Exploring Nonstatistical Reactions: Computational Challenges and Design Opportunities

Tomislav Rožić

This thesis has been submitted
to the PhD School of The Faculty of Science,
University of Copenhagen

Adviser: Gemma C. Solomon

January 2025

Abstract

A cornerstone of theoretical modeling for chemical reactivity is transition state theory (TST). In its simplest form, it allows us to reduce the complex dynamics of molecules to a model that focuses on the free energy barrier the molecule needs to surpass, as defined by its transition state geometry. Expansions of this theory may also take into account the shape of the reaction barrier, allow re-crossing it, as well as include contributions from quantum tunneling across the barrier. As such, it is not only a tool to explain experimental findings, but also the basis of massive automated discovery of reaction networks, used to gauge the relevance and feasibility of each generated elementary reaction step.

When this approach fails to recreate experimental reaction rates and product ratios, the reaction may be labeled as exhibiting nonstatistical effects or dynamics. One notable cause of nonstatistical effects is an excess of energy that does not dissipate before the molecule continues to react, violating the basic assumption of TST – that the molecule is in its thermal equilibrium. Furthermore, depending on dynamics that preceded the nonstatistical reaction step, this chemical activation may be localized to certain parts of the molecule. How well it relaxes and whether it affects the reaction coordinate of interest is structure-dependent and will vary from molecule to molecule. These types of thermally activated reactions will be the focus of my thesis.

In the first half, I describe a modeling strategy composed of three parts: quantifying vibrational energy localization, applying a vibrational energy relaxation model to predict the time evolution of this energy localization, and using this data to generate indicators for comparative ranking of vibrationally activated molecules. The strength of this approach is its modular nature, as focusing on a single hot intermediate/product at a time, we can stay within the frame of its vibrational modes and swap out any of the models used in the three steps without affecting the others. I discuss whether we can completely avoid the full dynamical treatment usually performed with *ab initio* molecular dynamics, and if we cannot – what will be its minimal application.

i

I then apply a proven vibrational relaxation model, including a master equation that provides a time evolution of the vibrational energy into the picoseconds range. With those results, I propose two simple indicators that aim to collapse all the obtained information into a single value per molecule. Using these we may screen for yield control opportunities by proposing temporary modifications to the structure or using alternative reactants, tuning nonstatistical effects while ensuring the reaction mechanism is conserved. Additionally, I show how the time-dependent density matrix renormalization group method (TD-DMRG) can be used with a vibrational Hamiltonian allowing us to efficiently apply quantum dynamics as an intramolecular vibrational energy redistribution model.

In the second half of the thesis, I present a nonstatistical reaction exploration performed in direct collaboration with a synthetic organic chemist. This reaction involves a much larger molecule, representative of the scale that could be relevant to the pharmaceutical industry. I discuss how the extent of reported nonstatistical effects is tied to the completeness of our statistical model; this is both in terms of the level of theory applied and the possibility of parallel reaction paths, as we discover that the reported thermal reaction can also be catalyzed with light.

Throughout this work I also present the current computational methods available for modeling chemical reactivity in general, focusing on those that are freely available (for academic use) and accessible to non-expert users. Finally, I present an outlook considering recent developments in theoretical chemistry relevant to this field and discuss the resurging interest in externally controlled mode-selective chemistry and mechanochemical mechanisms.

Dansk Resumé

En hjørnesten i den teoretiske modellering af kemisk reaktivitet er overgangstilstandsteori (TST). I sin enkleste form tillader den os at reducere molekylers komplekse dynamik til en model, der fokuserer på den frie energibarriere, molekylet skal overvinde, som defineret ved dets overgangstilstandsgeometri. Udvidelser af denne teori kan også tage hensyn til reaktionsbarrierens form, tillade genkrydsning af den samt inkludere bidrag fra kvantetunneling på tværs af barrieren. Som sådan er det ikke kun et værktøj til at forklare eksperimentelle resultater, men også grundlaget for omfattende automatiseret opdagelse af reaktionsnetværk, der bruges til at vurdere relevansen og gennemførligheden af hvert elementært reaktionstrin, der bliver genereret.

Når denne tilgang ikke formår at genskabe eksperimentelle reaktionshastigheder og produkt- forhold, kan reaktionen betegnes som udvisende ikke-statistiske effekter eller dynamik. En bemærkelsesværdig årsag til ikke-statistiske effekter er overskydende energi, der ikke dissiperer, før molekylet fortsætter med at reagere, hvilket bryder TST's grundlæggende antagelse om, at molekylet er i termisk ligevægt. Desuden kan denne kemiske aktivering være lokaliseret til visse dele af molekylet, alt afhængig af den bevægelse, der gik forud for det ikke-statistiske reaktionstrin. Hvor godt denne energi relakserer, og om den påvirker reaktionskoordinatet, afhænger af molekylets struktur og vil variere fra molekyle til molekyle. Disse typer termisk aktiverede reaktioner vil være fokus for min afhandling.

I den første halvdel beskriver jeg en modelleringsstrategi bestående af tre dele: kvantificering af vibrationsenergilokalisering, anvendelse af en vibrationsenergirelaksationsmodel til at forudsige denne energilokaliserings udvikling over tid og brug af disse data til at generere indikatorer til komparativ rangering af vibrationsaktiverede molekyler. Styrken ved denne tilgang er dens modulære natur; ved at fokusere på ét varmt mellemprodukt ad gangen kan vi holde os inden for rammerne af dets vibrationsmodi og udskifte enhver af de mod- eller, der bruges i de tre trin, uden at påvirke de andre. Jeg diskuterer, om vi helt kan undgå den fulde dynamiske behandling, som normalt udføres med *ab initio* molekylær dynamik, og hvis vi ikke kan, hvad der så vil være dens minimale anvendelse.

Jeg anvender derefter en velprøvet vibrationsrelaksationsmodel, der inkluderer en masterligning, som giver vibrationsenergiens udvikling over tid ned til picosekundet. Ud fra disse resultater foreslår jeg to enkle indikatorer, der sigter mod at sammenfatte al den op- nåede information til en enkelt værdi pr. molekyle. Ved at bruge disse kan vi screene for muligheder for at kontrollere udbytte ved midlertidigt at foreslå ændringer i strukturen eller bruge alternative reaktanter, mens ikke-statistiske effekter finjusteres og reak- tionsmekanismen bevares. Derudover viser jeg, hvordan den tidsafhængige densitetsmatrix- renormaliseringsgruppe-metode (TD-DMRG) kan bruges med en vibrationshamiltonian, så vi effektivt kan anvende kvante- dynamik som en VER-model.

I den anden halvdel af afhandlingen præsenterer jeg en undersøgelse af ikke-statistiske reaktioner udført i direkte samarbejde med en syntetisk organisk kemiker. Denne reaktion involverer et meget større molekyle, der er repræsentativt for den skala, der kan være relevant for medicinalindustrien. Jeg diskuterer, hvordan omfanget af rapporterede ikke-statistiske effekter er knyttet til fuldstændigheden af vores statistiske model; dette gælder både niveauet af anvendt teori og muligheden for parallelle reaktionsveje, idet vi opdager, at den rapporterede termiske reaktion også kan katalyseres med lys.

Gennem dette arbejde præsenterer jeg også de nuværende computermetoder, der er tilgængelige for modellering af kemisk reaktivitet generelt, med fokus på dem, der er frit tilgængelige (til akademisk brug) og tilgængelige for ikke-ekspertbrugere. Endelig giver jeg et fremtidsperspektiv med fokus på nylige udviklinger inden for teoretisk kemi, der er relevante for dette felt, og diskuterer den fornyede interesse i eksternt kontrolleret modespecifik kemi og mekanokemiske mekanismer.

Acknowledgments

I wholeheartedly dedicate this thesis to the people who have helped me reach this point of my life: to my parents, for, among other things, the first computer I sat down at.

To my elementary and high school chemistry teachers, Marija Šumanovac and Ivan Petrović, who supported me through national competitions and showed me that it is not only chemistry that is interesting, but the people who are passionate about it. The brilliant professors Nađa Došlić and Tomica Hrenar at my home university, and the talented team of the Theoretical Chemistry Group of Ruđer Bošković Institute in Zagreb. To my university colleagues, including those from PRIMUS, with whom I embarked on this adventure.

To Saša, Lovro, and Vlatko, who have remained a close-knit group of friends for over 12 years and counting. To Dominik, a steadfast friend since elementary school; Jacob for being my first and closest Danish friend & everyone else from Mensa who kept me company in a brand new country. The people I worked with in UCAPS, PAND, and the 4EU+ Alliance, and who added plenty of flavor to a PhD in the "hard sciences".

To the Reiher Group which welcomed me in Zurich, especially Nina and Alberto. To my supervisor, Gemma C. Solomon, and fellow group members – Matthew, William, Lea, Susanne, Joseph, Louise, Raka, and Michael. To Matea, my only Croatian work colleague turned friend – turned collaborator; and to Yuxuan, who shared the challenges and excitement of this project with me.

Finally, to my Mélodie; your love and constant support have (quite literally) crossed many borders and carried me through the final stages of this journey.

List of Publications

The following list of publications is arranged in chronological order. These papers are presented in the dissertation and included in the appendices.

1. <u>Tomislav Rožić</u>, Matthew S. Teynor, Nađa Došlić, David M. Leitner, and Gemma C. Solomon. A Strategy for Modeling Nonstatistical Reactivity Effects: Combining Chemical Activation Estimates with a Vibrational Relaxation Model. *The Journal of Chemical Theory and Computation*, **2024** 20 (20), 9048-9059.

2. <u>Tomislav Rožić</u>, Yuxuan Hou, Lea Kjærgaard Northcote, Christian Markus Pedersen, and Gemma C. Solomon. Tutorial Review on Modeling Nonstatistical Reactivity: an Example of Light and Heat in the Garratt–Braverman/[1,5]-H Shift of Ene-diallenes. *in preparation*

In addition to the papers presented in the thesis, I have also participated in work leading to the following publications. These are not included in the appendices.

3. <u>Tomislav Rožić</u>, Majdi Hochlaf, Ridha Ben Said, and Nađa Došlić. A Computational Approach to Nontraditional Intrinsic Luminescence: Vibrationally Resolved Absorption and Fluorescence Spectra of DABCO. *The Journal of Physical Chemistry A* **2022** 126 (7), 1094-1102.

4. Josip Draženović, <u>Tomislav Rožić</u>, Nađa Došlić, and Nikola Basarić. Excited State Intramolecular Proton Transfer (ESIPT) from -NH2 to the Carbon Atom of a Naphthyl Ring. *The Journal of Organic Chemistry* **2022** 87 (14), 9148-9156.

5. Matea Sršen, Stephan K. Pedersen, <u>Tomislav Rožić</u>, Gemma C. Solomon, Arianna Lanza, and Michael Pittelkow. Four-layered [27]Helicenoids. *in preparation*

Preface

This thesis presents research that has been conducted under the supervision of Professor Gemma C. Solomon at the Department of Chemistry, University of Copenhagen and the PhD School at SCIENCE, from June 2021 to January 2025.

The objective of this thesis was to computationally explore opportunities for reaction control and design in nonstatistical reactions by leveraging the structure dependence of vibrational energy redistribution. This synopsis-based thesis has been divided into six chapters; Chapter 1 presents an introduction to the field, Chapters 2 and 3 discuss the modeling strategy proposed in the first appended paper, and Chapter 4 summarizes the second manuscript, which is currently in preparation. Chapter 5 is the product of my change of research environment, which was performed at the group of Professor Markus Reiher at ETH Zurich. Finally, Chapter 6 offers an outlook into other areas that intersect with nonstatistical reactivity and represent my own research interests.

The work performed in this thesis has been presented as a poster at ICQC 2023 (Bratislava), CECAM PIQM 2023 (Tel Aviv), TACC 2023 (Sapporo), the 6th ICReDD (Sapporo), PQCC 2024 (Zurich), and our department's PhD Seminar. Additionally, I have given talks at the Excitation Energy Transport in Physical, Chemical, and Biological Systems workshop (2023, Split) and at the Theoretical Section of the Croatian Chemistry Society (2024, Zagreb).

x

Contents

1

The concept of nonstatistical reactivity

During my Master's thesis in Zagreb, I was tasked with characterizing the nontraditional luminescence of DABCO (1,4-diazabicyclo[2.2.2]octane), a relatively popular organic molecule, often used as a ligand or linker. Due to its small size and relevance in literature, it was a fitting model system for the wider field of NTIL (non-traditional intrinsic luminescence) which was a blanket term for luminescence that could not be explained by standard photochemistry/photophysics. In the end, I managed to show that there was nothing exceedingly mysterious about the NTIL of DABCO – the molecule was following all the "basic rules" of photophysics by absorbing through higher oscillator strength singlet transitions, resulting in fluorescence from the lowest one, S_1. It was just difficult to model as the states involved were of Rydberg nature, as opposed to more standard valence excitations, and required diffuse basis functions to treat properly.

This project also introduced me to ORCA, which had its roots in theoretical molecular spectroscopy and has since evolved to be one of the most popular quantum chemistry packages. With the path integral approach of ORCA's excited state dynamics module, I was able to compute absorption and fluorescence spectra with their vibronic structure and match them

to experiment, something that would have been prohibitively expensive just some years ago. I often found ORCA introducing features relevant to me as my PhD project progressed, and its robustness with large molecules allowed for less constrained collaborations with synthetic chemists. I feel this is important to note, given the software is free for academic use, is increasingly well integrated with other codes such as reaction rate calculators, and has an overall very helpful online community forum.

In the end, it is hard to quantify whether properties should stay labeled nontraditional or unexpected once we can fully model them with theory. This experience prepared me well for a project on nonstatistical effects, arguably an even more chaotic term in literature. Per definition, these are effects on rates and/or product ratios that do not match predictions of transition state theory (TST). Figure 1.1 shows a cartoon depiction of some common motifs of nonstatistical reactivity and their underlying cause. In the first two cases, the localization of energy in the molecule dictates how the next reaction step occurs, violating the fundamental assumption of TST: that the molecule is in its thermal equilibrium. With dynamic matching, the reactive dynamics are pre-determined by the previous step. In the more general case thermally activated or hot products and intermediates can also exhibit nonstatistical effects even without matched reaction coordinates when vibrational energy relaxation (VER) is slow. Leveraging the nonstatistical/asymmetric nature of VER, specifically intramolecular vibrational energy redistribution (IVR), will be the main topic of this thesis.

In the third example of Figure 1.1, the result is completely decided by dynamics, as there is no clear intermediate structure at which the system may stay long enough to thermalize. In some cases, this may present itself as going "uphill" on the PES, as is also seen in the third diagram of Figure 1.2. While this thesis focuses on unimolecular reactions, the preceding exothermic step may well be a bimolecular reaction, examples being the formation of the ozonide intermediate preparing the nonstatistical ozonolysis of vinyl ethers[1] and the hot intermediates involved in the hydroboration of alkenes.[2]

1.1 State of the art: *ab-initio* molecular dynamics and coordinate matching algorithms

The term dynamic matching was brought into prominence by Professor Barry Carpenter, an exceptional physical organic chemist whose work has become a cornerstone of this area. In addition to his research contributions, he was also a strong advocate for applying computa-
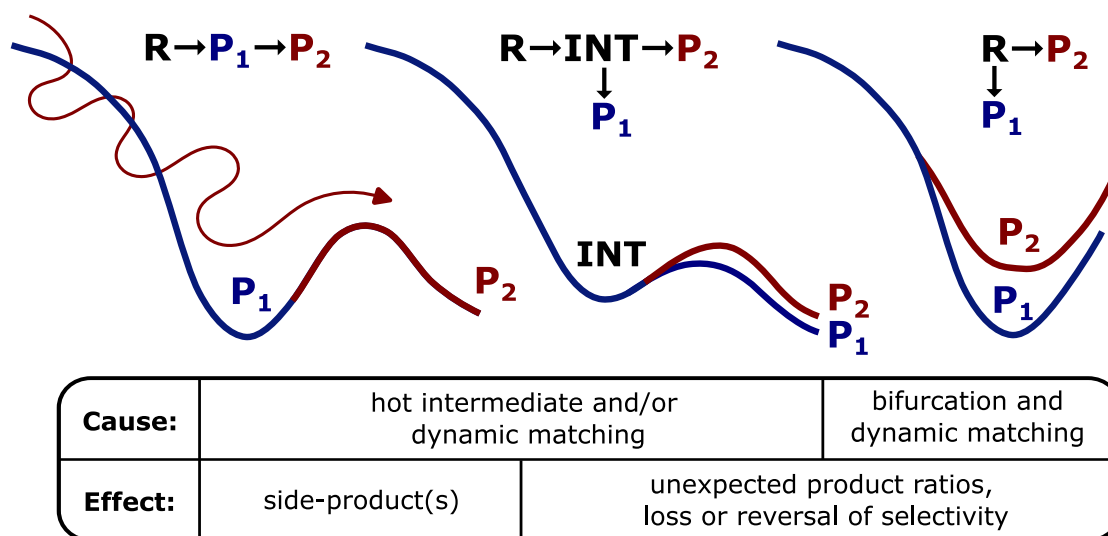
Cause:	hot intermediate and/or dynamic matching		bifurcation and dynamic matching
Effect:	side-product(s)	unexpected product ratios, loss or reversal of selectivity	

Figure 1.1: Three examples where a larger amount of the second, nonstatistical product (P_2 in red) is obtained, shown on a cartoon representation of the potential energy surface along the reaction coordinate(s). In the first two cases, the preceding reaction (from reactant R) produces a thermally activated product (P_1) or intermediate (INT). In the third example, underdamped dynamics from the previous step may result in the less energetically stable P_2.

tional simulations to explain these effects. This serves as a reminder of how important it is to also validate theoretical methods, lest they are delayed from reaching general use. After all, newer generations may take for granted that we are able to explicitly model much of chemical reactivity with the use of *ab initio* molecular dynamics (AIMD), followed by trajectory analysis, and instead focus mainly on the fact it is expensive to do so. An editorial on his career, including a full publication list that is invaluable for a more thorough understanding of dynamic effects in chemistry, may be found at Ref. 3.

These papers set off the literature search at the start of my project and so most of the theoretical and experimental work done to characterize nonstatistical mechanisms that I have been exposed to and cite in this thesis has been in solvent-based organic chemistry. At Ref. 4 a 2016 Perspective titled "The Study of Reactive Intermediates in Condensed Phases" explains why these effects survive even in solvated systems and lays out a summary of both the computational models and experimental techniques used to study them.

Much of the theoretical work in this area is supported by AIMD simulations using density functional theory (DFT) as the method of choice for the electronic structure of the modeled molecules. Both of these are widely in use in computational chemistry and are out of the scope of this synopsis-based thesis, though a handy 2010 review of this specific area of physical or-

ganic chemistry, titled "Molecular dynamics simulations and mechanism of organic reactions: non-TST behaviors", can be found in Ref. 5.

Of course, these phenomena may occur in other areas under different names. As AIMD has also been a powerful tool for catalyst design, I will just mention the terms of "dynamically coupled" reactivity and the concept of "ballistic" trajectories discussed by Daniel H. Ess and his group. These ballistic trajectories are reactive trajectories that skip intermediate structures, effectively turning a 2-step mechanism into a single-step one; they are also differentiated from unrelaxed trajectories, those that we may consider a thermally activated intermediate, as presented in Ref. 6.

It is the *ab initio* part that makes AIMD both versatile and expensive, as the potential energy surface (PES) that molecules move on is obtained through quantum chemical methods and allows for general chemical reactivity to be modeled. However, the dynamics of the simulation are still set in classical physics. At first glance, this is not a deal-breaker as the energy localization we will discuss can certainly occur in a set of coupled classical oscillators. Discrepancies may arise due to the quantum nature of molecules, especially with very light atoms, such as hydrogen. From my experience with the literature, the lack of included quantum tunneling, zero-point energy (ZPE) leakage, and inadequate sampling have been key talking points throughout the application of AIMD.

With ZPE leakage, we face the problem of not being able to add ZPE to more accurately capture the movement of molecules without it eventually "leaking" from higher frequency vibrational modes into lower frequency ones – which also makes it an issue for accurate modeling of vibrational relaxation. Sampling is a more general issue, as capturing reactivity requires different initial conditions: more sampling for larger, flexible molecules, worsening the scaling of the computational cost to model these systems properly, and longer simulations, which brings us back to the issue of ZPE leakage, especially when dealing with vibrationally hot molecules. A recent perspective on these issues in IVR modeling may be found in Ref 7. It is also a good entry point into quantum theories for IVR that are more complete than the model I apply in this project.

Instead, let us mention several approaches proposed to replace or accompany these costly simulations. The first one, dubbed the VRAI (Valley Ridge Augmented Implementation) selectivity algorithm requires no more data than those required to perform TST calculations or an AIMD study – the geometries of reactants, intermediates, products, and the transition states connecting them.[8] The algorithm then determines a dimensionality reduction of the

full coordinate system to a 2D PES that best describes the bonding differences between two possible products whose selectivity (ratio) is not described well by TST. The projection of the imaginary eigenvector of the preceding TS onto this 2D coordinate system then provides a selectivity estimate that matches well those of AIMD for several published examples. The several scenarios where this approach may be applied are laid out in Figure 1.2.
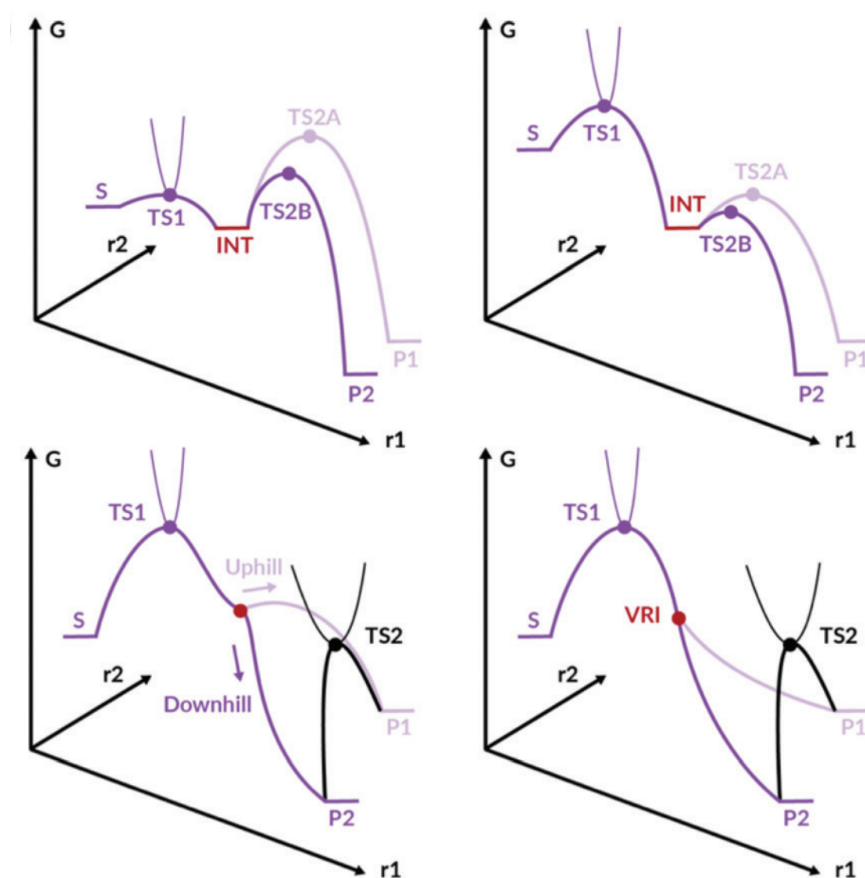
Figure 1.2: A summary of the cases that can be modeled by the Valley Ridge Augmented Implementation Selectivity algorithm; where G represents the Gibbs free energy, TS are transition states and P are products. The coordinates r1 and r2 are chosen by identifying key bond differences between the two possible products. Reproduced from Ref. 8 under CC BY license.

Another approach on the trail of dynamic matching is the Sudden Vector Projection Model, demonstrated on uni- and bimolecular reactions in the gas phase and with gas-surface reactions.[9] It determines the coupling of a reactant mode with the reaction coordinate at the transition state by simply projecting one onto the other. It is likely the simplest way of capturing the directionality imposed by the location of the reaction barrier in the "sudden limit"

when reactivity occurs much faster than IVR.

The approach of focusing on slow IVR, reaction coordinates, and specific vibrations that match them ties well into one of the many "holy grails" presented in chemistry: the desire to command reactivity at will with external control, its flagship being the concept of mode-specific chemistry. This idea, which is much older and made even more popular by the rapid development of laser technology in the 1980s, imagined we could deposit energy into specific bonds (via vibrational modes) and force them to react. This would essentially allow us to rise above traditional chemistry and depending on the level of available external control, increase by several orders of magnitude the number of available reaction mechanisms to use in chemical reaction design.

In a way, the wider area of nonstatistical reactivity has taken up this torch, as it is home to much of the theory required to understand the feasibility of mode-specific chemistry – from discussing the asymmetric flow of localized energy in molecules to how dynamic effects dominate reactivity, given an initial stimulus. In turn, it is also reactions presenting as highly nonstatistical that offer some amount of control, as we try to eliminate or exacerbate these effects. While this thesis will focus on localized energy created by the reaction mechanism itself, at the very end we will briefly mention recent developments in external control.

Another major reason for my focus on solvent-based organic chemistry is that our project included funding for a synthetic organic postdoc, with which I have been directly collaborating for the last year of my project. The goal from the start was to push this research in a direction where interesting theoretical predictions could be immediately tested in the lab. The underlying work package was directly inspired by examples of asymmetrical energy flow in molecules and the potential of using the unique vibrational dynamics of molecules to affect their reactivity in cases where these timescales overlap, that is in cases that inherently violate TST. Figure 1.3 shows a cartoon example in which slower vibrational energy relaxation leads to a greater impact of preceding dynamics and in general, a wider exploration of the PES, allowing for more reactive events.

As IVR is dominated by resonances of the vibrational modes, it can be considered a form of control through molecular interference, a continuous research theme of my supervisor. Although these effects may also occur in classical physics, the accurate electronic structure data required to model them will inherently call for quantum chemical methods, as will the TST models we compare them to.
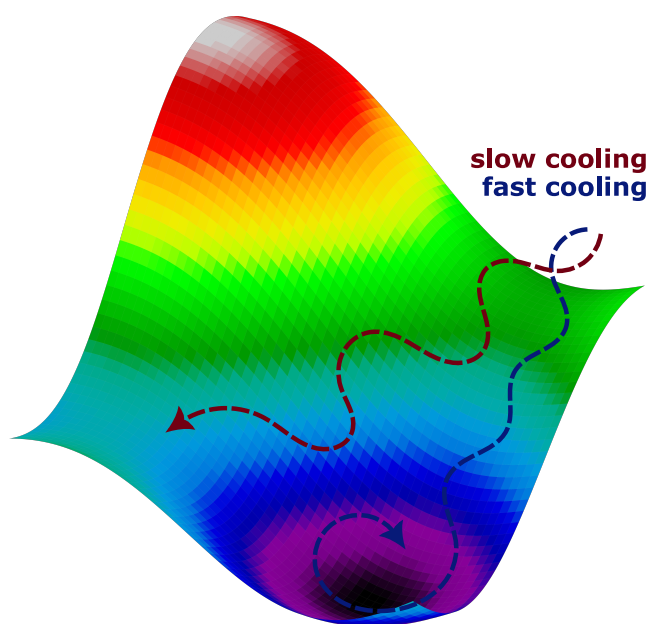
Figure 1.3: When relaxation is slow, both the initial energy and energy released during the reaction allow the system to explore more of the potential energy surface (PES). Staying longer in high-energy regions can allow for new reactions to occur (including tunneling).

1.2 TRANSITION STATE THEORY (TST) RATES AND RATIOS

It is implied that when we talk about nonstatistical reactivity, we have exhausted the capabilities of standard practice transition state theory models – and therefore need to be aware of them. Many of these methods have been established rate theories for decades, so I will provide a brief overview using only the most recent reviews and publications, as to show their current common use and current directions of development.

First off, we have to address the difference between an incomplete theory and an incomplete model. In the perspective by Glowacki et al., titled "Taking Ockham's razor to enzyme dynamics and catalysis",[10] a seemingly complex nonstatistical reaction involving enzymes is made simple again just by including another conformer of the enzyme complex to the model. At the same time, the basic TST rate equations are improved upon by including a thermally averaged tunneling transmission coefficient. The resulting two-state TST model is still as simple as it gets while already providing a near-perfect fit to the experimental data for several different enzyme systems. Therefore, once a reaction path and/or energy diagram for a reaction is obtained, it is important not to become tunnel-visioned toward other, similar reaction

paths occurring in parallel.

Some TST approaches take this into account explicitly, for instance, multiconformer TST,[11] ensemble-averaged variational TST, and multistructural variational TST.[12] Variational transition state theory (VTST) has been one of, if not the most, widely used evolution of TST since the 1980s. In short, the variation comes from adjusting the location of the dividing surface, usually assumed to be located at the saddle point (TS). This may become expensive, as it requires a free energy profile along the minimum energy path, obtained by performing many frequency calculations on geometries displaced along the reaction coordinate. Additionally, the location of the dividing surface is dependent on temperature and the VTST calculation must be performed again when it changes. However, as the underlying electronic structure calculations may be reused, this process is not as demanding. The recent review of VTST in Ref. 12 also includes a brief history of the theory of chemical kinetics; the reader may follow the literature trail to find further publications on how modern TST methods developed during the last few decades.

A tunneling transmission coefficient is included in these rates, which leads me to mention semiclassical TST (SCTST) approaches, which aim for accurate rates when quantum effects dominate. This occurs with low temperatures in general, hydrogen transfer reactions for which tunneling may remain relevant at higher temperatures, but also for some examples of heavy atom tunneling.[13] In a very recent perspective, semiclassical approaches in general have been described as a way to include quantum effects for all degrees of freedom, as well as their anharmonic coupling to the reaction coordinate.[14] They emphasize SCTST applications now reach systems with over a hundred degrees of freedom, limited mainly by the cost of electronic structure calculations required to obtain anharmonic constants.

Improvements upon basic TST often stem from moving away from their heavy reliance on the transition state geometry, as obtained by saddle point optimization algorithms on an electronic potential energy surface. For instance, we may compute the Gibbs free energy of the barrier, as defined by the TS, but that does not mean the TS is also a saddle point of the free energy surface. Due to the entropy term ($T\Delta S$) this discrepancy will be temperature dependent. Other than adjusting the dividing surface to account for this, we should be aware the free energy surface may assume a completely different shape near the reaction barrier, as is discussed through the concept of "entropic intermediates". In the 2019 review (Ref. 15) on these intermediates, the label of dynamically concerted and unconcerted reactions is also suggested. In a dynamically concerted reaction, all bonding changes are complete in less than

60 fs, and while this is considered too fast to exhibit entropic intermediates, it also does not allow for proper IVR, making this another likely label or class of nonstatistical reactivity.

Additionally, in 2014 and in the broader area of biological chemistry, non-equilibrium transition state theory was presented for rates of stem cell differentiation. With a path integral approach, a minimal action most probable path is identified, and a new saddle point is chosen as the "global maximum along the dominant path".[16] It is derived from Kramer's theory, a form of TST that accounts for the stochastic dynamics of realistic systems by leveraging the Langevin equation to include both random forces representing thermal fluctuations as well as deterministic force as defined by the PES. The use of the Langevin equation inherently allows for barrier recrossing and a friction term is included to bring the model closer to reactivity occurring in solutions.

Finally, when talking about excited intermediates, it is important to mention Rice-Ramsperger-Kassel-Marcus (RRKM) theory, specifically designed for high-energy uni-molecular reactions.[17] This theory is still fundamentally statistical, as the energy is assumed to be randomized (due to fast IVR) among all the degrees of freedom. However, in some modeling scenarios, energy is manually added to specific vibrational modes and the accompanying discussion on intrinsic RRKM versus non-RRKM dynamics, ergodicity and bottlenecks in phase space offers plenty of insight and direct overlap with nonstatistical dynamics.

2

Quantifying a vibrationally hot molecule

I mentioned some systems require full treatment with molecular dynamics, the development of which is an ongoing effort for chemistry in the broadest sense. While I will discuss where AIMD should be used, the aim of this PhD project was to draw upon the knowledge of vibrational energy localization present in chemical physics and argue whether it can be applied to the more general problem of chemical reactivity.

The ideal application cases will be those systems where we can still place the model, and ourselves as the observer, into the framework of vibrational normal modes. This allows us to talk about specific modes, for example, stretching and bending motions of certain functional groups. Besides making available the plethora of molecular vibration models based on the quantum harmonic oscillator, it also makes it simpler to collaborate with synthetic chemists as the theory behind vibrational normal modes is part of any introduction to physical chemistry and is further reinforced by their connection to infrared (IR) spectroscopy. A chemist who regularly uses IR tables would likely be quicker to suggest modifications to the molecule that change VER properties, once the modeling strategy is presented to them.

The theoretical prerequisites for normal modes are the common Born–Oppenheimer approximation and the more limiting harmonic approximation. For the harmonic approximation, we must model a molecule in one of its equilibrium geometries and assume that the potential energy around this minimum has the shape of a harmonic (parabolic) potential whenever displaced along the normal modes. As chemical bonds resemble Morse potentials, this is a fair approximation, and departures from this approximation, termed anharmonicity, are often manageably low. For example, computed harmonic normal mode frequencies are regularly scaled down around 5%, depending on specific method and basis set benchmarks, to match experimental results from spectroscopy.[18] Notable exceptions of very anharmonic modes are bound to appear in larger molecules, especially those with much conformational flexibility, as these modes lead to other conformational minima instead of bond breaking. The same goes for modes that correspond well to reaction coordinates with low activation barriers, as instead of a parabolic potential wall, the PES tapers off toward the reactant or the next product minima. The treatment of these modes is likely the greatest challenge of this framework.

In other words, ideal candidates to apply VER models are vibrationally excited or "hot" products and intermediates, where any further reactivity depends on the localized vibrational energy of a single structure. As we saw in the literature on nonstatistical reactivity, this is not an uncommon motif. Additionally, the concept of elementary reaction steps and the use of TST rates are building blocks of many approaches to automated reaction discovery. If we wished to implement a way of flagging potential nonstatistical effects and generating alternate structures where these effects would be toned down (or exaggerated) without modifying the underlying mechanism, we would also start with identifying elementary steps that produce potentially "hot" products. The overall motivation of this thesis is therefore to build toward predicting and quantifying nonstatistical reactivity fully *ab initio* and ahead of comparison to experimental yields and ratios.

Chapters 2 and 3 will provide context and additional information on the modeling strategy presented in the first paper: *A Strategy for Modeling Nonstatistical Reactivity Effects: Combining Chemical Activation Estimates with a Vibrational Relaxation Model*. We will start out by describing a way to quantify thermal chemical activation.

The model system chosen for this project was 2-methylbicyclo[2.1.0]pent-2-ene (2-MeBCP), in which a bond of the strained four-membered ring breaks, bonds are re-arranged, and 2-methyl-1,3-cyclopentadiene (2-MeCP) is formed, as shown in 2.1. As this releases a large amount of energy (up to 300 kJ mol^{-1}), another product (1-MeCP) is formed following a hydrogen transfer before the initial, hot product (2-MeCP) fully thermalizes. The reasoning behind this choice was that the molecule is small, the mechanism discussed in experimental literature, and most importantly, it has been previously modeled computationally, using AIMD, by Goldman, Glowacki, and Carpenter.[19]
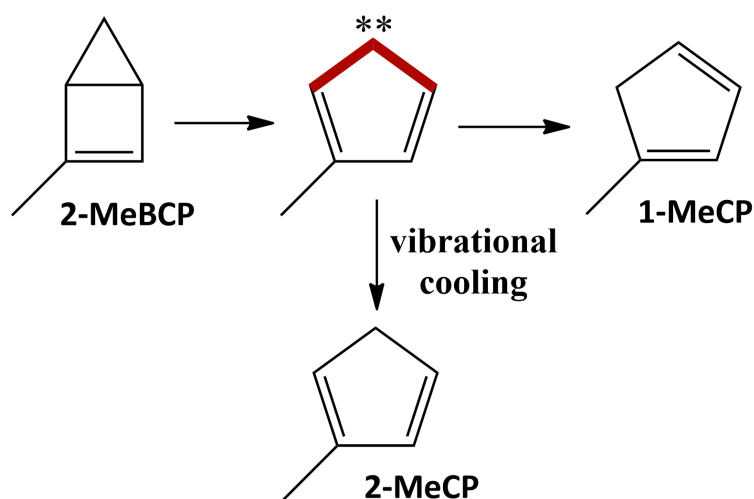

Figure 2.1: Reaction Scheme for the Ring Opening of 2-MeBCP. Figure and caption reproduced from Paper 1.

Despite the small size of the system, it already showcases several challenges of theoretical modeling in chemistry. The first reaction step is a bond rearrangement and at (and around) its TS geometry we can expect some contribution of singlet diradical character to the electronic structure of the molecule. This allowed me to discuss the use of broken-symmetry DFT, an affordable approach when static correlation is present but relatively localized. Secondly, the nonstatistical reaction step that follows is a hydrogen transfer, and as we will see in the second half of the thesis, the contribution of quantum tunneling cannot be ignored. Finally, the methyl group on the ring behaves as a rotor, allowing for faster relaxation by drawing away vibrational energy[20] – not only is this not directly included when using a VER model based

solely on vibrational modes, but the normal modes that describe this rotation are bound to be anharmonic.

2.2 Direct geometry projections

The simplest approach to estimate the chemical activation of the hot intermediate would be to directly compare its geometry with that of the ring-opening transition state. This difference can then be expressed in terms of the vibrational normal mode displacement of the hot product and its vibrational excitation estimated as:

$$V_\alpha = \frac{\omega_\alpha^2 Q_\alpha^2}{2} \tag{2.1}$$

Where V_α is the estimated potential energy in mode α with its frequency ω_α and displacement Q_α. In essence, we are imagining the TS to be a very stretched version of the product. As the normal modes assume a harmonic potential, this approach will encounter problems with very anharmonic modes or large displacements in general. Despite its crudeness, this approach has been used in areas where the geometry differences are small enough, one example being photochemistry, where it can be used to gauge whether a molecule excited to its Franck-Condon region has enough energy to reach a nearby conical intersection.[21]

The main danger with this approach is that even small errors in displacing high-frequency bond stretches could result in a large amount of added estimated vibrational energy. As vibrational relaxation occurs more easily down the vibrational ladder, this may indirectly supply other modes with unreasonable amounts of energy during the time propagation of any VER model.

