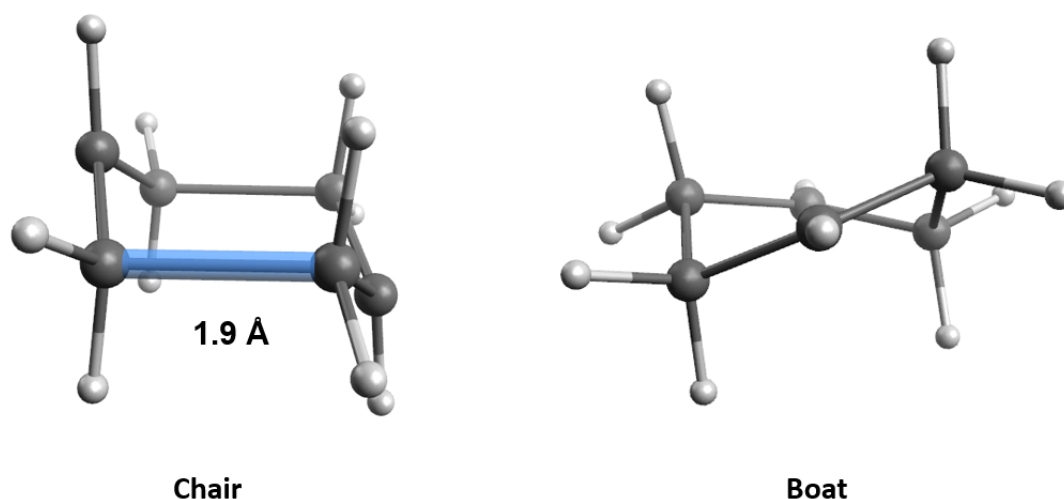
Fig C15. CASPT2 optimized S_0 structures starting from chair and boat conformations.

revPBE optimization starting from S_0-chair diradical eventually converge to the HD product. If we start from S0-boat conformation, however, it will end up at BCH product. All SA-CASSCF optimizations converge to BCH product. SS-CASSCF and DFT (M06-2X and B3LYP) computations with ORCA all located a diradical structure.

Surface Hopping Structures

As is shown in Fig C16, C17 and C18, surface hopping structures mainly populate around the diazinyl radical structure, which is at the 4-fold crossing region.
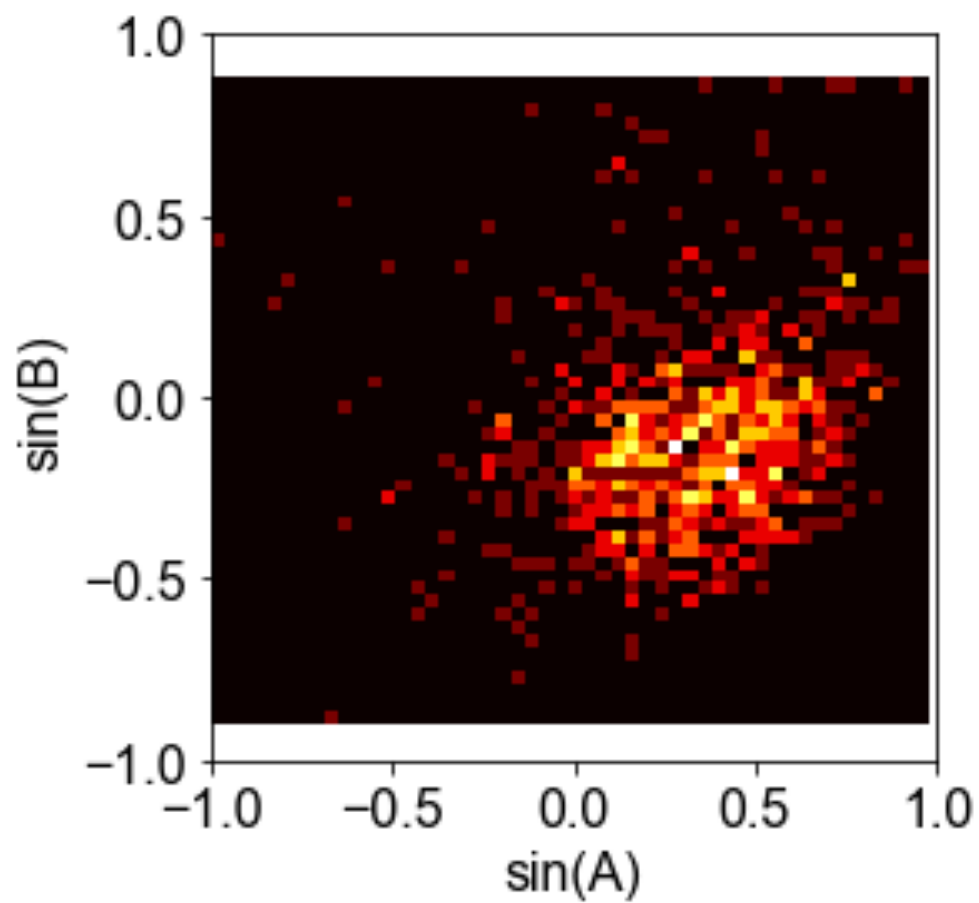
Fig C16. Population density of dihedral angle A and B of hopping point for *(Z,E)*-HD forming

NAMD trajectories.
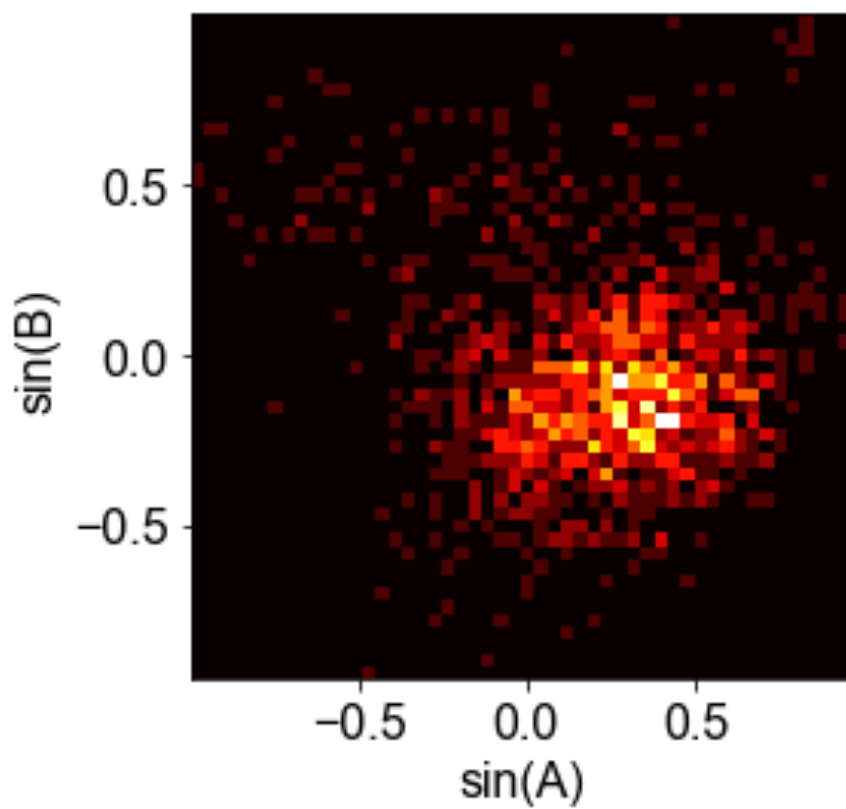
Fig C17. Population density of dihedral angle A and B of hopping point for *exo*-BCH forming

NAMD trajectories.
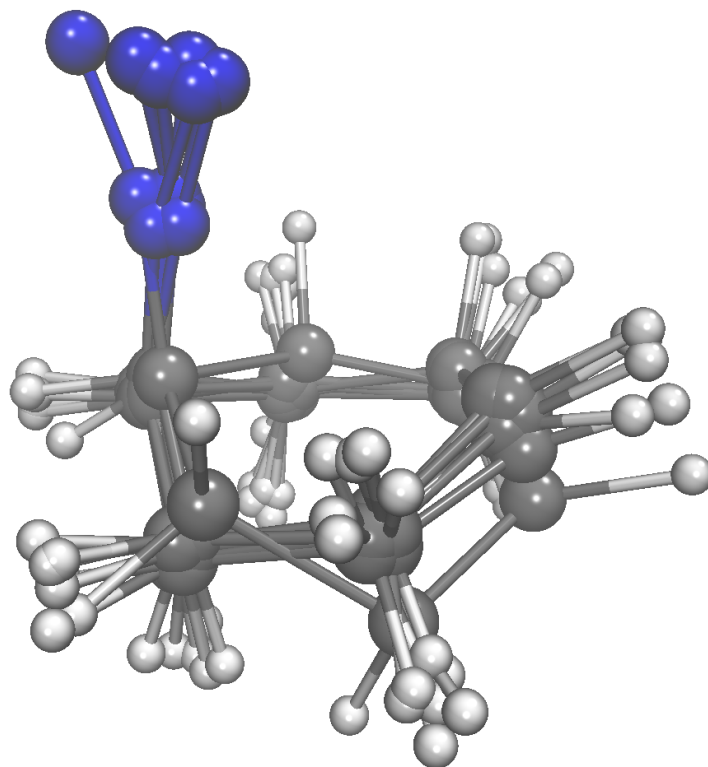
Fig C18. Overlay of surface hopping structures.

Additional Population Difference Plot
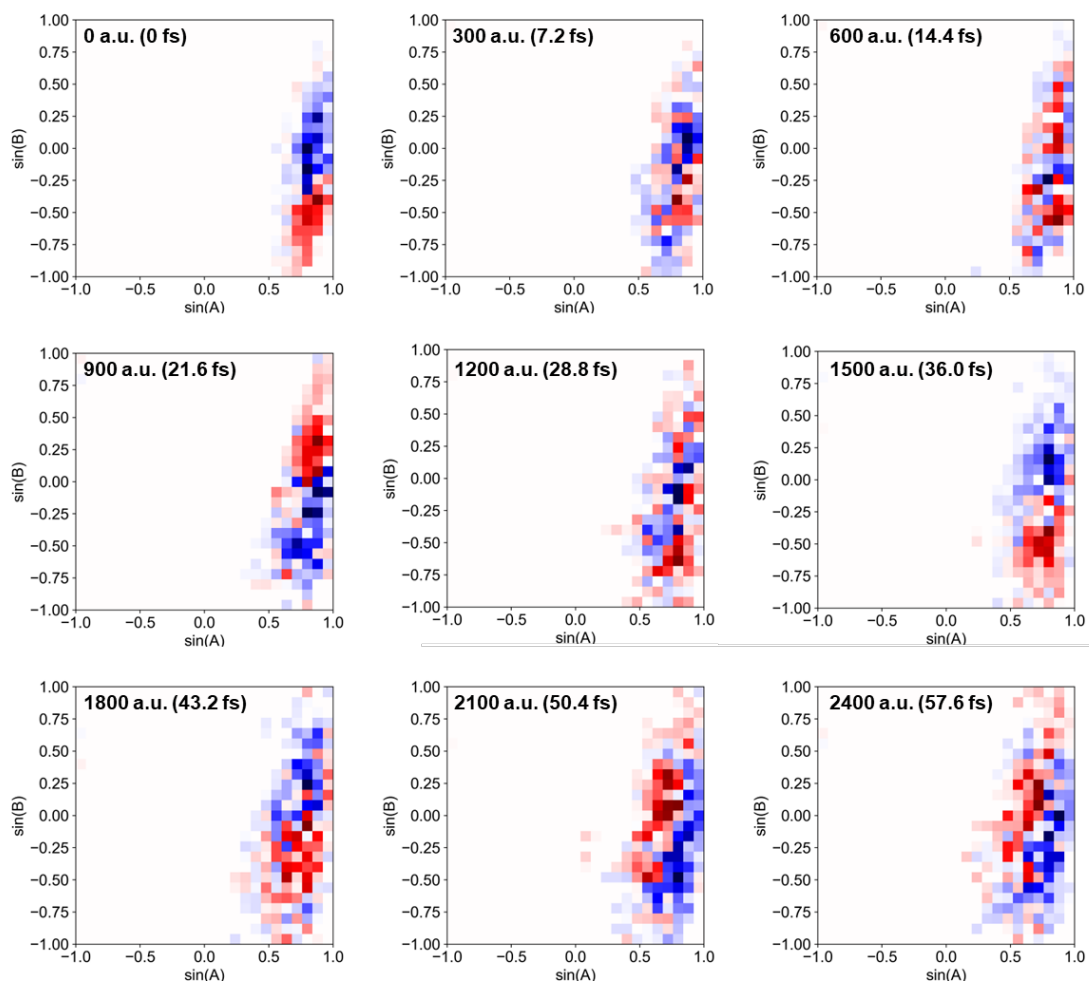
Fig C19. Population difference of the dihedral angle for the NAMD trajectories within 57.6 fs.

Verification of Functional, Basis Set and Active Space

Table C4. Benchmark of functional and basis set, an active space of (8e,8o) is used for all calculations.

level of theory	Vertical Excitation Energy
ftrevPBE/ANO-S-VDZP	87.3
trevPBE/ANO-S-VDZP	87.2
ftpbe/ANO-S-VDZP	87.5
tpbe/ANO-S-VDZP	87.5

ftblyp/ANO-S-VDZP	87
tblyp/ANO-S-VDZP	86.9
trevPBE/cc-pVDZ	96.7
trevPBE/cc-pVTZ	97.1
trevPBE/ANO-L-VTZP	86.2
trevPBE/ANO-L-VQZP	86.2
trevPBE/6-31G(d)	97.3
trevPBE/6-31G(d,p)	97.3
Experiment	76.4
CASPT2(8,8)/ANO-S-VDZP//revPBE(8,8)/ANO-S-VDZP	80.5

Table C5. Benchmark of active space. revPBE functional of MCP-DFT is used for all calculations.

	Vertical Excitation Energy	S_1-TS barrier
(2e,2o)	85.3	
(4e,4o)	86.7	
(6e,6o)	85.1	13.0
(8e,8o)	87.2	9.4

Product Distribution of ML-NAMD with Different Thermostat
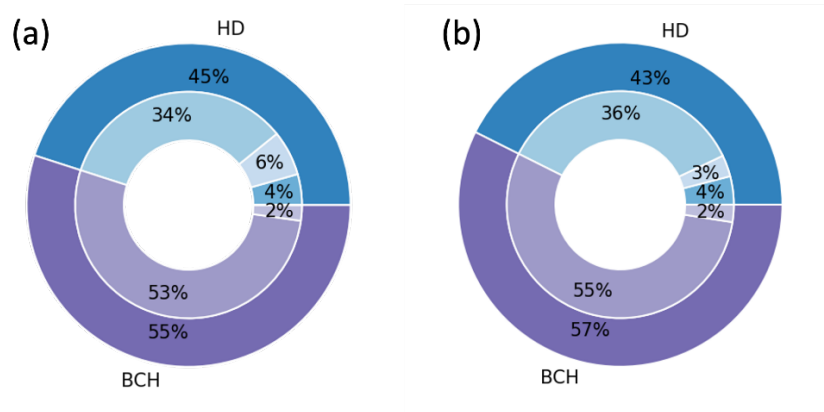
Figure C20. (a) NVE ensemble in S_1 state and NVT in S_0 state. (b) NVE ensemble for both states.