Once the indicators are defined in Chapter 3, we will see that a lot of nuance in the results is lost with this method, even for a molecule this small. For systems that have very similar VER, predictive power could be lost. As the molecule goes from a strained chair shape to a cyclopentadiene ring, this structural change is already difficult to directly project onto normal modes, and as it is not an uncommon one, prevents us from using it in a general, black-box approach. Despite this, the mode matching/projection methods mentioned in Chapter 1 may still be a suitable complementary method as they aim to predict selectivity based on the shape and position of the transition state(s), as opposed to my motivation of modifying the IVR properties while keeping the TS unchanged.

In conclusion, we will first need to reach the harmonic basin of the PES around the hot

product, where the harmonic projection scheme becomes valid. As the simplest way to do that now is by running a molecular dynamics simulation, we will look for the minimal application necessary.

2.3 AIMD ENSEMBLE VS. SINGLE GLIDING TRAJECTORY

The standard approach for ab initio molecular dynamics is to generate an ensemble with different starting conditions and simulate classical dynamics trajectories. Newton's equations of motion are applied with a numerical propagation algorithm such as Verlet or Velocity Verlet and forces computed by quantum chemical or semiempirical methods. For reactivity modeling, quasiclassical AIMD is often applied. Its name implies that the dynamics are still fully classical, but the initial sampling gives the system both ZPE and at temperatures above 0 K, randomly sampled thermal excitations.

We do this by obtaining a Hessian from a frequency calculation at the TS geometry and then applying either Wigner sampling, where both velocities and displacements along normal modes are generated (and correlated), or the simpler sampling of velocities only from a Maxwell-Boltzmann thermal distribution. Running a large ensemble of trajectories can become prohibitively expensive; at the same time with longer simulations, we encounter the previously mentioned issue of ZPE leakage.

Larger systems with more degrees of freedom will in turn require more sampling, leading to unfavorable scaling, but a more specific issue here is that a system crossing, re-crossing, or quantum tunneling through a reaction barrier might hold a specific dynamic signature, with the barrier acting as a filter. This might be difficult to capture with standard AIMD starting and being sampled at the TS, despite the size of the ensemble. Additionally, moving down a very steep PES will require shorter timesteps to properly model the fast vibrational dynamics of higher frequency modes. We used a relatively conservative constant timestep of 0.2 fs, but Goldman et al. relied on an algorithm that varies the timestep based on the current gradient value that suggested an even lower average value of just 0.1 fs (as explained in the SI of Ref. 19).

With the IVR model I use in Chapter 3 and with the intermediate size (10 – 100 atoms) molecules I work with, the cost of a single AIMD simulation is roughly on the order of obtaining the electronic results needed for the IVR model and running the model itself. Therefore, being able to benefit from a single trajectory as opposed to a full ensemble would entail a

significant speed-up, no matter the price of the underlying electronic structure used.

We chose to initiate this single trajectory with no ZPE to ensure only the kinetic energy gained from the drop is obtained, but add 0.5 kcal mol^{-1} (just above 2 kJ mol^{-1}) in to nudge it down the PES. This saves time by cutting down on the steps required to leave the very top of the barrier and also ensures the correct phase (direction) of the imaginary mode, towards the product. We designate it as a single direct trajectory, or more specifically a "0 K pushed trajectory".
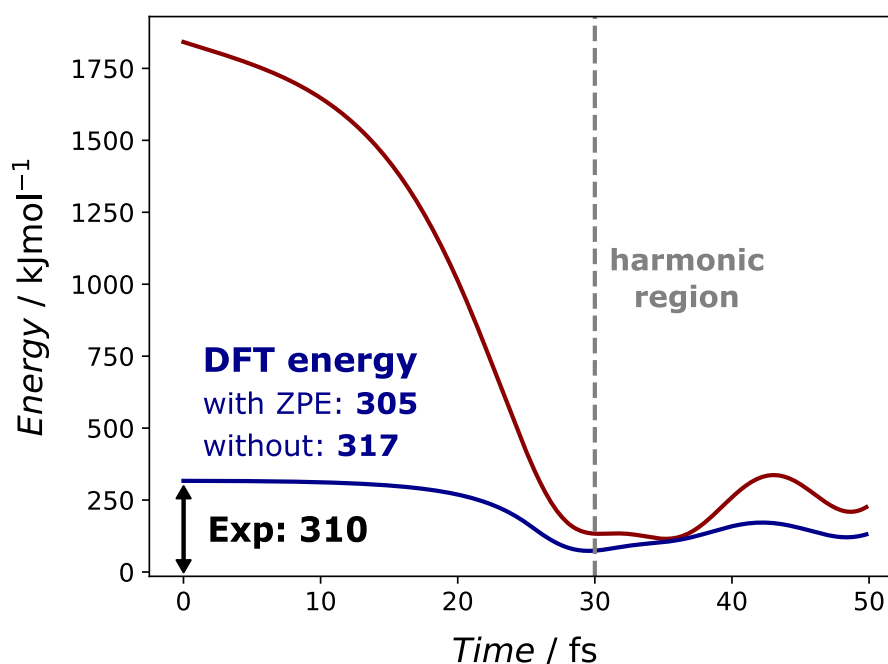

Figure 2.2: Figure shows a comparison of the computed electronic energy (dark blue) and the sum of the estimated energies from the harmonic projection scheme (dark red) with reference to the CP minimum during a "0 K pushed" trajectory started at the BCP transition state. The gray dotted line points out the chosen time step, at which the system is in the harmonic basin near its minimum, though it does not pass directly over it. If the discrepancy in energies is too large, the trajectory needs to be rejected. Figure and caption reproduced from Paper 1.

Then, we need a criterion to pick a final step of the trajectory, whose atomic coordinates and velocities would be used for the projection. Since we already have the electronic energy of the TS relative to the product of interest, we may simply perform the projection at each step and pick the one with the least error. I chose to minimize the sum of both the total projected and actual electronic energy (red and blue on Figure 2.2), as a lower value of both suggests a more accurate projection point, while minimizing their difference may give us steps where they cross on accident. In this case, all the simulations ran for 50 fs, so I would be able

to explore whether the trajectories will come closer to the exact local minimum. From my experience with this model reaction, the second or third "bounce" did not yield better results, so it would likely be efficient to stop the simulation when the sum reaches a predetermined value – or when it starts to increase again, as opposed to more arbitrary stopping criteria such as bond lengths or simulation time.

The full decomposition details may be found in the Supporting Information of Paper 1. How these choices affected results will become apparent in the next chapter, but in summary, it seems that with a robust choice of metric, a single trajectory is enough to capture the relevant information.

3

Predicting nonstatistical effects

3.1 A STRUCTURE-BASED INTRAMOLECULAR VIBRATIONAL ENERGY REDISTRIBUTION MODEL

Once we have established where the chemical activation is located in terms of normal modes, we require a VER model to predict its time evolution. The model I use originated from phonon scattering theory but has since evolved to model VER in molecules, where the quantized excitations of vibrational modes are sometimes referred to as vibrons. The model's success in predicting relaxation lifetimes of vibrational modes extends beyond small molecules to complex systems, such as peptides and proteins. [22,23] Specifically, it is an IVR model, while the decay to the solvent environment is included as an empirical constant, in this case equal for all modes.

In Figure 3.1 I show one-half of the model equations, those that describe the dominant decay pathway. Without going into the details of the model described in Paper 1 and its references, there are two conditions to be met for energy to flow. First, the frequencies ω of the vibron which decays into two smaller ones, or in turn, the two smaller vibrons that collide to

populate a higher energy mode, must be in or near resonance (yellow). Secondly, there needs to be significant coupling between all three of these modes, as defined by the cubic anharmonic coupling Φ (red). Both symmetry rules and the physical distance between the atoms participating in the different modes will be reflected in this coupling, leading to the nonstatistical nature of IVR in molecules. The contributions to the total relaxation rate of a mode (Γ, in blue) are essentially Lorentzian curves, and since the decay pathway only relies on the population of the decaying mode it is generally more significant than the collision process, which requires two modes to be populated.

$$\boxed{\Gamma_\alpha^{\text{decay}}} = \frac{\hbar}{16\omega_\alpha} \sum_{\beta,\gamma} \frac{\boxed{|\Phi_{\alpha\beta\gamma}|^2}(1 + n_\beta + n_\gamma)}{\omega_\beta\omega_\gamma} \times$$

$$\frac{\boxed{(\Gamma_\alpha + \Gamma_\beta + \Gamma_\gamma)}}{(\omega_\alpha - \omega_\beta - \omega_\gamma)^2 + \frac{1}{4}\boxed{(\Gamma_\alpha + \Gamma_\beta + \Gamma_\gamma)^2}}$$
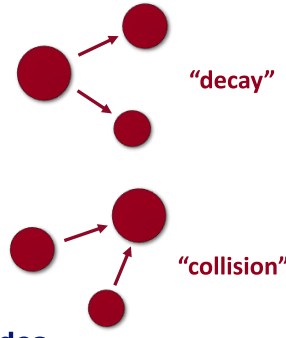
Figure 3.1: The equation that describes the total relaxation of mode α through the decay process (left), followed by a diagrammatic representation of the two sets of equations (right), the decay and collision processes. Constant decay into the solvent is added explicitly to the master equation, but not included in the linewidth (blue) at this stage, while the frequencies ω are ideally computed anharmonically and using an implicit solvent model.

In the full model, the width of this Lorentzian is defined by the sum of relaxation rates Γ of all modes involved, making it a self-consistent set of equations. Additionally, the anharmonic frequencies of the modes would shift based on their relaxation rate, resulting in modes moving in and out of resonances of differing widths until a steady state is achieved.[24] Solving this set of equations would be difficult for large systems, especially as I am now applying it to reacting molecule, as opposed to individual excited modes; this implies the Bose-Einstein mode populations n of many modes will also significantly vary throughout the simulation and the set of equations must be solved more than once. Previous use of the model suggests that using a fixed total linewidth between 16 to 48 cm^{-1} already provides a reasonable estimate of the relaxation rates.

Since applying this IVR model is comparatively cheap to the rest of the procedure, several linewidths may be used to compare how they affect the final results. Our electronic structure

method of choice will likely exhibit errors in the computed anharmonic (or scaled harmonic) vibrational frequencies at the same order of magnitude as the chosen linewidth. Therefore, if we expect a higher error, we need to lower our selectivity by increasing the linewidth used. When our confidence in the exact frequency values is lower, we are allowing more flexibility to include combinations with strong coupling, while also toning down the effect of accidental resonances. In the end, strongly localized IVR causing nonstatistical effects is likely to be significant enough even when caught off-center – we just need to be careful our chosen width includes it.

As the user chooses the linewidth Γ, we only require the vibrational mode frequencies and couplings, which both result from a GVPT2 (generalized vibrational perturbation theory to the second order) calculation, a methodology present in many quantum chemistry packages and used to compute accurate infrared spectra.[25,26] Besides spectroscopy, the efficient computation of these couplings has also been recently investigated for their application in statistical rate theories, such as SCTST.[13]

In this work, I utilized couplings computed at the exact same level of theory, but it is reasonable to expect that a hybrid approach where we combine more accurate and costly frequencies with cheaper couplings would yield similar results. A mismatched resonance decides whether VER occurs at all or not, while the exact coupling value only scales the rate while being dominated by the symmetry of the modes; a property that is preserved despite eg. the size of the basis set used. The harmonic frequencies obtained at a higher level of theory could then be scaled down to match experiment based on benchmarked data (one example being the the NIST database).[18] Otherwise, for the same reason described above, we may also aim to obtain anharmonic frequencies by applying GVPT2 corrections using couplings computed at the more affordable level of theory. This is already available in the software package iGVPT2, which also offers the possibility of QM/MM treatment for very large systems.[27]

The end result, when the decay and collision contributions are summed up, is a matrix of VER rates between the modes, as visualized in Figure 3.2. This matrix is then used in the master equation to propagate the estimated vibrational energy from the previous step. For the final results of the paper, I settled on updating the matrix every step until reaching 1 ps, after which it was updated every 20 steps; the step size being 0.01 fs. In practice, it is likely that less updates and larger step sizes would provide the same qualitative results. However, we should be careful to remember that the master equation loses any resemblance to molecular dynamics and traditional AIMD step sizes of 0.5 to 1 fs would lead to far too much numerical
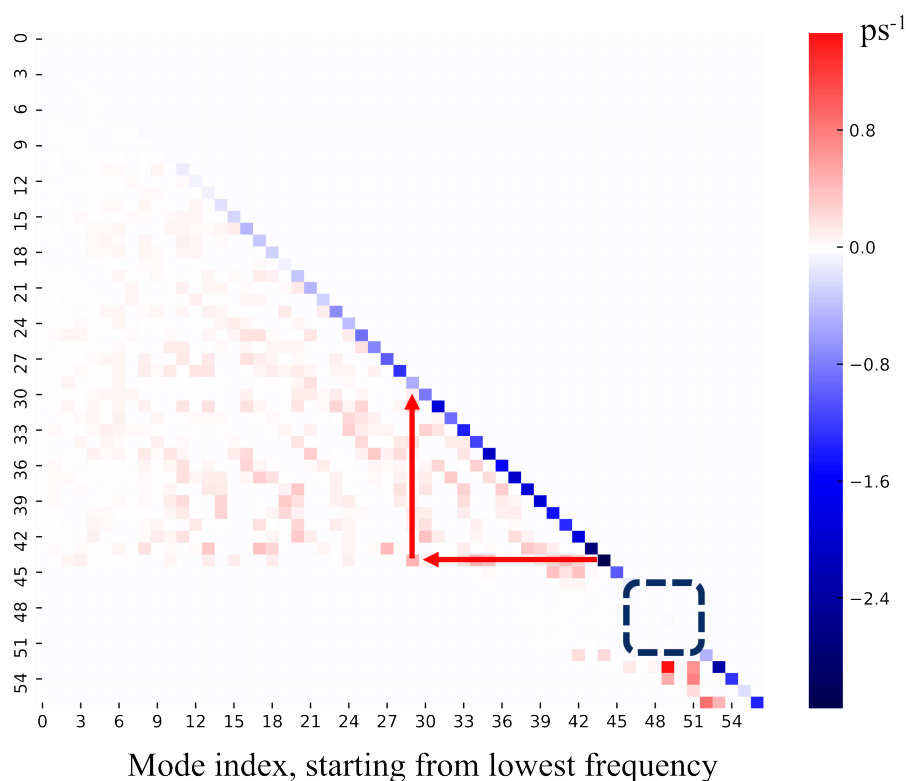
noise.

Figure 3.2: An example transfer rate matrix (labeled k_{mat} in Paper 1) of one of my benchmark systems, isopropylbenzene, using only the decay process described in Figure 3.1. The red arrows show one of the stronger decay pathways from mode 44 into mode 29, an intermediate frequency mode with a lower decay rate. The dark blue box shows a set of isolated C-H stretching modes of the isopropyl substituent, which do not couple well to the rest of the molecule – energy deposited into these modes will stay localized for a long time in the simulation, as it can only relax into the solvent. On closer inspection, most of the C-H stretching mode form a self-contained tier.

3.2 Single-number indicators for reaction design

Finally, I wish to collapse all the obtained information into a single-valued metric or indicator that could be used as a measure of nonstatistical effects. We are now presented with several choices; the main one being whether or not to perform the time propagation with the master equation. Even though in my use case the cost of the VER model is negligible we may imagine a scenario where frequencies and couplings are much cheaper, for example from a semiempirical or machine learning potential. On the other hand, we need to decide if we will use the data generated from the entire molecule or just the part that is involved in the nonstatistical

reaction. Both the estimate of the activation and the VER model will ensure we are looking at vibrational dynamics unique to this reaction, but it is also possible that not all of it is relevant information.

The first indicator I propose is to simply weigh the quanta of the initial estimated vibrational energy with the computed relaxation lifetime of each mode, as obtained by inverting the relaxation rate; shown in Eq. 3.1.

$$Indicator = \frac{\sum_i \frac{E_i}{\omega_i} \cdot \frac{1}{\Gamma_i}}{\sum_i \frac{E_i}{\omega_i}} \qquad (3.1)$$

We can compute this value using all the modes of the molecule or a given set, based on their frequency or locality on the molecule. In Paper 1 I chose a set of modes belonging to the region formed by the CH_2 group as well as its neighboring two carbon atoms, labeled as **3C**, As this set of modes was determined by projecting the normal mode vectors onto these atoms, we may go a step further and use values from all the modes of the molecule, but weighted by their participation in **3C**, instead of just fully including those that cross a certain threshold. We could use projections of the vibrational modes onto the reaction coordinate and define (or even weigh) modes of interest in this way, but this raises the concerns from the previous chapter on whether the reaction coordinate(s) of the next reaction step is well described at the geometry of the reactive intermediate.

As can be seen in Figure 3.3, I find that the regional results of this indicator are actually quite similar to the indicator computed with all the modes. This is especially true in the most comprehensive comparisons shown in Fig 3.3a, which qualitatively matches Figure 6 in the main text of Paper 1, the only difference being the use of harmonic versus anharmonic mode frequencies. This value tells us how well the localized energy begins to relax, but not the full picture as no time propagation is actually performed. It follows it would be more correct for very fast reactions that occur at the timescale of the first "tier" of IVR – relaxation from those modes that are excited into their direct resonances.

In the second approach, the master equation is applied and a lifetime is obtained as the time when the total energy in **3C** reaches $\frac{1}{e}$ of its original value. Of course, as some of the modes now have energy flowing into them before they decay into other modes and the solvent, their relaxation curves are no longer exponential plots. This is especially visible in both the lifetime value and the plot for the CH2Cl substituent (Fig. 8f of Paper 1). The choice of $\frac{1}{e}$ becomes slightly ambiguous here but is still useful as the long-term relaxation is expected to follow the
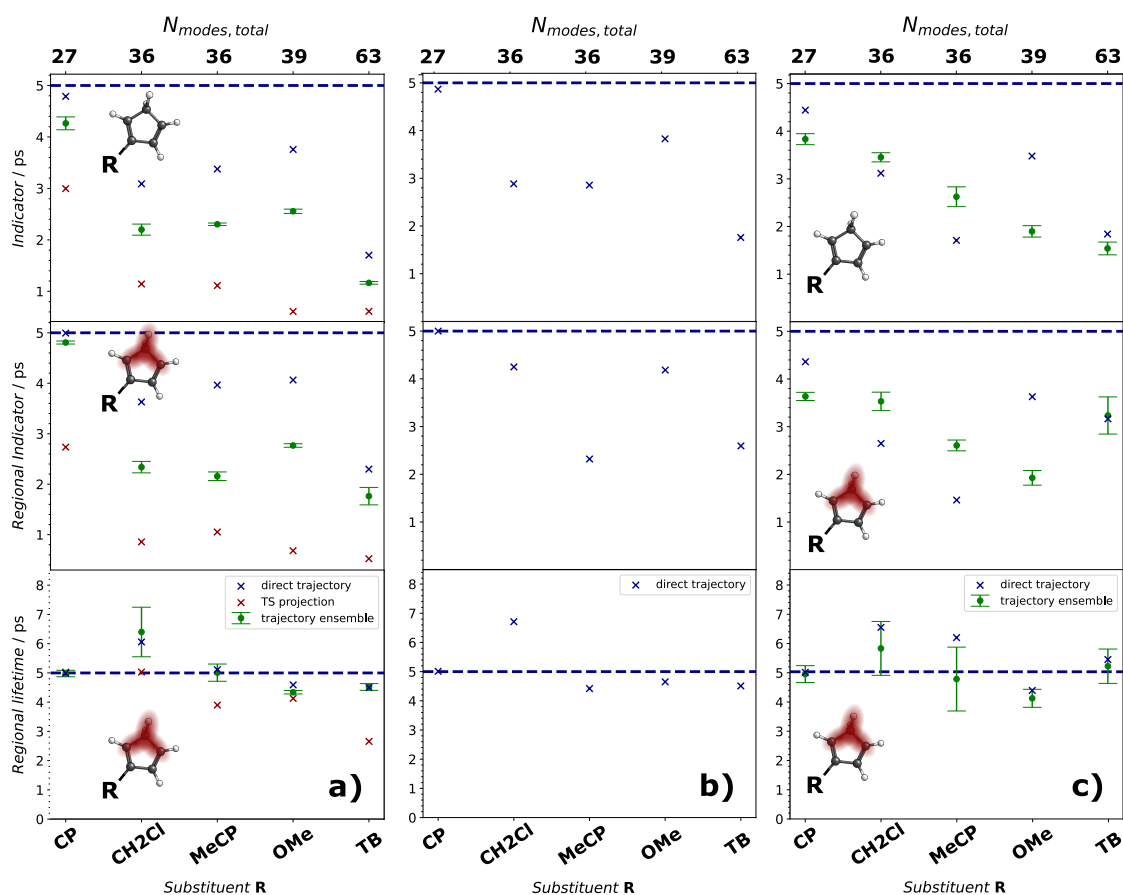
Figure 3.3: Computed indicator, regional indicator and regional lifetime values for the five differently substituted systems using a) harmonic CAM-B3LYP/def2-TZVP D3BJ frequencies with no GVPT2 correction b) harmonic frequencies at the same DFT level of theory using implicit solvent modeling for tetrahydrofuran and c) GVPT2 corrected frequencies at the MP2/cc-pVTZ (for CP–OMe) and MP2/cc-pVDZ (TB) level of theory. Figure and caption reproduced from the SI of Paper 1.

exponential decay to solvent.

As discussed in the paper, the trends of the indicator and lifetime values may begin to match for large molecules where most of the energy is relaxing away from the reaction center. Currently, I see no reason not to perform the time propagation as it can identify modes that stay populated on long timescales due to a lack of IVR pathways leading away from them and other pathways depositing energy into them. Without going into depth about the discussion presented in Paper 1 and its SI, both the use of anharmonic frequencies and an implicit solvent model yield similar results.

There is a discrepancy between the single trajectory and ensemble results, which is expected

as the ensemble includes ZPE, but their trends are very well reproduced by the direct, "0K pushed" single trajectory. There are only two cases where the single trajectory result behaves as an outlier: CH_2Cl with anharmonic frequencies and MeCP with the implicit solvent model, both for the regional indicators. For me, this only encourages the use of the master equation and the lifetime (or other metrics) obtained from it.

Additionally, the relatively small standard deviations of the ensemble results suggest that 50 trajectories were already more than enough to obtain a reasonable mean value. The standard deviations also do not seem to correlate with how different their mean will be from the direct trajectory result, nor does it strictly increase with the number of degrees of freedom. In essence, if the available resources allow it, the most careful approach would be to use a small number of quasiclassical AIMD trajectories with short timesteps. In case large standard deviations make it difficult to compare similar systems, the results can always be updated by running some additional simulations.

Fig 3.3c was used to gauge the effect of method mixing, as I use DFT dynamics to get to the harmonic region, as with a) and b), but use MP2 (Møller–Plesset perturbation theory of the second order) GVPT2 results for the IVR model. This in turn also requires the DFT trajectory to be projected onto MP2 modes. Standard deviations of the ensemble results are larger, and while the indicator trends are very different from the purely DFT results, the regional lifetimes remain better conserved across methods.

To conclude, I propose performing the IVR model's master equation, but a large AIMD ensemble should not be necessary. In competing reactions (attempting to reproduce A/B ratios), it might also make sense to combine the master equation result directly with a TST or mode projection one; attempting to produce a mixed indicator. In general, TST may predict more selectivity than experiment in cases where the energy of a hot intermediate lies well above two different barriers, but also less selectivity in cases where the system is dynamically matched to cross one of two similar barriers. A measure of vibrational relaxation such as the regional lifetime should point towards a structure more likely to match statistical predictions, which is something I hope to explore in the future.

4

Nonstatistical reactivity in the real world

In the second part of this PhD project, I received the opportunity to work directly with a postdoc in synthetic organic chemistry, employed on the same work package. As the systems discussed in previous chapters were small and difficult to functionalize, we set out to find solvent-based nonstatistical reactions that are both relevant to the development of current organic chemistry and allow for structural modifications that would not affect the reaction mechanism. From my side, an additional motivation behind this chapter was to push the limits in terms of system size and determine the scope of computational chemistry methods available to treat this complex reactivity. An important part was determining how these methods could be used to plan for (or adjust) lab experiments, as opposed to the common route of justifying experimental results *a posteriori*. For this reason, Paper 2 is written in the form of a tutorial that complements our new findings on the reaction mechanism with the procedure and reasoning of the computational and experimental work performed continuously and in parallel. This chapter will give a brief overview of Paper 2 and discuss some details that were too specific to be included in the final manuscript.

As the manuscript itself introduces this class of reactions, I will only reiterate key points and reproduce the final computed MEP and relevant structures with Fig 4.1. In this two-step reaction, in the first step, a ring is formed from a double allene structure, releasing a considerable amount of energy and preparing a hot intermediate. In the second step, a hydrogen transfer occurs that forms a double bond in two different ways, resulting in either an E or Z product. Since the TST product ratio varies from the experimental one, which also exhibits very strict temperature independence, the reaction was classified as a nonstatistical one. Furthermore, AIMD that the authors performed on a very simplified version of this structure resulted in a very fast (<50 fs) hydrogen transfer when started at the first barrier.[28]
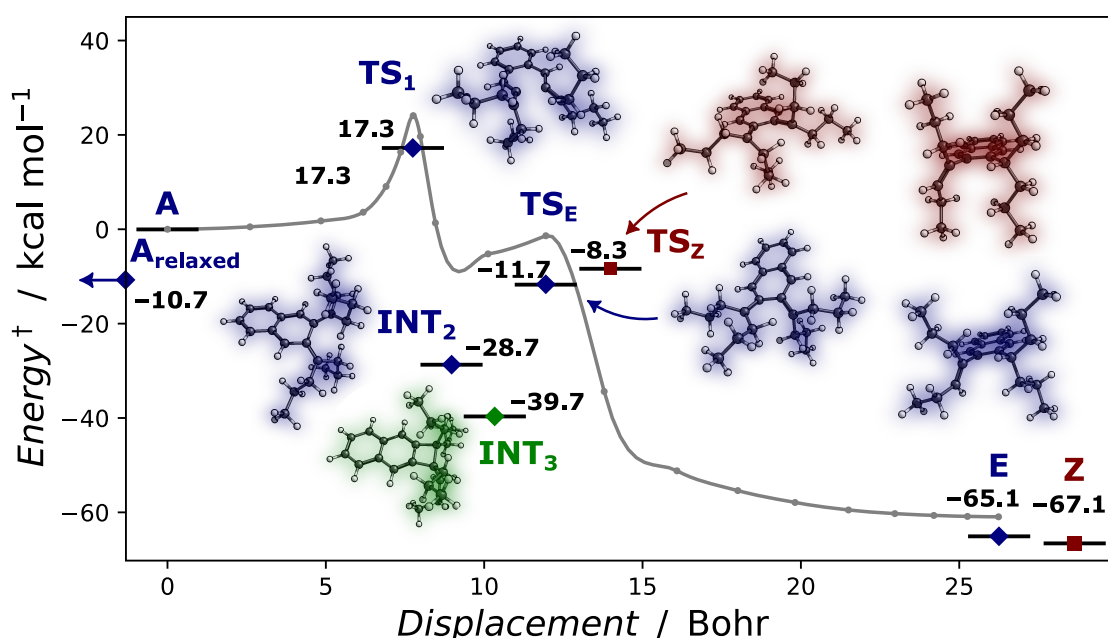
Figure 4.1: Calculated free energies of relevant geometries (blue) overlaid on a full BS-DFT NEB minimal energy reaction path (gray) from the initial structure **R** to the **E** product, confirming a reaction path with two steps passing over the two-ring intermediate **INT**$_2$. For comparison, we include the Z product and its transition state (red). In our exploration, the lower energy intermediate **INT**$_3$ (green) shares the same hydrogen transfer transition states **TS**$_Z$ and **TS**$_E$. †:Note that the reaction path in the background does not include a free energy correction. Figure and caption reproduced from Paper 2.

Our initial goal was to gauge how the nonstatistical E/Z product ratio changes as we make modifications to the structure that preserve the mechanism but add or remove relaxation pathways from the vibrationally hot region. The simplest change was to extend the length of the

alkyl chains that formed the "legs" of the molecule from propyl to butyl chains. This simple change adds 16 atoms or 48 degrees of freedom. We already know from the discussion in Paper 1 that such a large increase in total vibrational modes might not make the reaction less nonstatistical in case the alkyl chains are bad heatsinks for the newly formed aromatic ring. I will note that it was possible to complete a GVPT2 calculation on the propyl system with the CAM-B3LYP functional and def2-SVP basis set, but not the significantly larger def2-TZVP. With the recent improvements in efficiency and the upgraded VPT2 module in ORCA6, larger basis sets are likely available.

However, the experimental product ratio for the butyl system surprised us, as it was actually less selective: dropping from around 10:1 to 6:1. The initial hypothesis was that if the statistical selectivity between the two transition states stayed consistent, the larger structure would either conserve the propyl experimental ratio if the added degrees of freedom were not consequental or the ratio would increase, as higher VER recovers the selectivity predicted by TST. This prompted me to re-model the reaction; but at the same time I started my calculations, we also observed a change in kinetics in the lab depending on how much light the reaction mixture was exposed to. We will briefly discuss this in the next subsection, but first focus on the electronic ground state.

As an example of increasingly available tools, I have started to rely on nudged elastic band (NEB) methods to robustly obtain minimal energy reaction paths (MEPs). These are not necessarily cheap as they may need many cycles to converge while requiring force calculations for each NEB image (point on the reaction path). But as discussed in the SI of the paper, shortcuts can be made to make them more affordable, and in this case, they even led to the discovery of a separate and lower energy intermediate, INT_3. This intermediate, in which the diradical is resolved by closing a third ring, was not identified in experiment, likely as the formation of the structurally closer INT_2 was better dynamically matched. Only after this computational discovery did we manage to find much older literature where different diallene structures result in an additional ring formation.

Regarding the electronic structure of choice, it is often said that the zoo of DFT functionals leaves people confused – an added zoo of DFT benchmarks has not necessarily made the situation much clearer. It is fair to say that in many cases good results can already be obtained by DFT functionals without Hartree-Fock exchange energy, which tend to be very fast when used with the resolution-of-identity (RI) approximation and nearly-free empirical dispersion corrections. A refinement at the hybrid or even double-hybrid level may then follow. [29,30]

The same logic follows for basis sets – though def2-SVP is the smallest of the Ahlrichs basis set family, many may not realize it is comparable in size to the outdated Pople type 6-31G(d), still regarded by many as an ideal and comparable option. If in some cases even those two basis sets are prohibitively large, it might be wise to start off with a semi-empirical method such as tight binding GFN2-xTB, which among many other use cases may be used to guess the initial path for the NEB methods in ORCA. Exceptions of course exist, such as conjugated systems requiring range-separated DFT functionals, anions and electron-rich systems requiring diffuse functions in the basis set, or in general tougher electronic structures such as the singlet diradicals encountered in this thesis, for which we should at least use BS-DFT. However, the raw amount of papers using these methods is massive, so simply explaining your chemical system to a large language model might already produce some useful warnings about which methods could be inadequate.

Besides these tutorial-minded points, my TST rates including Eckart tunneling and DFT free energies with a larger basis set actually hit the experimental ratio spot-on for the propyl system. TST with Eckart tunneling even predicts a small temperature range where the ratio would be constant, as tunneling contributions taper off relative to the base TST rate. While it might be unlikely that a more accurately computed tunneling contribution would be significant enough to produce a temperature-independence at such a wide range and up to 160°C , it does raise the question of how nonstatistical a reaction has to be to warrant the label. Furthermore, if there is an evolution of TST available that can model the reaction rates and ratios – but it is very costly and is not enough of a black-box approach to be used for automated reaction discovery or modeling in a general organic chemistry lab, does this reaction remain nonstatistical for all intent and purpose?

4.2 INTERPLAY OF LIGHT AND HEAT

During our discussions on the coupling of lab and computational discovery, we came to the conclusion that light reactivity is one of the hardest things to predict for a synthetic chemist – but from our side, theoretical photochemistry has become quite precise and affordable, and often inspired by advances in computational spectroscopy, where accuracy is key.

Perhaps surprisingly, a computed spectrum is in many cases more useful than a measured one, including with the reaction in Paper 2, as the compound already reacts while the UV-VIS spectrum is being taken. It is possible to determine whether this is the case by switching the

direction the instrument is scanning, from high wavelengths to low and vice versa, and then compare whether there is a noticeable difference. A computed spectrum then helps determine which reactants, intermediates, and product make up the rapid changes in the measured spectrum.

An unnoticed photochemical reaction may also add difficult confusion to any mechanistic exploration, as molecules coming down from an excited electronic state can end up participating in the same reaction steps as ground state thermal reactions, but at the same time contribute their own specific dynamics. Even if no additional mechanisms of selectivity are introduced on the excited surface, for example, due to a lack of barriers or a close correspondence to the curvature of the ground state, the very act of crossing through a conical intersection to return to the ground state may imbue the system with a certain vibrational excitation – the same way we discuss reaction barriers as being filters for dynamical features. This analogy was discussed in great detail last year by Feng et al in Ref. 31. When we talk about hot intermediates or ballistic trajectories being "above" the PES, the "above" may well result from electronic excitation and not just a high ground state barrier.

In a way, we were lucky that with this reaction, the dark and light experimental conditions produced different kinetics but identical (or near identical) ratios. While I cannot really speculate without obtaining a conical intersection geometry, it is likely located near the first barrier and as such results in a similar enough dynamical signature (and excess energy). After all, it is just a matter of how we will "pay" to disturb the allene π system(s), after which the second ring is bound to form.

In a more general case, it would not be out of the question that the reaction proceeds all the way to the final product(s) while on an excited electronic surface, including the possibility of crossing to the triplet surface. Since we have also performed multi-reference calculations, this is something we will be able to comment on in the final version of the manuscript. These calculations are costly and user-intensive, and effectively brings us back to the energy diagram framework, considering only key geometries. Much can be missed without modeling the nonadiabatic dynamics of the system.

5

Efficient quantum vibrational dynamics with TD-vDMRG

When I was planning my change of scientific environment at the group of Professor Markus Reiher, my initial intention was to use their vibrational density matrix renormalization group (vDMRG) method to obtain more accurate anharmonic frequencies and the resulting mode resonances. In vDMRG the vibrational wave function is parametrized in the form of a matrix product state (MPS), which in turn allows the full configuration interaction (FCI) expansion coefficients to be decomposed as products of matrices with a chosen maximum dimension (bond dimension). This process is analogous to singular value decomposition (SVD) or principal component analysis (PCA) and results in a very efficient reduction from exponential to polynomial scaling. [32,33]

Additionally, in the canonical form of the second-quantized Watson Hamiltonian, which is applied here, the reference coordinate system is Cartesian normal coordinates. This means each site of the resulting DMRG lattice corresponds to a vibrational mode, and I may define the PES for this calculation directly with the cubic and quartic anharmonic couplings I al-

ready have from GVPT2 calculations. With more confidence in the computed anharmonic frequencies, I would be able to use stricter criteria for locating resonances in the IVR model and in that way obtain overall more accurate predictions.

Despite the approach being used for the rather large SarGly$^+$ dipeptide, I learned that it would be difficult to automatically converge the frequency of each vibrational mode for many molecules at once, especially those of practical size, with a double to triple-digit number of degrees of freedom. However, at the same time, the Reiher Group was also working on the time evolution of Hamiltonians in this matrix product operator (MPO) form, allowing for large-scale quantum dynamics with the time-dependent (TD-DMRG) method. While this TD-DMRG method was published with examples of exciton dynamics and vibronic spectra, the same approach may be used with the purely vibrational Hamiltonian.[34]

This effectively provides me with a fully quantum IVR model that is both in the same vibrational normal mode framework and relies on the same electronic structure results, couplings from GVPT2. As nearly the whole cost of the IVR model in Chapter 3 is in obtaining the couplings, this enables us to do both at the same time if desired, with quantum dynamics for the short-term and the resonance-based master equation for long-term propagation. The approach also fits with the tiered nature of VER, but also with experimental determination of energy localization, which struggles to determine the nature of very fast VER/IVR. Take for instance para-difluorobenzenes, where experimental vibrational relaxation data is fitted to two-time series or two "tiers" of IVR – a short one faster than one picosecond, and another, longterm process. In the longterm process, fitted lifetimes vary between 7 and 23 picoseconds based on the structure (symmetry), despite ortho, meta, and para-substituted difluorobenzes having the same number of degrees of freedom.[35]

It is also possible to go a step further than defining the PES using GVPT2 couplings and build a custom *ab initio* PES using an *n*-mode expansion, more suitable for difficult and very anharmonic cases.[33] This formalism provides a fully flexible and intrinsically anharmonic modal basis for the vDMRG method and the code is soon expected to support it for the TD-vDMRG method as well.

Other than the MPS compression to a certain bond size, it is important to mention that in the time-dependent simulation, we also pick the maximum excitation degree (n_{max}) any mode may have, thus limiting the basis for each site. Using a higher number also warrants a better quality PES, as higher excitations may spread out further. For instance, with a PES defined by only GVPT2 couplings, using a value of n_{max} larger than 6 would not even be plausible. I chose
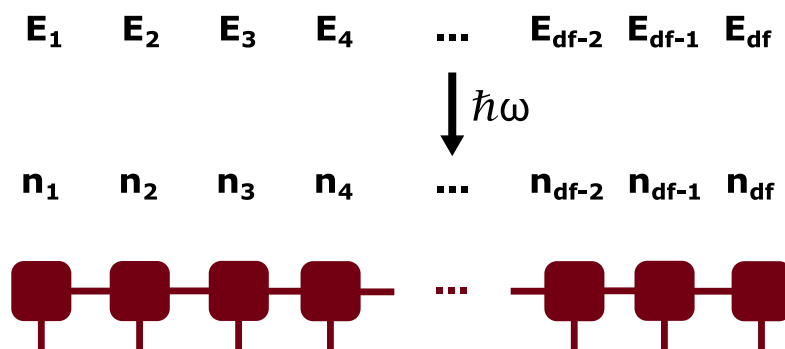
Figure 5.1: E_{kin} from velocities and E_{pot} from geometry displacements from a preceding classical simulation are projected to a normal mode basis, from which occupation numbers of each mode (n_i may be estimated and subsequently mapped onto sites on the vDMRG lattice (red).

a conservative value of 4, as I expect most IVR to occur through relatively simple resonances and as we are unlikely to estimate a higher vibrational state in our mode projection scheme. It is still possible that some unlikely IVR pathways proceed through even higher vibrational steps, though this approach is still a massive step up in completeness from classical dynamics or a Golden Rule model.

As proof of concept, Figure 5.1 shows a schematic view of how a short classical simulation from the top of a reaction barrier could slot into a TD-vDMRG simulation at the harmonic basin of the hot intermediate. In this manner, it could be used in parallel with a longterm IVR model to determine the behavior at short timescales, where quantum effects such as entanglement or interference may be more relevant, before further thermalization and decoherence occur. The initial matrix product state for the time-dependent simulation would then be a coherent superposition of the vibrational states we estimate are excited by the previous reaction step. The motivation behind this chapter is therefore to show some preliminary benchmarks on the challenges and costs involved in using this method.