Training procedure
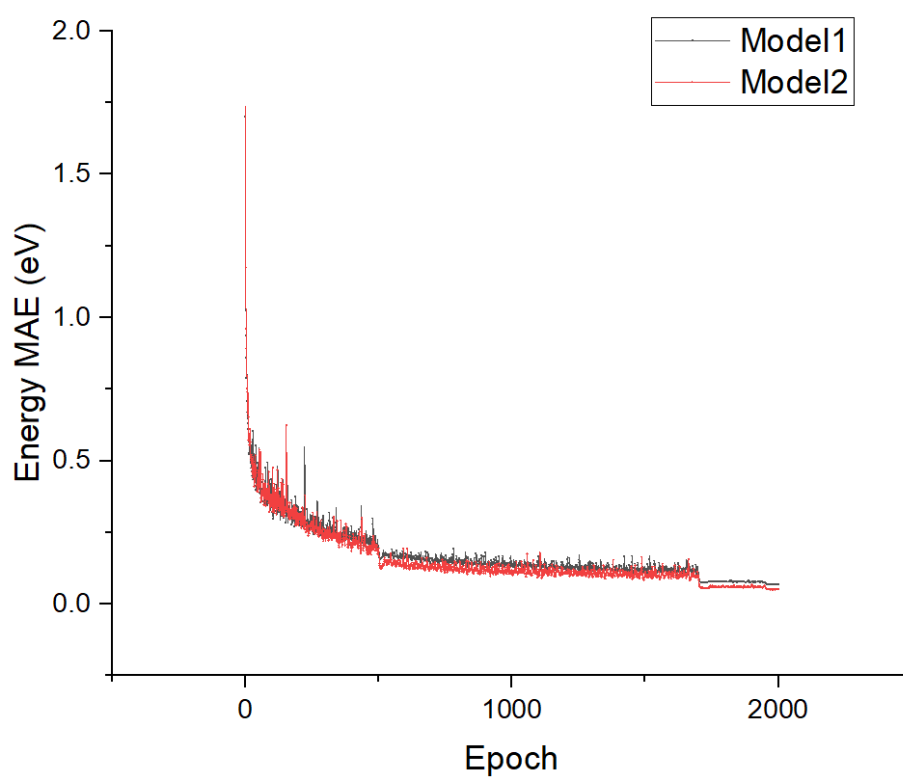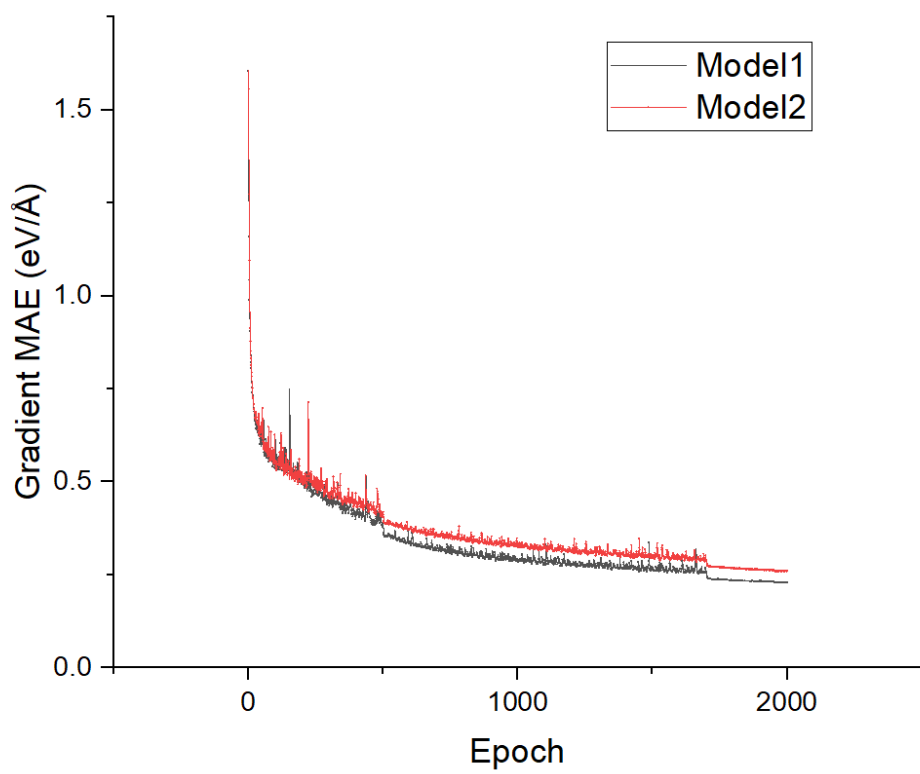
Figure S22. Training procedure of the ML potential.

D. Supporting Information for chapter 5

Approach of HSO₃⁺ from the concave face of allopupukeanane

When considering the approach of HSO_3^+ from the bottom (concave) face of allopupukeanane **7**, the 3D structure of **7** indicates that this is energetically unfavorable due to hindrance by the adjacent isopropyl group (Fig D1).
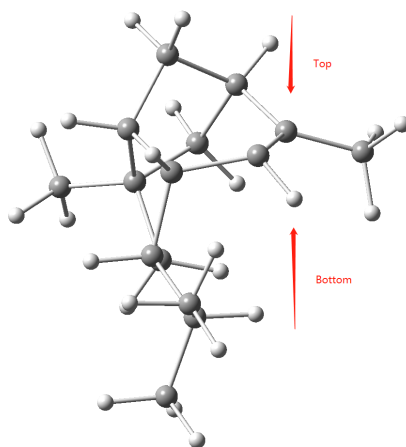
Figure D1. Two face attack of the allopupukeanane **7.**

This is indirectly supported by the relative energies of the two isomers (Scheme D1).

2A	***epi*-2A**	**2A**	***epi*-2A**
0.0 kcal/mol	+3.9 kcal/mol	0.0 kcal/mol	+3.9 kcal/mol
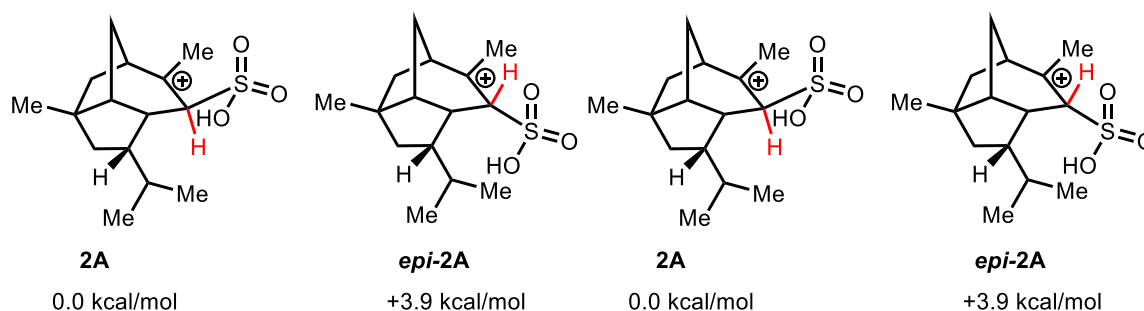
Scheme D1. Two isomers of the allopupukeanane **7.**

During our computational analysis of the approach of HSO_3^+ to **7**, an alternative pathway emerged (Scheme D2).
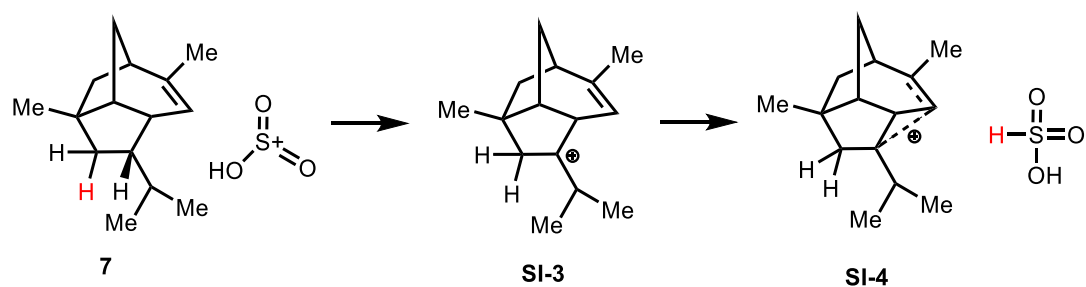
Scheme D2. An alternative H-shift pathway.