For my benchmark calculations, I chose the fundamental excitation of the highest frequency C-H stretch of para-difluorobenzene (initial vibrational state |100...0>). The PES was obtained from an MP2/cc-pVTZ GVPT2 calculation. During the simulation, the population/excitation of each site as a function of time can be extracted from the average of the number operator. This process, termed here as a measurement of the sites, is not negligible in cost, so it makes sense not to perform it at each propagation step. In this case, I chose to measure the population or degree of excitation every 5 femtoseconds, as this provided enough

resolution to spot details in the vibrational dynamics of high-frequency modes while minimizing overhead on the TD-vDMRG calculation. This measurement can also be parallelized, resulting in a modest speed-up when the number of sites (modes) is at or near a factor of the available CPUs.

When using the small bond sizes (<100) that have been recommended for converging vibrational state energies with vDMRG, the cost of the time propagation is comparable, or up to an order of magnitude higher, to the cost of performing *ab initio* molecular dynamics. From my limited experience with quantum dynamics, this is a significant advance, assuming the full CI wavefunction stays well-compressed during the time evolution. However, during my initial benchmarking of the several difluorobenzene molecules, I noticed the total energy was not being conserved, with the energy of each sweep decreasing, as well as an accompanying loss of population. By several picoseconds of simulation, the energy and excitation degree return back to the vibrational ground state.

This drop always occurs between 1 and 2 picoseconds, depending on the parameters used, suggesting that as the vibrational dynamics advance, the system becomes ill-described – most likely as the required bond dimension that would be required to efficiently compress the evolved wavefunction rapidly increases. Another cause could be a deficiency in the PES as the system explores more of it over time, however, a comparison with a DFT PES with slightly different modes and ZPE still exhibited identical behavior.

When comparing the sweep energy across several bond dimensions and simulation timesteps in Figure 5.2, we see that both a smaller timestep and a larger bond dimension matter for energy conservation. Simply performing more sweeps (timesteps) increases cost linearly, unlike increasing bond dimension, so we might believe that we should reach longer timescales by prioritizing a smaller timestep. However, if we look at the actual measured excitation over time in Figure 5.3, we see that for a given, small bond dimension of 50, going from 0.50 to 0.10 fs does not qualitatively change the results. Furthermore, it seems that a stepsize of 0.25 fs actually produces a longer-lasting simulation before the excitations begin to rapidly decay. This may also be seen in the previous plot of energies if we notice that the inflection point or onset of the faster drop in sweep energy appears sooner for 0.10 fs, even if it drops slower overall.

Let us now explore what happens with much larger bond dimensions. In both Figure 5.2 and Figure 5.4 we see that increasing the bond size conserves the energy and excitation degree for a longer time. This has a direct effect on the vibrational redistribution as well, as the
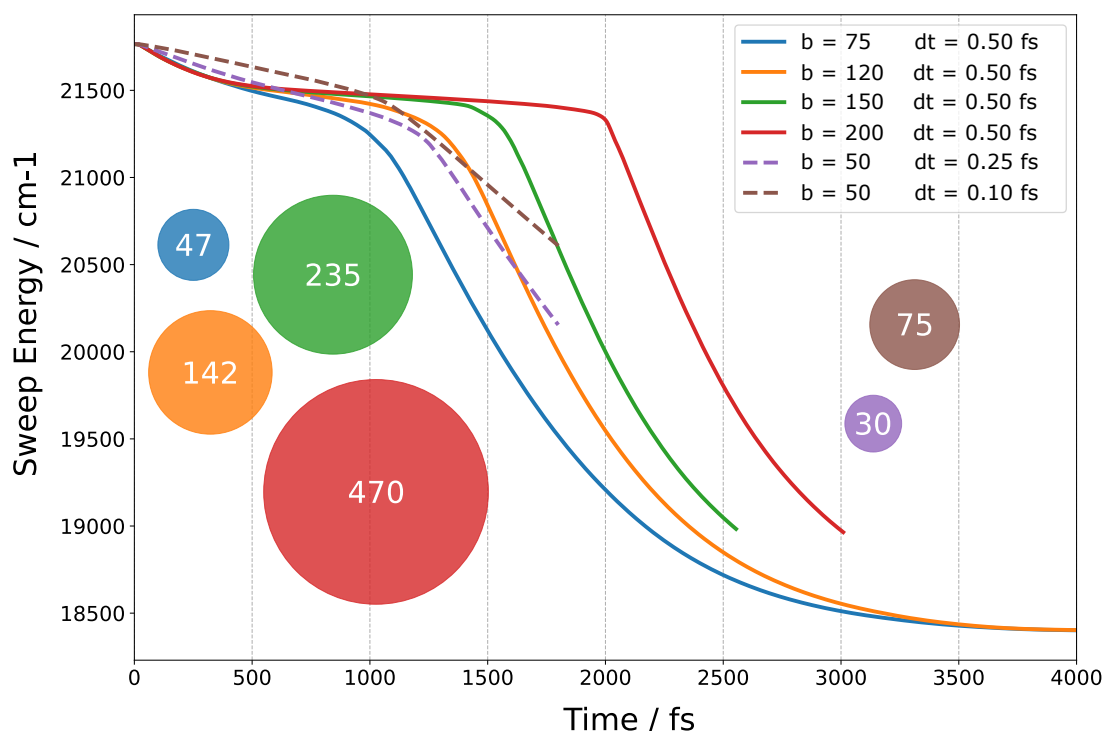
Figure 5.2: Sweep Energy during a TD-vDMRG calculation of the fundamental excitation of the highest frequency C-H stretching mode in para-difluorobenzene. Full lines for calculations with different bond dimensions and a 0.50 fs timestep, and dotted lines for a smaller bond dimension combined with a smaller timestep. The colored circles and their relative size represent the computational cost of each simulation: the number inside the circle is the time (in seconds) required for 0.50 fs of a simulation on 48 CPU cores of an Intel Xeon Gold 6248R compute node.

decay seems to occur faster as if it were "squeezed" by the impending inadequate compression. More structure is also seen in the decay of the excited Mode 30, though between bond dimensions 150 and 200, a fit to the plot would likely produce a very similar lifetime value – and the experiment is not accurate enough to discriminate between these results. According to preliminary results, a somewhat larger bond dimension is required for TD-vDMRG over vDMRG calculations, but the simulation of fast IVR in the 1−5 picosecond range is within reach. For condensed phase examples, such as solvent reactivity, any simulation attempting to reach 5 ps or further would also need to take into account vibrational energy decay into the environment. Further testing also needs to be performed on more complex, chemically realistic, coherent initial states involving several hot modes.

An added bonus of getting acquainted with DMRG-based methods is their direct connection to quantum information theory and algorithms, especially those being developed for

quantum computing for chemical systems. Tensor network techniques like DMRG have been driven for their efficient modeling of strongly correlated systems, similar to the initial focus of possible chemical applications in quantum computing. Whether or not the added completeness of this theory compared to more established approaches such as Fermi's golden rule models or (quasi)classical AIMD are required to guide experiments, it is a useful educational exercise. The concepts of vibrational anharmonicity, resonance, and directional energy relaxation are directly related to vibrational entanglement. This etanglement may exist between vibrational modes, but also between the different basis states of a single mode. This can be explored with vDMRG, as when time propagation is not required, no predefined excitation rank (n_{max}) needs to be defined allowing for a systematic convergence to the full vibrational configuration (VCI) limit. Furthermore, from these calculations, Fermi resonances can be directly identified and visualized through both their single-modal entropy and diagrams of mutual modal information.[36] We will also touch upon another avenue towards quantum dynamics in the final chapter.
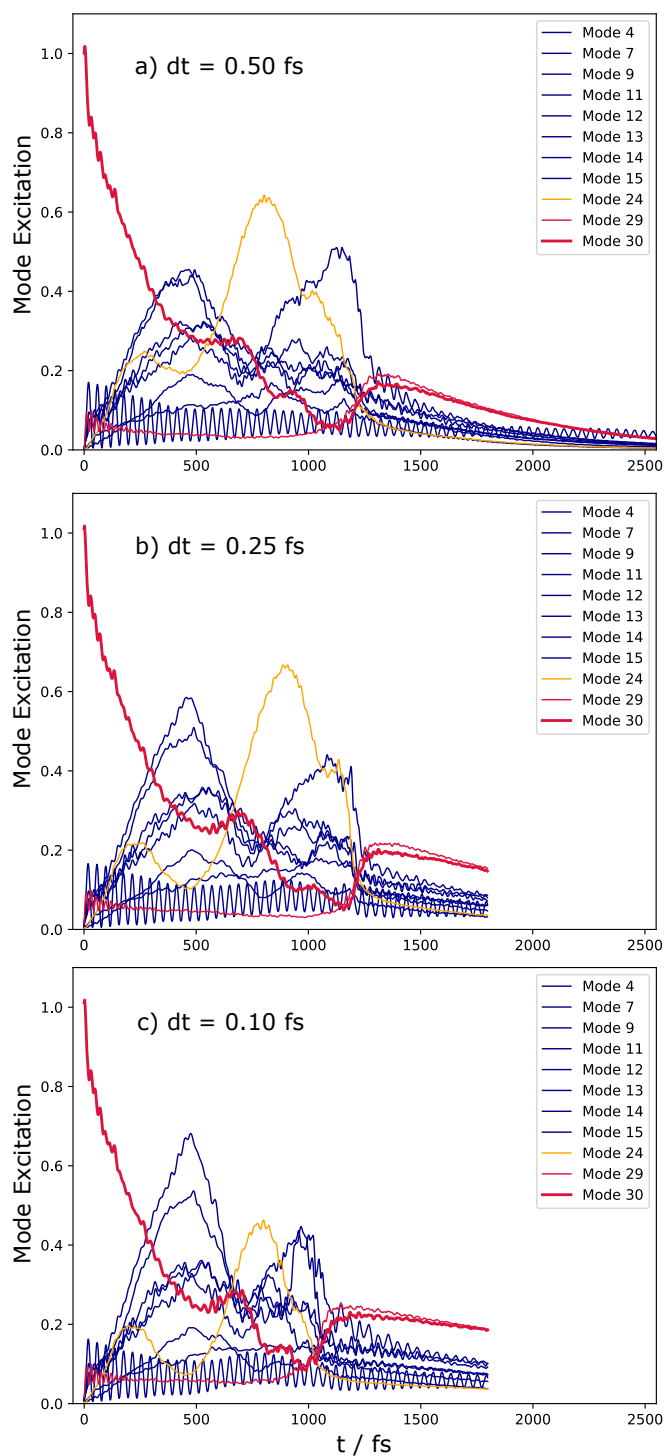
Figure 5.3: Measured mode excitation during a TD-vDMRG calculation after exciting the highest frequency C-H stretching mode in para-difluorobenzene (in bold red) with varying timestep size. Bond dimension is kept small and constant at b = 50. Only those modes whose excitation exceeds 0.15 during the first 1 ps are included. Red lines correspond to C-H stretches (>3000 cm^{-1}), yellow to intermediate modes (above 1000 cm^{-1}), and blue to the rest (below 1000 cm^{-1}).
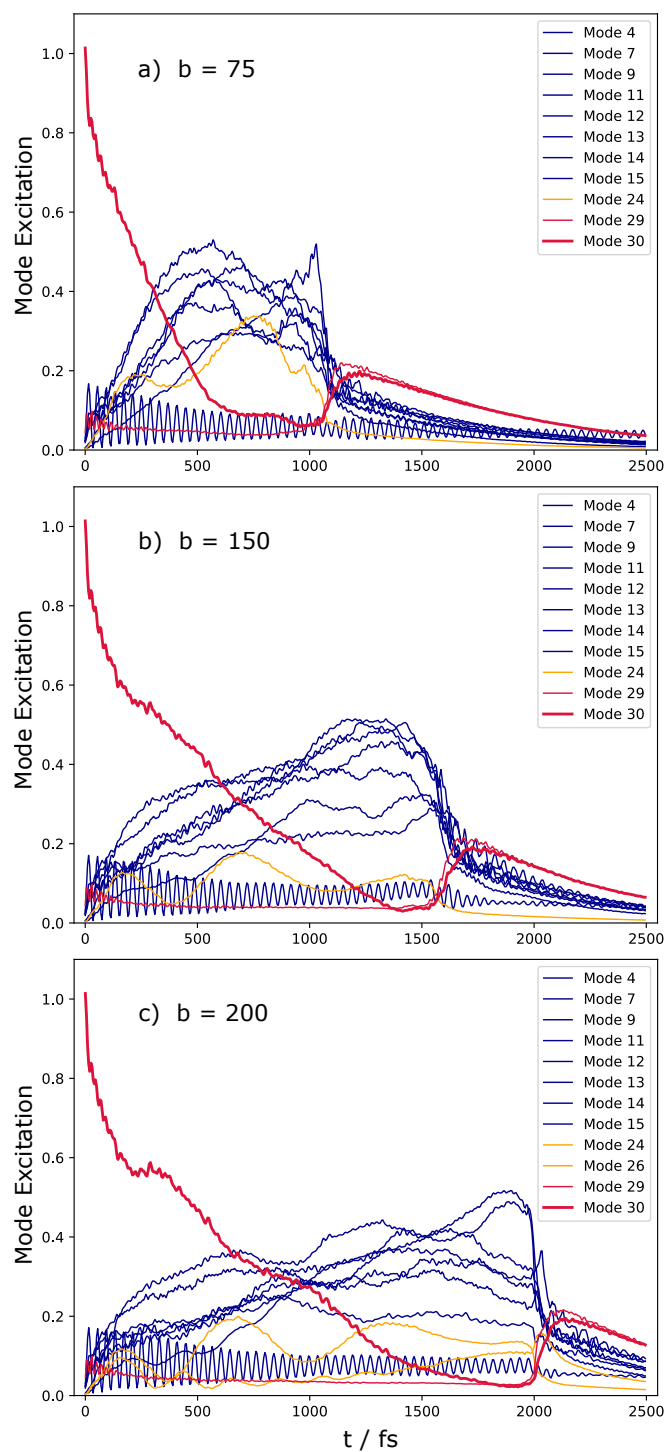
Figure 5.4: Measured mode excitation during a TD-vDMRG calculation after exciting the highest frequency C-H stretching mode in para-difluorobenzene (in bold red) with varying bond dimensions. The timestep is kept constant at dt = 0.5 fs. Only those modes whose excitation exceeds 0.15 during the first 1 ps are included. Red lines correspond to C-H stretches (>3000 cm^{-1}), yellow to intermediate modes (above 1000 cm^{-1}), and blue to the rest (below 1000 cm^{-1}).

6
Outlook

6.1 Quantum simulators and path integral quantum mechanics

In 2017 Markus Reiher and Matthias Troyer, who have both been in some way affiliated with quantum computing programs at Copenhagen University, wrote a perspective on the use of quantum computers in modeling reaction mechanisms.[37] Right at the start of their introduction, they generalize chemical kinetics to energy differences of minima and transition states and then proceed to focus on the problem of electron correlation. They argue that the most impactful use case for quantum computing would be computing the correlated energy of a system, while the remainder of the steps, such as kinetic modeling and entropic corrections, would still be done on a classic computer.

Though that might be the case, it is also an example of how reactivity that does not conform to Eyring rate theory might be easily disregarded, given a traditional textbook introduction to even cutting-edge projects. As such, nonstatistical dynamics would lag behind unless schemes that profit from single-point energies computed at extreme accuracy are developed. At the same time, very different approaches are being presented for dynamical processes. In this case,

we step away from qubits to analog quantum simulators, built up from bosonic modes and oscillators, from which we may draw parallels to molecular vibrations represented as a set of vibrons.

From my own experience in the past three years, papers that discussed quantum advantage in dynamics focused on systems of harmonic oscillators, one example being the 2018 Nature paper by Sparrow et al.[38] If adding anharmonic coupling becomes simpler to implement, we could also benefit from the inherent locality of vibrational dynamics by only including those strong couplings as determined by a screening process; this screen might as well be performed with a classical computer. As mentioned in Chapter 5, the entanglement of vibrational modes in molecules could be determined through vibrational DMRG calculations.[36]

The inherent quantum nature of molecular vibrations also opens doors to other areas, now considered vital to quantum computing, such as information theory. In the recent paper by Zhang, Wolynes, and Gruebele, intramolecular vibrational energy redistribution is explored as quantum information scrambling in molecules.[39] Both quantum and classical dynamics are compared and their deviations due to quantum interference effects are discussed. Since this interference limits the number of participating states, scrambling is in general slower for the quantum case, especially in the intermediate and longer time regimes. They note that classical simulation may not be enough, even for molecules at room temperature and dozens of degrees of freedom. This hints at a layered problem for the theory of nonstatistical reactivity – as energy localization that can be modeled classically already poses an issue for modeling many reactions, a certain number might also suffer (or benefit) from quantum interference effects.

The path integral formalism of quantum mechanics offers us a way of performing approximative quantum dynamics "on the fly", with the flexibility of classical AIMD. This approach was termed Path Integral Molecular Dynamics (or PIMD) and sparked the creation of i-Pi, a Python 3 force engine.[40] This package serves as a home for many path integral methods (and other advanced algorithms) and its developers have made it clear, including in the communication at Ref. 40, that they are aware of how serious of a bottleneck implementation and ease-of-access is to state-of-the-art computational tools. One way to encourage the use of new methods is for the developer community to come together and organize workshops such as the CECAM Flagship School on Path Integral Quantum Mechanics I attended two years ago. Already then, simulating condensed phase molecular dynamics was fast, efficient, and, most importantly, easy to use through the simple modification of existing Jupyter Python notebooks.

For chemists, the most interesting methods based on this formalism are likely centroid molecular dynamics (CMD) and ring polymer molecular dynamics (RPMD). Both are based on theory that shows the quantum dynamics of a molecule may be described through the classical dynamics of a fictitious ring polymer composed of a certain number of beads (copies of the system) connected with harmonic springs – more beads allowing for a closer correspondence to quantum dynamics. The approaches differ on whether they focus on the centroid (average position of all the beads) or the dynamics of the whole ring polymer.

A 2016 Feature Article (Ref. 41) on RPMD chemical reaction rates explains how the method is able to take into account tunneling, recrossing, and anharmonicity explicitly while preserving zero point energy; it also points to the software package RPMDrate used to perform these rate calculations. The focus remains on small molecules where tunneling dominates, and their application on larger systems will need to be tested. However, the cost of the method scales linearly, as we simply require on-the-fly gradients computed for all the beads. At the moment this value seems to be between 16 and 200, with higher temperatures requiring fewer beads, effectively placing it one to two orders of cost above quasiclassical AIMD.[41] Additionally, the use of contraction schemes lowers or even fully eliminates this overhead.[42]

While I have not yet used this method for my research, I will share insights from the discussions I had at the workshop. In the case of RPMD, the frequencies of the oscillators connecting the fictitious beads may show up in the overall vibrational dynamics when spurious resonances appear with the molecular vibrations. For this reason, the simulations are usually heavily thermostatted, and in practice rate calculations might start with a thermalization/equilibration simulation before the rate constants or diffusion rates are computed. This is especially true for "thermostatted RPMD" or TRPMD in which a thermostat is also applied to the internal modes of the ring polymer.[43] On the other hand, CMD focuses only on the centroid molecule, making it more accurate for long-term dynamics from which infrared spectra may be obtained, inherently also making it more accurate as an IVR model. TRPMD might offer a middle-ground, but it remains an open question whether it is the best choice for polyatomic long-lived hot intermediates such as those involved in nonstatistical chemistry, including the trade-offs involved with the different thermostat settings. Finally, as (T)RPMD is well suited for short-time quantum dynamics effects, it would not include, for example, interference effects.[41]

As mentioned in Chapter 1, mode-specific chemistry has been an open field for many decades, but it has certainly not gone cold. For instance, the idea of IR lasers exciting specific modes has expanded to include electronic excitation in the UV/VIS range of the spectrum, targeting specific chromophores of the molecule and in turn, their vibrational modes. In 2019, Rather et al. [44] photoexcited a molybdenum complex in which two molybdenum atoms were bridged by a dinitrogen molecule (N_2). When they used ^{15}N-^{15}N, a stretching mode of the dinitrogen was lowered enough in frequency to become nearly resonant with one of the modes of the molybdenum chromophore. This resulted in better energy transfer to the nitrogen molecule, an interesting target for chemical activation, even if an isotopic one.

Using two-dimensional electronic spectroscopy they were able to argue the vibrational wavepacket including these two excited modes originates from the singlet electronic excitation, but also dephases on the same timescale as the intersystem crossing to the triplet state (around 2 ps), suggesting the vibrational wavepacket can survive the relaxation of a short-lived electronic state. They refer to this coupling as a vibronic type of Fermi resonance and propose using it to open channels of quantum energy flow.

At the time of writing this thesis, the field of vibrational polariton chemistry is exciting yet at the same time hotly debated among theorists. Radiation modes present inside optical cavities (in the infrared spectrum) are being used to alter the reactivity of a solvated system within the cavity. In the 2023 Nature Communication by Lindoy et al, the experiments are presented as still lacking a firm theoretical explanation. [45] To summarize, it seems the phenomenon stems from the coupling of the cavity mode to the vibrational modes of the reactant when they are near-resonant and not the barrier mode; that the chemical reactivity may both be suppressed and enhanced and is sensitive to a narrow frequency range (under 100 cm^{-1}), signaling sharp resonances. The effect needs to be modeled using a fully quantum model while taking great care of including both solvent-molecule interactions and cavity loss – dissipation of the cavity mode into the environment.

In another very recent article, Mondal and Keshavamurthy discuss the effects an optical cavity can have on IVR dynamics, and in turn chemical reactivity, within the regime of vibrational strong coupling (VSC). [46] In a way that very much matches the theme of this thesis, when a cavity-molecule resonance exists, the involved vibrational modes effectively add pathways of relaxation but also a bridge to redistribute energy between modes, depositing energy

into modes that might would have otherwise been more isolated. To directly cite Ref 46: "This highlights the point that VSC with modes other than the reactive mode can potentially increase the rate of the reaction provided that the network of anharmonic resonances connects to the reactive mode."

In the end, the use case of a cavity coupling to a small number of molecular vibrations might seem modest, but we can imagine pairing it with a strategy such as the one proposed in this thesis. When damping or bridging a certain set of modes would greatly affect reactivity, we can identify which resonances we would need to introduce. Some of them may be added (or removed) by sampling structural changes, but if no change is possible without disturbing the reaction center – a cavity could be tuned to exactly fill in the gap.

6.3 Force activated reactivity and mechanochemistry

According to the 2011 review "The many flavors of mechanochemistry and its plausible conceptual underpinnings" found at Ref.[47] "in mechanochemistry, the rate at which a material is deformed effects which and how many bonds break". This intentionally broad definition applies to many approaches to synthesis, ranging from grinding powders (with or without added solvent) in ball mills to processes involving ultrasonic waves as in sonochemistry. Sonochemistry, sometimes thought as distinct from traditional mechanochemistry, involves the use of a sonicator device and the phenomenon of acoustic cavitation.

The scale of the systems involved in mechanochemical reactions ranges from small molecules to large polymer structures, with the smallest unit that exhibits force activated reactivity being termed a "mechanophore". Trying to theoretically model how the bulk strain might present at a scale small enough to be treatable with quantum chemical methods is complex, but often boils down to a model composed of a fragment, for example, a short segment of a polymer chain, that contains the suspected mechanophore. This fragment is then strained (stretched) and a model as simple as Hooke's law for springs is used. Naturally, transition state theory is also utilized, for example, to estimate reactivity rates using the energy differences of the equilibrium and "force activated" states of the reactive site / mechanophore.[47]

Besides the fact TST is often applied, this whole approach is inherently statistical – it looks to best explain the bond most likely to break and in that way capture the lion's share of the experimentally observed reactivity. This is can be seen in the tension activated carbon-carbon

bond model, presented just last year.[48] In this model, two criteria are defined as relevant for a structure-reactivity relationship of mechanophores with C–C bond cleavage: the effective force constant of the bond, as computed through several stretched single point energies and with Hooke's law, and the force-free energy difference between the electronic ground state and its diradical intermediate. This gap represents the cost needed to break the bond and is computed with broken-symmetry DFT.

It is still an open question to what extent exactly nonstatistical effects appear in general solvent organic chemistry, or better to say what amount of deviation for statistical rates we would deem significant. However, I believe it is not a stretch to claim that mechanochemistry would be a breeding ground for hot reactants and intermediates for whom general VER into the environment would not be a viable way to thermalize in time. I find the collapsing cavitation bubbles in sonochemistry most exciting, as they are said to create short bursts of intense pressures and temperatures, which in turn may result in localized heating and the appearance of sonolytically generated radicals, both underexplored and unaccounted for contributions.[47]

The concepts of vibrational energy redistribution and vibrational state coupling are rarely mentioned in the works I've read, even when discussing the mechanisms of these reactions. A notable exception comes from the recent work performed by the groups of Professors Yan Xia and Todd J. Martinez, who discuss dynamic effects in the cascade unzipping of ladderanes, and have since discussed the nonstatistical dynamics for several similar polymer structures. The experimental findings in these works are supported by *ab initio* steered molecular dynamics, a form of quasiclassical, Wigner sampled AIMD where pulling forces are introduced to the mechanophore.

Professor Martinez and his group also develop the *ab initio* nanoreactor, a GPU-accelerated code package that performs AIMD on a massive scale, capable of reaching over a nanosecond of simulation with hundreds of atoms.[49] While still very costly, such AIMD reactors may serve as a sort of "Swiss knife" when the TST-based reaction discovery approach fails, especially when the capabilities of the reactor are expanded to include photochemical reactivity, as is the case with the non-adiabatic nanoreactor.[50] Going even further with the advent of GPU computing, an explosion in the availability of machine learning based methods has marked the last three years of my project. My closest encounter has been with AIQM1 and its recent upgrade, AIQM2 – a method parametrized for general organic chemistry with the C, H, N, and O atoms. Both are implemented in the free and open source Python-based MLatom package, which as other good examples, strengthens their documentation with example use case

Jupyter notebooks. Within it, last year, surface hopping with AIQM methods was released and the developers currently hint at a fully active learning approach that could be tailored to more complex systems.[51]

Finally, nonstatistical chemistry is demanding for any method – it requires accuracy far from equilibrium structures as the system quickly reacts and reaches high vibrational temperatures. Even with tools such as Fermi Golden Rule type models, we rely on a PES that is accurate enough to provide numerical third and fourth derivatives. Following the main crux of nonstatistical effects, it is not enough to train to reproduce barrier heights and vibrational properties of well-defined minima. Given that we've achieved the required level of accuracy, the inherent speed of these approaches, coupled with approximative quantum dynamics methods that can capture tunneling or electronic surface-hopping, would revolutionize our ability to explore non-TST chemistry.

Acronyms

SI   supporting information

PES   potential energy surface

VER   vibrational energy relaxation

IVR   intramolecular vibrational energy redistribution

ZPE   zero-point energy

VSC   vibrational strong coupling

DFT   density functional theory

UV VIS   ultraviolet and visible (light)

AIMD   ab initio molecular dynamics

MP2   Møller–Plesset perturbation theory of the second order

PIMD   path integral molecular dynamics

CMD   centroid molecular dynamics

RPMD   ring polymer molecular dynamics

A

Paper 1

Article

A Strategy for Modeling Nonstatistical Reactivity Effects: Combining Chemical Activation Estimates with a Vibrational Relaxation Model

Tomislav Rožić, Matthew S. Teynor, Nađa Došlić, David M. Leitner, and Gemma C. Solomon*

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The kinetics of many chemical reactions can be readily explained with a statistical approach, for example, using a form of transition state theory and comparing calculated Gibbs energies along the reaction coordinate(s). However, there are cases where this approach fails, notably when the vibrational relaxation of the molecule to its statistical equilibrium occurs on the same time scale as the reaction dynamics, whether it is caused by slow relaxation, a fast reaction, or both. These nonstatistical phenomena are then often explored computationally using (quasi)classical ab initio molecular dynamics by calculating a large number of trajectories while being prone to issues such as zero-point energy leakage. On the other side of the field, we see resource-intensive quantum dynamics simulations, which significantly limit the size of explorable systems. We find that using a Fermi's golden rule type of model for vibrational relaxation, based on anharmonic coupling constants, we can extract the same qualitative information while giving insights into how to enhance (or destroy) the bottlenecks causing the phenomena. We present this model as a middle ground for exploring complex nonstatistical behavior, capable of treating medium-sized organic molecules or biologically relevant fragments. We also cover the challenges involved, in particular quantifying the excess energy in terms of vibrational modes. Relying on readily available electronic structure methods and providing results in a simple master equation form, this model shows promise as a screening tool for opportunities in mode-selective chemistry without external control.
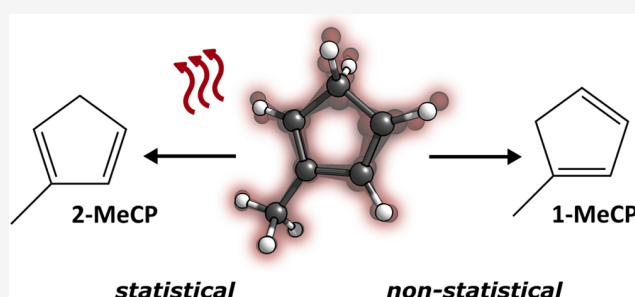
INTRODUCTION

Nonstatistical reactivity has grown to become a broad term in the literature that describes chemical systems undergoing dynamics that do not conform to the well-established statistical transition state theory (TST). Transition state theory, also known as the activated complex theory, was developed in the 1930s with ideal gas reactions in mind and has since seen widespread use in efficiently computing reaction rates.[1] The simplest forms of TST, taught in chemical thermodynamics courses and used throughout computational organic chemistry is based on the Eyring-Polanyi equation:

$$k = \frac{\kappa k_{\mathrm{B}} T}{h} e^{-\Delta G^{\ddagger}/RT} \tag{1}$$

To calculate the reaction rate k one only needs to provide the temperature of the system and estimate the Gibbs energy of activation, $\Delta G^{\ddagger}$. The transmission coefficient κ accounts for how often the system recrosses the transition state and is often assumed to be 1. The standard TST calculation will also simplify away from the Gibbs energy and express this equation via the (adiabatic) reaction barrier ΔV.[2]

The simplicity of these models is quite powerful, as it reduces the problem to comparing only the energies of single points of the multidimensional potential energy surface (PES)

— minima for reactants, products, or intermediates and first-order saddle points for transition states (TS). Once these points are located, an additional calculation can be performed to obtain the entropic correction required for Gibbs energies. Many quantum chemistry software packages also provide relatively "black box" methods that aid in the search for transition states.

Further developments of this theory include, among others: defining a surface in the PES called a diving surface (variational TST)[3] or better describing the transition state region by redefining a classical Hamiltonian at the TS geometry (semiclassical TST).[2] An approach that goes hand in hand with these theories is predicting product ratios of chemical reactions based on the relative computed energies of the major products. This can also be done with Rice−Ramsperger−Kassel−Marcus (RRKM) theory, which is more suitable than

standard TST for reactions that do not reach thermal equilibrium.

Nonstatistical Reactivity. When this statistical approach to reactivity fails, the reaction scheme could be incomplete or not modeled rigorously enough (or both). In a publication by Glowacki et al.[4] we see an informative example of the former, as a seemingly complex enzyme reaction is made simple again once the authors include just one more conformer to their statistical model. Therefore, we need to be careful not to mislabel reactions in which there are other competing mechanisms present.

In this work, we choose to focus on nonstatistical effects caused by the complex dynamics of the reacting molecule(s). Simply put, these effects can be attributed to overlapping time scales, usually assumed to be separate. For example, during a chemical reaction, the system can find itself vibrationally excited. The excess energy could originate from a molecular collision, the absorption of a photon or it could be left over from a previous reaction step (for instance, a bond forming or breaking). In those cases, the molecule is often termed chemically activated and labeled in reaction schemes with a *.

If vibrational relaxation is slow (or the reaction occurs very quickly) the system does not reach a thermal equilibrium that can be described statistically (thermalize) before the reaction moves forward and the two processes cannot be separated. A cartoon example is shown in Figure 1, though if we were
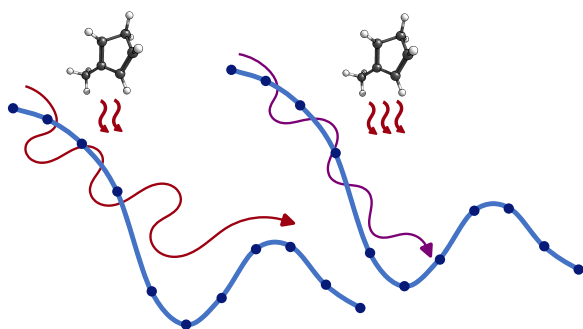

Figure 1. While it is hard to visualize in one dimension, it helps to remember most vibrations are oscillations from the proposed reaction path. Higher levels of excitation correspond to larger oscillations and higher reactivity if there are other minima nearby to explore. If the drop down the PES is steep enough and relaxation slow (as shown on the left example), some parts of it could be hopped over, including nearby barriers. In that case, molecular vibrations are underdamped. As relaxation becomes faster, it is as if the molecule moves on the PES with more friction (right).

drawing an accurate PES, the excess energy would also place the reaction dynamics "higher" above the path, which in turn makes the landscape (local minima and barriers) less restrictive. In other words, the separate reaction steps cannot be considered as a Markov chain as a system that does not thermalize in time is not completely "memoryless". This goes against the main assumptions of TST, where the activated complex is in thermal equilibrium and its vibrations are populated according to the Boltzmann distribution, as well as RRKM where such coupling is not included.

The vibrational energy/population of a molecule can then influence the reaction and result in experimental product ratios significantly different from TST predictions.[5] In reactions facilitated by collisions, reactivity can be enhanced by reactant

and product vibrational modes matching the reaction coordinate.[6] In extreme cases, one can find reactions that seem to proceed "uphill" on the PES.[7] Sometimes intermediates that would be relevant steps for TST could end up being less explored, or even skipped completely in cases of "ballistic" movement.[8] We could even think of these effects as an extreme case of kinetic versus thermodynamic reaction control, presenting at the scale of a single molecule.

A very common approach to model chemical reactivity (both statistical and nonstatistical) is to sample a larger part of the Born–Oppenheimer PES with ab initio classical molecular dynamics (AIMD).[9] The chemical system is propagated classically, using Newton's laws of motion for the nuclei, though the forces that act on them are computed through quantum mechanics (QM), or if that is too resource-intensive, with predefined force fields. Another alternative is to propagate an atom-centered density matrix. Conclusions are then drawn from a number of trajectories containing information about the position and velocity of the nuclei over time.[9−14]

Looking for these dynamic effects in the literature can be demanding as they are often described independently in different research fields. Along with examples cited above, we point out two very descriptive terms: dynamic matching,[15] implying some molecules are filtered through a preceding transition state and already possess dynamics matched to the reaction coordinate, and flyby reaction trajectories[16] where extrinsic force enables skipping certain side reactions. A perspective on dynamic effects in organic chemistry, including those not caused by vibrationally hot molecules, can be found in ref 17.

All these topics touch upon localized energy or directional movement on the PES. However, intramolecular vibrational energy redistribution (IVR) itself is also directional and energy does not move equally from one vibrational mode to another.[18,19] The simplest way to put it is that not all degrees of freedom will be equal when it comes to exchanging energy, and up to a certain time scale the energy can at the same time be distributed unevenly and also flow in predetermined directions. There are also doubts about using AIMD trajectories for vibrational relaxation modeling due to a lack of correspondence with actual quantum dynamics such as the issue of zero-point energy leakage. For further reading, we refer to a recent chemical perspective on IVR modeling in ref 20.

Furthermore, dynamic effects can be present even if the separation of time scales required for applying TST is present. For example, relaxation can be faster than the overall reaction rate, yet still slower than the barrier crossing, resulting in a reduced reaction rate ($\kappa < 1$ in eq 1). For a broader review of vibrational energy flow and its effects on chemical reactivity, we suggest reading the review in ref 21, which includes the well-studied examples of cyclohexane ring inversion in solution and stilbene photoisomerization in molecular beams and gas phase, reaching high-density gases.

Substitutions that do not change the overall mechanism could then lead to some amount of control over the selectivity of desired products.[22] In a recent work by Chen et al., different pathways to the same products can result in different vibrational populations and affect subsequent reaction steps, potentially allowing for reaction control at more than one point by exciting or modifying the vibrational modes involved.[23]

This concept has led to plenty of research over the years in the hope of mastering mode-specific laser-driven chemistry,

where reactions can be tailored and guided by pumping energy into certain vibrational modes, for instance using an infrared laser pulse to climb the vibrational ladder[24] until dissociation occurs and "quantum control"[25] strategies in general. Recently, the field saw a surge of activity in modifying reactivity by coupling molecular vibrational modes with an optical cavity, termed polariton chemistry.[26,27] In this approach, the cavity in which the reaction takes place is tailored for an interaction with a vibrational mode such that it fortuitously influences the reaction.

If it is not probable we can deposit energy in all the right places, for example with lasers, we wish to use the overall direction of vibrational energy flow to guide the chemical excitation we already have available to useful regions of the molecular system. Or in case of fortuitous "mode-matching" present from the start, we wish to keep it localized for longer. Constructing a special environment, such as an optical cavity, will not always be viable if several modes need to be coupled simultaneously or the reaction conditions do not allow it.

Therefore, we propose an efficient methodology that aims to predict how tuning the molecular structure of a system changes this vibrational scaffolding while still preserving the overall reaction mechanism.

■ THE WORKFLOW

In this work, we propose a workflow that combines estimating the vibrational excitation of the reacting system with a vibrational relaxation model that relies on (third-order) anharmonic coupling constants to describe the structure-based directionality of IVR and identify key relaxation pathways.

In the first half of the article, we will walk through the process (and obstacles) of obtaining all the required data through an example of a known nonstatistical chemical reaction, after which we will introduce the IVR model itself in detail and the conclusions we draw from it. An overview of the procedure discussed in this article is given in Figure 2.
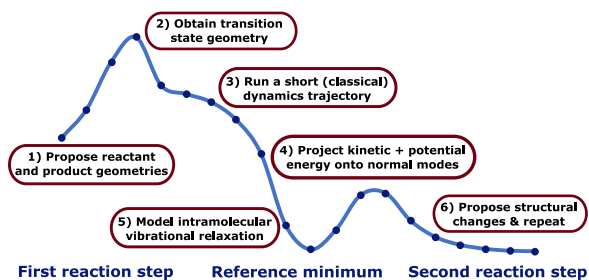

Figure 2. A summary of the modeling scheme required to explore a two-step reaction. In the background: energy profile of the first and second step of the example reaction (for cyclopentadiene) obtained with NEB-CI at the BS UCAM-B3LYP+D3BJ/def2-TZVP level of theory. Note that computing the reaction path is not required if the TS is easily obtained.