From an optimized trajectory where HSO_3^+ resides on the concave face, HSO_3^+ would first abstract a hydride (in red) from **7**, leading to a 1,2-H shift to give tertiary carbocation **SI-3**. Interaction of the carbocation with the proximal C=C bond leads to non-classical carbocation **SI-4**. These findings imply that an entirely different reaction could occur when HSO_3^+ approaches the bottom face. This reaction is very exothermic, and a very early transition state is expected, thus it is very difficult to locate the transition state using standard techniques. We employed the nudged elastic band (NEB) algorithm to depict the minimal-energy path (MEP) and a TSS with an electronic energy barrier of 0.3 kcal/mol was located (ωB97X-D3/def-SV(P),CPCM(water)), shown in Fig D2.
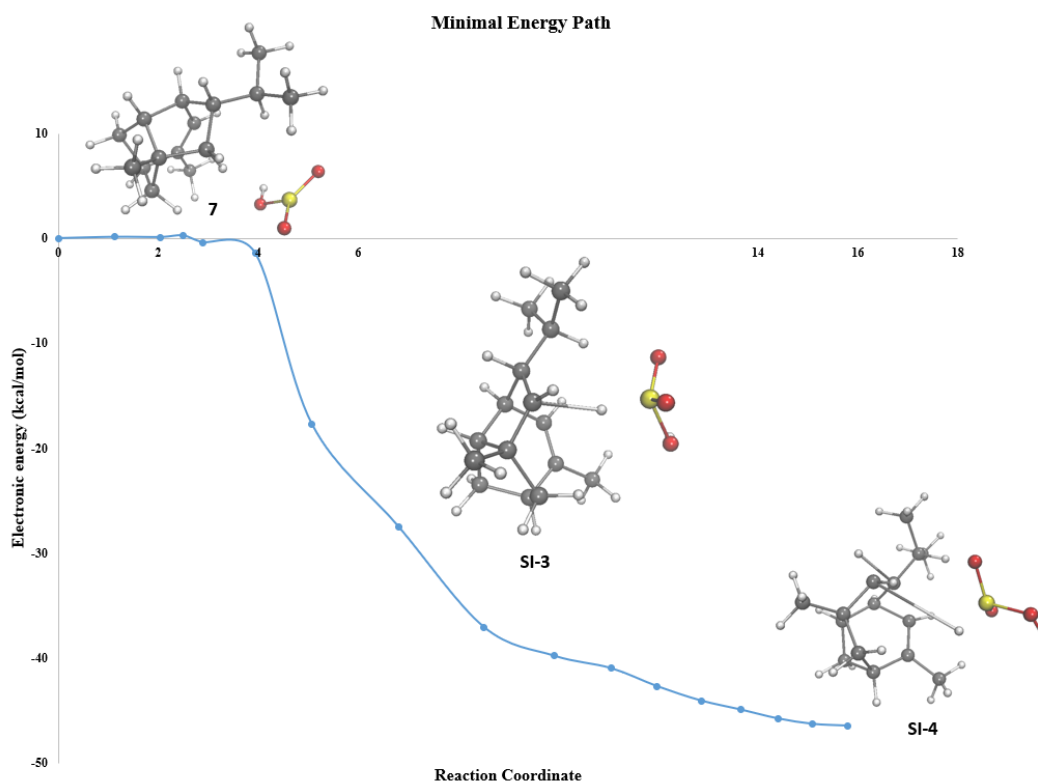
Figure D2. MEP connecting allopupukeanane **7 with SI-4** derived from NEB calculation, energies are in kcal/mol.

Both this pathway to non-classical carbocation **SI-4** and the pathway leading to pupukeanyl sultone **15** are very exergonic and barrierless. While the pathway to **15** is more exergonic, both pathways could potentially occur. It is almost impossible to enumerate all pathways, and, in order to predict the predominant pathway, one would need to carry out ab initio molecular dynamics simulations. Our results demonstrate the complexity of the reaction (which is consistent with the 38% yield of **15**) and a possible pathway that leads to the observed product.

Direct capture of the tertiary carbocation to form the C–O bond

We also considered the possibility of the pathway beginning with the direct formation of the C–O bond. However, this is expected to be reversible under the reaction conditions. The

energetics of the following reaction indirectly support the reversibility of this transformation, depicted in Scheme D3.
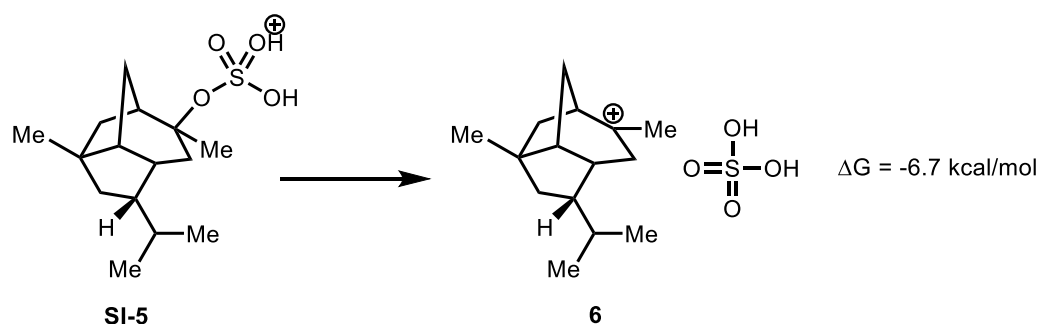
Scheme D3. Direct formation of C-O bond.

Once the corresponding sulfonation product (**SI-5**) is protonated, the dissociation is exothermic and carbocation **6** is calculated to form easily.
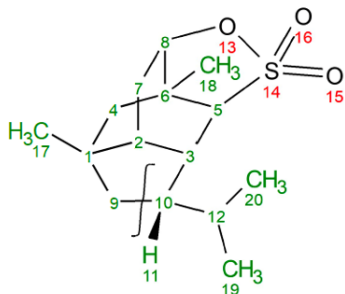
Capture of the sulfonated hydrocarbon as a β-sultone
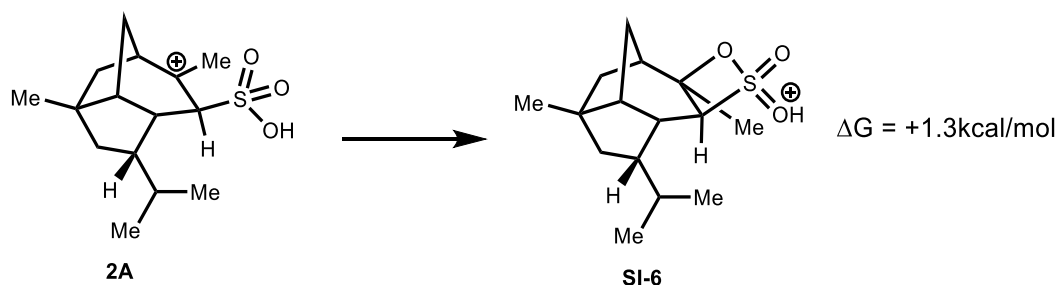
Scheme D4. Intramolecular direct capture of the tertiary carbocation **2A**.