We will consider a common theme in nonstatistical reactions: a reaction with several steps in which the first step releases a relatively large amount of energy. We assume that at each subsequent reaction step, there is an excess of energy (compared to thermal equilibrium) stored in molecular vibrations. We will choose the first "activated" product (or intermediate) as our reference, which means we will use its vibrational normal modes when describing energy localization.

The question is then whether the energy stays long enough in the reactive mode(s) that match the next step of the reaction, or if the energy dissipates quickly to the rest of the molecule— and the solvent. In this context, the reactive mode(s) could also be those enabling some yield-decreasing side reaction. Whether these modes are populated right away or depend on IVR to reach them would depend on the energy released during the previous step.

Our example reaction, shown in Scheme 1, is a ring-breaking reaction followed by a rearrangement. The rearranged product,

Scheme 1. Reaction Scheme for the Ring Opening of 2-MeBCP[a]
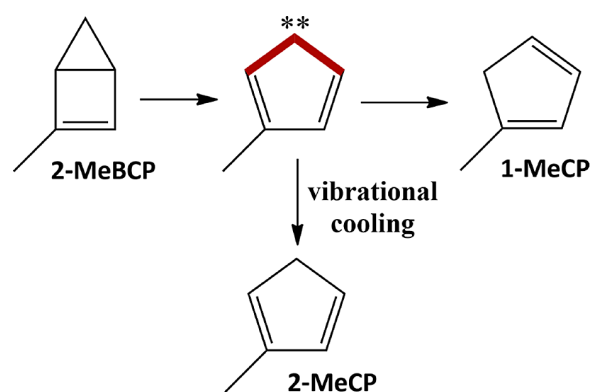
[a]After the initial ring opening, the vibrationally excited product (affected atoms highlighted in red) can undergo a further rearrangement, a process which, according to ref 31, competes with the collisional cooling of the initial product.

1-methylcyclopentadiene (**1-MeCP**), is observed in larger quantities than expected compared to the initial product, **2-MeCP**. The rearrangement is reported to occur on the same time scale as the vibrational relaxation of the molecule[28] and the surprising product ratio is therefore attributed to the nonstatistical dynamics the initial product undergoes while it is still chemically activated.

While the mechanism of the ring opening of bicyclo[2.1.0]-pent-2-ene(s) has attracted a fair amount of attention,[28−30] the main reason we chose this reaction is that it is computationally nontrivial and explored in the more recent work by Goldman et al.[31] In the article they model a reduced system (without the methyl group) using nearly two thousand AIMD trajectories at very high temperatures of 1000−2000 K and at the BS-UO3LYP/3-21G level of theory.

Excitation Estimation. Since we are dealing with a reaction and not a laser experiment, the first piece of the puzzle is quantifying what being "chemically activated" means in this context. We assume that the activation lies in excess vibrational energy right after the bond breaking. We then choose to stay in the frame of reference of the initially formed product, **2-MeCP**, by using its calculated vibrational normal modes and coupling constants for the vibrational relaxation model.

The first step is to estimate where this excess energy is located, that is how much of it is initially deposited in the populations of each of the product's vibrational normal modes.

We see three ways of approaching this problem, of which we will explore the first two:

1. Run a shorter AIMD simulation from the TS geometry and decompose the resulting geometry and kinetic energy of the nuclei within the normal mode picture. The work by Goldman et al.[31] does this, using bond lengths as stopping criteria.

2. Compare key geometries along the reaction path (such as the TS) in terms of displacements along the normal modes and using only these displacements as estimates of vibrational excitation.

3. "Eye-ball" one or several modes that best match the proposed reaction step (or use a computed intrinsic reaction coordinate) and inspect their decay separately. This is the simplest approach but puts us at risk of missing out on information that is not as visually obvious (for example, bond stretching as a result of the bond order changing).

In all three cases, a reasonable transition state geometry is useful (or mandatory). For this purpose, we locate the transition state (**TS**) between the original two-ring structure (**2-MeBCP**) and the first product (**2-MeCP**).

We utilize the climbing image variant of the nudged elastic band method (NEB-CI)[32] as implemented in ORCA[33] for a relatively black-box approach to obtaining both the minimum energy path (MEP) of the reaction and potential transition states and intermediates along that path. After a converged NEB-CI calculation, a TS optimization is performed at the same level of theory. The result can be seen in Figure 3.
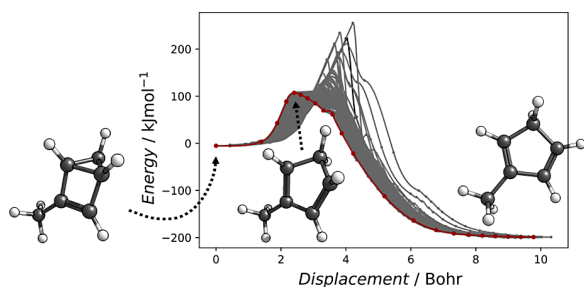

Figure 3. Final minimal energy path (red) for the MeCP ring-opening reaction with only the initial (left) and final (right) molecular geometries given as inputs. The converged MEP provides a transition state guess (middle) that is only 3 subsequent optimization steps from a converged TS geometry. In the background, in gray, are the previous iterations.

Additionally, we perform an NEB calculation between the TS and the second product, **1-MeCP** to gauge barrier sizes (see Figure S2 for discussion). In short, the 100 kJ mol^{-1} (24 kcal mol^{-1}) barrier between the two products is not easily overcome at room temperature but is roughly a third of the energy released during the ring opening.

At this point, we should warn that dealing with reactions where bonds break can lead to areas along the proposed reaction path where the molecule is not well-defined with a single-reference electronic structure method. In our case, singlet diradical character is present, from a configuration where the ring-forming single bond has been broken, but the electrons have not yet rearranged to form new double bonds. Our transition state geometry lies exactly in this area, something that can be quantified with the T_1 diagnostic (see Figure S3).

This failure of the chosen electronic structure method would also, at least in our experience, cause failure to converge the

NEB-CI path. We find, however, that using an unrestricted formalism of DFT we obtain reasonable MEPs. All that needs to be done to confirm the proposed TS is to perform a subsequent geometry optimization and frequency calculation using a broken-symmetry approach at the same level of theory.

This success is not surprising as broken-symmetry unrestricted DFT with a hybrid functional has been shown to provide an adequate description of organic molecules with diradical character, including transition states,[13,34,35] polarizabilities[36] and bond breaking in general.[37,38] In the work by Hamaguchi et al.,[39] we see an example of a reaction both involving diradicals and dynamic effects modeled with unrestricted DFT. As we are interested in geometries and not precise energies, we see no need to apply additional spin projection correction schemes. We do note a warning from ref 40. that discrepancies in calculated broken-symmetry geometries could be larger when dealing with systems where the radical character is very delocalized (for example, due to π conjugation). With these warnings in mind, we consider the potential discrepancies to be small enough for our use case.

This detour still keeps our approach relatively black-box, as only simple keywords for the SCF procedure (such as "brokensym 1,1" in ORCA or "guess = mix" in Gaussian) are required to obtain a broken-symmetry wave function solution. Now that we have obtained a transition state geometry, we will address the first two proposed ways of quantifying the vibrational excitation of a "hot" molecule.

In the normal mode picture of approach (ii), we stay in the frame of reference of the **2-MeCP** product. The difference between this minimum geometry and the TS geometry is then expressed in terms of dimensionless normal coordinates. First, the two geometries are optimally aligned according to the Kabsch algorithm. The differences in the positions of each atom are then projected onto the normal modes of the product to obtain the dimensionless mode displacements Q. A detailed look at the procedure can be found in the SI. The potential energy is then estimated as

$$V_\alpha = \frac{\omega_\alpha^2 Q_\alpha^2}{2} \tag{2}$$

This assignment is a common approach in spectroscopy, for example when mapping the PES between the initial Franck–Condon (vertical) excitation and a nearby conical intersection[41] or to estimate rearrangement energy in electron transfer calculations.[42] Since it is based on the harmonic approximation, the approach will fail if the geometries are too different and/or if the reaction coordinate is very anharmonic.

An example of this can be seen in Figure 4, which shows an AIMD trajectory of the system, starting at the transition state geometry. At the first point, the projection onto normal modes (red) estimates an unrealistically high potential energy compared to the relative electronic energy obtained by BS DFT. This would result in a large overestimation of the initial energy present in the vibrational modes. However, in case all the energies are overestimated similarly, the results would not change. For this reason, we will not immediately discard approach (ii).

Since the relatively affordable broken-symmetry DFT calculation provides a good estimate of the barrier height, we are encouraged to move on to approach (i) and perform an AIMD simulation. The goal now is to determine whether a single "drop" down the PES, which we choose to term a "0 K
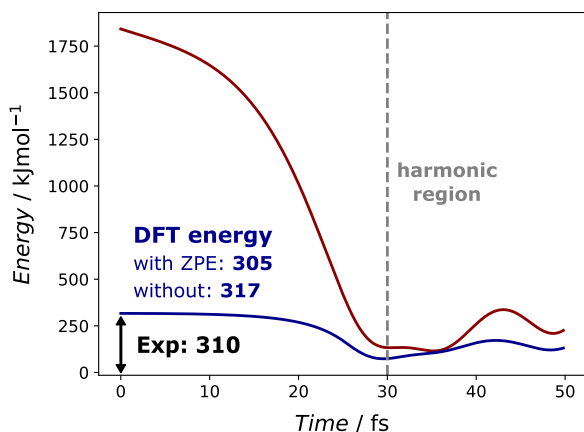
Figure 4. Figure shows a comparison of the computed electronic energy (dark blue) and the sum of the estimated energies from the harmonic projection scheme (dark red) with reference to the **CP** minimum during a "0 K pushed" trajectory started at the **BCP** transition state. The gray dotted line points out the chosen time step, at which the system is in the harmonic basin near its minimum, though it does not pass directly over it. If the discrepancy in energies is too large, the trajectory needs to be rejected.

pushed trajectory", provides similar results to a much more costly ensemble sampled at the experimental temperature of 303 K. The 0 K pushed trajectory will start without any additional energy (not even zero point energy) other than 0.5 kcal mol^{-1} (around 2 kJ mol^{-1}) in the imaginary, reaction-driving mode.

This slight "push" saves some computational time and ensures the correct direction of the trajectory. For our purpose we will aim to run the trajectories just long enough for the molecule to "drop" down into the harmonic area of the PES to quantify which modes of the product end up vibrationally excited in the process. In short, we are performing the same projection based on normal mode displacements, but closer to where it is theoretically sound to do so.

The appropriate point in all trajectories can be determined automatically. We chose to look for points at which the sum of the calculated electronic and projected potential energies is the smallest. This ensures picking a trajectory step that is near the minimum, where the potential energy projection is more accurate. Ideally, this would be the point closest to the minimum before its first "bounce", even if the trajectory approaches closer later on. The energy profile of an example trajectory is shown in Figure 4. Note that despite "bouncing" quite high up the PES, this trajectory is not reactive and either unmanageably long simulations or unphysically high temperatures would be required to directly sample the next reaction step.

Both the projections of the kinetic and potential energies onto the normal modes are then used as the initial energies for the time evolution of a more simplified vibrational relaxation model.

Modeling Intramolecular Vibrational Relaxation. The IVR model we use (eqs 3 and 4) originates from phonon scattering theory[43] and estimates the flow of excess vibrational energy Γ from mode α into two other modes (represented by β and γ), depending on how well they match in frequencies ω (that is, energy), their mean Bose–Einstein occupation numbers n and their third order anharmonic coupling constant $\Phi_{\alpha\beta\gamma}$.

$$\Gamma_\alpha^{\text{decay}} = \frac{\hbar}{16\omega_\alpha} \sum_{\beta,\gamma} \frac{|\Phi_{\alpha\beta\gamma}|^2 (1 + n_\beta + n_\gamma)}{\omega_\beta \omega_\gamma}$$
$$\times \frac{(\Gamma_\alpha + \Gamma_\beta + \Gamma_\gamma)}{(\omega_\alpha - \omega_\beta - \omega_\gamma)^2 + \frac{1}{4}(\Gamma_\alpha + \Gamma_\beta + \Gamma_\gamma)^2} \quad (3)$$

$$\Gamma_\alpha^{\text{collision}} = \frac{\hbar}{8\omega_\alpha} \sum_{\beta,\gamma} \frac{|\Phi_{\alpha\beta\gamma}|^2 (n_\beta - n_\gamma)}{\omega_\beta \omega_\gamma}$$
$$\times \frac{(\Gamma_\alpha + \Gamma_\beta + \Gamma_\gamma)}{(\omega_\alpha + \omega_\beta - \omega_\gamma)^2 + \frac{1}{4}(\Gamma_\alpha + \Gamma_\beta + \Gamma_\gamma)^2} \quad (4)$$

The model has been used previously to explain experimental findings of asymmetrical vibrational energy transfer,[44] as well as vibrational energy transfer in proteins.[45,46] A graphical representation of the three-mode processes is depicted in Figure 5.
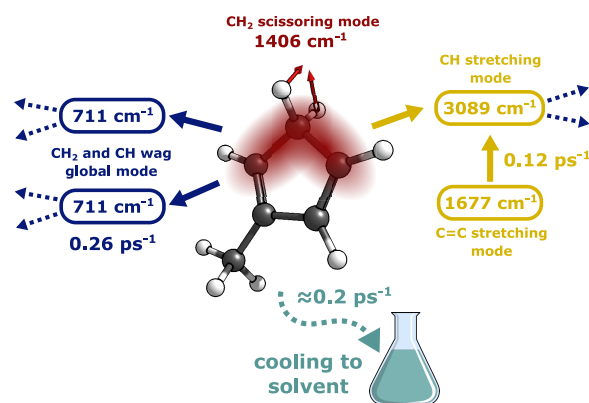
Figure 5. An example of two possible energy pathways leading from a vibrationally excited mode (red). On the left, in blue, the mode splits into two modes it is coupled with—corresponding to the "decay" process. On the right, in yellow, the mode "collides" with another populated mode, moving the energy upward.

The anharmonic constants $\Phi_{\alpha\beta\gamma}$ can be calculated using ab initio electronic structure methods readily available in several quantum chemistry packages as they are primarily used in the context of computational (infrared) spectroscopy. They also make up almost the entirety of the required resources for this step of the workflow, though in our case they are still more affordable than a single AIMD trajectory.

We obtained them using Gaussian's implementation of GVPT2,[47] which also includes the treatment of frequency shifts due to both anharmonicity and resonances of the vibrational states. This allows us to bypass the similarly formulated second set of self-consistent equations which would describe the frequency shifts[43] and therefore to focus only on the decay Γ.

Computational Details. All electronic structure calculations are performed using either the Gaussian 16[48] (geometries, GVPT2 frequencies, coupling constants, and forces) or ORCA 5[33] (transition state optimization with NEB-CI) software packages. AIMD trajectories were propagated using the Milo v1.1.0[49] dynamics program. Geometries in the figures were visualized with ChemCraft.

We showcase two levels of theory: a computationally less demanding DFT method in the form of the range-separated

hybrid density functional CAM-B3YLP[50] with the def2-TZVP[51,52] basis set and D3BJ empirical dispersion,[53] a combination which should provide numerically stable anharmonic constants[54] and good value for computation time with medium-sized systems,[55] and second-order Møller–Plesset perturbation theory (MP2)[56] with the cc-pVTZ basis set[57] as our practical limit for obtaining anharmonic couplings. As previously discussed, we were limited to broken-symmetry unrestricted DFT calculations when obtaining the transition state and running AIMD simulations.

Insight into the accuracy of obtaining both frequency and coupling constants can be taken from benchmarks already present in the field of computational spectroscopy and examples, including the application of GVPT2, can be found in a recent review by Barone et al.[58] It is worth mentioning that the cost of the relatively expensive anharmonic calculation can be efficiently reduced via a hybrid approach—calculating harmonic frequencies at a higher level of theory and obtaining their anharmonic corrections with a more affordable approach. We have observed that the results depend much more on differences in computed vibrational frequencies as they may change the ordering of the modes and move them in or out of resonance. Their respective couplings, when strong enough to be significant, would stay large even at different levels of theory.

Since the decay rate depends on itself (Γ_α) as well as the decay rates of the other two modes (Γ_β, Γ_γ), this set of equations needs to be solved self-consistently. However, the dependence is small and it is also reasonable[44] to assume the Γ (in the right side of eq 3) add up to a line width of 24 cm^{-1}, putting the total Γ near the empirical value of 1 ps^{-1}, appropriate for organic molecules of this size. This effectively fixes the width of the Lorentzian curve that "searches" for resonances. By leaving a decent width available to each mode we will lose quantitative accuracy, but ensure we do not miss resonances due to the difficulties of calculating accurate vibrational frequencies.

The Master Equation. After obtaining the decay rates, we can then construct the matrix of transfer rates from one mode to another, **k**. In this work, we examine these transfer rates as a map of possible pathways for relaxation from one mode to another.

The **k** matrix is built up as shown in eq 5. The relaxation rate $\Gamma_{\alpha\beta}$ that makes up the off-diagonal elements is a subset of the sum of all terms from eqs 3 and 4, in which we take those terms that contain both the modes α and β. Additionally, the lifetime is obtained simply as the reciprocal value of the total decay rate of a vibrational mode, Γ_α.

$$k_{\alpha\beta} = \frac{\omega_\beta}{\omega_\alpha}\Gamma_{\alpha\beta} \quad k_{\alpha\alpha} = -\left(\Gamma_{\alpha,\text{solvent}} + \sum_{\beta \neq \alpha} k_{\alpha\beta}\right) \tag{5}$$

Constant decay into the solvent is included in the diagonal elements at a value of $\Gamma_{\alpha,\text{solvent}} = 0.2$ ps^{-1}. The results themselves are not very sensitive to this value, it simply ensures energy is removed from the system. However, a larger solvent decay will bring down the value of all estimated lifetimes. The resulting lifetime of 5 ps is the lower bound of values reported for molecules of similar size in organic solvents.[19,59]

The matrix **k** is used in a master equation approach to follow how a given energy distribution changes over time, with a simulation time step of 0.01 fs and the approximation noted in eq 6.

$$\mathbf{E}(t) = e^{\mathbf{k}t}\mathbf{E}_0 \approx (1 + \mathbf{k}t)\mathbf{E}_0 \tag{6}$$

Since the energy assigned to each mode changes over time we used occupation numbers n directly from our initial projected energies (as $n_\alpha = E_\alpha/\hbar\omega_\alpha$) and updated them every step for the first 1 ps of the simulation and every 20 steps afterward. Recomputing the **k** matrix adds cost while allowing for more collision pathways to be available. Since these are usually not the dominant ones, updating the **k** matrix less often yields very similar results.

We note that there may be interesting energy dynamics among some vibrational states that may not appear in the results of the master equation simulations. For example, energy may be trapped temporarily among a subset of modes and exhibit resonant oscillations, which the master equation would not predict. Nevertheless, bottlenecks to energy relaxation due to strong coupling among such modes and weak coupling to the others would be captured by the master equation approach, at least to the extent that the interactions among the modes are largely due to third-order anharmonic coupling.

Single-Valued Indicators. We will now expand the discussion to include the nonsubstituted **CP** ring structure explored in ref 31, as well as three other substituents: a methoxy group (**OMe**), chloromethyl group (**CH2Cl**) and tertbutyl group (**TB**). This provides us with examples of both different system sizes and vibrations of different character.

The final challenge of using this information to aid reaction design is the data analysis. Molecules have many vibrational modes that "talk to each other" and mapping them out visually becomes impossible. For example, any in-depth analysis will involve looking for major pathways for vibrational modes that either start with the most energy or point toward the rearranged product, as well as plotting the energy contained in certain fragments or bonds of the molecule.

This becomes a problem if we wish to quickly sample and compare the effects of changing the structure. Ideally, a single number indicator that quantifies how likely the molecule is to exhibit nonstatistical behavior would not only help explain current nonstatistical reactions but also aid in reaction discovery—which is currently dominated by either assumptions of statistical dynamics and barrier estimations or massive applications of ab initio molecular dynamics. For this purpose, we will attempt to devise an indicator that aims to show how quickly the initial excitation can relax. We formulate it as eq 7:

$$\text{Indicator} = \frac{\sum_i \frac{E_i}{\omega_i} \cdot \frac{1}{\Gamma_i}}{\sum_i \frac{E_i}{\omega_i}} \tag{7}$$

where E is the estimated initial energy of the mode (from any projection scheme), ω is its frequency (in matching energy units) and Γ is the calculated IVR rate. Simply put, we multiply the estimated initial energy in each mode by its lifetime according to the IVR model and normalize it by the total estimated quanta of excitation in the region/molecule. As the decay rate Γ includes the empirical solvent decay of 0.2 ps^{-1}, the maximum lifetime of any mode is capped at 5 ps, which in turn also ensures this indicator will always stay in a range between 0 and 5 ps.

If a reaction is exhibiting useful nonstatistical behavior due to energy localization, we would then aim for structural changes that increase this value (as close to the limit set by the solvent). On the other hand, if we wished to prevent nonstatistical effects, we would aim to add relaxation pathways

and lower the value of this indicator. In essence, a low value represents well-damped modes that lead to statistical results, while those closer to the solvent limit are underdamped.

To extract more relevant information from the data, we may also choose to define a group of modes of interest or a group of atoms that make up the reaction center and compute the indicator and plot only for those modes. In this case, we will choose the CH$_2$ group along with its two neighboring carbon atoms (highlighted in Figure 6) as our reaction center and focus on only modes localized to that region (details of this assignment can be found in the SI). We will refer to the region as "**3C**".

A more biased approach, picking out modes that specifically position the ring for a hydrogen rearrangement, might be more fruitful, but this way we hope to include all of them without adding manual work to the workflow. As this specific reaction may involve some tunneling, it is also important to include all modes whose combined displacements help better position the CH$_2$ hydrogen atoms. Figure 7 shows which modes of **2-MeCP** fall into this grouping.

Finally, since propagating the master equation has negligible cost compared to our previous steps, we wish to extract information from the actual time evolution. The simplest way to do this (for any set of modes) is to determine the time when there is only $\frac{1}{e}$ (around 36.79%) of the initial energy remaining. As energy can flow both in and out of the chosen set, this value may go above 5 ps. The plots also provide a graphical way to interpret the results and we show some representative plots in Figure 8. The first four plots, grouped by mode frequency, clearly show the common "tiered" nature of IVR as most of the energy flow will occur by the decay mechanism toward lower frequency modes. When looking at custom regions of interest such as in (e) and (f), the intermediate or steady-state might or might not appear, depending on whether energy flows into the region.

■ RESULTS AND DISCUSSION

From the plots in Figure 8, we see that some systems have much larger energy estimations, for example, **MeCP** in (c) compared to **OMe** in (d), even though they are of similar size. This is mainly due to the trajectory not coming as close to the local minimum of the PES on its first pass into the well. We ran trajectories for 100 fs and the first pass would always occur within the first 50 fs, however, running for longer and picking the second or third "bounce" did not guarantee a relevant improvement.

The resulting overestimation is mainly reflected in the highest energy modes (such as C−H stretches), where even small displacements contribute a lot of additional energy and resemble the (dotted) estimation obtained straight from the transition state. However, in terms of the single-valued indicators, we did not notice a much larger standard deviation of the ensemble result nor a larger difference between the single trajectory result and the ensemble mean. The results should stay robust, as long as a criterion is present to reject trajectories that completely miss the minimum.

Moving on to the computed indicators shown in Figure 6, we see that the values obtained for the **3C** region (b) are similar compared to those for the whole molecule (a). As the regional indicator (b) contains fewer modes overall, the differences between similar systems are more pronounced. As expected, the overall values decrease as the system size
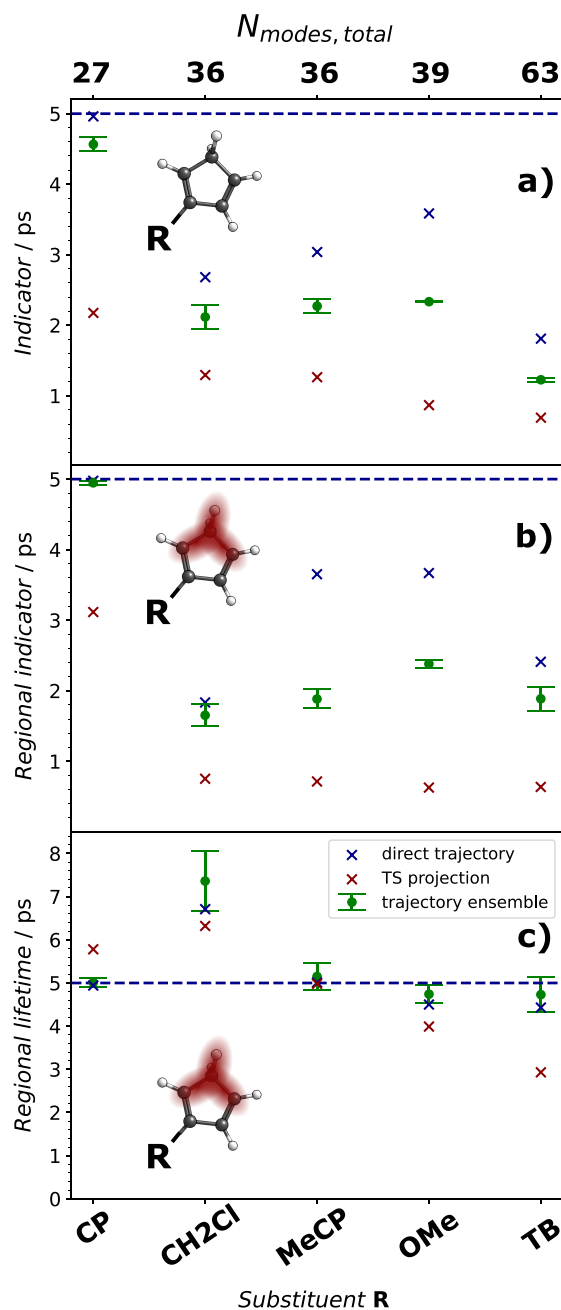
Figure 6. Indicator values for the relaxation of all modes (a) and those localized to the **3C** region (b) in differently substituted (R) systems; lifetimes computed from the master equation for the **3C** region modes (c). The dashed blue line represents the solvent decay limit to which all values would collapse with no intramolecular vibrational relaxation present. For each substituent change we use three different ways of estimating the initial excitation. The trajectory ensemble mean values (green circles) are capped above and below by their standard deviations.

increases and more relaxation pathways are added. There is, however, a clear dip in the three substituents (**CH2Cl**, **MeCP** and **OMe**) that are of similar size but different makeup, during which this trend actually reverses.

This is a typical example of how vibrational relaxation relies on a local density of states. Adding a substituent onto such a small ring creates many more pathways, but their effect is measured by how well they couple with existing ones. In the

374 cm⁻¹   835 cm⁻¹   987 cm⁻¹   1142 cm⁻¹
372 cm⁻¹   825 cm⁻¹   974 cm⁻¹   1113 cm⁻¹
3.46 ps   4.70 ps   1.61 ps   4.03 ps

1321 cm⁻¹   1434 cm⁻¹   3045 cm⁻¹   3070 cm⁻¹
1291 cm⁻¹   1406 cm⁻¹   2920 cm⁻¹   2924 cm⁻¹
1.64 ps   1.61 ps   4.58 ps   5.00 ps

Figure 7. By defining the **3C** region using three carbon atoms and the two hydrogen atoms of the CH$_2$ group we focus on the vibrational relaxation of 8 out of a total of 36 vibrational modes in the example of **MeCP**. Modes obtained with DFT are shown here on displaced geometries (displacement scale of 0.5) with their scaled displacement vectors, harmonic frequency (gray), GVPT2 corrected frequency (blue), and computed relaxation lifetime (red). The computed lifetimes include the empirical solvent decay value and therefore have an upper bound of 5 ps.
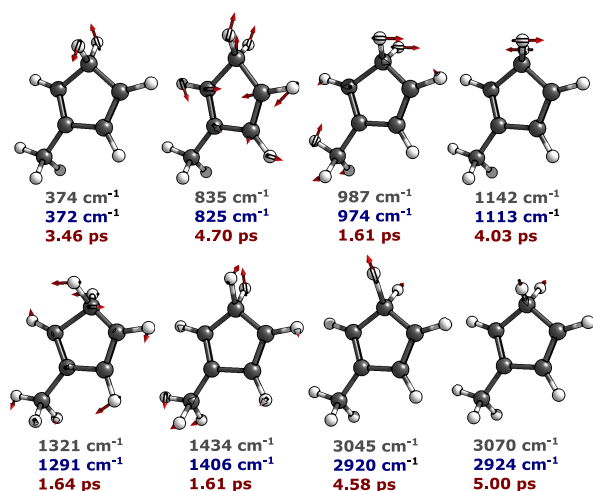
extreme example of **TB** the system now has more than double the original number of modes. In the indicator value for the whole molecule (a) this large difference is still visible in the

form of a second dip, but the "payoff" is less than switching out no substituent for a small one. For the values based on only the **3C** region of interest (b and c), **TB** is outcompeted by the heavy-atom effect in **CH2Cl**, as many of the modes added by the *tert*-butyl group are localized to its three branched methyl groups and simply too far away.

The dip at **CH2Cl** is also present in the computed lifetime of the **3C** region's modes (c), but in this case, it seems to point to the complete opposite of the indicator value. This can be explained by considering that the indicator values only describe the presence of usable vibrational relaxation pathways between the reaction center and the rest of the molecule. However, the lifetime value is computed directly from the master equation and as the coupling is present, energy can also flow into the region, depending on its initial distribution. In those cases, the lifetime can go above the imposed solvent relaxation limit of 5 ps.

For reference, the **3C** region contains 8 out of the total 36 modes in both the **MeCP** and **CH2Cl** systems (as shown in Figure 7). But changing one hydrogen to a chlorine atom more than triples the mass of the small substituent, changing some of the frequencies and couplings inside the ring itself and also how the system gains energy while dropping from the transition state. There are now 158 resonances that contribute more than 0.01 ps⁻¹ to the total decay, an inclusion criterium we use to simplify and reduce noise, as opposed to 95 in **MeCP**. However, as the indicator estimates "first tier" relaxation and the **3C** modes do not change much, it only reports a slightly lower value. A good example is the highest frequency mode belonging to the **3C** region (the CH$_2$
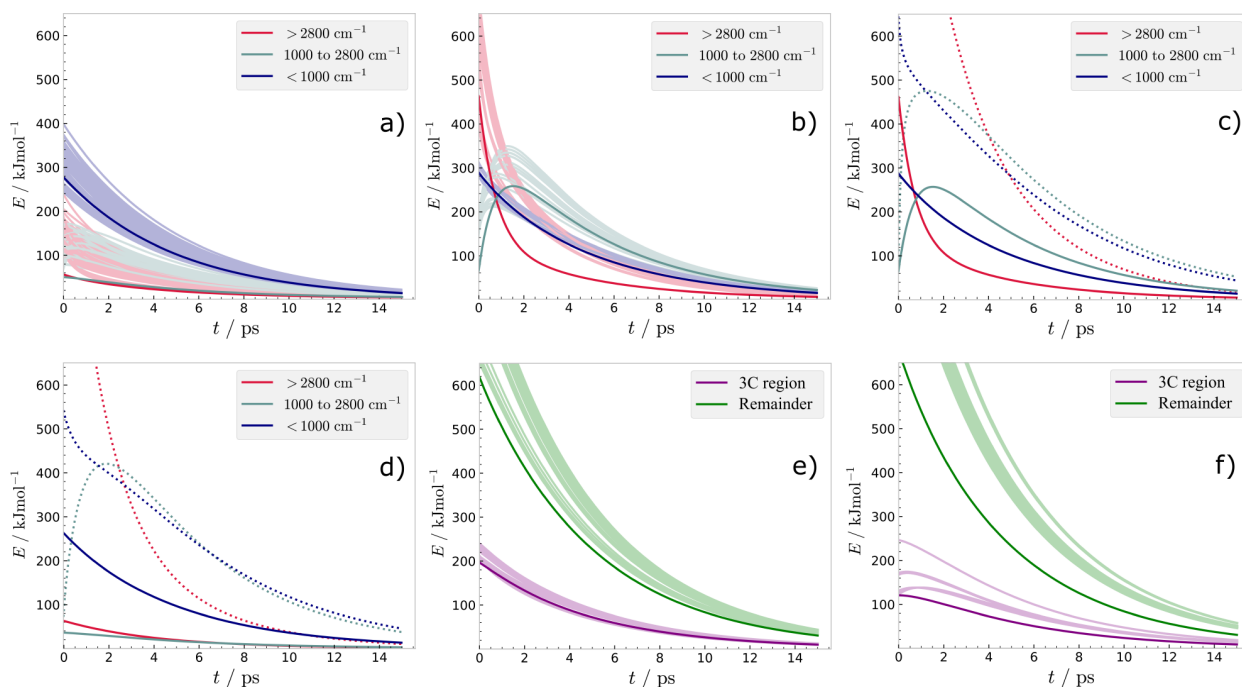
Figure 8. Model allows us to follow the energy in any group of vibrational modes, the obvious choices being groups by frequency range and those localized to a region/functional group of the molecule. While most of the IVR from the highest frequency modes (red) occurs very quickly, the tiered (bottlenecked) nature of IVR can lead to several time scales or steps being observed experimentally. The population of vibrations in the medium energy range (pastel green) peaks in the first few picoseconds, containing most of the "first tier" modes. We show (a) **CP** and (b) **MeCP** results from 50 trajectories (transparent) and from the single 0 K pushed trajectory (solid); followed by 0 K pushed trajectory results (solid line) compared to TS projection results (dotted) for (c) MeCP and (d) OMe. The last two plots show the time evolution by grouping the modes into two regions for (e) **MeCP** and (f) **CH2Cl**.

Table 1. Computed BS-DFT Single-Point Energy (ΔE, in kJmol^{-1}) and Gibbs Energy (ΔG, in kJmol^{-1}) Differences for the First Barrier (TS, Reactant to Transition State) and the Drop towards the First Product (1-TS, Transition State to First Product) along with the Imaginary Frequency at the Transition State (ν_{TS}); in Vacuum and with Implicit Solvent Modeling for Tetrahydrofuran

substituent	ΔE_{TS}	ΔE_{1-TS}	ΔG_{1-TS}	ν_{TS}/cm^{-1}	$\Delta E_{TS,THF}$	$\Delta E_{1-TS,THF}$	$\Delta G_{1-TS,THF}$	$\nu_{TS,THF}$/cm^{-1}
CP	112.6	316.7	301.5	−688	44.5	250.5	235.8	−668
CH2Cl	113.8	314.1	301.0	−653	114.2	296.9	287.6	−641
MeCP	113.6	311.0	298.2	−659	113.4	312.1	299.4	−642
OMe	116.8	318.5	303.7	−637	117.8	318.3	303.7	−646
TB	115.0	308.4	295.1	−633	115.0	310.2	296.8	−636

asymmetric stretch at 2928 and 2924 cm^{-1} for **CH2Cl** and **MeCP** respectively), as it has no viable pathways of relaxation in either system.

Only by considering the master equation results of individual modes, we can see two representative modes that exhibited the longer lifetime steady state for **CH2Cl** - a symmetric global mode at 1377 cm^{-1} and a ring stretching mode at 826 cm^{-1} that resembles the ring-breaking mode of the TS. Interestingly, the frequencies of these vibrational modes are not changed, as they are 1379 and 825 cm^{-1} respectively in the **MeCP** system. However, they do rely on a small number of pathways created by the low-energy modes—which do change more, an example being the CH$_2$ (and ring) wagging mode at 384 cm^{-1} in **CH2Cl** that is very distorted compared to the same wagging motion in **MeCP** at 372 cm^{-1}, also resulting in a smaller coupling constant Φ for one of these pathways (decay: 1377 to 384 + 972 falling from 11.1 to 6.5 cm^{-1}). As many changes like this one can add up, it would be prudent not to skip the master equation, especially since running it is orders of magnitude faster than obtaining the couplings.

As all of the systems are still relatively small, we have effectively modulated the localization of energy by adding a heavy atom to the ring substituent. As a result, IVR of the reactive region in **CH2Cl** has actually slowed down, even though the overall density of states of the system increased. In a larger molecule, where we would imagine the majority of the energy moving away from the nonstatistical reaction center, we should see the indicators and lifetimes point in the same direction. This heavy atom effect is known from classic works by R. A. Marcus, William L. Hase, and co-workers. Inherently the effect is caused by frequency changes and mode localization, though a decrease in the number of useful resonances also makes IVR more sensitive to changes in couplings. For functional groups bridged by heavier atoms, a heavy atom blocking effect can arise and further isolate energy localization.[60,61]

As for excitation estimation, the single 0 K pushed trajectory method gives overall the same qualitative behavior and insight as the 50 trajectory ensemble. Since no zero-point energy is provided, the 0 K pushed trajectory results in a much more localized energy profile, resulting in higher indicator values (that is, less of the coupling reported as significant). The single-valued indicator results from a TS projection are not far from the dynamical picture from one or more trajectories, as the results depend on the common vibrational pathway structure. However, we risk losing any nuance between similar systems as we are essentially providing an excess of energy to the highest energy modes whose subsequent decay may mask any relevant energy localization in the midenergy or "tier two" modes. Overall, going down to a single trajectory approach will

bring down the cost by one or 2 orders of magnitude, while choosing to use no AIMD at all would only halve it from there, as couplings are still required.

It is also worth noting that including anharmonic corrections for DFT frequencies does not always move them closer to experimental energy levels and resonances.[62] Since the GVPT2 calculation is costly and provides the resonance shift treatment, we prefer to use the provided anharmonic fundamentals in the model as well. For our results, switching back to harmonic frequencies does not change the results qualitatively (Figure S1), however for any larger exploration, we would advise checking benchmarks of the chosen electronic structure method.

On the other hand, MP2 results should generally benefit from GVPT2 corrections, but in our case, it would not be a reasonable choice of PES for the AIMD simulations. This led us to project coordinates and velocities from BS-DFT trajectories onto MP2 vibrational modes. Results from this mixing of methods are discussed in the Supporting Information. In short, the ordering of the substituents based on their calculated regional lifetime gives the same qualitative results. The indicator values differ, suggesting that they could be less robust metrics than directly obtained lifetimes, giving us another argument to propagate the master equation at little added expense.

Additionally, we performed implicit solvent calculations in tetrahydrofuran, as used in the experiment, for the single trajectory method, as well as the frequencies and couplings. With the unclear exception of the unsubstituted CP in solvent, barrier heights did not change much from substitution or solvation. An overview can be seen in Table 1. For CP, the first barrier is lower relative to both the CP and BCP structures, though it does not change the result obtained without solvation.