One could also envision a direct capture of the tertiary carbocation **2A** to form *β*-sultone **SI-6**. We calculated that this possibility was energetically disfavored and would be easily reversible were it to occur.

NMR calculations

#	sigma	shift computed	shift experimental	Deviation
1	155.0208	41.069	38.62	-2.45
2	152.5574	43.602	41.16	-2.44
3	152.2353	43.933	41.62	-2.31
4	148.6375	47.632	47.86	0.23
5	130.7949	65.977	58.91	-7.07
6	154.1009	42.015	39.18	-2.83
7	166.4933	29.273	27.32	-1.95
8	112.1147	85.184	88.41	3.23
9	147.4829	48.819	47.11	-1.71
10	148.2337	48.047	47.76	-0.29
12	163.0614	32.802	30.39	-2.41
17	170.4153	25.241	25.89	0.65
18	172.7064	22.885	23.93	1.05
19	174.772	20.761	21.23	0.47
20	174.8424	20.689	21.58	0.89
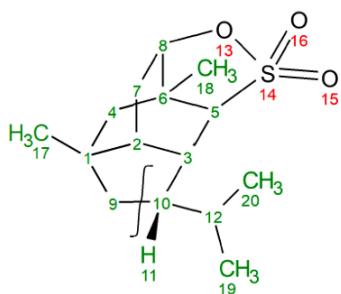
| #	sigma	shift computed	shift experimental	Deviation
2	29.975	1.600	1.59	-0.01
3	28.6311	2.844	3.07	0.23
4	30.1454	1.442	1.51	0.07
4"	30.1775	1.412	1.42	0.01
5	28.4856	2.978	3.01	0.03
7'	29.5231	2.018	2.04	0.02
7"	29.7381	1.819	1.96	0.14
8	27.3561	4.024	4.43	0.41
9'	29.6753	1.877	1.77	-0.11
9"	30.445	1.165	1.16	0.00
11	29.872	1.695	1.77	0.08
12	30.201	1.390	1.42	0.03
17	30.6693	0.957	1.04	0.08
18	28.7745	2.711	1.42	-1.29
19	30.7763	0.858	0.94	0.08
20	30.7089	0.920	0.97	0.05