For all other systems, the differences in barrier heights are lower than the expected error of the energy projection scheme. The imaginary frequency at the top of the first barrier slightly decreases as the substituent size is increased. The results from implicit solvation, which are shown and discussed in the SI (Figure S1), qualitatively differ overall only in the regional indicator value.

It is important to note that in protic solvents a hydrogen-bonded solvent molecule could modify the vibrations of its bonded site, as well as draw some of the heat more efficiently. In those cases, one should consider including explicit solvent molecules in the electronic structure calculations and/or treat the bath (solvent) environment in a way where the cooling rates are not equal for all modes.

Additionally, coupling between vibrational and rotational (or torsional) motion may accelerate IVR, something that we do not capture with our current model. This is explained well in

the work on the methyl rotor in *p*-fluorotoluene by Goodfellow and Parmenter,[63] where the increase in IVR (compared to difluoro benzene) is so large it cannot be justified by added modes and lowered symmetry alone. In our case, it is very likely the reported **MeCP** exhibits faster IVR, even if a purely vibrational model suggests it is similar to bulkier substituents.

Finally, we emphasize that these metrics aim to directly compare similar systems undergoing the same reaction mechanism. While we cannot propose direct product ratios, it is not guaranteed that a more exhaustive AIMD study would be able to either. In the previous study, only 2% of the BS-DFT trajectories at a sampling temperature of 1000 K were reactive, rising to 12% at 2000 K, while the experimental ratio of rearrangement was close to 20%.[31] But given this experimental data point for **MeCP** we can suggest that there are other substituents (such as CH_2Cl) where the rearrangement ratio would be higher. On the other hand, we would not expect a much smaller (more statistical) ratio despite switching out for a heteroatom group such as the methoxy, or a much larger tertbutyl group.

For completeness, in Scheme 2 we present a summary of the suggested procedure in which the steps match the proposed

Scheme 2. A Summary of the Recommended Modeling Procedure

- **1a** Identify vibrationally hot intermediate/product and optimize its geometry with very tight convergence and SCF criteria

- **1b** Perform a GVPT2 calculation to obtain vibrational modes and their couplings and anharmonic frequencies

- **2a** (Optional) Obtain reaction path and TS guess using the NEB-CI method

- **2b** Optimize the geometry of the transition state preceding the hot intermediate/product

- **3** Start one or more molecular dynamics simulations at the TS geometry and run for 50-100 fs or stop automatically using criteria from Step 4

- **4** Project trajectory onto normal modes of intermediate/product and pick an appropriate point to estimate the vibrational excitation (implemented as Jupyter Python notebook)

- **5** Using the excitation estimate and GVPT2 output, run the IVR model (implemented in Python)

- **6a** Using plots and calculated lifetimes obtained from the IVR master equation, quantify the extent to which a certain set or region of vibrational modes is bottlenecked

- **6b** Use these insights to modify the structure to enhance or suppress non-statistical reactivity effects and repeat the procedure

modeling scheme in Figure 2. Recently, GVPT2 anharmonic frequencies are also available in ORCA6, which we use alongside Gaussian16, though other quantum chemistry packages also offer the method. For dynamics, we provide a Jupyter notebook written for Milo, which is currently only interfaced with Gaussian, but which should be simple to adapt for other trajectories.

In terms of improvements, the approach would overall benefit from a more accurate vibrational relaxation model, though its cost would need to be comparable. At present, we are looking into approximative methods of propagating

quantum dynamics of the vibrational Hamiltonian. Approaches to performing semiclassical AIMD that reduce or remove zero-point energy leakage do exist and coupled with cost-efficient gradients could become viable for massed nonstatistical reaction exploration.

CONCLUSIONS

While AIMD simulations remain a popular tool for studying nonstatistical effects, computational and experimental techniques that quantify them are not in common use.[31] For this reason, we propose a combination of methods that estimate chemical excitation in terms of excess vibrational energy and then follow its (de)localization using an intramolecular vibrational relaxation model.

As the likelihood of nonstatistical chemical dynamics increases with greater energy localization, predicting the relaxation rates between reactive modes and the rest of the molecule sets us up to estimate the extent these effects could have on a chemical reaction. We achieve this by mapping out the restricted, tier-based energy transfer pathways.

This mode-centered approach has a clear relation to the molecular structure and remains less demanding than long time scale AIMD in terms of computational resources and postprocessing, though it lacks its exploratory nature for events occurring far away from the reference geometries. It is suitable for medium-sized organic molecules, but due to the local nature of these effects, structures that prove too large for anharmonic calculations can likely be cut down without loss of information.

In our example nonstatistical reaction we see that increasing system size does not guarantee either faster relaxation or less localization, and systems of identical size but distinct vibrational modes behave differently as new resonances emerge. Adding a tertbutyl substituent more than doubles the size of the base molecule, yet shows comparable results to the smaller substituents. From a reaction design standpoint, this suggests that the nonstatistical behavior could hold even as part of a much larger carbon backbone - while at the same time being sensitive to nearby heteroatoms.

In the future we aim to explore chemical systems of varying structure, using this model as a screening tool to guide chemical reactivity—by either introducing fragments that behave as nearby "heatsinks" or by breaking up existing, but unwanted, energy transfer pathways. The most significant predictions would then be tested in a synthetic lab. In this way, we hope to expand the applicability of this model from an analytical tool for difficult problems to a method of discovery for synthetic procedures in which the asymmetry of vibrational energy flow guides us toward a desired product.

ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.4c01011.

> Python IVR model, MEP and transition state geometries, GVPT2 results, AIMD trajectories, and their vibrational analysis; and more detailed information on the region assignment and trajectory velocity decomposition schemes (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Gemma C. Solomon − *Nano-Science Center and Department of Chemistry and NNF Quantum Computing Programme, Niels Bohr Institute, University of Copenhagen, DK-2100 Copenhagen, Denmark;* ● orcid.org/0000-0002-2018-1529; Email: gsolomon@chem.ku.dk

Authors

Tomislav Rožić − *Nano-Science Center and Department of Chemistry, University of Copenhagen, DK-2100 Copenhagen, Denmark;* ● orcid.org/0000-0001-7591-1624

Matthew S. Teynor − *Nano-Science Center and Department of Chemistry and NNF Quantum Computing Programme, Niels Bohr Institute, University of Copenhagen, DK-2100 Copenhagen, Denmark;* ● orcid.org/0000-0002-6981-4809

Nađa Došlić − *Department of Physical Chemistry, Ruđer Bošković Institute, HR-10000 Zagreb, Croatia;* ● orcid.org/0000-0001-6535-9020

David M. Leitner − *Department of Chemistry, University of Nevada, Reno, Nevada 89557, United States;* ● orcid.org/0000-0002-3105-818X

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.4c01011

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Levine, I. N. *Physical Chemistry*, 6th ed.; McGraw-Hill, 2009.

(2) Shan, X.; Burd, T. A.; Clary, D. C. New Developments in Semiclassical Transition-State Theory. *J. Phys. Chem. A* **2019**, *123*, 4639−4657.

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

(4) Glowacki, D. R.; Harvey, J. N.; Mulholland, A. J. Taking Ockham's razor to enzyme dynamics and catalysis. *Nat. Chem.* **2012**, *4*, 169−176.

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

(6) Guo, H.; Liu, K. Control of chemical reactivity by transition-state and beyond. *Chemical Science* **2016**, *7*, 3992−4003.

(7) Lee, S.; Goodman, J. M. VRAI-selectivity: Calculation of selectivity beyond transition state theory. *Organic and Biomolecular Chemistry* **2021**, *19*, 3940−3947.

(8) Teynor, M. S.; Scott, W.; Ess, D. H. Catalysis with a Skip: Dynamically Coupled Addition, Proton Transfer, and Elimination during Au- and Pd-Catalyzed Diol Cyclizations. *ACS Catal.* **2021**, *11*, 10179−10189.

(9) Tantillo, D. J. *Beyond transition state theory—Non-statistical dynamic effects for organic reactions*; Academic Press Inc., 2021; Vol. 55; pp. 1−16.

(10) Kabadi, V. N.; Rice, B. M. Molecular Dynamics Simulations of Normal Mode Vibrational Energy Transfer in Liquid Nitromethane. *J. Phys. Chem. A* **2004**, *108*, 532−540.

(11) Rice, S. A.; Dinner, A. R. *Advancing Theory for Kinetics and Dynamics of Complex, Many-Dimensional Systems: Clusters and Proteins*; Komatsuzaki, T.; Berry, R. S.; Leitner, D. M., Eds.; Advances in Chemical Physics; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2011; Vol. 145.

(12) Lourderaj, U.; Park, K.; Hase, W. L. Classical trajectory simulations of post-transition state dynamics. *Int. Rev. Phys. Chem.* **2008**, *27*, 361−403.

(13) Rush, L. A.; Gallo, K. F.; Stumetz, K. S.; Rodríguez-Pérez, I. A.; Cremeens, M. E. Non-statistical dynamics for the allene oxide to cyclopropanone conversion. *J. Phys. Org. Chem.* **2022**, *35*, No. e4385.

(14) Oda, K.; Tsutsumi, T.; Keshavamurthy, S.; Furuya, K.; Armentrout, P. B.; Taketsugu, T. Dynamically Hidden Reaction Paths in the Reaction of CF_3^+ + CO. *ACS Physical Chemistry Au* **2022**, *2*, 388−398.

(15) 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.

(16) Liu, Y.; Holm, S.; Meisner, J.; Jia, Y.; Wu, Q.; Woods, T. J.; Martinez, T. J.; Moore, J. S. Flyby reaction trajectories: Chemical dynamics under extrinsic force. *Science* **2021**, *373*, 208−212.

(17) Tantillo, D. J. Dynamic effects on organic reactivity—Pathways to (and from) discomfort. *J. Phys. Org. Chem.* **2021**, *34*, No. e4202.

(18) Pein, B. C.; Sun, Y.; Dlott, D. D. Controlling Vibrational Energy Flow in Liquid Alkylbenzenes. *J. Phys. Chem. B* **2013**, *117*, 10898−10904.

(19) Pein, B. C.; Dlott, D. D. Modifying Vibrational Energy Flow in Aromatic Molecules: Effects of Ortho Substitution. *J. Phys. Chem. A* **2014**, *118*, 965−973.

(20) Karmakar, S.; Keshavamurthy, S. Intramolecular vibrational energy redistribution and the quantum ergodicity transition: a phase space perspective. *Phys. Chem. Chem. Phys.* **2020**, *22*, 11139−11173.

(21) Leitner, D. M. Quantum ergodicity and energy flow in molecules. *Adv. Phys.* **2015**, *64*, 445−517.

(22) 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.

(23) Chen, Q.; Zhang, S.; Hu, X.; Xie, D.; Guo, H. Reaction Pathway Control via Reactant Vibrational Excitation and Impact on Product Vibrational Distributions: The O + HO2 → OH + O2 Atmospheric Reaction. *J. Phys. Chem. Lett.* **2022**, *13*, 1872−1878.

(24) Witte, T.; Hornung, T.; Windhorn, L.; Proch, D.; de Vivie-Riedle, R.; Motzkus, M.; Kompa, K. L. Controlling molecular ground-state dissociation by optimizing vibrational ladder climbing. *J. Chem. Phys.* **2003**, *118*, 2021−2024.

(25) Kohler, B.; Krause, J. L.; Raksi, F.; Wilson, K. R.; Yakovlev, V. V.; Whitnell, R. M.; Yan, Y. Controlling the Future of Matter. *Acc. Chem. Res.* **1995**, *28*, 133−140.

(26) Dunkelberger, A. D.; Simpkins, B. S.; Vurgaftman, I.; Owrutsky, J. C. Vibration-Cavity Polariton Chemistry and Dynamics. *Annu. Rev. Phys. Chem.* **2022**, *73*, 429−451.

(27) Xiong, W. Molecular Vibrational Polariton Dynamics: What Can Polaritons Do? *Acc. Chem. Res.* **2023**, *56*, 776−786.

(28) Flowers, M. C.; Frey, H. M. Hot molecule effects in the thermal isomerization of methylbicyclo[2.1.0]pent-2-enes. *J. Am. Chem. Soc.* **1972**, *94*, 8636−8637.

(29) Golden, D. M.; Brauman, J. I. Thermal unimolecular isomerization of bicyclo [2.1.0]pent-2-ene. *Trans. Faraday Soc.* **1969**, *65*, 464.

(30) Farneth, W. E.; D'Amore, M. B.; Brauman, J. I. The isomerization of bicyclo[2.1.0]pent-2-enes. *J. Am. Chem. Soc.* **1976**, *98*, 5546−5552.

(31) Goldman, L. M.; Glowacki, D. R.; Carpenter, B. K. Nonstatistical dynamics in unlikely places: [1,5] Hydrogen migration in chemically activated cyclopentadiene. *J. Am. Chem. Soc.* **2011**, *133*, 5312−5318.

(32) Ásgeirsson, V.; Birgisson, B. O.; Bjornsson, R.; Becker, U.; Neese, F.; Riplinger, C.; Jónsson, H. Nudged Elastic Band Method for Molecular Reactions Using Energy-Weighted Springs Combined with Eigenvector Following. *J. Chem. Theory Comput.* **2021**, *17*, 4929−4945.

(33) Neese, F. Software update: The ORCA program system—Version 5.0. *WIREs Comput. Mol. Sci.* **2022**, *12*, No. e1606.

(34) Isobe, H.; Takano, Y.; Kitagawa, Y.; Kawakami, T.; Yamanaka, S.; Yamaguchi, K.; Houk, K. N. Systematic Comparisons between Broken Symmetry and Symmetry-Adapted Approaches to Transition States by Chemical Indices: A Case Study of the Diels-Alder Reactions. *J. Phys. Chem. A* **2003**, *107*, 682−694.

(35) Balcioglu, N.; Özgür Özsar, A. Thermal conversion of 1,3-hexadien-5-yne to benzene: a revisited theoretical study. *Journal of Molecular Structure: THEOCHEM* **2004**, *677*, 125−132.

(36) Oviedo, M. B.; Ilawe, N. V.; Wong, B. M. Polarizabilities of π-Conjugated Chains Revisited: Improved Results from Broken-Symmetry Range-Separated DFT and New CCSD(T) Benchmarks. *J. Chem. Theory Comput.* **2016**, *12*, 3593−3602.

(37) Skraba-Joiner, S. L.; Johnson, R. P.; Agarwal, J. Dehydropericyclic Reactions: Symmetry-Controlled Routes to Strained Reactive Intermediates. *Journal of Organic Chemistry* **2015**, *80*, 11779−11787.

(38) Kedziora, G. S.; Barr, S. A.; Berry, R.; M?ller, J. C.; Breitzman, T. D. Bond breaking in stretched molecules: multi-reference methods versus density functional theory. *Theor. Chem. Acc.* **2016**, *135*, 79.

(39) Hamaguchi, M.; Nakaishi, M.; Nagai, T.; Nakamura, T.; Abe, M. Notable Effect of an Electron-Withdrawing Group at C3 on the Selective Formation of Alkylidenecyclobutanes in the Thermal Denitrogenation of 4-Spirocyclopropane-1-pyrazolines. Nonstatistical Dynamics Effects in the Denitrogenation Reactions. *J. Am. Chem. Soc.* **2007**, *129*, 12981−12988.

(40) Malrieu, J.-P.; Trinquier, G. A Recipe for Geometry Optimization of Diradicalar Singlet States from Broken-Symmetry Calculations. *J. Phys. Chem. A* **2012**, *116*, 8226−8237.

(41) Mališ, M.; Loquais, Y.; Gloaguen, E.; Jouvet, C.; Brenner, V.; Mons, M.; Ljubić, I.; Došlić, N. Non-radiative relaxation of UV photoexcited phenylalanine residues: probing the role of conical intersections by chemical substitution. *Phys. Chem. Chem. Phys.* **2014**, *16*, 2285.

(42) Barbara, P. F.; Meyer, T. J.; Ratner, M. A. Contemporary Issues in Electron Transfer Research. *J. Phys. Chem.* **1996**, *100*, 13148−13168.

(43) Kuzmin, M.; Letokhov, V.; Stuchebryukhov, A. Threshold energy dependence of the intramolecular vibrational relaxation rate tor an isolated polyatomic molecule. *Soviet Phys.—JETP* **1986**, *63*, 264−271.

(44) Leitner, D. M.; Pandey, H. D. Asymmetric energy flow in liquid alkylbenzenes: A computational study. *J. Chem. Phys.* **2015**, *143*, 144301.

(45) Yu, X.; Leitner, D. M. Vibrational energy transfer and heat conduction in a protein. *J. Phys. Chem. B* **2003**, *107*, 1698−1707.

(46) Cui, Q.; Bahar, I., Eds. *Normal Mode Analysis*; Chapman and Hall/CRC, 2005; Chapter 15−16.

(47) Barone, V. Vibrational zero-point energies and thermodynamic functions beyond the harmonic approximation. *J. Chem. Phys.* **2004**, *120*, 3059−3065.

(48) 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 16, Revision A.03*; Gaussian Inc.: Wallingford CT, 2016.

(49) Teynor, M. S.; Wohlgemuth, N.; Carlson, L.; Huang, J.; Pugh, S. L.; Grant, B. O.; Hamilton, R. S.; Carlsen, R.; Ess, D. H. *Milo, Revision 1.1.0*; Brigham Young University: Provo, UT, 2022.

(50) Yanai, T.; Tew, D. P.; Handy, N. C. A new hybrid exchange−correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51−57.

(51) 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.

(52) Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057.

(53) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(54) Sitkiewicz, S. P.; Zaleśny, R.; Ramos-Cordoba, E.; Luis, J. M.; Matito, E. How Reliable Are Modern Density Functional Approximations to Simulate Vibrational Spectroscopies? *J. Phys. Chem. Lett.* **2022**, *13*, 5963−5968.

(55) Mitra, H.; Roy, T. K. Comprehensive Benchmark Results for the Accuracy of Basis Sets for Anharmonic Molecular Vibrations. *J. Phys. Chem. A* **2020**, *124*, 9203−9221.

(56) Head-Gordon, M.; Head-Gordon, T. Analytic MP2 frequencies without fifth-order storage. Theory and application to bifurcated hydrogen bonds in the water hexamer. *Chem. Phys. Lett.* **1994**, *220*, 122−128.

(57) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.

(58) Barone, V.; Ceselin, G.; Fusè, M.; Tasinato, N. Accuracy Meets Interpretability for Computational Spectroscopy by Means of Hybrid and Double-Hybrid Functionals. *Front. Chem.* **2020**, *8*, No. 584203.

(59) Park, S.-M.; Nguyen, P. H.; Stock, G. Molecular dynamics simulation of cooling: Heat transfer from a photoexcited peptide to the solvent. *J. Chem. Phys.* **2009**, *131*, 184503.

(60) Stuchebrukhov, A. A.; Marcus, R. A. Theoretical study of intramolecular vibrational relaxation of acetylenic CH vibration for v = 1 and 2 in large polyatomic molecules (CX3)3YCCH, where X = H or D and Y = C or Si. *J. Chem. Phys.* **1993**, *98*, 6044−6061.

(61) Swamy, K. N.; Hase, W. L. The heavy-atom effect in intramolecular vibrational energy transfera). *J. Chem. Phys.* **1985**, *82*, 123−133.

(62) Nejad, A.; Meyer, E.; Suhm, M. A. Glycolic Acid as a Vibrational Anharmonicity Benchmark. *J. Phys. Chem. Lett.* **2020**, *11*, 5228−5233.

(63) Longfellow, R. J.; Parmenter, C. S. How a (nearly) free methyl rotor accelerates intramolecular vibration relaxation. Theory and Experiment. *Journal of the Chemical Society, Faraday Transactions* **1988**, *84*, 1499.

Supporting Information:

A strategy for modeling non-statistical reactivity effects: combining chemical activation estimates with a vibrational relaxation model

Tomislav Rožić,[†] Matthew Teynor,[†,‡] Nađa Došlić,[¶] David M. Leitner,[§] and Gemma C. Solomon[*,†,‡]

[†]Nano-Science Center & Department of Chemistry, University of Copenhagen, DK-2100 Copenhagen, Denmark

[‡]NNF Quantum Computing Programme, Niels Bohr Institute, University of Copenhagen, DK-2100 Copenhagen, Denmark

[¶]Department of Physical Chemistry, Ruđer Bošković Institute, HR-10000 Zagreb, Croatia

[§]Department of Chemistry, University of Nevada, Reno, Nevada 89557, USA

E-mail: gsolomon@chem.ku.dk

Contents

1   Data repository

All relevant data is hosted on ERDA - the Electronic Research Data Archive at the University of Copenhagen. A frozen repository can be found at: `https://erda.ku.dk/archives/1e66701d9bc1c695d792b1de909921aa/published-archive.html`

In this data repository, we include output files for all used Gaussian GVPT2 calculations (frequencies and couplings on BS-DFT level with and without solvent as well as vacuum MP2), trajectories from BS-DFT AIMD simulations with Milo + Gaussian, and the BS-DFT NEB-CI reaction paths obtained with ORCA. We also include the Jupyter notebook used for trajectory analysis and energy estimation, the .csv files containing these energy estimates for the different approaches (TS projection, single 0K trajectory, and ensemble of trajectories), and the Python code that contains the IVR model - including the master equation, indicator/lifetime and different plotting options.

All uploaded data is also split into three compressed archives (easily found with sort by Size): all_md_data.zip, ivr_model_final.zip and orca_bs_dft.tar.gz

2   Computed indicator and lifetimes at different levels of theory

2.1   Harmonic frequencies and implicit solvent modeling

Using harmonic frequencies obtained by BS-DFT, instead of the GVPT2 corrected ones at the same level of theory, did not qualitatively change the results. We then included implicit solvation for tetrahydrofuran for the 0K pushed trajectory approach only, as we deemed a full ensemble of trajectories not necessary to compare to vacuum results, as the ensemble results already resemble single trajectory ones.

We note only one major difference, the **MeCP** system where the trend for the three middle systems is disrupted in the case of the regional indicator (Figure 1b; second plot).
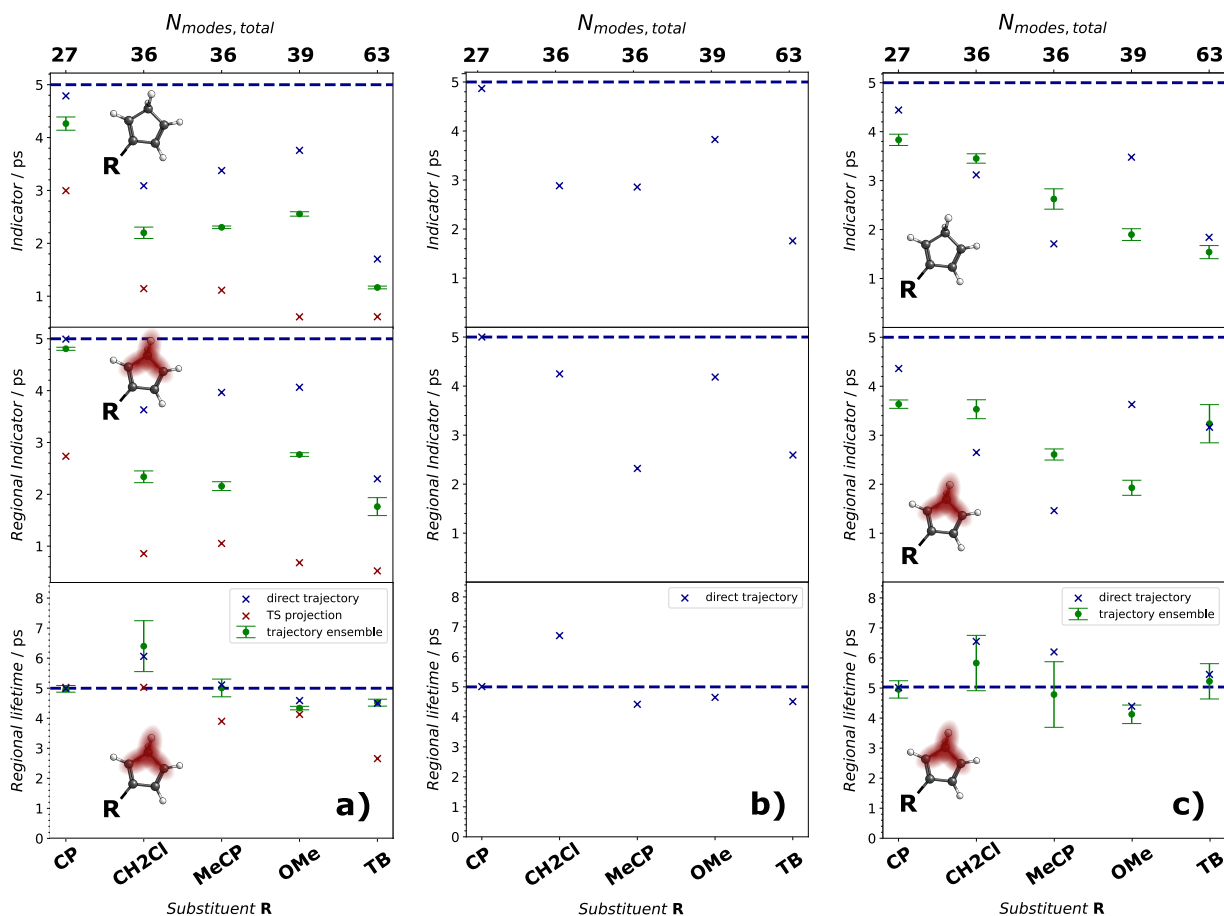
Figure S1: Computed indicator, regional indicator and regional lifetime values for the five differently substituted systems using a) harmonic CAM-B3LYP/def2-TZVP D3BJ frequencies with no GVPT2 correction b) harmonic frequencies at the same DFT level of theory using implicit solvent modeling for tetrahydrofuran and c) GVPT2 corrected frequencies at the MP2/cc-pVTZ (for CP–OMe) and MP2/cc-pVDZ (TB) level of theory.

We initially thought this discrepancy could be caused by the different assignments of modes to the regions, as including the implicit solvent model not only shifts the vibrational frequencies but may change the normal modes themselves. For example, a mode that we would assign to the **3C** region may have some contribution from the substituent atoms as well. Depending on the size of the substituent, this contribution may increase to the point that over 50% of the norm of the vector lies outside the chosen region and the mode would be excluded. In the **CH2Cl** example, we found one such mode whose norm dropped to 45%, however explicitly

including it back did not change the results in any substantial way. In the **MeCP** system, which shows the most significant difference, there was no mismatch and both calculations had the same 8 modes assigned to the **3C** region, meaning this difference compared to vacuum results does follow from the difference in computed modes and couplings. Using a more flexible measure of region assignment or even a mode localization scheme could help eliminate potential doubts in the future.

2.2 Vacuum MP2 results

As performing AIMD at the MP2/cc-pVTZ level was not feasible, both due to cost concerns and the method failing near the TS geometry, we used existing (vacuum) AIMD results from UKS DFT for the energy estimation. We projected these existing trajectories onto MP2/cc-pVTZ normal modes and couplings (MP2/cc-pVDZ for tert-butyl, TB, due to its size). Accordingly, when choosing the trajectory point at which the vibrational energy projection would be performed we only consider the smallest projected vibrational energy within 50 fs and not the sum of the projected estimate and computed electronic potential energy, as was the case previously.

While the GVPT2 corrected frequencies and corresponding couplings obtained with MP2/cc-pVTZ should be more accurate, the mismatch of electronic structure methods does lead to a larger standard deviation of the ensemble results and a poorer match between the ensemble and single 0K pushed trajectory results, making it less reliable. We still see it as a useful exercise of the robustness of mixing methods, especially as the system size determines which half of the procedure becomes the computational bottleneck (the scaling of GVPT2 will at some point become more expensive than a 50-100 fs AIMD trajectory).

Compared to the DFT results, the all-atom and regional indicators (first two plots of Figure 1c) both show a lower value for **MeCP**, similar to the solvent results (though no solvation was applied). However, these indicator values are sensitive to changes to major

pathways, whether the pathways are "used" or not – but they do not necessarily translate to actual IVR. If we were to use lifetime values from the actual master equation simulations (last row of plots), the conclusion would stay the same as from DFT: the energy stays localized the longest in CH2Cl, followed by MeCP, TB, and finally OMe, with CP sitting tightly at the solvent limit. That is, in the context of "hits", both cases suggest exploring OMe for a more statistical reaction and CH2Cl for one that behaves even more non-statistically.

3 Minimal energy path (MEP) to the second product and associated barrier heights
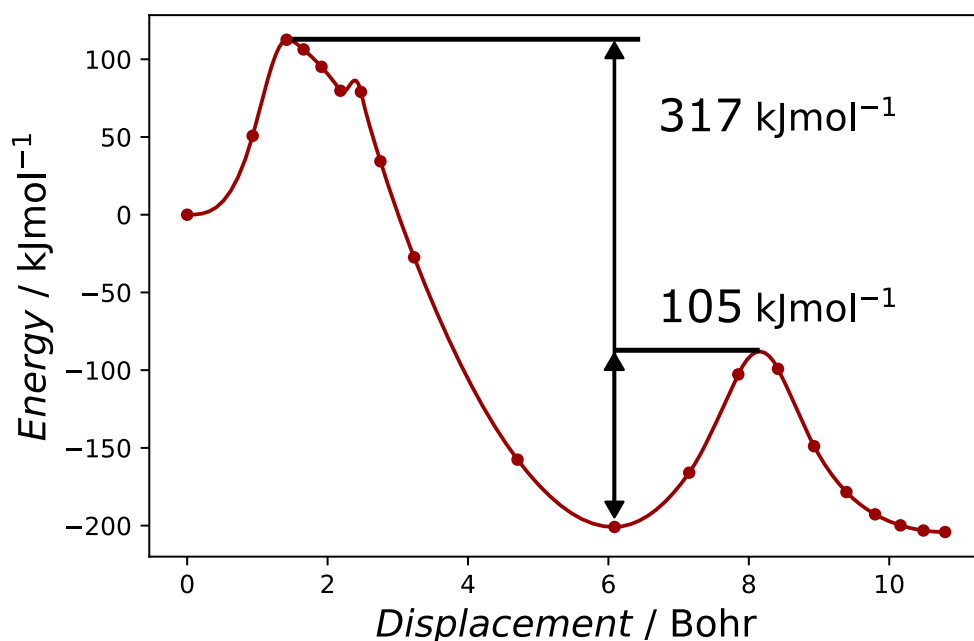
Figure S2: The final minimal energy path for the full ring-opening reaction followed by the hydrogen rearrangement with no transition state or intermediate guess provided to the algorithm, calculated at the UCAM-B3LYP D3BJ def2-TZVP level of theory.

While calculating a MEP directly to the second product (**1-MeCP**), we again obtain the MEP to **2-MeCP**, followed by the smaller re-arrangement barrier to the second product, as seen in Figure S2. This re-assures us that **2-MeCP** is a clear intermediate product in this

two-step reaction and that our choice of method for the potential surface of the molecular dynamics is satisfactory as both barriers are well reproduced: the drop from the ring opening barrier transition state with an experimental[S1] value of 310 kJ mol^{-1} is calculated to be 317 kJ mol^{-1} without the zero-point energy correction and 305 kJ mol^{-1} with the correction. The barrier for the hydrogen rearrangement has an experimental enthalpy of 99 kJ mol^{-1}, which we obtain as 105 kJ mol^{-1} without and 107 kJ mol^{-1} with the zero-point energy correction.

We additionally compute unrestricted DLPNO-CCSD(T)-F12 energies with the cc-pVTZ-F12 basis set and a broken-symmetry guess at the MEP geometries obtained by DFT. The default ORCA5 RI-J options and convergence criteria were used along with the cc-pVTZ-F12-CABS cc-PVQZ/C auxiliary basis sets. Accompanying the energy profile we also obtained the associated T_1 diagnostic, which serves as a useful tool for diagnosing difficult cases.

From Figure S3 we see that the relative heights of the barriers are replicated and the T_1 diagnostic does increase near the first barrier, believed to have some singlet diradical character. However, only one of the points of our MEP crosses the dotted red line signifying a "rule of thumb" value of 0.02, where a single reference result can become unreliable. We concede the shoulder in the BS-DFT PES (at around 2.5 Bohr in Figure 2) could be erroneous. This could influence the AIMD trajectories, though the deviation seems small compared to the size of the barrier. We would consider this limitation of the PES a greater issue in cases where the reaction was enabled by dynamic matching/mode matching, which does not seem to be the case for this hydrogen migration.

4 Definition of a region

To automatically select a grouping of vibrational modes that may be of interest to the reaction we first select a group of atoms that form the reaction center/region of the molecule.
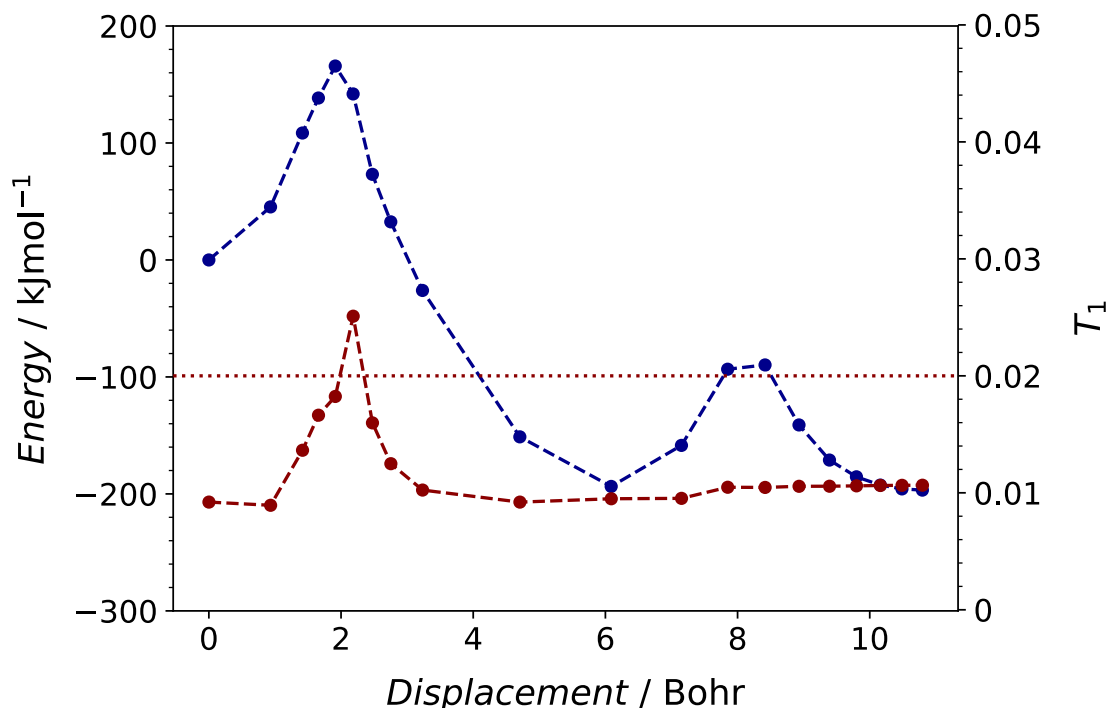
Figure S3: DLPNO-CCSD(T)-F12 / cc-pVTZ-F12 energies (blue) at the geometries from Figure S2 and the computed T_1 diagnostic (red). The dotted red line represents the rule of thumb value of 0.02 for when the result could become unreliable.

In our case, we chose the three carbon atoms that formed the smaller ring that breaks and the two hydrogen atoms that can re-arrange to form the second product, as highlighted in Figure 6 in the main text.

We then sum up the norm contributions at these atoms of interest for each vibrational normal mode vector ($\boldsymbol{v}_\alpha$):

$$norm_{\alpha,region} = \sum_{atom\ i}^{region\ atoms} (\boldsymbol{v}_{\alpha,i,x}^2 + \boldsymbol{v}_{\alpha,i,y}^2 + \boldsymbol{v}_{\alpha,i,z}^2) = \sum_{atom\ i}^{region\ atoms} \boldsymbol{v}_{\alpha,i} \cdot \boldsymbol{v}_{\alpha,i} \tag{S1}$$

Since the full norm (including all atoms) of a given mode is 1, we chose 0.5 (50%) as our cut-off value for considering the mode localized to the region of choice. If "global" modes were to be excluded as a special group of its own, this criterion should be higher (eg. 80%).

The approach could be made more automatic with a robust procedure of matching modes to the reaction coordinate. One could also consider using these norms as a scaling factor - to compute indicators or plots of interests using the norm value, as opposed to simply completely including (100%) or excluding (0%) the given mode's result.

5 Velocity Decomposition Overview

For a given step of a trajectory, we want to know how much kinetic and potential energy are in each vibrational mode. To calculate this, we must first remove the translational and rotational motion from the atomic positions and velocities of the trajectory step. Then we can project the new atomic positions and velocities onto the normal mode displacement vectors to get the mode positions and velocities. This step can be thought of as a change of basis from Cartesian coordinates per atom to normal mode coordinates, which is possible because the normal modes form an orthonormal basis. We can then calculate the kinetic energy per mode from the reduced masses and mode velocities, as well as the harmonic potential energy from the force constants and mode positions.

This analysis relies on the normal mode displacement vectors, which are only valid for stationary point geometries. Unfortunately, trajectories are very unlikely to sample stationary points on high dimensional potential energy surfaces, especially for reactive trajectories where we start at a transition state, so this analysis is only approximate. We can choose the best point in the trajectory for this analysis by comparing the potential energy given by the electronic structure program to the harmonic potential energy calculated by this analysis. The best step will be the one with the lowest percent error between these values.

6 Velocity Decomposition Details

6.1 Removing translational motion

We can remove the translational motion by simply calculating the center of mass position and center of mass velocity and subtracting them from the positions, $\boldsymbol{r}$, and velocities, $\boldsymbol{v}$. So the centered positions, $\boldsymbol{r}^m$, and rovibrational velocities, $\boldsymbol{v}_{rovib}$, for atom i are given by

$$\boldsymbol{r}_i^m = \boldsymbol{r}_i - \frac{1}{M}\sum_{j=1}^{N} m_j \boldsymbol{r}_j \, , \text{ and} \tag{S2}$$

$$\boldsymbol{v}_{rovib,i} = \boldsymbol{v}_i - \frac{1}{M}\sum_{j=1}^{N} m_j \boldsymbol{v}_j \, , \tag{S3}$$

where M is the total mass and N is the number of atoms.