Figure D3. NMR confirmation of the product
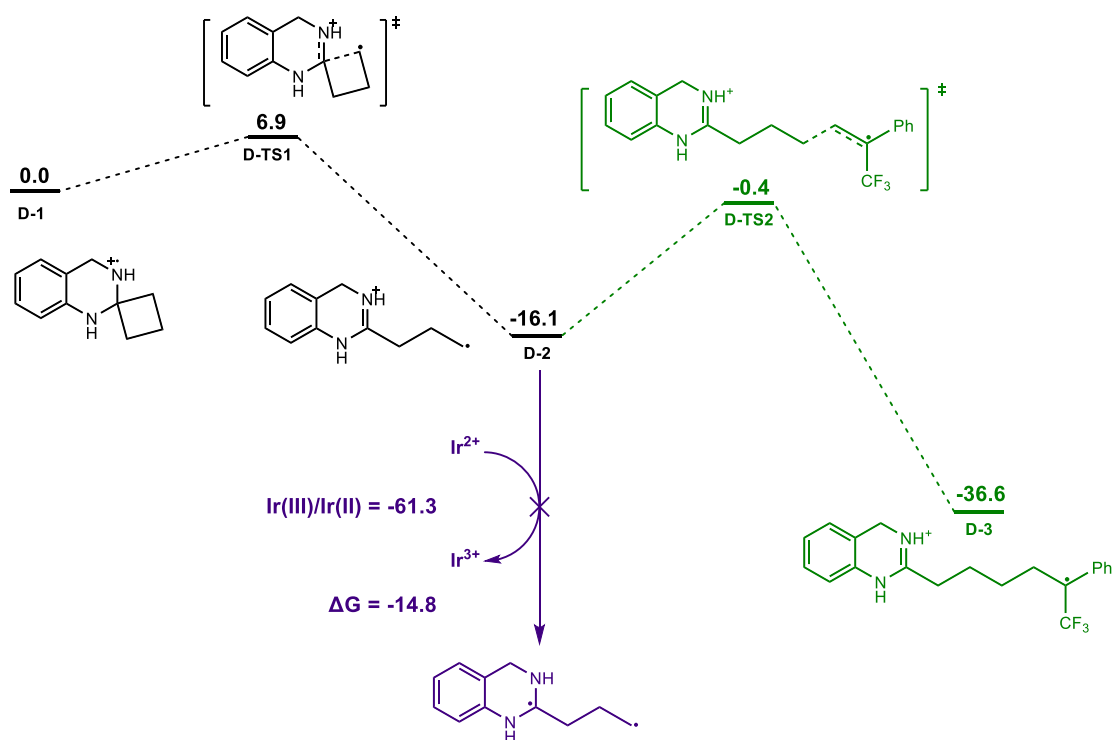
E. Supporting Information for chapter 6
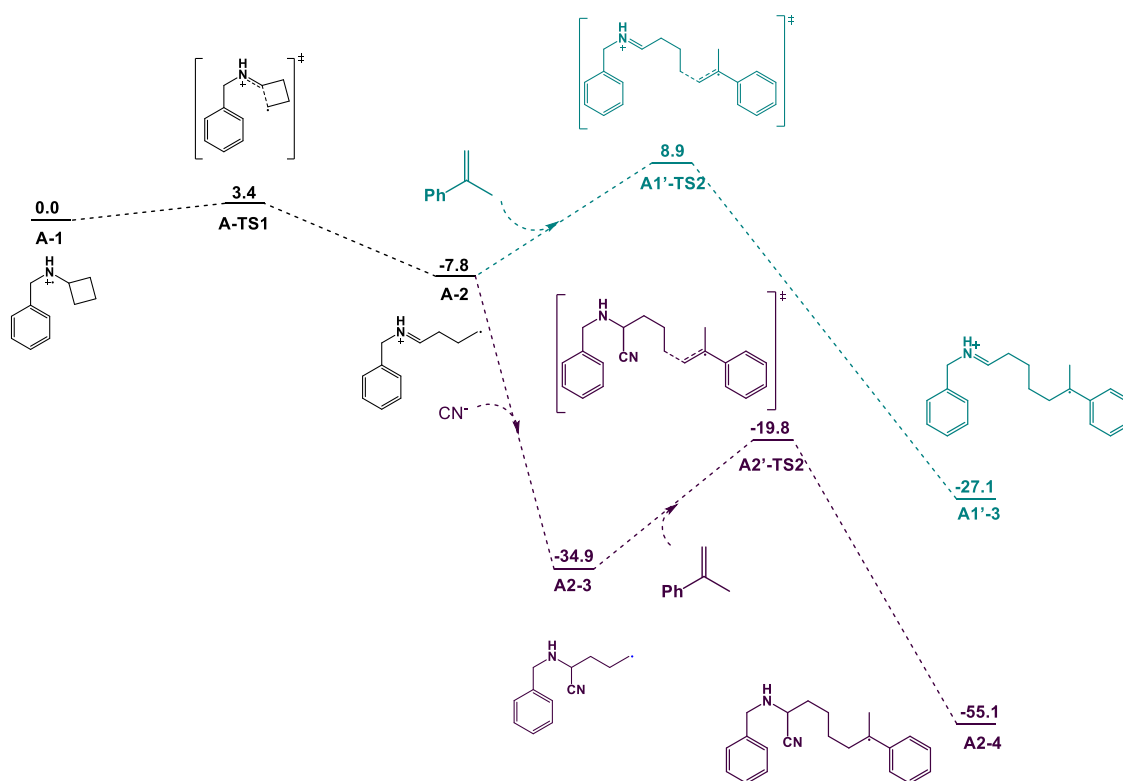
Scheme E1 Reaction coordinate of non-spirocyclic N-benzylcyclobutanamine with styrene. Energies are in kcal/mol.
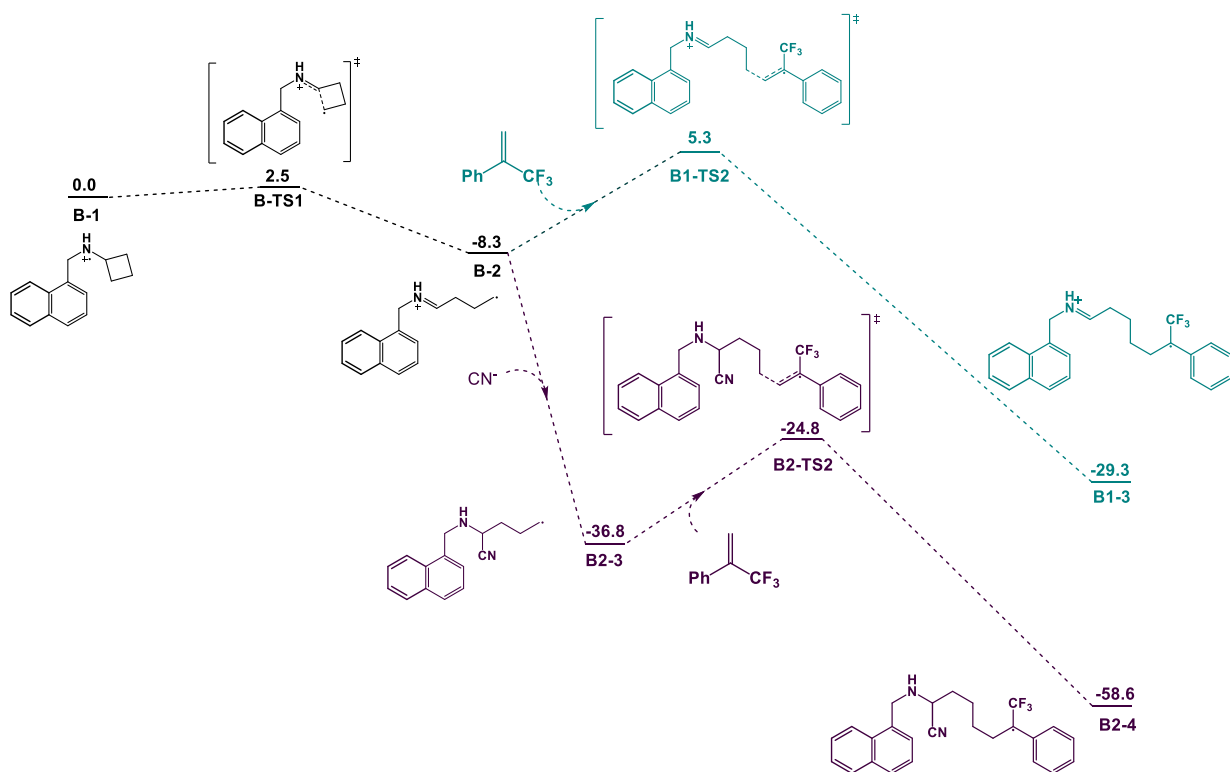
Scheme E2 Reaction coordinate of non-spirocyclic N-(naphthalen-1-ylmethyl) cyclobutanamine with CF$_3$-styrene. Energies are in kcal/mol.
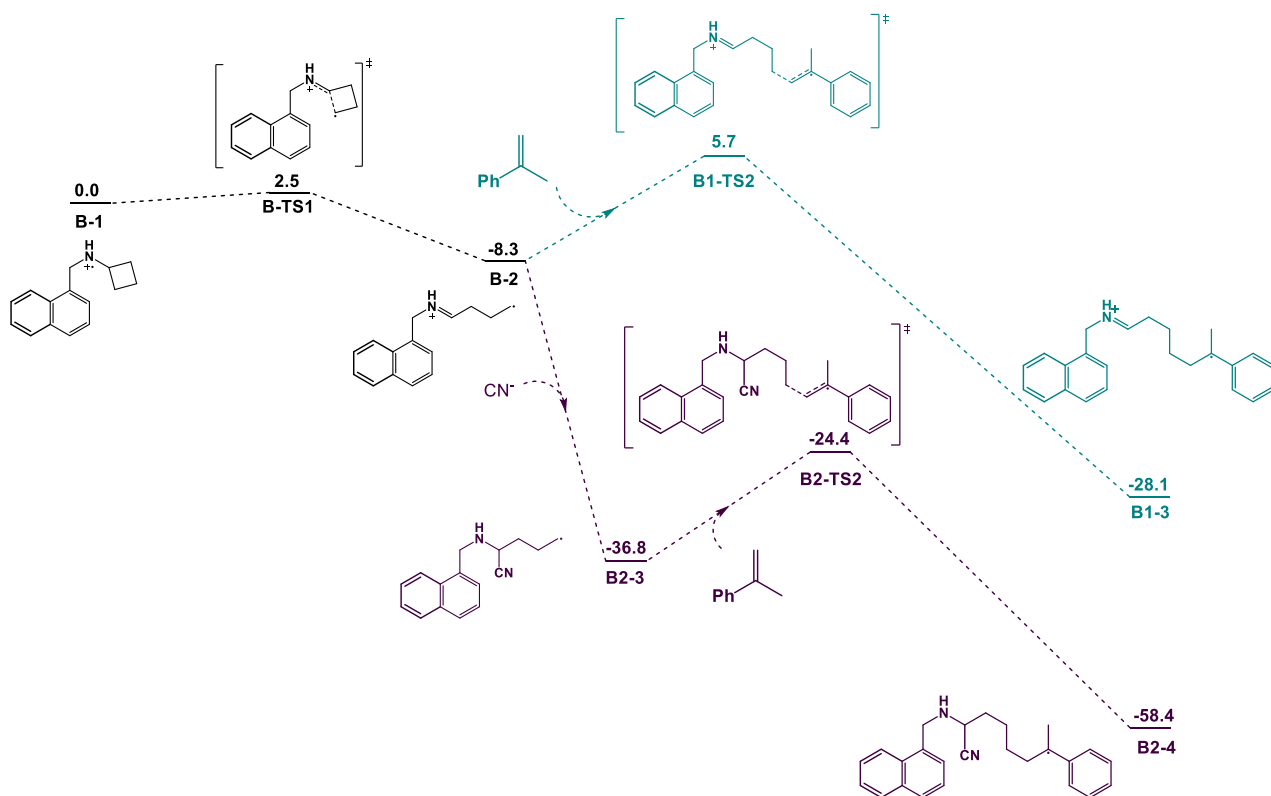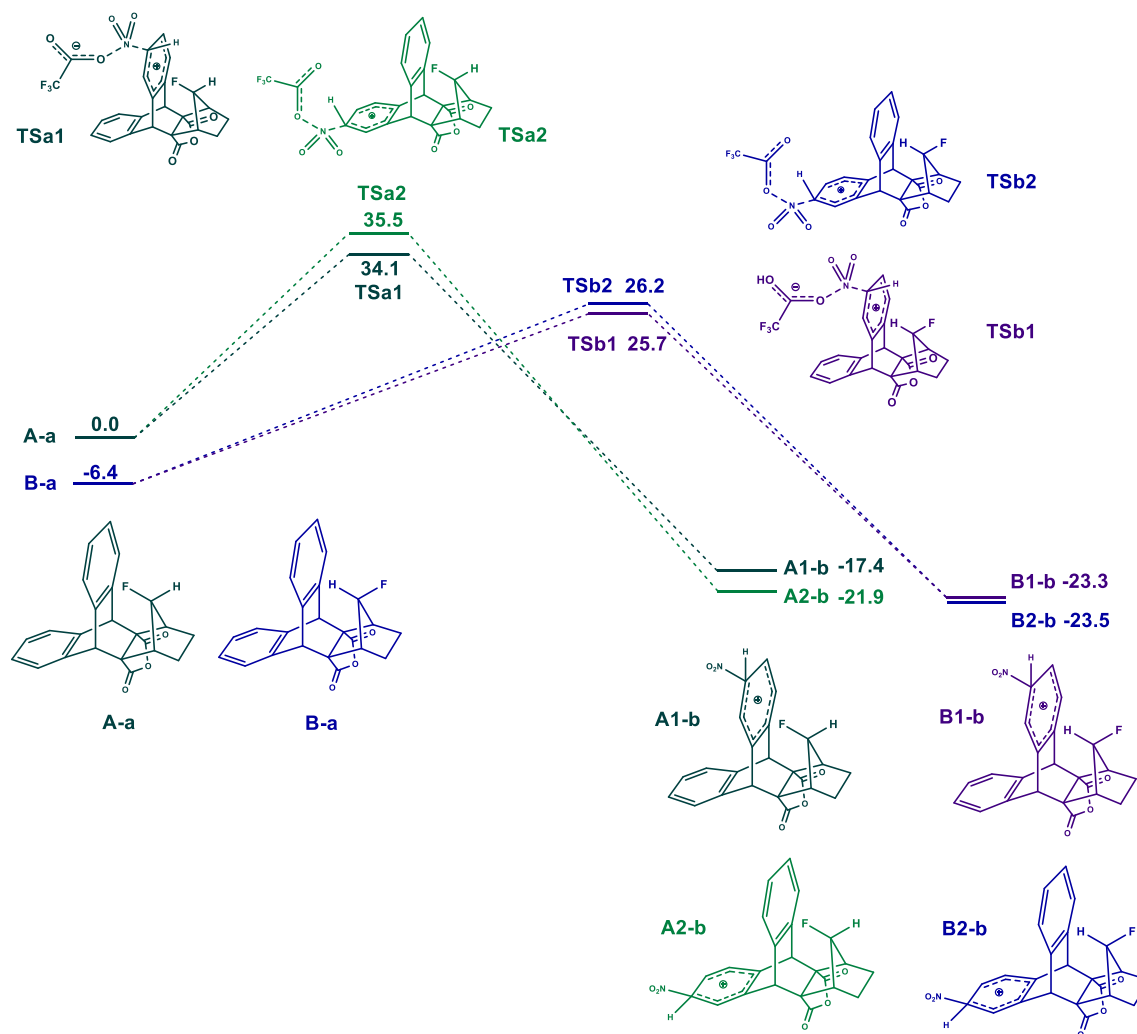
Scheme E3 Reaction coordinate of non-spirocyclic N-(naphthalen-1-ylmethyl) cyclobutanamine with

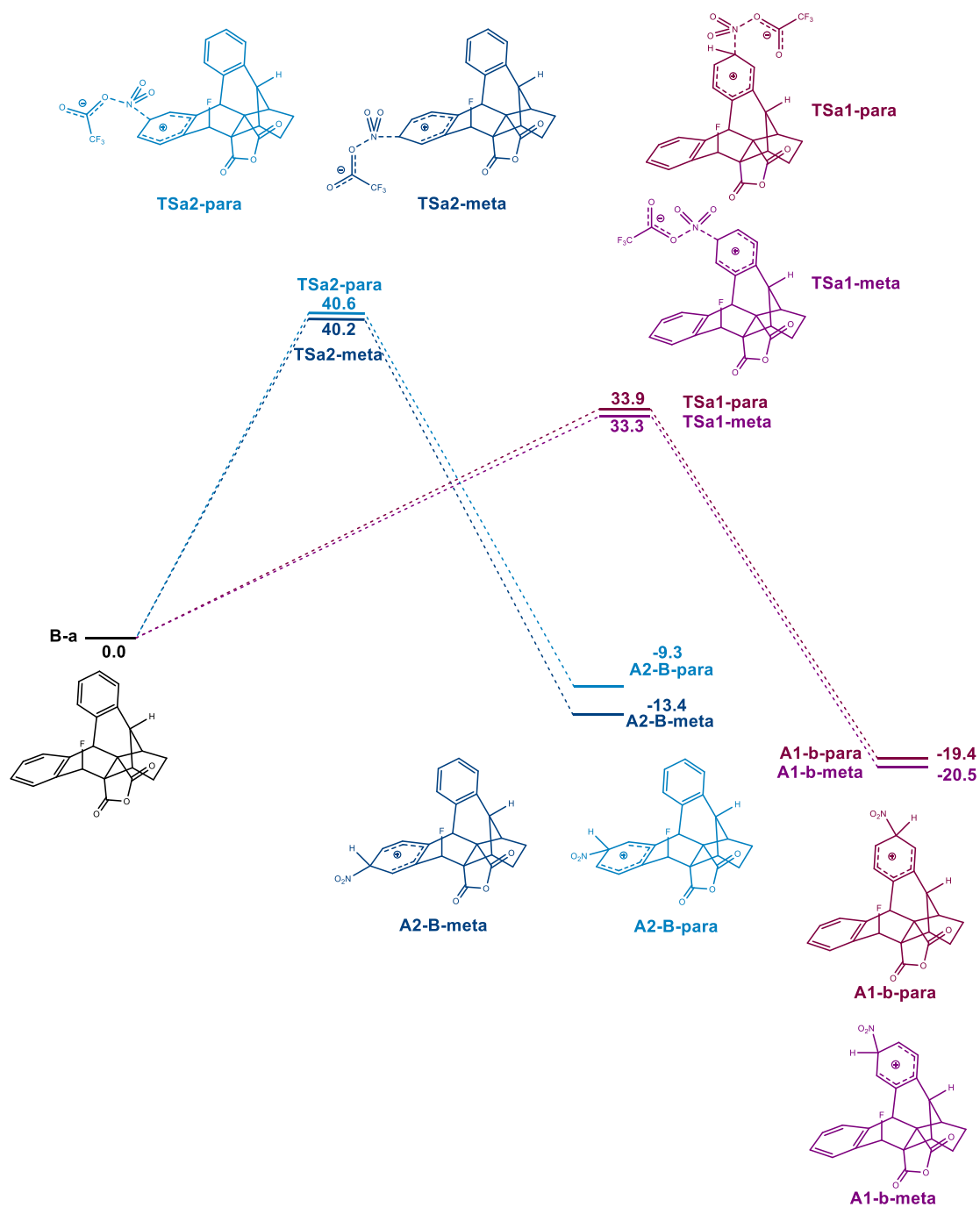-styrene. Energies are in kcal/mol.