6.2 Removing rotational velocity

To calculate the rotational velocity, we first calculate the angular momentum by

$$\boldsymbol{L} = \sum_{j=1}^{N} m_j \boldsymbol{r}_i^m \times \boldsymbol{v}_j \, . \tag{S4}$$

We then calculate the moment of inertia tensor

$$\boldsymbol{I} = \sum_{j=1}^{N} m_j \begin{bmatrix} y_j^2 + z_j^2 & -x_j y_j & -x_j z_j \\ -x_j y_j & x_j^2 + z_j^2 & -y_j z_j \\ -x_j z_j & -y_j z_j & x_j^2 + y_j^2 \end{bmatrix} \, . \tag{S5}$$

This allows us to solve the linear equation

$$\boldsymbol{L} = \boldsymbol{I}\boldsymbol{\omega} \, . \tag{S6}$$

for the angular velocity, $\boldsymbol{\omega}$. We can then calculate the rotational velocity,

$$\boldsymbol{v}_{rot,i} = \boldsymbol{\omega} \times \boldsymbol{r}_i^m \,, \tag{S7}$$

and finally the desired vibrational velocity,

$$\boldsymbol{v}_{vib,i} = \boldsymbol{v}_{rovib,i} - \boldsymbol{v}_{rot,i} \,. \tag{S8}$$

6.3   Removing rotations from positions (Kabsch algorithm)

We now need to align the positions for the desired trajectory step with the centered positions of the reference geometry, $\boldsymbol{r}_{ref}$, that the normal modes are based on. This is equivalent to removing accumulated rotational motion from the positions. We start by calculating the 3×3 covariance matrix

$$\boldsymbol{H} = \boldsymbol{r}_{ref}{}^{\boldsymbol{T}} \boldsymbol{r}^m \,, \tag{S9}$$

where $\boldsymbol{T}$ denotes the transpose. We then decompose the covariance matrix through singular value decomposition,

$$\boldsymbol{H} = \boldsymbol{U\Sigma V}^{\boldsymbol{T}} \,. \tag{S10}$$

This allows us to calculate the rotation matrix that aligns the positions of the trajectory steps to the reference positions,

$$\boldsymbol{R} = \boldsymbol{V} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & d \end{bmatrix} \boldsymbol{U}^{\boldsymbol{T}} \,, \tag{S11}$$

where d is chosen to be 1 or -1 to ensure the coordinate system is right-handed. The aligned positions, $\boldsymbol{r}^a$, are then given by

$$\boldsymbol{r}^a = \boldsymbol{r}^m \boldsymbol{R}\,. \tag{S12}$$

We also need to align the vibrational velocities by

$$\boldsymbol{v}^a_{vib} = \boldsymbol{v}_{vib}\boldsymbol{R}\,. \tag{S13}$$

6.4 Changing basis from atomic Cartesian to normal mode coordinates

We need to define the change of basis from Cartesian coordinates per atom to normal mode coordinates. Let the normal mode displacement vectors be defined in the matrix

$$\boldsymbol{M} = \begin{pmatrix} x_{1,1} & x_{2,1} & \cdots & x_{3N-6,1} \\ y_{1,1} & y_{2,1} & & y_{3N-6,1} \\ z_{1,1} & z_{2,1} & & z_{3N-6,1} \\ x_{1,2} & x_{2,2} & & x_{3N-6,2} \\ y_{1,2} & y_{2,2} & & y_{3N-6,2} \\ z_{1,2} & z_{2,2} & & z_{3N-6,2} \\ \vdots & & \ddots & \vdots \\ x_{1,N} & x_{2,N} & & x_{3N-6,N} \\ y_{1,N} & y_{2,N} & & y_{3N-6,N} \\ z_{1,N} & z_{2,N} & \cdots & z_{3N-6,N} \end{pmatrix}\,. \tag{S14}$$

If we flatten the aligned coordinate matrix (either position or velocity) into a single vector,

$$Q_{atomic} = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ x_2 \\ y_2 \\ z_2 \\ \vdots \\ x_N \\ y_N \\ z_N \end{pmatrix}, \tag{S15}$$

and define the normal mode coordinates as

$$Q_{mode} = \begin{pmatrix} \nu_1 \\ \nu_2 \\ \vdots \\ \nu_{3N-6} \end{pmatrix}, \tag{S16}$$

then we have the relation

$$Q_{atomic} = M Q_{mode}. \tag{S17}$$

We can then solve this equation for the normal mode coordinates using a least squares regression. It is worth noting that because the normal mode matrix $\boldsymbol{M}$ is unitless, the normal mode coordinates will have the same units as the atomic coordinates.

6.5 Calculating vibrational energy per normal mode

We then use this coordinate transfer to get the normal mode displacements, $\boldsymbol{d}_{mode}$, and velocities, $\boldsymbol{v}_{mode}$, according to

$$(\boldsymbol{r}^a - \boldsymbol{r}_{ref})^F = \boldsymbol{M}\boldsymbol{d}_{mode}\,, \text{ and} \tag{S18}$$

$$(\boldsymbol{v}^a_{vib})^F = \boldsymbol{M}\boldsymbol{v}_{mode}\,, \tag{S19}$$

where $()^F$ denotes flattening the matrix into a one-dimensional vector. The energy for vibrational mode i is then

$$E_i = \frac{1}{2}k_i\boldsymbol{d}^2_{mode,i} + \frac{1}{2}\mu_i\boldsymbol{v}^2_{mode,i}\,, \tag{S20}$$

where k_i and μ_i are the corresponding force constant and reduced mass.

References

(S1) Goldman, L. M.; Glowacki, D. R.; Carpenter, B. K. Nonstatistical dynamics in unlikely places: [1,5] Hydrogen migration in chemically activated cyclopentadiene. *Journal of the American Chemical Society* **2011**, *133*, 5312–5318.

7 Computed frequencies

TB				
CAM-B3YLP/def2-TZVP D3BJ		DFT + CPCM(THF)	MP2/cc-pVDZ	
harmonic	anharmonic	harmonic	harmonic	anharmonic
3245	3128	3243	3268	3127
3242	3084	3239	3266	3167
3217	3097	3216	3240	3110
3127	2995	3124	3173	3031

3126	3001	3123	3173	3038
3123	3003	3121	3170	3022
3122	2983	3121	3167	3034
3116	2963	3116	3161	3016
3116	2995	3114	3160	3023
3070	2924	3070	3118	2967
3058	2923	3056	3069	2960
3051	2897	3049	3068	2933
3050	2895	3048	3065	2957
3045	2918	3045	3064	2953
1704	1669	1699	1647	1603
1617	1581	1612	1568	1528
1527	1490	1519	1513	1472
1511	1475	1504	1497	1459
1511	1473	1503	1496	1456
1497	1457	1490	1481	1442
1494	1457	1488	1477	1440
1487	1446	1484	1471	1433
1439	1414	1434	1416	1384
1432	1393	1424	1410	1378
1414	1384	1410	1395	1352
1409	1382	1406	1385	1353
1407	1380	1404	1385	1353
1343	1311	1341	1330	1292
1299	1266	1297	1284	1249
1289	1258	1288	1264	1224
1253	1219	1250	1250	1216
1244	1215	1242	1230	1201
1161	1139	1157	1148	1124
1143	1115	1140	1111	1082
1105	1089	1102	1090	1074
1059	1037	1057	1046	1023
1053	1031	1051	1044	1022
1007	991	1002	1010	993
980	972	985	965	948
971	959	971	959	941
963	949	962	954	938
959	941	958	953	935
951	938	950	924	905
946	921	946	911	896
925	911	923	909	884
836	827	835	826	811

830	816	830	810	800
797	788	796	758	751
737	730	737	705	688
656	648	655	632	622
585	587	586	576	571
473	471	472	465	461
428	430	430	412	413
374	381	375	356	365
345	348	346	341	340
335	335	335	328	328
320	314	321	317	325
296	313	296	303	292
274	269	277	288	278
228	204	231	237	220
211	216	211	208	209
141	138	142	137	136
67	68	69	67	64

CP					
CAM-B3YLP/def2-TZVP D3BJ		DFT + CPCM(THF)		MP2/cc-pVTZ	
harmonic	anharmonic	harmonic	anharmonic	harmonic	anharmonic
3250	3126	3248	3122	3271	3162
3244	3105	3241	3105	3264	3143
3226	3106	3224	3103	3247	3123
3217	3099	3215	3092	3237	3112
3074	2928	3075	2929	3110	2964
3048	2921	3049	2921	3067	2937
1676	1641	1670	1636	1614	1571
1593	1554	1588	1549	1536	1491
1431	1401	1425	1395	1418	1385
1416	1377	1407	1367	1404	1358
1340	1310	1339	1308	1321	1289
1290	1257	1290	1254	1264	1223
1144	1128	1139	1106	1129	1110
1141	1113	1138	1121	1123	1091
1124	1107	1117	1099	1109	1089
1026	1008	1023	1004	1033	1011
984	969	988	965	989	971
984	972	982	966	947	929
977	973	979	963	944	927
942	927	939	925	941	927
931	910	929	906	912	884
832	821	831	819	811	800
829	820	828	818	807	796
733	727	733	716	706	697
689	680	685	668	681	670
534	531	536	532	516	512
353	353	355	355	334	332

CH2Cl					
CAM-B3YLP/def2-TZVP D3BJ		DFT + CPCM(THF)		MP2/cc-pVTZ	
harmonic	anharmonic	harmonic	anharmonic	harmonic	anharmonic
3248	3107	3246	3099	3266	3147
3233	3103	3234	3111	3250	3117
3224	3106	3221	3099	3241	3116
3169	3030	3182	3041	3186	3044
3108	2996	3119	3005	3117	3001
3075	2929	3076	2930	3108	2961
3049	2923	3050	2923	3065	2935
1711	1673	1705	1667	1649	1605
1620	1583	1615	1576	1560	1516
1494	1457	1489	1451	1491	1451
1433	1405	1429	1399	1420	1387
1419	1377	1410	1368	1408	1362
1331	1299	1331	1299	1312	1283
1310	1286	1307	1283	1297	1265
1284	1255	1284	1247	1264	1230
1243	1215	1242	1214	1237	1206
1177	1154	1175	1143	1172	1146
1145	1116	1143	1111	1124	1094
1116	1099	1111	1090	1103	1083
1010	994	1007	996	1013	995
982	972	986	967	957	941
959	941	957	944	942	936
955	944	954	934	942	920
945	925	944	923	918	894
913	898	910	891	910	891
836	826	834	827	816	805
802	792	801	791	785	771
741	729	738	730	735	723
722	709	702	691	709	697
676	667	674	658	665	652
583	576	582	570	570	564
382	384	382	389	368	370
334	331	332	351	323	319
286	284	289	294	279	275
127	127	126	122	124	123
69	68	70	82	70	68

MeCP					
CAM-B3YLP/def2-TZVP D3BJ		DFT + CPCM(THF)		MP2/cc-pVTZ	
harmonic	anharmonic	harmonic	anharmonic	harmonic	anharmonic
3242	3102	3240	3097	3261	3144
3228	3096	3225	3104	3246	3093
3211	3089	3209	3085	3229	3106
3134	2998	3133	2997	3168	3029
3097	2959	3097	2959	3138	3000
3070	2924	3070	2924	3106	2959
3049	2956	3049	2952	3066	2967
3045	2920	3046	2919	3064	2934
1716	1677	1711	1674	1658	1615
1618	1580	1612	1576	1559	1518
1499	1463	1490	1451	1506	1464
1489	1454	1480	1439	1496	1456
1434	1406	1429	1396	1421	1369
1429	1385	1421	1384	1415	1388
1409	1379	1405	1368	1396	1362
1321	1291	1319	1288	1296	1262
1282	1253	1281	1249	1261	1227
1235	1207	1232	1204	1230	1196
1142	1113	1140	1108	1123	1091
1116	1099	1110	1092	1103	1082
1079	1055	1077	1050	1067	1039
1032	1016	1029	1010	1028	1008
987	973	984	969	986	969
978	971	983	964	954	937
957	942	955	939	940	926
942	922	942	915	918	901
927	913	925	910	915	888
835	825	833	822	815	804
769	759	768	745	745	734
720	711	719	702	698	686
629	624	628	620	619	612
588	582	588	583	566	561
374	372	376	374	355	354
327	329	327	330	316	321
236	257	236	259	230	258
161	124	163	106	158	116

OMe					
CAM-B3YLP/def2-TZVP D3BJ		DFT + CPCM(THF)		MP2/cc-pVTZ	
harmonic	anharmonic	harmonic	anharmonic	harmonic	anharmonic
3263	3136	3262	3139	3283	3156
3250	3105	3247	3100	3270	3158
3229	3110	3226	3107	3247	3122
3162	3027	3167	3029	3202	3065
3091	2958	3101	2962	3131	2996
3067	2923	3068	2924	3103	2957
3043	2919	3044	2921	3062	2934
3033	2833	3041	2951	3054	2924
1706	1668	1697	1658	1656	1615
1625	1587	1618	1580	1575	1533
1513	1476	1505	1464	1527	1487
1495	1454	1488	1437	1506	1467
1486	1452	1485	1448	1480	1446
1443	1417	1435	1412	1432	1402
1432	1390	1424	1381	1422	1377
1331	1300	1325	1291	1308	1276
1287	1255	1283	1248	1274	1239
1282	1256	1281	1256	1260	1223
1217	1194	1215	1186	1214	1190
1188	1165	1186	1156	1193	1167
1146	1119	1143	1115	1128	1097
1122	1102	1114	1096	1108	1084
1079	1058	1070	1049	1072	1049
1020	1004	1015	997	1025	1007
971	972	974	970	953	935
955	940	953	937	934	925
945	932	942	930	933	917
938	918	935	915	911	885
825	816	823	815	804	795
790	781	788	776	761	752
704	694	703	680	675	663
661	653	660	652	649	641
633	625	632	620	620	610
470	464	468	462	466	460
395	393	394	393	380	381
291	318	290	314	287	322
250	264	250	242	253	247
213	169	213	131	223	177
125	118	123	100	128	122

B

Paper 2

<u>Tomislav Rožić</u>, Yuxuan Hou, Lea Kjærgaard Northcote, Christian Markus Pedersen, and Gemma C. Solomon. Tutorial Review on Modeling Nonstatistical Reactivity: an Example of Light and Heat in the Garratt–Braverman/[1,5]-H Shift of Ene-diallenes. *in preparation*

Tutorial Review on Modeling Nonstatistical Reactivity: an Example of Light and Heat in the Garratt–Braverman/[1,5]-H Shift of Ene-diallenes

Tomislav Rožić[a],[†],[‡] Yuxuan Hou[a],[†] Lea Kjærgaard Northcote,[†],[¶] Christian Marcus Pedersen,[†] and Gemma C. Solomon[*],[†],[‡],[¶]

†*Department of Chemistry, University of Copenhagen, Denmark*
‡*Nano-Science Center, University of Copenhagen, Denmark*
¶*NNF Quantum Computing Programme, Niels Bohr Institute, University of Copenhagen, Denmark*

E-mail: gsolomon@chem.ku.dk

[a] These authors contributed equally.

Abstract

Broad terms in the literature, such as nonstatistical reactivity or nontraditional luminescence, emerge when standard theories fail to explain experimental results. In the case of nonstatistical and dynamic effects, reaction rates and product ratios may vary wildly from transition state theory (TST) predictions, and are often signaled by a lack of temperature dependence.

In this tutorial, we explain how to use modern and freely available computational chemistry tools to model a reported nonstatistical reaction of a relatively large structure, the thermal Garratt–Braverman/[1,5]-H shift of an ene-diallene, such that a non-expert could easily do the same following our steps.

As a team of synthetic organic chemists and computational chemists, we also hope to encourage the use of preparatory computational work that may aid in reaction design during the experimental process, not just as complementary data to finished experimental studies.

Through this approach, we discover that the thermal Garratt–Braverman/[1,5]-H shift exhibits a parallel light-enabled reaction that bypasses the rate-limiting first step. Additionally, when tunneling effects are accounted for, TST predictions return to realistic values, only to be disproved again by careful variable temperature experiments.

As the motifs of reactive $\pi - \pi^*$ absorptions, hydrogen transfers, and diradical intermediates are quite common, the points made in this paper are general and indicative of the underlying complexity behind many chemical reactions that exhibit unexpected rates and ratios. The failure of TST also serves as a warning to massive reaction network discovery schemes that heavily rely on calculated ground state activation energies and the dangerous simplicity of the conventional free energy diagram.

Introduction

Computational chemistry has served as a perfect aid for organic chemists to understand and explain their experimental results for decades. Yet it is our impression that it has not been leveraged enough while planning for experiments or refining the synthetic strategy. Additionally, as we apply these tools, we often fall prey to picking those that best fit the experimental results, starting with the tools that we are most used to.

On the other hand, the new and improved

computational models offered to us by method developers often do not reach widespread use, as the improvement in accuracy or cost-efficiency does not match the implied cost of adoption for non-expert users. Another major filter in the spread of these new computational approaches is their adoption in the top quantum chemistry software packages.

A representative example is the "zoo" of available functionals for density functional theory (DFT). To add to the confusion, many of these functionals perform similarly. It is possible to pick out cases where a simpler and computationally cheaper approach can result in considerable savings in computational time and a more accurate result.[1,2] Small basis sets, such as the popular 3-21G and 6-31G* Pople basis sets are still in common use, likely due to inertia and especially in fields dealing with large structures, as well as due to a historical notion that a smaller basis might lead to fortuitous error cancellations.

Here it is important to note that modern DFT functionals perform best with larger basis sets, in the same manner as the default integration grids which have grown over the years; a starting example being the triple zeta def2-TZVP and pc-2 sets and their twins with added diffuse functions, def2-TZVPD and aug-pc-2.[2]

In addition to following best practices for calculating the electronic structure of molecules, we also make textbook assumptions about how the reaction proceeds. Transition state theory (TST) has been one such cornerstone assumption, used to estimate reaction rates and product ratios.[3] It is simple, as it only requires us to determine the transition states (TS) for each relevant reaction step, eliminating the complexity of the highly dimensional and complex potential energy surface (PES) that they inhabit. Therefore, it is a reasonable first choice when modeling the kinetics of a chemical reaction.

However, it is not sufficient for even some of the most basic reactions, one example being the hydroboration of alkenes, which gives a higher ratio for the Markovnikov product experimentally (anti-Markovnikov product:Markovnikov product = 90:10) compared to a TST prediction (99:1–98:2).[4] In these cases, labeled as nonstatistical reactions, we observe kinetics and product ratios that do not match the results of statistical models, such as TST or Rice–Ramsperger–Kassel–Marcus (RRKM) theory. In these theories, it is assumed either that the molecule is thermalized (equilibrated) before any reaction step proceeds (for TST) - or if the molecule is vibrationally "hot", that this vibrational excitation is distributed statistically across the molecule (for RRKM). Therefore, with such reactions, no amount of increased cost and accuracy in calculating the transition state geometries and their activation energies will provide us with a model that explains what we observe in an experiment.

Failures of these theories are interesting, as many reactions are reproduced well with TST. They also serve as a warning that the free energy diagrams we all rely on, while immensely useful as a visual summary of a reaction, do not automatically correspond to the reaction dynamics – especially when TST estimates from those values differ significantly from observed yields and ratios. However, the true extent of the deviation from statistical theories is very hard to estimate, as it blends into the many other probable causes for unexpected reaction yields and product ratios in the lab.

For further reading, we suggest the review on reactive intermediates by Barry Carpenter,[5] who made significant contributions exploring these effects in solvent-based organic chemistry. It is paired well with a more recent and broader look at the theory behind these effects as can be found in the review by Jayee and Hase.[6]

A reaction could also be primarily statistical, but the applied model fails to account for side reactions caused by light, oxygen, or the presence of water in solvent that is not dry enough. In theory, many of these practical factors in the lab can be already accounted for by careful re-consideration of the computational results. Computational modeling can guide organic chemists in determining whether the compound is sensitive to light or if a side reaction could be mediated by water. With these precautions, more accurate experimental results can be obtained with less time and resources in the lab.
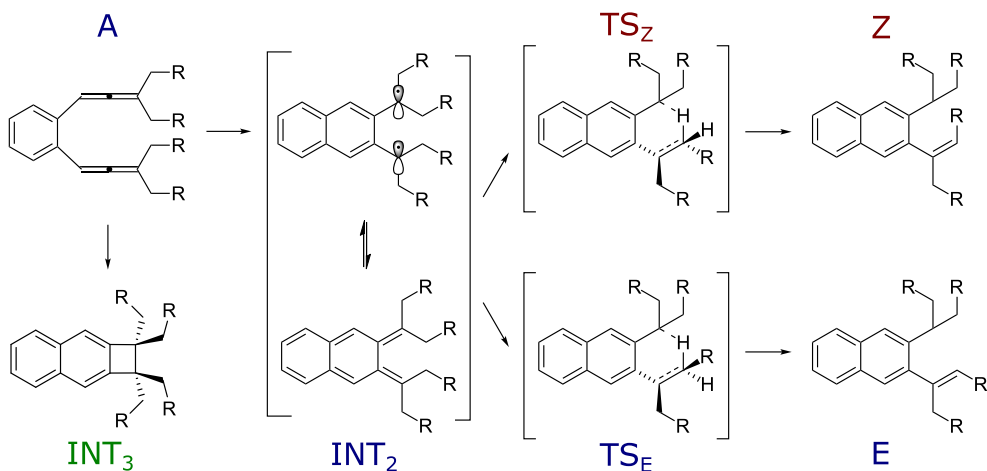
Scheme 1: General scheme of the thermal GB/[1,5]-H shift reaction

However, some factors, such as mysterious impurities in reagents, glassware, or stir bars, indeliberate carelessness of the chemists in the lab, are impossible to plan for, and there will always be unexpected uncertainty in the experimental results. The more possibilities we may screen for computationally, we demystify odd cases of chemical reactivity and allow for greater reproducibility and reaction control.

In this work, we aim to show that even a reported nonstatistical reaction involving a relatively large molecule (50–100 atoms) can be readily mapped out and modeled using well-established DFT methods and optimization algorithms. The ease of use has improved rapidly in recent years, both due to modern approximations bringing the cost of these calculations down considerably, but also due to open data practices and the availability (and documentation) of quantum chemistry software packages. For this exact reason, we also provide all our input and output files, which can be used with minimal modifications for a wide range of structures.

While many flavors of nonstatistical effects exist, in this case, we have the common motive of a two-step reaction, in which the first step releases a large amount of energy, so much that the short-lived intermediate is likely not thermalized on the timescale of the next reaction step. Among all the research, the thermal Garratt-Braverman (GB)/[1,5]-H shift of an ene–diallene reaction from Schmittel and

coworks[7] drew our attention (Scheme 1). They observed constant product ratios at different temperatures regardless of the energy difference between two products based on DFT calculations, which could not be explained by TST. The temperature-independent selectivity was attributed to nonstatistical dynamics.

GB cyclization is an important type of reaction for C—C bond formation,[8,9] which has been used widely in organic synthesis.[10–13] Mechanistic studies have been conducted both experimentally and computationally and it often produces a mixture of isomers depending on the nature of starting materials and reaction conditions.[14–16] The generally accepted mechanism involves the formation of a diradical intermediate,[17,18] especially for the cases where cyclobutane is formed, which cannot be accessed through [2+2] cycloaddition from a o-xylylenic intermediate under thermal conditions.[14,19] Thus, understanding the underlying mechanism of the selectivity of GB cyclization, furthermore, controlling the ratio of the final products is fascinating.

According to the original work,[7] the activation free energy for the hydrogen transfer is 2.3 kcal mol^{-1} higher in the case of Z isomer, matching the surplus of the E product as observed in the experiment. However, the actual obtained ratio differs significantly from the calculated TST values ($E/Z = 34.9{:}1$–$22.7{:}1$), while also unexpectedly remaining constant at around 10:1 for the wide temperature range

3

from 60 to 140 °C. Therefore, the reaction was declared nonstatistical.

We recently proposed a strategy for modeling nonstatistical effects that focuses on thermally activated reactions such as this one. Our model suggested a way to identify the vibrational energy localization and rank how well similar molecules redistribute it, providing us a way to suggest structure modifications that do not change the underlying reaction mechanism but enable control over product selectivity.[20] Therefore, we aimed to use the reported nonstatistical GB/[1,5]-H shift of the ene-diallene reaction, which has a relevant mechanism and considerable molecule size, to test the limits of our computational size.

Using this approach, we wished to modify the molecule in a way that would tone down the nonstatistical effects, resulting in higher product selectivity. We started out with a relatively modest change of extending the alkyl chains one carbon longer, from the original four n-propyl groups (**Diallene-nPr**) to n-butyl groups(**Diallene-nBu**). The goal was to see if this change would allow the thermally excited intermediate to relax faster and therefore result in a higher E/Z ratio.

However, we soon realized that the reaction was proceeding even at room temperature, despite the reported high barrier of almost 31 kcal mol^{-1} for the initial ring-forming step.[7] Additionally, the reaction sped up when exposed to light at room temperature – while maintaining the constant product ratio. This prompted us to explore this reaction further, including revisiting the original n-propyl derivative **Diallene-nPr**.

Results and discussion

The ene-diallene with four $n$butyl groups (**Diallene-nBu**) was synthesized by adapting the procedure used to form the n-propyl analog **Diallene-nPr** (Scheme 2).[7] Briefly, addition of lithiated $n$butylacetylene to a solution of phthalaldehyde in dry tetrahydrofuran (THF) at –78 °C, followed by in situ trapping of the resulting alkoxide with acetic anhydride, affords the bispropargylic acetate as a mixture of diastereomers, which was carried on without further purification for the formation of diallene by a combination of nBuMgBr, LiBr, and CuI. The successful formation of the desired ene-diallene **Diallene-nBu** has been confirmed by the vinylic proton signal at 6.47 ppm in ^{1}H spectrum (Figure S1) and allene carbon signal at 203.2 ppm in ^{13}C NMR spectrum (Figure S2).
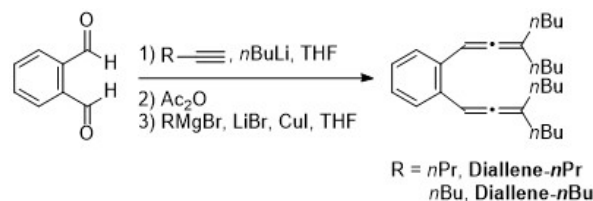
Scheme 2: Synthesis of **Diallene-nPr** and **Diallene-nBu**

The Garratt-Braverman/[1,5]-H shift of **Diallene-nBu** to afford E/Z isomers was examined by dissolving the sample in dry and degassed PhMe and heated at 140 °C in a sealed tube. As shown in Figure 1b, the disspearance of signals from **Diallene-nBu** and arising of signals from E/Z isomers indicate the successful thermolysis. However, it was noticed that the **Diallene-nBu** was not stable in solution under ambient conditions, indicated by a ^{1}H NMR spectroscopic analysis performed after keeping the original NMR sample under ambient conditions for a day (Figure 1c).

The new signals for this sample are consistent with the thermolyzed products (Figure 1b). Therefore, we wondered if the thermolysis could happen at ambient temperature despite the high reported computed barrier of 31 kcal/mol[7] or if an additional reaction pathway was present.

As such a significant change in the kinetics of the reaction was not expected by the change of alkyl chain, the effect of temperature of selectivity of GB/[1,5]-H shift of **Diallene-nBu** was experimentally examined by the procedure developed by Samanta et al for the original $n$propyl derivative **Diallene-nPr**. Then, the sealed tubes were wrapped in aluminum foil during the reaction to eliminate the effect of
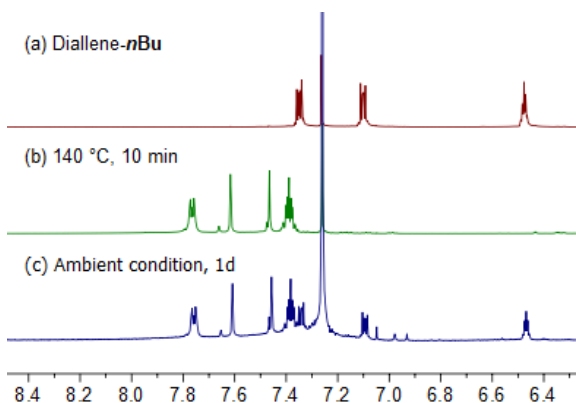
Figure 1: Stacked plot of the partial region of ^{1}H spectrum of (a) **Diallene-nBu**, (b) mixture of E/Z isomers after thermolysis at 140 °C, and (c) original NMR sample of **Diallene-nBu** kept under ambient condition after 1 day

light. In the dark, the reactions did not complete at lower temperatures within the indicated time range (Table S1).

For example, in the original paper, **Diallene-nPr** fully thermolyzed at 60 °C in 60 min, while we observed up to 10% of **Diallene-nBu** left even after 16 h. When irradiating the NMR sample in a quartz cuvette with 254 nm UV light at room temperature, the reaction was significantly sped up, full conversion was achieved in 2 h when using a 10 W LED UV light. It was noticed that the E/Z ratio remained constant at around 6.5:1, no matter the reaction conditions.

At this point, we decided to replicate the synthesis of original **Diallene-nPr** and thermolysis experiments, both to rule out any unexpected uncertainties in the experimental processes and for computational simplicity, as both the qualitative mechanism and the characteristics of its light absorption should not depend on the extension of these alkyl chains. We replicated the same E/Z ratio at around 10.3:1 for **Diallene-nPr** (Table S1). Similar to the thermolysis of **Diallene-nBu**, the reaction did not complete within the indicated time at lower temperatures for **Diallene-nPr** in the dark condition.

Made possible by light

As a first step in computationally determining whether a compound is light-sensitive, we wish to obtain an optimized (electronic ground state) geometry and compute a certain number of singlet vertical excitations. The term "vertical" here states the energies are obtained at the same geometry without any relaxation of the structure, setting them apart from "0–0" transition energies.

The exact number will depend on the size of the molecule, as for large molecules there could be several sets of similar, low energy but low oscillator strength transitions which contribute little to the full structure of the compound's UV-VIS spectrum. It is simple enough to plot the resulting spectrum in a program built to analyze the outputs of various quantum chemistry software packages, some examples being Multiwfn[21] or Chemcraft,[22] and see if enough of the spectrum is covered. If not, the calculation can be restarted with a higher requested number of excited states, saving some computational time.

This simple approach is often enough as a screening tool, especially for molecules in solvent at ambient or higher temperatures, as the vibronic structure of their spectra is obscured by very broad peaks. If a significant part of the intensity of a computed UV-VIS is missing compared to the experiment, it would be wise to consult the manual of the chosen quantum chemistry package for a more complete spectroscopy method (eg. the Excited State Dynamics module in ORCA). This is especially true for fluorescence spectra, which usually depend on the structure of a single electronic transition.

In this procedure, we need to allow for a margin of error, as popular TD-DFT functionals exhibit mean absolute errors between up to 0.3 eV.[23] In practice, this means the computed UV-VIS spectrum may need to be shifted by up to 20 nm, and it might not be possible to line up the first experimental peak at the same time as the highest intensity one. A qualitative UV-VIS spectrum can be measured experimentally to be compared with the computed result. However, if the rate of reaction under light is faster or

at the timescale of scanning, the obtained spectrum would not be able to reflect the true absorption of the compound. On the other hand, NMR spectroscopic analyses for light/dark control experiments are more reliable and practical in most cases.

Accounting for solvent effects has become easier as implicit solvent models are widely available and do not add a substantial computational overhead to the calculation. Additionally, we've seen advances in solvators, tools such as the one recently implemented in ORCA6, which use a docking approach to automatically add and position a certain number of solvent molecules before the calculation starts.

It is important to note that conformationally flexible molecules may exhibit different optical properties in different local minima. While tools for global optimizations and conformer searches exist, one example being the GOAT (global geometry optimizer and ensemble generator) in ORCA6, these tools may be prohibitively expensive, especially as any study we are performing would need to be repeated at any conformer deemed relevant. Here, chemical intuition may save us a lot of work - by hand-picking only those conformers that we expect to exhibit different optical properties or reaction pathways.

In our specific case, while there is plenty of conformational freedom provided by the four alkyl chains, we know the reaction cannot proceed without the proper positioning of the two allene groups. This leaves us to consider only two starting structures; the first structure $\mathbf{A}$ is a conformation of the initial compound where the allene groups are parallel and close to each other, in the plane of the aromatic ring. This starting point makes it easier to obtain the first transition state (TS), but we should be aware that there is a less sterically hindered, lower energy conformer. It is the lowest energy conformer that needs to be our baseline when estimating the activation energy. Later on, we will refer to this fully relaxed structure as $\mathbf{A}_{\mathrm{relaxed}}$.

We optimized the geometry of the electronic ground state of this conformer and confirmed the minimum with a frequency calculation using the hybrid, range-separated CAM-B3LYP

functional[24] with D3BJ empirical dispersion correction[25] and the ma-def2-SVP basis set.[26] The solvent was taken implicitly into account through the CPCM model for toluene.[27] All calculations in this article that involve DFT (and TD-DFT) were performed in ORCA 5.0.4 using its default RIJCOSX approximations.[28] The Tamm–Dancoff approximation (TDA) for TD-DFT is also a default in this package, and it is the combination of the resolution of identity (RI), the chain of spheres integration for the exchange energy (COSX), and the TD approximation that together makes modeling systems of this size viable.

Here we chose the "minimally augmented" ma-def2-SVP basis as we wished to include diffuse functions when modeling excited electronic states. We do this to ensure excitations that might move the electron density further from the molecular skeleton, such as low-lying Rydberg states, are included, and this was the largest basis set we could use without encountering any convergence issues. These convergence issues are a common occurrence with adding diffuse functions, especially to larger systems, and if the initial calculation shows there is no need for them - the calculation can be repeated with a larger, triple zeta basis set such as def2-TZVP to ensure more accurate transition energies. Therefore, for posterity, in Table S3 we compare 10 singlet excitations computed with the larger def2-TZVP and ma-def2-TZVPP basis sets and see no qualitative change in the ordering of the states.

A natural transition orbital (NTO) analysis was then performed using Multiwfn and ORCA's output, though NTOs can also be obtained directly from ORCA with an additional keyword (DoNTO True in the %tddft input block). We show the dominant NTOs for the first singlet transition, and the transition with the highest oscillator strength, in Figure 2. For visualization, we use the free Gabedit interface program.[29] As expected, the excitations were mainly of delocalized $\pi - \pi^*$ character. Even the energy of the lowest excited singlet state lays higher than the reported first barrier, implying that if light was involved in the reaction, the price for breaking the allene bonds would

be paid in advance.

It is interesting to note the presence of helical orbitals (Figure 2b). These orbitals are not surprising for allenes in general and are formed from combinations of different π or π^* orbitals, though their unique shape and optical properties are sensitive to structural changes once the allene is substituted or incorporated into other structures.[30] We will also note that saturated substituents, such as the chains we have here, will hyperconjugate into the π system,[31] a factor that will become relevant again near the end of this article and is also, to an extent, visible in the NTOs.

a)

b) 92.5%

c) 52.0% 41.5%

$S_0 \rightarrow S_1$ $S_0 \rightarrow S_7$
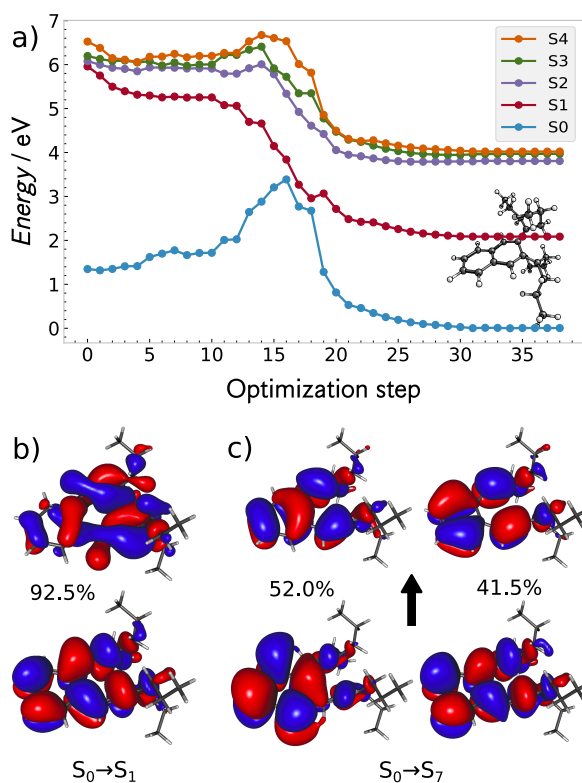
Figure 2: The result of a CAM-B3LYP/ma-def2-SVP geometry optimization on the S_1 surface (a) from the reactant structure, where the propyl legs are parallel (**A**), proceeds barrierless to an S_1 minimum which closely resembles the two-ring intermediate structure (**INT**$_2$). The dominant natural transition orbitals for the first (b) and seventh (brightest) (c) singlet excitations are shown with 90% of their total electronic density and their accompanying eigenvalue.

Of course, being able to absorb light does

not mean the molecule will be reactive due to its absorption, as it can relax from the excited electronic state by fluorescence or non-radiative transitions. The simplest exploratory step forward would be to perform a geometry optimization of the first excited singlet state. The first state might not be "bright", meaning its associated oscillator strength is low – but a fair assumption in photochemistry is that molecules reach the lowest excited state, S_1 quickly by internal conversion through conical intersections. We should be aware that as with TST, this is another textbook assumption. Still, a reaction that behaves nonstatistically in the ground state may very well exhibit traditional photophysics in its excited states. If we expect that the first reaction barrier simply prepares the system to react nonstatistically in the next step, basic chemical intuition should still hold at the first, rate-limiting step.

When we ran the S_1 geometry optimization (shown in Figure 2a) we observed a barrierless path to a structure that closely resembles our intermediate containing two aromatic rings, **INT**$_2$. In essence, the price for disturbing the double allene structure is pre-paid by the $\pi - \pi^*$ excitation, after which the system easily crosses "above" the transition state of the first reaction step. In this region, the S_0 and S_1 surfaces become nearly degenerate.

For us, this suggests there could be a conical intersection in this region that funnels the molecule back into the electronic ground state, ready to react further. In fact, we are increasingly aware that reactions proceeding from a conical intersection resemble post-transition-state bifurcations and may directly influence product selectivity.[32] This is easily explained if we consider that the molecule is both vibrationally excited as it "drops" down the PES and that the shape of the conical intersection may serve as a filter for a specific localization of the energy.

We did not pursue obtaining a conical intersection geometry as it wouldn't add further value to our experimental process. This is because the product ratio stays consistent whether the reaction is sped up with light or with heat and in darkness, suggesting it might

not matter how the rate limit barrier is passed. We propose that the conical intersection lies near the first transition state, which, as we discuss in the next section, has a multireference character. Because of this, it is likely TD-DFT would not reproduce an exact geometry of the conical intersection, though insights gained about the reaction path leading into and past this region should still hold.