Scheme E4 Reaction coordinate of spirocyclic cyclobutanamine with CF$_3$-styrene. Energies are in
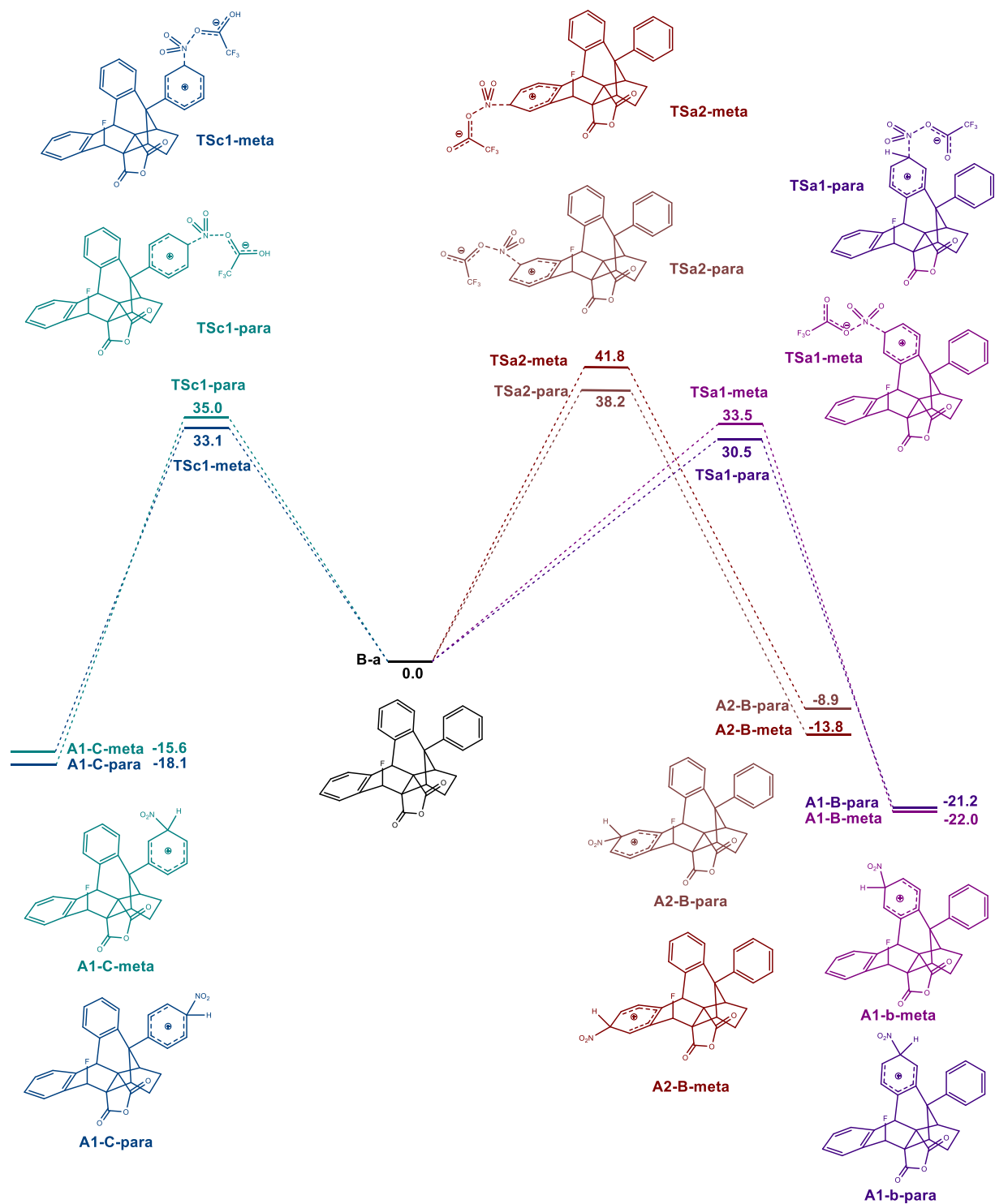
kcal/mol.

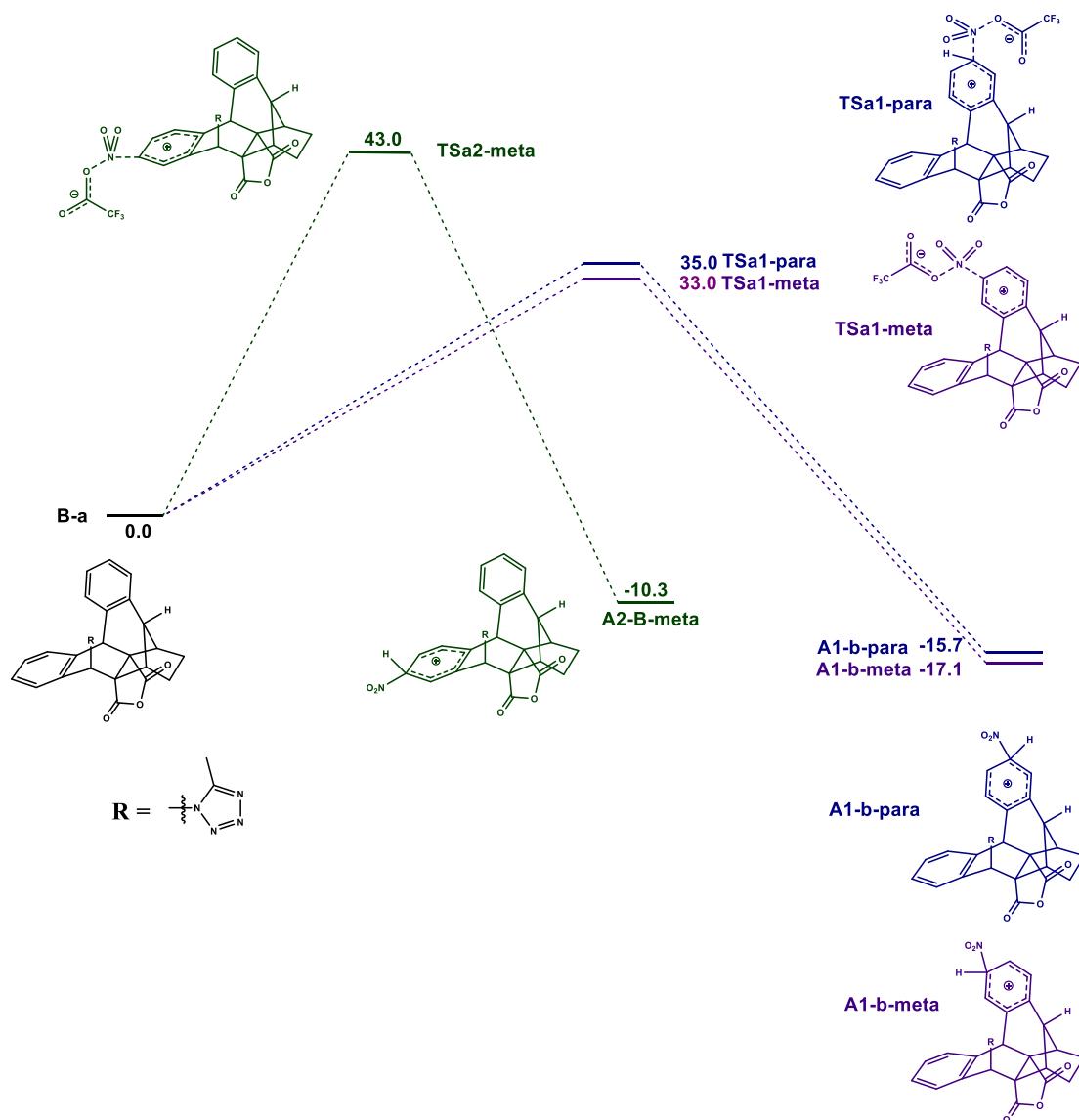# F. Supporting Information for chapter 7

Fig F1. Reaction coordinate of Scaffold I with nitronium ion under M06-2X(SMD)/def2-TZVPD//M06-2X/6-311G(d) level of theory. Energies are in kcal/mol.

Fig F2. Reaction coordinate of **6** with nitronium ion under M06-2X(SMD)/def2-TZVPD//M06-2X/6-311G(d) level of theory. Energies are in kcal/mol.

Fig F3. Reaction coordinate of **10** with nitronium ion under M06-2X(SMD)/def2-TZVPD//M06-2X/6-311G(d) level of theory. Energies are in kcal/mol.

Fig F4. Reaction coordinate of **18** with nitronium ion under M06-2X(SMD)/def2-TZVPD//M06-2X/6-311G(d) level of theory. Energies are in kcal/mol.