In other cases, the product ratios of dark and irradiated reactions might differ, suggesting different dynamics in the excited state(s). A small barrier in the excited state might be easily thermally accessible, but would not be discovered in an optimization calculation. These cases would warrant a more expert approach, from computing scans/cuts along proposed reaction paths to simulating excited state molecular dynamics, for example with one of the popular surface hopping approaches.[33] Here we refer to recent reviews on surface hopping,[34] as well as the state-of-the-art of computational photochemistry in general.[35]

Obtaining a reaction path

A robust yet somewhat expensive method for mapping out a minimal energy reaction path (MEP) is the nudged elastic band (NEB) method.[36] Two geometries are provided for the reactant and product, while a guess for the transition state geometry is optional. The algorithm then interpolates a path between these structures and starts iteratively converging to the (nearest) MEP connecting these geometries.

In its climbing image variant (NEB-CI) the highest energy point (image) is declared as the climbing image and increased in energy – eventually resulting in an excellent starting guess for a transition state optimization. As in our case, this approach can also identify an intermediate structure and converge over several reaction steps, although only the highest barrier would be refined as a climbing image.

We use the unrestricted formalism (UKS or unrestricted Kohn-Sham) of DFT, which is more suitable for the treatment of systems with unpaired electrons, as bonds could be broken and reformed. It is even more so relevant as

the literature on this class of reactions mentions diradical intermediates, as we also indicate in Scheme 1,

However, the broken-symmetry approach for DFT is in common use with organic singlet diradicals,[37–40] and despite its limitations still provides a black-box approach to treat molecules with a small to intermediate diradical character, which would otherwise need to be modeled with multireference methods. In ORCA, this calculation is switched on using a simple keyword (BrokenSym 1,1) in the input file and no further user input is necessary.

For this reason, we first perform an NEB-CI calculation using UKS DFT with the B3LYP functional, a def2-TZVP basis set, D3BJ empirical dispersion, and a CPCM implicit solvent model for toluene. We can then recompute the energies along this path with the BS-DFT approach. The whole path, saved as a ".allxyz" file can be fed into a single calculation but may also be used to reconverge the NEB-CI at BS-DFT level using the "Restart_ALLXYZFile" keyword. The resulting reconverged MEP is shown in the background of Figure 3.

The final climbing image is then used as a starting point for a BS-DFT transition state optimization of $\mathbf{TS}_1$, confirmed with a frequency calculation with only one imaginary frequency vibrational mode. Despite not being a climbing image, the top of the second barrier of the MEP is also a suitable starting geometry to obtain $\mathbf{TS}_E$. The second transition state $\mathbf{TS}_Z$ is obtained through a separate NEB-CI calculation, but as the first reaction step remains the same, we initiate the second NEB between the intermediate $\mathbf{INT}_2$ and the $\mathbf{Z}$ product.

It is best not to "waste" points on conformational changes that are not relevant to the reaction mechanism, which is the reason why we initiate the first NEB at the $\mathbf{A}$ geometry, as opposed to the lower energy $\mathbf{A}_{\text{relaxed}}$. Results may also be obtained much faster by performing the NEB-CI calculation with a smaller basis set, such as def2-SVP or def2-SVPD, and then restarting it with the preferred one.

If we were expecting solvent to take part in the reaction, for example as in water-mediated hydrogen transfers, we might want to explic-
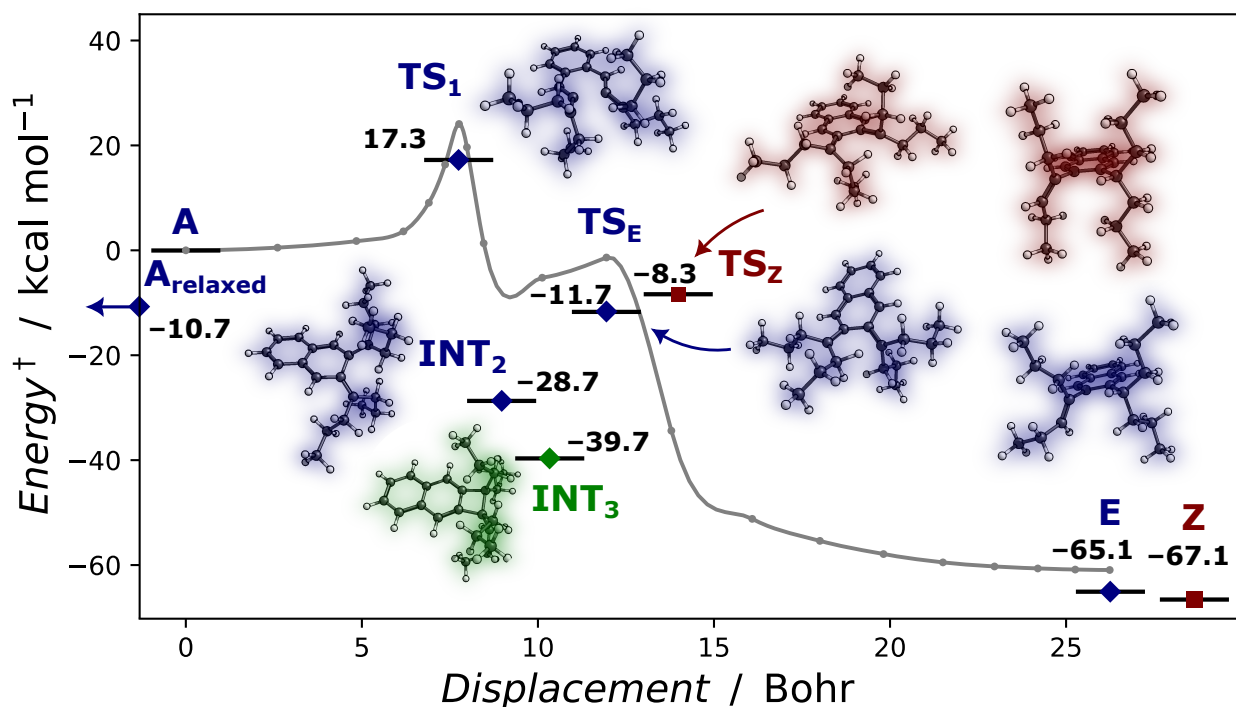
Figure 3: Calculated free energies of relevant geometries (blue) overlaid on a full BS-DFT NEB minimal energy reaction path (gray) from the initial structure **R** to the **E** product, confirming a reaction path with two steps passing over the two-ring intermediate INT_2. For comparison, we include the Z product and its transition state (red). In our exploration, the lower energy intermediate INT_3 (green) shares the same hydrogen transfer transition states TS_Z and TS_E. †:Note that the reaction path in the background does not include a free energy correction.

itly include molecules of solvent. This is also the case if we suspect moisture in an otherwise dry solvent. The process of docking solvent molecules was discussed in the previous section and the same logic applies as with eliminating conformational changes – optimal locations of the solvent might differ from reactant and product, which would result in the solvent unnecessarily roaming across the molecule. Therefore, some manual intervention when drawing the geometries could be needed.

In our case, adding a molecule of water between the two sites resulted in a water-mediated hydrogen transfer, but also a TS much higher in energy. However, in one of these MEPs another intermediate appears, lower in energy both with and without the presence of water. The unpaired electrons bond in this intermediate and a third ring is formed, giving a naphtho-cyclobutene structure. We label it INT_3 and show it in green in Figure 3. A NEB spanned between INT_2 and INT_3 results in a 20.2 kcal

mol^{-1} barrier (Figure S7).

This would mean that if the reaction proceeded into both intermediates, INT_3 might be likely to persist longer, which is why we will include it in our kinetics study. As they share TS_1, we would need to simulate ab initio molecular dynamics in order to predict whether both intermediates are formed easily. As these are costly and not definitive, it would likely be more efficient to seek evidence of the intermediate forming through experiment. Although the formation of INT_3 was presented in the literature for a similar substrate before, we did not observe it in our system, as there was no cyclobutene carbon signal ($\sim$50 ppm) observed in ^{13}C NMR spectra.[41,42] A closer inspection of the TS_1 also shows the structure is well-positioned for a hydrogen transfer, and a separate barrier of comparable energy is not as relevant if it is not dynamically accessible.

Finally, a fractional occupation number weighted density (FOD)[43] analysis was per-

formed on the whole reaction path in Figure 3. It is an efficient way of determining whether there could be significant static correlation in a system, providing an estimate of the strongly correlated electrons, as well as the ability to plot their spatial distribution on the molecule. In short, the analysis suggests the system does contain significant static correlation at $\mathbf{INT}_2$ ($N_{FOD} = 1.97$), $\mathbf{TS}_E$ ($N_{FOD} = 1.77$), and some at $\mathbf{TS}_1$ ($N_{FOD} = 0.57$) supporting the idea that this reaction proceeds through a diradical intermediate.

The values at each point are plotted in Figure S8 and two density plots are included in Figure 5. As discussed in the method paper,[43] if static correlation is high, but localized, it may still be treatable without multireference methods. One of the examples given is hybrid DFT functionals with low Hartree-Fock exchange. The functional we chose, B3LYP, fits into this category while still performing relatively well in benchmarks.

In the very last section, we will consider the simplest reaction kinetics model we can form using these results and show whether an expert approach to the electronic structure of these molecules was necessary.

Statistical rate theories and multireference treatment

Many software packages focused on thermochemistry and chemical kinetics are able to compute TST rate constants. We chose EyringPy as it is simple, can read ORCA output files, and is still in active development.[44,45] A summary of different programs and their capabilities can be found in the original EyringPy paper.[46]

As this reaction involves the transfer of a hydrogen atom, tunneling could play a significant role. For this reason, the rate constants were computed using transition state theory in combination with an Eckart tunneling contribution. Using BS-DFT frequency calculations of the intermediates, transition states, and products for both the E and Z transfer we obtain the rates k_E and k_Z and plot their ratio in Figure 4. Compared to previously reported values obtained
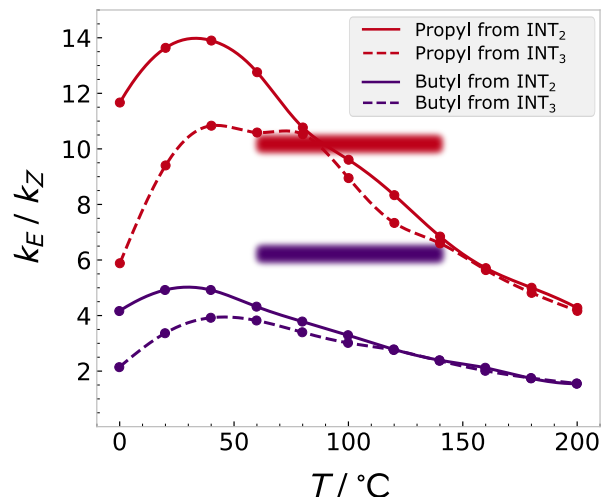

Figure 4: When a tunneling correction is included the TST ratios obtained from BS-DFT results match the experimental product ratios (shaded area) exceptionally well at lower temperatures for propyl (red) and reproduce the lower selectivity trend for butyl (purple). However, they still predict a temperature dependence large enough to be observable in the experiment.

with only TST, when tunneling is included the predicted product ratios begin to overlap with the experimental one. However, this rate theory still predicts that a temperature dependence of the E/Z ratio to such an extent that would be unavoidable in the experiment, with the selectivity dropping from nearly 14:1 at 40 °C to 6:1 and decreasing above 160 °C .

Something that could be considered counterintuitive is that even if the reaction were proceeding from the much deeper potential energy well of $\mathbf{INT}_3$, the result is nearly identical once the reaction is heated. Only around room temperature do we see lower predicted selectivity, along with a temperature-independent plateau. This plateau of the product ratio is caused by the decreasing impact of the tunneling contribution, which strongly favors the Z product. With an imaginary mode at -1402 cm^{-1} , $\mathbf{TS}_Z$ has a narrower and higher barrier, as opposed to -876 cm^{-1} for $\mathbf{TS}_E$. Product selectivity peaks when the temperature is high enough that the pure TST rate dominates and then decreases.

This also serves as a warning that temperature independence, or dependence smaller than

can be determined by NMR study, might persist for certain temperature ranges in perfectly statistical reactions. However, as is discussed in the EyringPy rate theory paper,[45] due to the exponential dependence on the Gibbs energy, a 1 kcal mol^{-1} error at room temperature will already result in a rate constant that is five times larger/smaller. While we may hope that these errors are smaller when producing ratios (that is, looking at relative Gibbs energies) we are still both near the limit of accuracy that can be expected from DFT and at risk of comparing the free energies of transition states that are not in their lowest energy conformations. The influence of conformational changes can be accounted for with added effort and computational resources, an example being ensemble-averaged variational TST and its application to enzyme kinetics.[47]

Due to the potential diradical character of the structure, we now proceed to compare the relatively black-box approach of (BS)-DFT with a more expert and user-intensive multireference treatment. We performed single-point energy calculations at the SA-CASSCF/MS-CASPT2 level of theory for the ground and first excited state for the singlet and the triplet.[48] All calculations were performed on neutral molecules with a PCM implicit solvent model for toluene, using previously obtained DFT geometries. For each structure, we computed two-root SA-CASSCF single-point wavefunctions and their energies.

In this multiconfigurational method the wave function is constructed as a linear combination of configuration state functions (CSFs). The number and type of available functions are determined by the active space (AS) and the overall spin. In CASSCF, you choose the number of electrons and orbitals which are included in the AS. The CSFs are built by distributing the active electrons into the active orbitals while maintaining the overall spin. Any occupation is allowed in CASSCF.

In the study presented here, all π-orbitals of the conjugated system are included in the AS because the electronic excitations are dominated by $\pi - \pi^*$ transitions. The two electrons of the diradical and the two p-orbitals on the carbon adjacent to the aromatic rings are included in the AS as well to follow the reaction. This gives a CAS(12,12) AS. Analysis of the orbitals in the AS was completed for all structures. As an example, we have included a visual representation of the AS orbitals for one of the intermediate structures in S6. We identified the π-orbitals, the bond-forming orbitals as well as the antibonding orbitals in the AS as we would expect.

All calculations employed an ANO-RCC basis set[49] with a double-zeta quality and were performed using OpenMolcas version 23.02.[50] The computed SA-CASSCF wave functions were used as reference configurations for subsequent MS-CASPT2 calculations.[51,52]

The first excited singlet state energies are also obtained, as shown in Table SS4, however, they do not bring us any closer to obtaining a conical intersection, as at $\mathbf{TS_1}$ the S_1 is nearly 90 kcal mol^{-1} above S_0. Unlike with TD-DFT, $\mathbf{TS_1}$ is higher in energy than $\mathbf{A_{relaxed}}$ in its excited state, but only by 3.7 kcal mol^{-1}, which is easily thermally accessible. However, the energy of $\mathbf{TS_E}$ is higher than $\mathbf{TS_Z}$. Though these geometries are not true transition states on S_1, it is still indicative that the reaction does not proceed for both steps on this surface, as the product ratio would change.

To perform reaction rate modeling with Gibbs energies composed of a CASPT2 electronic energy and DFT thermochemical corrections, we simply modified the DFT frequency calculation output file with the CASPT2 single point energy.

The kinetics difference shown in Figure 5c is caused only by the E/Z TS gap increasing from 2.9 to 3.5 kcal mol^{-1} gap, moving us away from the experimental result. This small change in relative energies is the same magnitude as could be caused by conformational changes. However, for a system where the high conformational flexibility is located outside the strict reactive space, NEB methods should still converge to optimal MEPs, as long as they are provided with reasonable end points – the reactant and product in a relaxed geometry. When the relaxed geometry must surpass a conformational barrier to react, the NEB can be restarted from
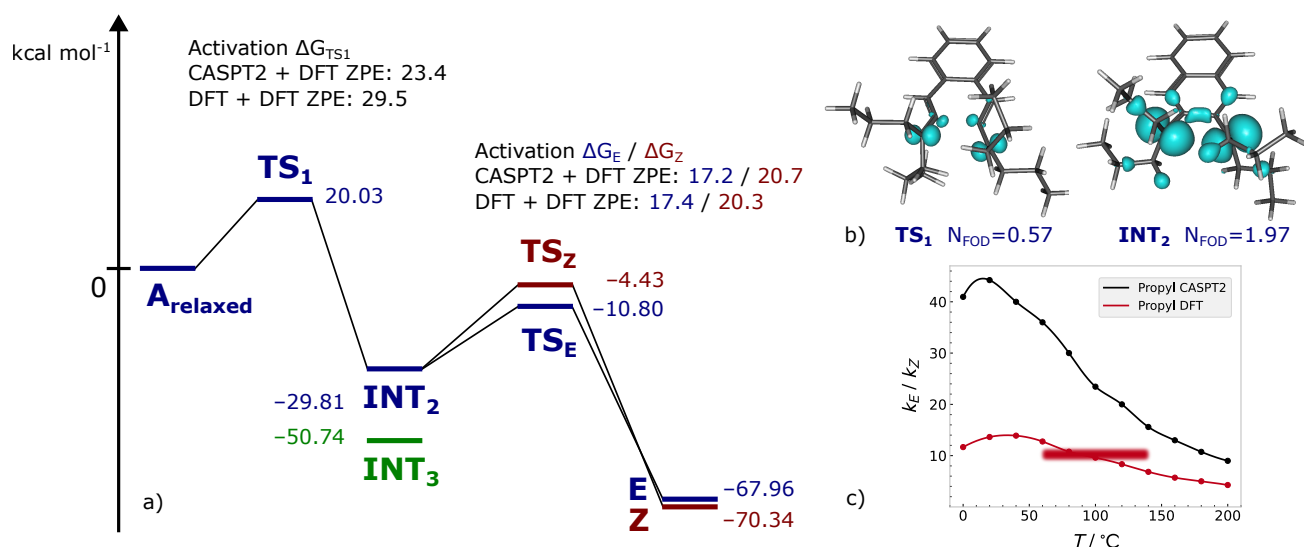
Figure 5: a) CASPT2 ground state electronic energies of key structures on the DFT reaction path, including activation Gibbs free energies computed at 293K. b) FOD plots c) predicted product ratio using CASPT2 single point energy and DFT thermochemistry

the more accessible conformation(s). If this barrier is easily surpassed thermally, the overall activation energy is still defined relative to the most populated, lowest energy conformer.

The worse fit of the multireference result to the experimental E/Z ratio could be a result of methodological inconsistency, as the single point energies were computed on DFT geometries. But as the activation energy of the first step is reproduced better with this hybrid approach than with DFT only, it is also possible that the DFT TST results cross over with the experimental value on accident.

Altogether, we may conclude DFT already provided a quantitively good enough story of the reaction mechanisms, most likely system does not have a strong and/or delocalized multireference character. This can also be seen through the fact the multireference improved significantly only our computed activation energy for the first barrier, likely the most difficult electronic structure to model.

While the previous experimental value for the first activation energy of a similar nMe structure[14] amounted to 19.8 kcal mol^{-1} , and our own vtNMR study for nBu gave us a value of 25.2 kcal mol^{-1} (see Figures S3 to S5), DFT free energies amounted to 29.5 kcal mol^{-1} (nPro) and 28.9 kcal mol^{-1} (nBu). The CASPT2 value

with DFT ZPE is closer at 23.4 kcal mol^{-1} in a way that would significantly affect the predicted speed of the reaction; but in the end, would not change our workflow – as the slightly larger DFT barrier already told us it was reacting too quickly in ambient conditions.

It is important to warn that in some cases multireference calculations could still reveal insights about the electronic structure that DFT cannot provide, and tools such as the FOD analysis may be used to identify such molecules. Furthermore, the significant tunneling contribution for all TST results in our work relies solely on the shape of the barrier obtained with BS-DFT. Refining the multireference calculations would require a delicate investigation of the basis set, AS, and SA choices, in addition to CASSCF geometry optimizations, which we believe is out of scope for this tutorial.

Conclusion

In Figure 6 we present a flowchart that collects the steps we performed, expecting them to be performed in parallel, or at least with an ample amount of communication between the experimental and computational teams.

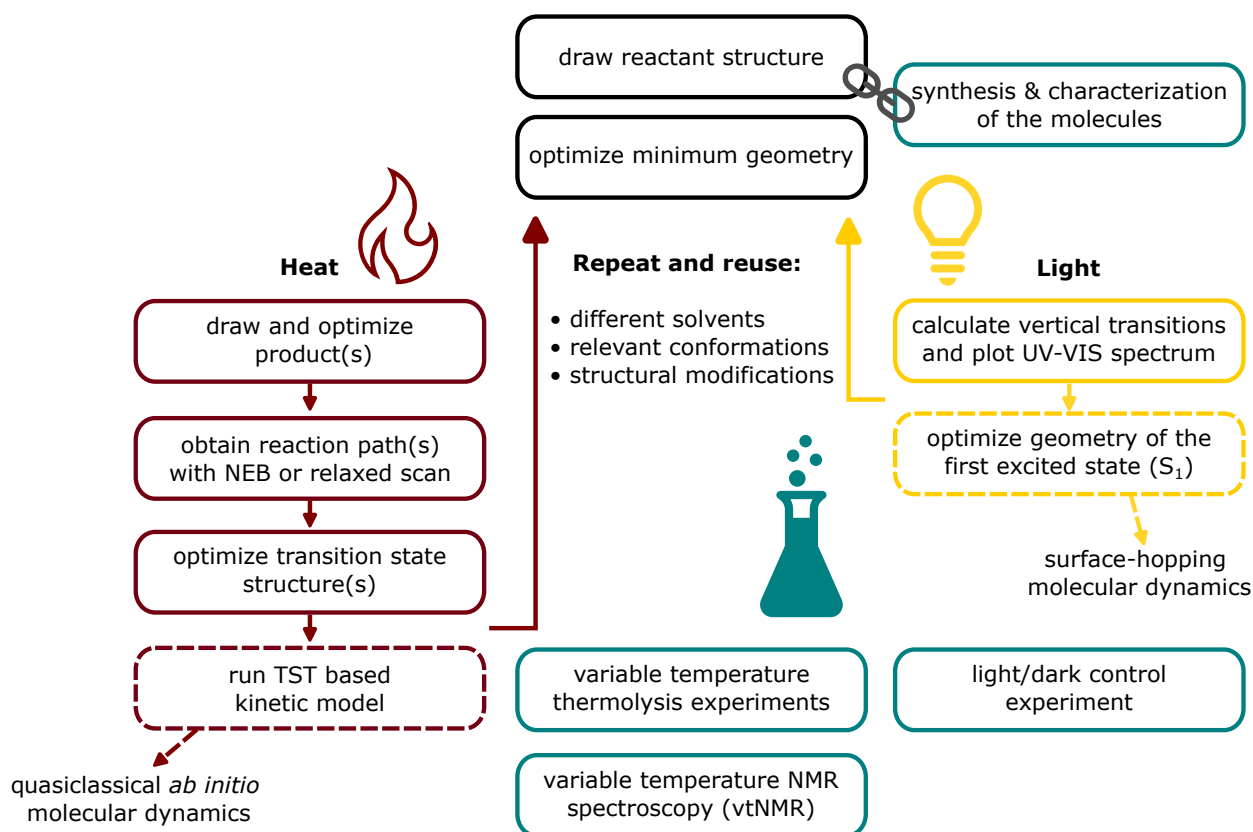Solid boxes represent well-established methods that are implemented in many quan-

Figure 6: A flowchart on concurrent computational and experimental reaction mechanism exploration. Boxes in red correspond to thermal, electronic ground state chemistry, boxes in yellow to photochemistry, while the blue boxes are experimental procedures. The black boxes are preparatory computational steps that are cycled back to once experimental conditions change or the structure is modified. Dotted boxes represent key decision steps in the computational flow, as if they are not conclusive, an experiment and/or expert theoretical treatment would follow.

tum chemistry packages and are still relatively black-box approaches, easily accessible to non-experts and aspiring computational chemists. If confusion arises, these types of calculations are discussed enough both in existing documentation and online forums, that consulting a large language model (LLM) would likely lead to expert-level advice. For more specific help, a PDF manual of the quantum chemistry package may also be provided to the LLM.

The dotted boxes are two-fold as they are both decision points and gateways to more advanced methods. As we hope we showed in this tutorial review, running a TST model is still relatively simple – however if it fails to predict experimental results, both more advanced experiments, such as variable temperature NMR and thermolysis, or costly computational treatment through molecular dynamics simulations

may need to follow. For these, more experienced practitioners should be consulted.

Based on our results, we still consider the Garratt–Braverman/[1,5]-H shift of ene-diallenes to be a nonstatistical reaction, though to a lesser extent than originally thought. This raises again the persistent question of how much of chemistry is statistical, and to which extent it can already be fully modeled by statistical models such as TST and its evolutions.

If we hope for massive, automated reaction discovery through generated reaction networks, as opposed to expensive full dynamical treatment, we first need to develop an understanding of the scope of nonstatistical reactivity. Once we are confident we have predictive measures of nonstatistical effects, once such reaction steps are flagged, they may be modeled appropriately – or handed back from the machine to

the chemist.

Supporting Information Available

The published version of this manuscript will also contain a link to an online data repository containing all the computational results discussed in the article, including some notes on choices made in the input files.

References

(1) Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions. *Physical Chemistry Chemical Physics* **2017**, *19*, 32184–32215.

(2) Mardirossian, N.; Head-Gordon, M. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Molecular Physics* **2017**, *115*, 2315–2372.

(3) Henriksen, N. E.; Hansen, F. Y. *Theories of Molecular Reaction Dynamics*; Oxford University Press, 2018; Vol. 1.

(4) 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.

(5) Carpenter, B. K. Energy Disposition in Reactive Intermediates. *Chemical Reviews* **2013**, *113*, 7265–7286.

(6) Jayee, B.; Hase, W. L. Nonstatistical Reaction Dynamics. *Annual Review of Physical Chemistry* **2020**, *71*, 289–313.

(7) Samanta, D.; Rana, A.; Schmittel, M. Nonstatistical Dynamics in the Thermal Garratt–Braverman/[1,5]-H Shift of One Ene–diallene: An Experimental and Computational Study. *The Journal of Organic Chemistry* **2014**, *79*, 8435–8439.

(8) Basak, A.; Mondal, S.; Mitra, T.; Mukherjee, R.; Addy, P. Garratt–Braverman Cyclization, a Powerful Tool for C–C Bond Formation. *Synlett* **2012**, *23*, 2582–2602.

(9) Bhattacharya, P.; Singha, M.; Das, E.; Mandal, A.; Maji, M.; Basak, A. Recent advances in Garratt-Braverman cyclization: Mechanistic and synthetic explorations. *Tetrahedron Letters* **2018**, *59*, 3033–3051.

(10) Zafrani, Y.; Cherkinsky, M.; Gottlieb, H. E.; Braverman, S. A new approach to the synthesis of 2-vinylthiophenes and selenophenes; competition between free radical and anionic cycloaromatization of bridged di- and tetrapropargylic sulfides and selenides. *Tetrahedron* **2003**, *59*, 2641–2649.

(11) Mitra, T.; Jana, S.; Pandey, S.; Bhattacharya, P.; Khamrai, U. K.; Anoop, A.; Basak, A. Asymmetric Garratt-Braverman cyclization: a route to axially chiral aryl naphthalene-amino acid hybrids. *J Org Chem* **2014**, *79*, 5608–5616.

(12) Bhattacharya, P.; Senapati, K.; Chattopadhyay, K.; Mandal, S. M.; Basak, A. Synthesis of benzochromenes and dihydrophenanthridines with helical motifs using Garratt-Braverman and Buchwald-Hartwig reactions. *Rsc Advances* **2015**, *5*, 61562–61574.

(13) Mandal, A.; Bhattacharya, P.; Das, A. K.; Basak, A. A Garratt-Braverman cyclization route towards the synthesis of phenanthridine derivatives and their DNA-binding studies. *Tetrahedron* **2019**, *75*, 1975–1987.

(14) Braverman, S.; Duar, Y. Thermal rearrangements of allenes. Sythesis and mechanisms of cycloaromatization of .pi. and heteroatom bridged diallenes. *Journal of the American Chemical Society* **1990**, *112*, 5830–5837.

(15) Jana, S.; Anoop, A. Does the Mechanism of the Garratt-Braverman Cyclization Differ with Substrates? A Computational Study on Bispropargyl Sulfones, Sulfides, Ethers, Amines, and Methanes. *J Org Chem* **2016**, *81*, 7411–8.

(16) Zhang, M.; Lu, H.; Li, B.; Ma, H.; Wang, W.; Cheng, X.; Ding, Y.; Hu, A. Experimental and Computational Study on the Intramolecular Hydrogen Atom Transfer Reactions of Maleimide-Based Enediynes After Cycloaromatization. *J Org Chem* **2021**, *86*, 1549–1559.

(17) Skraba, S. L.; Johnson, R. P. A Computational Model for the Dimerization of Allene. *The Journal of Organic Chemistry* **2012**, *77*, 11096–11100.

(18) Hopf, H.; Markopoulos, G. The chemistry of bisallenes. *Beilstein J Org Chem* **2012**, *8*, 1936–1998.

(19) Inanaga, J.; Sugimoto, Y.; Hanamoto, T. A novel method for the generation of 2,3-naphthoquinodimethanes utilizing samarium(II) iodide-promoted allene synthesis. *Tetrahedron Letters* **1992**, *33*, 7035–7038.

(20) Rožić, T.; Teynor, M. S.; Došlić, N.; Leitner, D. M.; Solomon, G. C. A Strategy for Modeling Nonstatistical Reactivity Effects: Combining Chemical Activation Estimates with a Vibrational Relaxation Model. *Journal of Chemical Theory and Computation* **2024**,

(21) Lu, T. A comprehensive electron wavefunction analysis toolbox for chemists, Multiwfn. *The Journal of Chemical Physics* **2024**, *161*.

(22) Chemcraft - graphical software for visualization of quantum chemistry computations. Version 1.8, build 682. https://www.chemcraftprog.com.

(23) Laurent, A. D.; Jacquemin, D. TD-DFT benchmarks: A review. *International Journal of Quantum Chemistry* **2013**, *113*, 2019–2039.

(24) Yanai, T.; Tew, D. P.; Handy, N. C. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chemical Physics Letters* **2004**, *393*, 51–57.

(25) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry* **2011**, *32*, 1456–1465.

(26) Zheng, J.; Xu, X.; Truhlar, D. G. Minimally augmented Karlsruhe basis sets. *Theoretical Chemistry Accounts* **2011**, *128*, 295–305.

(27) Barone, V.; Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *The Journal of Physical Chemistry A* **1998**, *102*, 1995–2001.

(28) Neese, F. Software update: The ORCA program system—Version 5.0. *WIREs Computational Molecular Science* **2022**, *12*.

(29) Allouche, A. Gabedit—A graphical user interface for computational chemistry softwares. *Journal of Computational Chemistry* **2011**, *32*, 174–182.

(30) Bro-Jørgensen, W.; Garner, M. H.; Solomon, G. C. Quantification of the Helicality of Helical Molecular Orbitals. *The Journal of Physical Chemistry A* **2021**, *125*, 8107–8115.

(31) Garner, M. H.; Corminboeuf, C. The fundamental relation between electrohelicity and molecular optical activity. *Physical Chemistry Chemical Physics* **2023**, *25*, 15200–15208.

(32) Feng, Z.; Guo, W.; Kong, W.-Y.; Chen, D.; Wang, S.; Tantillo, D. J. Analogies between photochemical reactions and ground-state post-transition-state bifurcations shed light on dynamical origins of selectivity. *Nature Chemistry* **2024**, *16*, 615–623.

(33) Travnikova, O. et al. Photochemical Ring-Opening Reaction of 1,3-Cyclohexadiene: Identifying the True Reactive State. *Journal of the American Chemical Society* **2022**, *144*, 21878–21886.

(34) Mai, S.; Marquetand, P.; González, L. *Surface Hopping Molecular Dynamics*; Wiley, 2020; pp 499–530.

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

(36) Ásgeirsson, V.; Birgisson, B. O.; Bjornsson, R.; Becker, U.; Neese, F.; Riplinger, C.; Jónsson, H. Nudged Elastic Band Method for Molecular Reactions Using Energy-Weighted Springs Combined with Eigenvector Following. *Journal of Chemical Theory and Computation* **2021**, *17*, 4929–4945.

(37) Rush, L. A.; Gallo, K. F.; Stumetz, K. S.; Rodríguez-Pérez, I. A.; Cremeens, M. E.

Non-statistical dynamics for the allene oxide to cyclopropanone conversion. *Journal of Physical Organic Chemistry* **2022**, *35*.

(38) Malrieu, J.-P.; Trinquier, G. A Recipe for Geometry Optimization of Diradicalar Singlet States from Broken-Symmetry Calculations. *The Journal of Physical Chemistry A* **2012**, *116*, 8226–8237.

(39) Tang, Z.; Jiang, Z.; Chen, H.; Su, P.; Wu, W. Energy decomposition analysis based on broken symmetry unrestricted density functional theory. *The Journal of Chemical Physics* **2019**, *151*.

(40) Saito, T.; Nishihara, S.; Yamanaka, S.; Kitagawa, Y.; Kawakami, T.; Yamada, S.; Isobe, H.; Okumura, M.; Yamaguchi, K. Symmetry and broken symmetry in molecular orbital description of unstable molecules IV: comparison between single- and multi-reference computational results for antiaromtic molecules. *Theoretical Chemistry Accounts* **2011**, *130*, 749–763.

(41) De Fonseka, K. K.; McCullough, J. J.; Yarwood, A. J. Products and transients in the photolysis of methylated 2-indanones. Observation and kinetics of o-xylylenes. *Journal of the American Chemical Society* **1979**, *101*, 3277–3282.

(42) McCullough, J. J. o-Xylylenes and isoindenes as reaction intermediates. *Accounts of Chemical Research* **1980**, *13*, 270–276.

(43) Bauer, C. A.; Hansen, A.; Grimme, S. The Fractional Occupation Number Weighted Density as a Versatile Analysis Tool for Molecules with a Complicated Electronic Structure. *Chemistry – A European Journal* **2017**, *23*, 6150–6164.

(44) E. Dzib, A. Quintal, F. Ortiz-Chi, G. Merino, Eyringpy 2.0, Cinvestav, Merida, Yucatan 2021.

(45) Dzib, E.; Quintal, A.; Merino, G. Enhancing Eyringpy : Accurate Rate Constants with Canonical Variational Transi-

Supporting Information:

Tutorial Review on Modeling Nonstatistical Reactivity: an Example of Light and Heat in the Garratt–Braverman/[1,5]-H shift of ene-diallenes

Tomislav Rožić[a],[†],[‡] Yuxuan Hou[a],[†] Lea Kjærgaard Northcote,[†],[¶] Christian Marcus Pedersen,[†] and Gemma C. Solomon[*],[†],[‡],[¶]

†*Department of Chemistry, University of Copenhagen, Denmark*

‡*Nano-Science Center, University of Copenhagen, Denmark*

¶*NNF Quantum Computing Programme, Niels Bohr Institute, University of Copenhagen, Denmark*

E-mail: gsolomon@chem.ku.dk

[a] These authors contributed equally.

Synthesis and Experimental Data

Reagents were purchased reagent grade from commercial suppliers and used without further purification. Dry tetrahydrofuran (THF) and toluene were obtained from a commercial solvent purification system (PS-MD-05, Innovative Technology INC) or by drying with 4 Å molecular sieves overnight. $MgSO_4$ was used as the drying reagent after aqueous work-up. ^{1}H and ^{13}C NMR spectra were recorded on a Bruker 500 MHz Ultra Shield Plus instrument with a non-inverse cryoprobe. NMR spectra were recorded at ambient probe temperature and referenced to the residual solvent signal ([1]H: $CDCl_3$: 7.26 ppm, [13]C: $CDCl_3$: 77.06 ppm) unless noted otherwise. The coupling constants of protons in ^{1}H spectra have been reported as pseudo first-order when possible, even though they can be higher-order (ABC, ABX, AA'BB') spin systems; coupling constants are reported as observed. High resolution mass spectra were obtained from a SolariX ESI/MALDI FTMS spectrum. Thin layer chromatography (TLC) analyses were carried out on VWR aluminum plates coated with silica F254 and visualized via UV light (254/364 nm). Compound **Diallene-nPr** was synthesized as previously reported.[S1]

Compound **Diallene-nBu**. The synthesis was carried out by adapting the procedure of Samanta et al. To a solution of 1-hexyne (135 mg, 1.64 mmol) in dry THF (5 mL), nBuLi (2.5 M in hexanes, 0.64 mL, 1.64 mmol) was added dropwise at –78 °C. The solution was stirred at –78 °C for 30 min and then added to a solution of 5,12-phthalaldehyde (100 mg, 0.746 mmol) in dry THF (10 mL) at –78 °C. The solution was warmed to rt and stirred for 4 h under an atmosphere of N_2. After cooling to 0 °C, acetic anhydride (0.70 mL, 7.46 mmol) was added. The solution was warmed to rt and stirred for 16 h. Water (50 mL) was added and stirred at rt for 30 min. The layers were separated, and the aqueous phase was extracted with EtOAc (2 × 50 mL). The organic phases were combined, washed with satd aq $NaHCO_3$ (3 × 100 mL), dried ($MgSO_4$), and filtered. After removing the solvent, the residue was dissolved in heptane and passed through a pad of silica gel with EtOAc/heptane = 1:2, followed by solvent removal in vacuo. The crude mixture was carried on through

the synthesis routine without further purification. A solution of $n$butylmagnesium bromide (nBuMgBr) was prepared by adding dropwise nBuBr (1.02 g, 7.46 mmol) to a mixture of Mg (181 mg, 7.46 mmol) in dry THF (10 mL). The reaction mixture was under self-reflux and cooled down to rt slowly. The solution was stirred at rt for an additional 30 min before adding dropwise into a suspension of LiBr (1.29 g, 14.9 mmol) and CuI (1.42 g, 7.46 mmol) in dry THF (10 mL) at 0 °C. After stirring at 0 °C for 15 min, the crude product mixture obtained before in dry THF (5 mL) was added. After stirring for 3 h at 0 °C under an atmosphere of N_2, satd aq NH_4Cl (150 mL) was added, followed by Et_2O (20 mL). The solution was stirred for 1 h at 0 °C. The flask was wrapped in aluminum foil during the reaction to limit light exposure. The layers were separated, and the aqueous phase was extracted with Et_2O (2 × 10 mL). The organic phases were combined, dried ($MgSO_4$), and filtered. After removing the solvent, the residue was dissolved in npentane and passed through a pad of silica gel with npentane, followed by solvent removal in vacuo, affording **Diallene-nBu** as a colorless oil (107 mg, 38% overall). $R_f = 0.60$ ($n$pentane). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.35 (dd, $J = 5.8, 3.4$ Hz, 2H), 7.10 (dd, $J = 5.9, 3.4$ Hz, 2H), 6.47 (quin, $J = 3.0$ Hz, 2H), 2.14–2.01 (m, 8H), 1.52–1.44 (m, 8H), 1.39–1.29 (m, 8H), 0.89 (t, $J = 7.3$ Hz, 12H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 203.2, 132.6, 127.6, 126.3, 107.8, 92.1, 32.6, 29.9, 22.6, 14.0.

Thermolysis Results

The sample was dissolved in dry and degassed PhMe and heated at different temperatures in sealed tubes. It is noted that the sealed tubes were wrapped in aluminum foil during the reaction to eliminate the effect of light on the dark controls.

NMR Kinetic Assays

Reaction kinetics of **Diallene-nBu** were tracked on aliquots that 19.0 mg of sample had been dissolved in 5.00 mL d_8-toluene and maintained at 248 K prior to reaction at 313
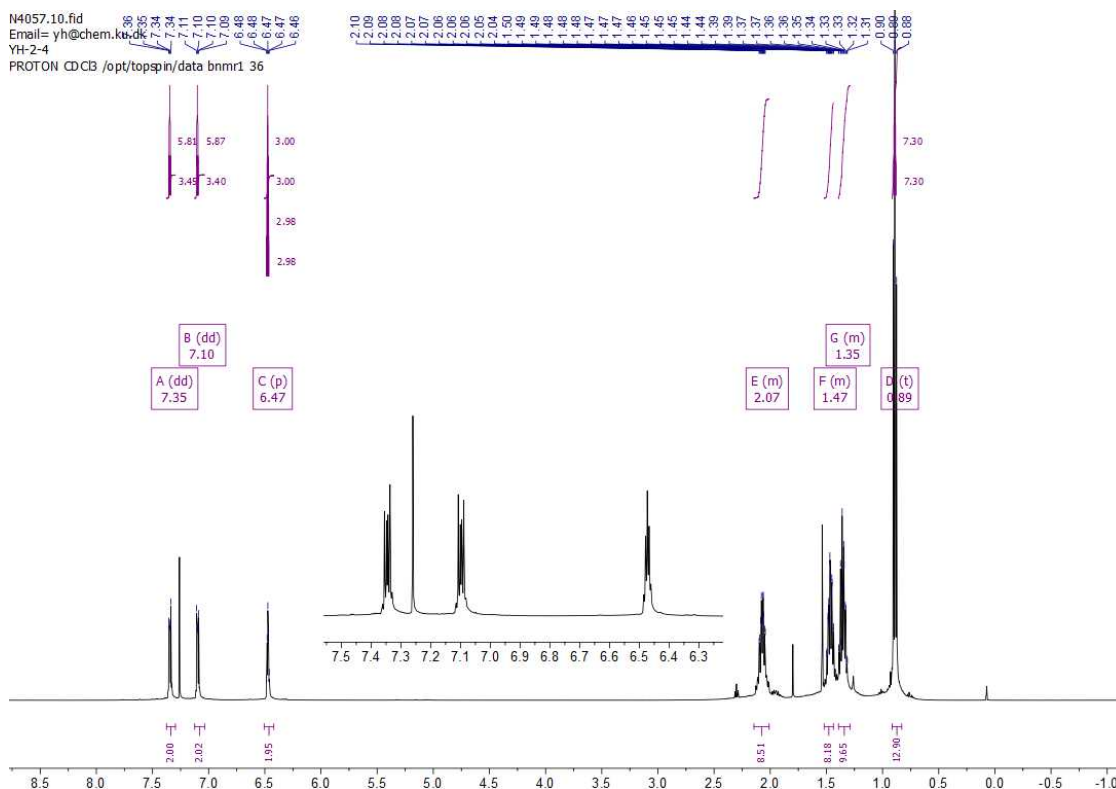
Figure S1: ^{1}H NMR spectrum (500 MHz) of compound **Diallene-nBu** in CDCl$_3$.

K, 333K, 353 K and 373 K for experimental determinations of the temperature-dependent rate constant and the activation energy. *In situ* NMR spectra were acquired on a Bruker (Fällanden, Switzerland) 600 MHz Avance III instrument equipped with a 14.1 T Ascend magnet and a BBO 600 MHz S3 5 mm Probehead. Kinetic assays were conducted using time-resolved reaction tracking by implementing a pseudo-2D experiment consisting of a repetition of one-dimensional ^{1}H NMR spectra. For each time point, two dummy scans were acquired prior to accumulating 16 scans using an inter-scan relaxation delay of 1.0 s and an acquisition time of approximately 1.37 s for 16384 complex data points. Spectra were acquired with a spectral width of 20 ppm around a carrier offset of 4.0 ppm. This setup resulted in a time resolution of approximately 43.2 s between the time points. A suitable number of time points were acquired to follow the decay of the reactant over time. Data were acquired using Bruker Topspin 3.5 pl6 and signal areas were obtained using the Dynamics module on the pseudo-2D data sets in the same software. Integrals for the reactant were
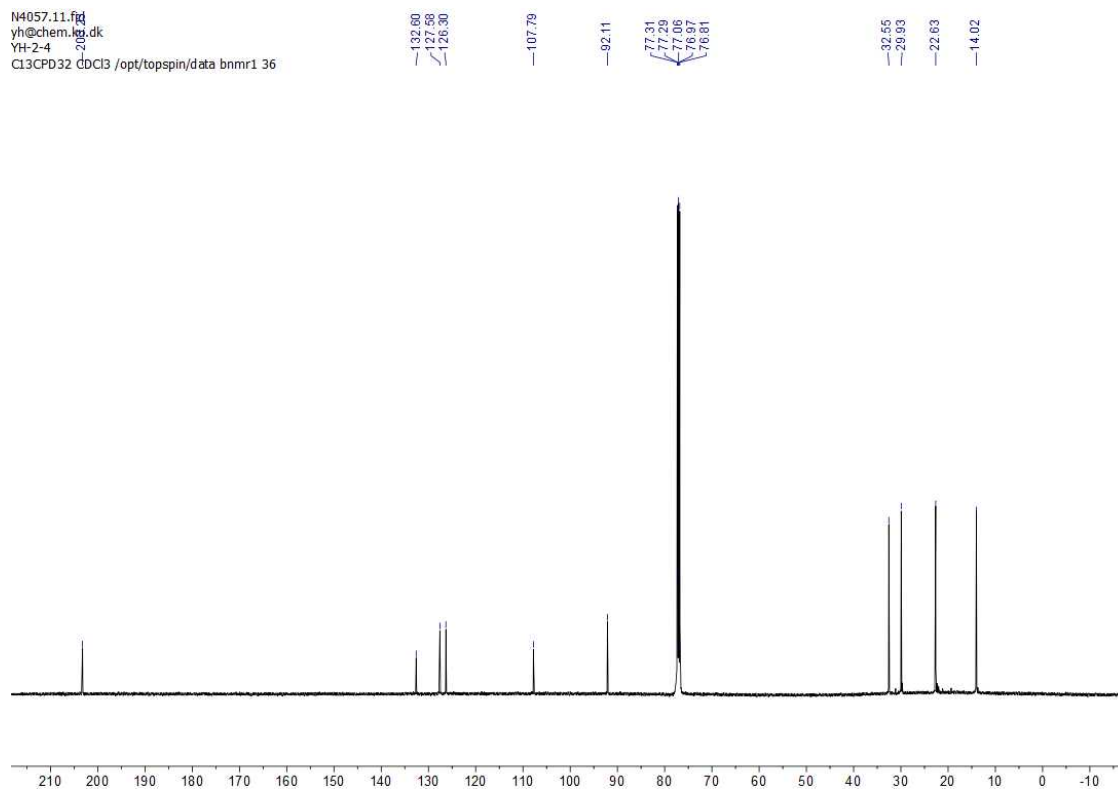
Figure S2: ^{13}C NMR spectrum (126 MHz) of compound **Diallene-nBu** in CDCl$_3$.

fitted to exponential decays using pro Fit 7 (Uetikon am See, Switzerland), assuming a non-reversible conversion of the starting material. Fits of the kinetic data were in excellent agreement with a complete conversion through a first-order reaction. The activation energy of the reaction was determined with a linear version of the Arrhenius equation in pro Fit 7, yielding a coefficient of determination $R^2 = 0.9987$.
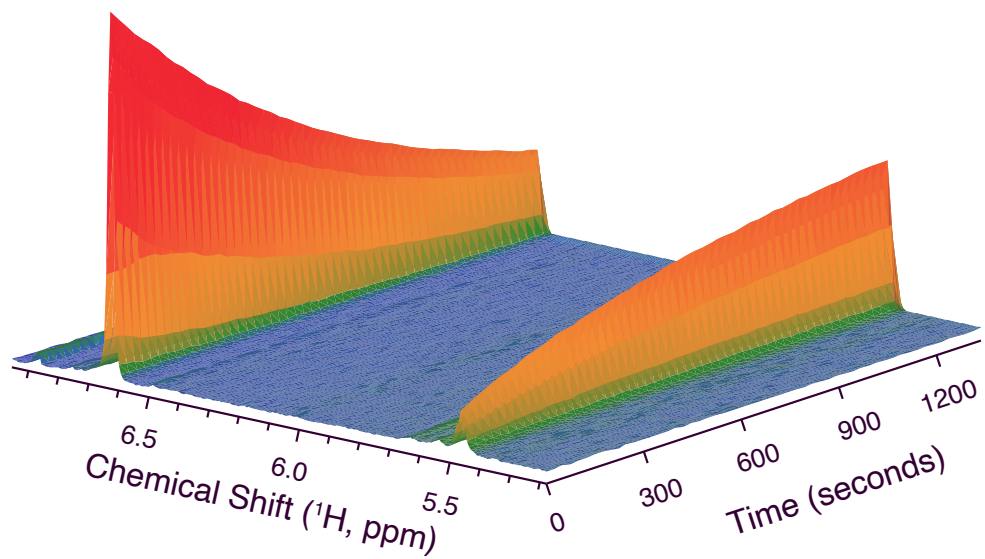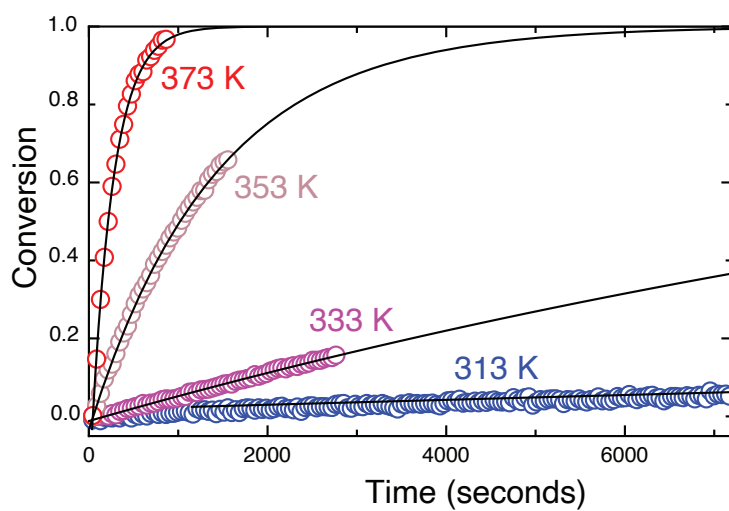
Figure S3: Reaction process at 353K

Figure S4: **Diallene-nBu** kinetics at different temperatures

Table S1: Summary of thermolysis results of **Diallene-nBu** in dark

Conditions			Conversion	E:Z ratio
Temperature	Time	Solvent		
140 °C	10 min	PhMe	100%	6:2:1
100 °C	20 min	PhMe	85%	6.7:1
100 °C	16 h	PhMe	100%	6.5:1
80 °C	40 min	PhMe	50%	6.5:1
80 °C	16 h	PhMe	100%	6.5:1
60 °C	60 min	PhMe	15%	6.6:1
60 °C	16 h	PhMe	90%	6.7:1
60 °C	40 h	PhMe	100%	6.4:1

Table S2: Summary of thermolysis results of **Diallene-nPr** in dark

Conditions			Conversion	E:Z ratio
Temperature	Time	Solvent		
140 °C	10 min	PhMe	100%	10.5:1
120 °C	15 min	PhMe	100%	10.3:1
100 °C	30 min	PhMe	100%	10.3:1
80 °C	1 h	PhMe	80%	10.4:1
80 °C	2 h	PhMe	85%	10.1:1
60 °C	2 h	PhMe	34%	10.2:1

UV-VIS spectrum

As the UV lamp we use in the lab emits a peak at 254 nm, we are looking for absorption in this region. According to TD-DFT results shown in Figure S6 (for the **A** structure), this would fall in the region between the first set of excitations and the absorption peak. Here we should remember that, as mentioned in the main text, an error up to 20 nm is not unlikely for TD-DFT and the lamp might be closer to either the first or second broad peak of the absorption spectrum.

Additionally, each transition will have a vibronic structure that is not captured by simply broadening the 25 computed vertical transitions and the full absorption spectrum is likely
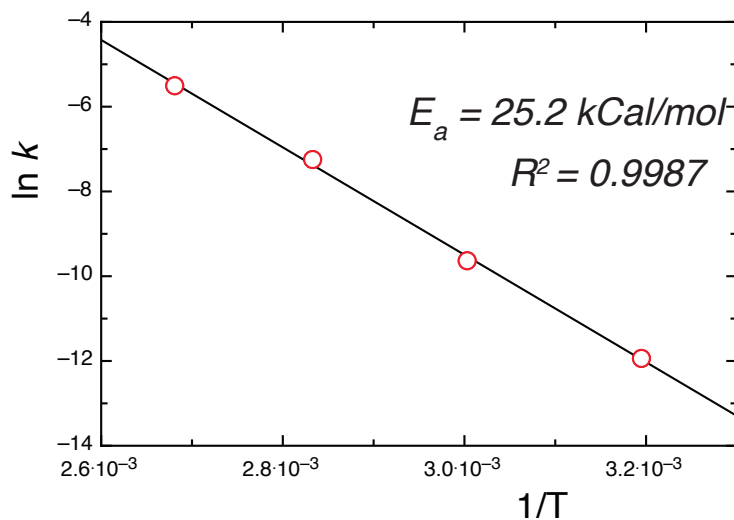
Figure S5: Arrhenius plot

to be more "filled-in". To showcase this, we compute the full spectrum for the S_1 state with the Herzberg-Teller effect included (inset of S6, with a 0-0 transition energy of 32961 cm^{-1}) and show its tail stretches into well into the 200 nm area). For a larger molecule in solvent, such as ours, computing a vibronic spectrum is most likely not practical or necessary, as up to 10 states would need to be included to fully cover the relevant region of the spectrum.

Table S3: TD-DFT transition energies and oscillator strengths for the first 10 excited singlet states at the **A** geometry

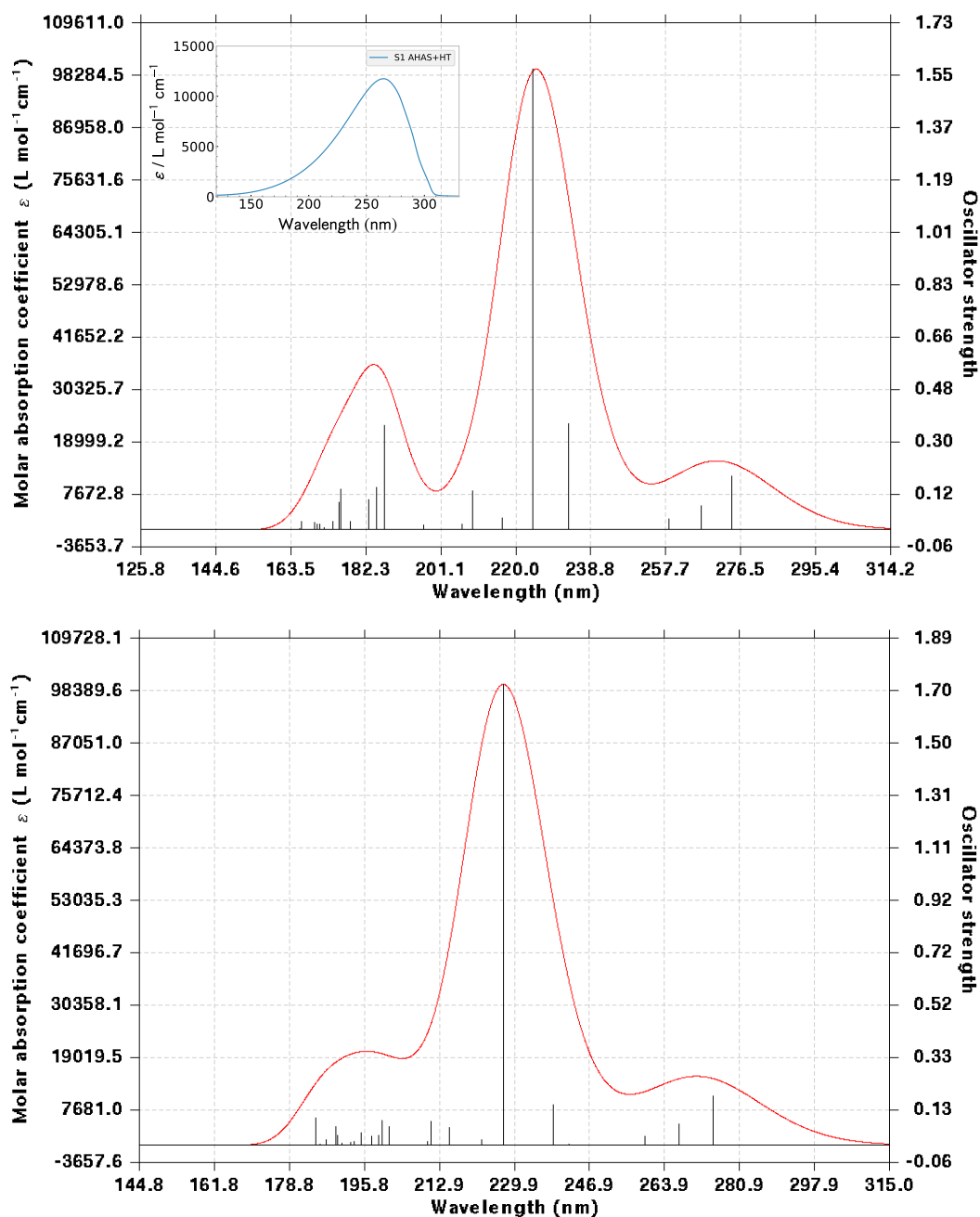|	CAM-B3LYP def2-TZVP			CAM-B3LYP ma-def2-TZVPP		
	E / cm^{-1}	λ / nm	f_{osc}	E / cm^{-1}	λ / nm	f_{osc}
S1	36467.3	274.2	0.18	36368.9	275.0	0.18
S2	37521.4	266.5	0.08	37423.7	267.2	0.08
S3	38673.4	258.6	0.04	38526.4	259.6	0.03
S4	41403.0	241.5	0.00	41292.8	242.2	0.00
S5	42885.6	233.2	0.36	41882.7	238.8	0.15
S6	44580.1	224.3	1.57	43993.1	227.3	1.72
S7	46173.9	216.6	0.04	44973.7	222.4	0.02
S8	47828.0	209.1	0.13	46473.5	215.2	0.07
S9	48430.2	206.5	0.02	47386.2	211.0	0.09
S10	50787.6	196.9	0.02	47568.4	210.2	0.02

Figure S6: Computed TD-DFT UV-VIS spectrum obtained for the first 25 singlet excitations with the CAM-B3LYP functional, implicit CPCM solvation for toluene, and with the def2-TZVP (top) or ma-def2-TZVPP (bottom) basis set; plotted with Multiwfn and a 0.5 eV Gaussian broadening. The inset on the top figure shows the vibronic structure of the S_1 transition, as computed with the adiabatic Hessian after a step (AHAS) approach in ORCA with the Herzberg-Teller effect included.
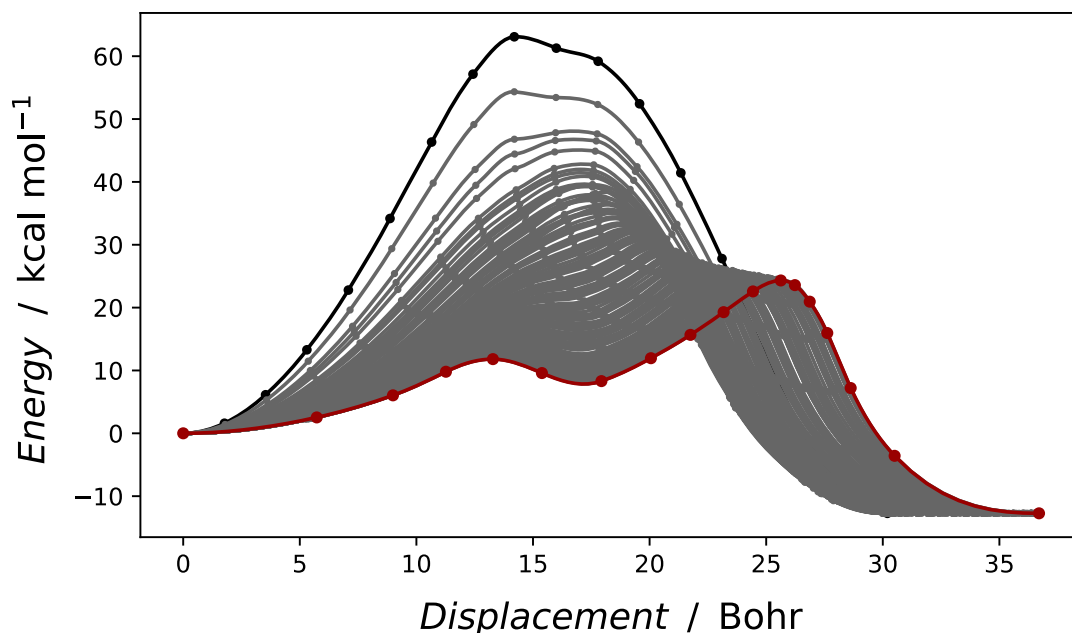
INT$_2$ to INT$_3$ Minimum Energy Path

Figure S7: Converged MEP (in red) from an UKS B3LYP/def2-TZVP D3BJ NEB-CI calculation initiated between the INT$_2$ structure and the INT$_3$ structure. Solvent effects are accounted for with an implicit CPCM model for toluene. The black curve indicates the initial guess for the path, while the grey curves represent the 136 iterations in-between.

Figure S7 shows the result of a MEP calculated between the two different intermediate structures of the **nPr** case. The first small barrier corresponds to conformational changes, torsions in the alkyl chains. The second barrier represents the third ring formation and is also automatically chosen to be the climbing image. The transition state was obtained by a TS optimization initialized at the converged climbing image; its Gibbs free energy compared to the first point amounted to 20.16 kcal mol^{-1} (84.4 kJ mol^{-1}).

At the same time, it serves as a lesson in computational efficiency, as we ran identical calculations with both the def2-SVP and def2-TZVP basis sets. Both NEBs converged in the same number of iterations (138 vs 137), which took 2 and a half days on 16 CPU cores for the smaller basis set calculations and 7 days and 17 hours for the more quantitative, larger one. When compared, the MEPs look qualitatively the same, meaning a transition state

optimization with the final/larger basis set should be performed starting from a cheaper NEB and its climbing image (CI). Even if the TS optimization and free energy calculation are performed with the def2-SVP basis set, the resulting 19.05 kcal mol^{-1} gives us very similar insight as the 20.16 kcal mol^{-1} from the def2-TZVP result.

More computational time may be saved with user intervention – once the MEP starts to exhibit a conformational minimum along the path, as signaled in the output ("possible intermediate minimum found"), the user may stop the calculation and restart the NEB starting at this structure (image). At the same time, fewer total images may now be used to successfully converge the MEP.

Multireference treatment

Fractional Occupation Number Weighted Electron Density (FOD)

A FOD analysis using the default settings in ORCA was performed on the full MEP as obtained by a NEB-CI calculation between **A** and the **E** product. As we can see in Figure S8, the N_{FOD} value, which is a size extensive indicator of static electron correlation, reaches a peak value of 1.97 in the intermediate region between the two transition states, suggesting a diradical system.

Computational details

All calculations employed an ANO-RCC basis set with a double-zeta quality and levshift = 0.1. The convergence thresholds for the SA-RASSCF calculations were set to 1e-8 for the energy, 1e-4 for the orbital rotation matrix, and 1e-4 for the energy gradient.

After a preliminary SCF calculation, the 1s orbitals of all non-hydrogen atoms, as well as the 2s and 2p orbitals of magnesium, were frozen in all following calculations. This amounted to 24 frozen orbitals, 59 inactive for CAS(12,12).

All MS-CASPT2 calculations employed an imaginary shift value of 0.1, with no deleted
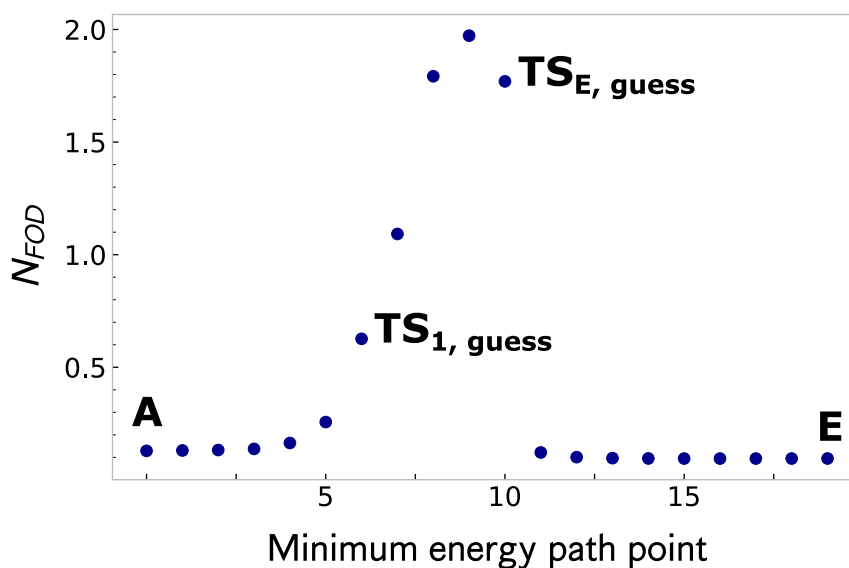
Figure S8: FOD analysis on the geometries of the MEP spanned between **A** and the **E** product

orbitals. Thresholds for the CASPT2 calculation were 1e-10 for the removal of zero-norm components in the first-order perturbed wavefunction and 1e-8 for the removal of near-linear dependencies in the first-order perturbed wavefunction.

All SA-CASSCF/MS-RASPT2 calculations were performed using OpenMolcas version 23.02. The calculations were performed on a dual processor (Intel Xeon Gold 6334 3.6G Dual CPU 1.6TB NVMe) and employed 10 GB of memory.

Multireference energies

In Table S4, the ground state and excited state energies are presented.

Discussion

The wave function configuration from the CAS(12,12) single point calculations in vacuum of all transition states shows significant contributions to the ground state from doubly excited orbitals. For all transition states, the ground state configuration only constitutes 67% of the wave function. The rest of the wave function is comprised of doubly excited orbitals. This

Table S4: SA-CASSCF/MS-CASPT2 ground, first singlet excited state and first triplet excited state energies (in kcal mol^{-1}) for the DFT-optimised structures. Implicit solvation for toluene is applied through PCM (CAS energies) and CPCM (DFT geometries).

Structure	Ground state energy	First excited state
A$_{relaxed}$	0	105.3449
TS$_1$	20.0276	109.0862
INT$_2$	-29.8075	44.4115
INT$_3$	-50.7386	49.4901
TS$_E$	-10.7953	69.3361
TS$_Z$	-4.4292	65.1360
E	-67.9617	31.8366
Z	-70.3384	28.6612

Reference ground state energy in Hartree: -932.5774

Table S5: SA-CASSCF/MS-CASPT2 ground and first singlet excited state (in kcal mol^{-1}) for the DFT-optimised structures in vacuum.

Structure	Ground state energy	First excited state
A$_{relaxed}$	0	105.0388
TS$_1$	18.5280	106.4284
INT$_2$	-31.2598	43.2931
INT$_3$	-50.4572	49.6465
TS$_E$	-12.4317	66.4512
TS$_Z$	-5.1441	64.0751
E	-59.1360	40.3841
Z	-60.7827	38.6924

Reference ground state energy in Hartree: -932.5734

indicates a multireference character for the transition states in the reaction.

The natural orbital occupation numbers showed multireference character for both the ground and excited states of the transition states. Especially the excited states show occupation numbers approaching 1, indicating a strong multireference character. The ground states were weakly to strongly multireference with two natural orbitals with occupations over 0.20 from fully occupied or unoccupied.

TS$_Z$ has a small triplet contribution to both its ground and first excited state, whereas TS$_1$ only has a minor triplet contribution for its ground state, and, finally, TS$_E$ has no contributions from triplet configurations in the wave functions for either state. Based on the occupation numbers, the system is primarily a singlet in these conformations. Therefore, no

calculations for triplet states were performed on the CAS(12,12)/CASPT2 level because the triplet excited state would be much higher in energy.

To ensure the validity of this assumption, a single-point triplet calculation was performed for TS_Z. It should be noted that we assume a coupled reaction pathway for the ground state and excited state reaction based on the common TS_1 and following intermediates and products. Investigation of the spin-orbit coupling (SOC) between the first excited singlet and triplet states for the first TS showed a weak to moderate SOC with a value of 1.075 cm^{-1} . Intersystem crossing is possible but not highly efficient. The three SOC states which derive from the triplet state are degenerate in energy. This illustrates the weak SOC interaction which is insufficient to significantly split the triplet components. The energy difference between the first excited state at TS_1 and the first excited triplet state is 2.16 eV. This suggests a low probability of intersystem crossing at this point.

References

(S1) Samanta, D.; Rana, A.; Schmittel, M. Nonstatistical Dynamics in the Thermal Garratt–Braverman/[1,5]-H Shift of One Ene–diallene: An Experimental and Computational Study. *The Journal of Organic Chemistry* **2014**, *79*, 8435–8439.

Table S6: Representation of the 12 orbitals included in the active space.

RAS1
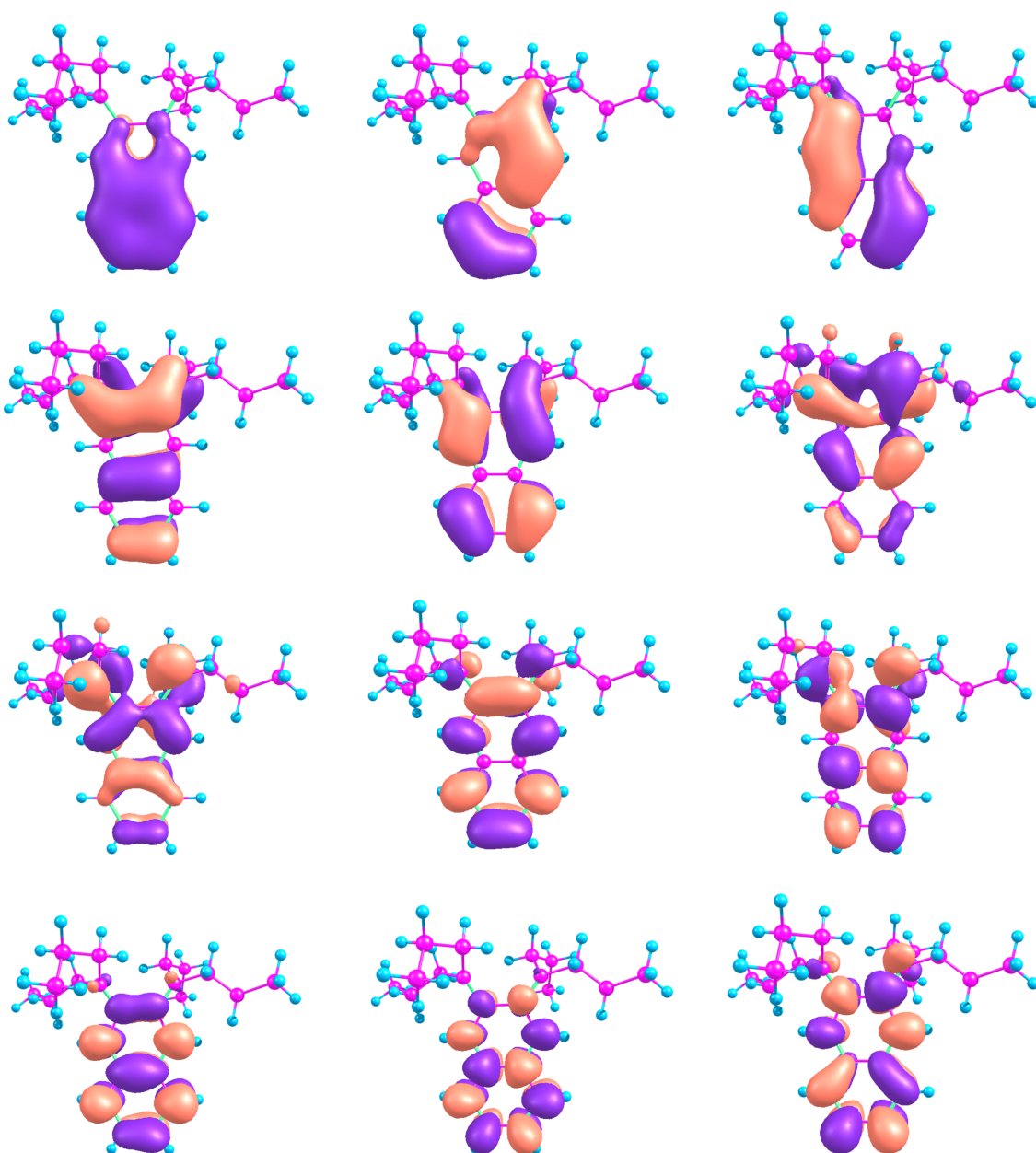
Bibliography

[1] 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. *Journal of the American Chemical Society* **2011**, *133*, 13824–13827.

[2] Glowacki, D. R.; Liang, C. H.; Marsden, S. P.; Harvey, J. N.; Pilling, M. J. Alkene Hydroboration: Hot Intermediates That React While They Are Cooling. *Journal of the American Chemical Society* **2010**, *132*, 13621–13623.

[3] Hare, S. R.; Pinhas, A. R.; Rehbein, J.; Tantillo, D. J.; Wiggins, S. Barry Carpenter—Pioneering physical organic chemist, teacher, mentor, and friend. *Journal of Physical Organic Chemistry* **2022**, *35*.

[4] Carpenter, B. K.; Harvey, J. N.; Orr-Ewing, A. J. The Study of Reactive Intermediates in Condensed Phases. *Journal of the American Chemical Society* **2016**, *138*, 4695–4705.

[5] Yamataka, H. In *Molecular dynamics simulations and mechanism of organic reactions: non-TST behaviors*; Richard, J. P., Ed.; Advances in Physical Organic Chemistry; Academic Press, 2010; Vol. 44; pp 173–222.

[6] Teynor, M. S.; Scott, W.; Ess, D. H. Catalysis with a Skip: Dynamically Coupled Addition, Proton Transfer, and Elimination during Au- and Pd-Catalyzed Diol Cyclizations. *ACS Catalysis* **2021**, *11*, 10179–10189.

[7] Karmakar, S.; Keshavamurthy, S. Intramolecular vibrational energy redistribution and the quantum ergodicity transition: a phase space perspective. *Physical Chemistry Chemical Physics* **2020**, *22*, 11139–11173.

[8] Lee, S.; Goodman, J. M. VRAI-selectivity: Calculation of selectivity beyond transition state theory. *Organic and Biomolecular Chemistry* **2021**, *19*, 3940–3947.

[9] Guo, H.; Jiang, B. The Sudden Vector Projection Model for Reactivity: Mode Specificity and Bond Selectivity Made Simple. *Accounts of Chemical Research* **2014**, *47*, 3679–3685.

[10] Glowacki, D. R.; Harvey, J. N.; Mulholland, A. J. Taking Ockham's razor to enzyme dynamics and catalysis. *Nature Chemistry* **2012**, *4*, 169–176.

[11] Zhao, Q.; Møller, K. H.; Chen, J.; Kjaergaard, H. G. Cost-Effective Implementation of Multiconformer Transition State Theory for Alkoxy Radical Unimolecular Reactions. *The Journal of Physical Chemistry A* **2022**, *126*, 6483–6494.

[12] Bao, J. L.; Truhlar, D. G. Variational transition state theory: theoretical framework and recent developments. *Chemical Society Reviews* **2017**, *46*, 7548–7596.

[13] Mandelli, G.; Aieta, C.; Ceotto, M. Heavy Atom Tunneling in Organic Reactions at Coupled Cluster Potential Accuracy with a Parallel Implementation of Anharmonic Constant Calculations and Semiclassical Transition State Theory. *Journal of Chemical Theory and Computation* **2022**, *18*, 623–637.

[14] Conte, R.; Aieta, C.; Cazzaniga, M.; Ceotto, M. A Perspective on the Investigation of Spectroscopy and Kinetics of Complex Molecular Systems with Semiclassical Approaches. *The Journal of Physical Chemistry Letters* **2024**, *15*, 7566–7576.

[15] Yang, Z.; Jamieson, C. S.; Xue, X.-S.; Garcia-Borràs, M.; Benton, T.; Dong, X.; Liu, F.; Houk, K. Mechanisms and Dynamics of Reactions Involving Entropic Intermediates. *Trends in Chemistry* **2019**, *1*, 22–34.

[16] Feng, H.; Zhang, K.; Wang, J. Non-equilibrium transition state rate theory. *Chem. Sci.* **2014**, *5*, 3761–3769.

[17] Kumar Paul, A.; Kolakkandy, S.; Pratihar, S.; Hase, W. L. *Reaction Rate Constant Computations: Theories and Applications*; The Royal Society of Chemistry, 2013.

[18] NIST Computational Chemistry Comparison and Benchmark DataBase. https://cccbdb.nist.gov/vsfx.asp.

[19] Goldman, L. M.; Glowacki, D. R.; Carpenter, B. K. Nonstatistical dynamics in unlikely places: [1,5] Hydrogen migration in chemically activated cyclopentadiene. *Journal of the American Chemical Society* **2011**, *133*, 5312–5318.

[20] Longfellow, R. J.; Parmenter, C. S. How a (nearly) free methyl rotor accelerates intramolecular vibration relaxation. Theory and Experiment. *Journal of the Chemical Society, Faraday Transactions 2* **1988**, *84*, 1499.

[21] Mališ, M.; Loquais, Y.; Gloaguen, E.; Jouvet, C.; Brenner, V.; Mons, M.; Ljubić, I.; Došlić, N. Non-radiative relaxation of UV photoexcited phenylalanine residues: probing the role of conical intersections by chemical substitution. *Physical Chemistry Chemical Physics* **2014**, *16*, 2285.

[22] Leitner, D. M.; Pandey, H. D. Asymmetric energy flow in liquid alkylbenzenes: A computational study. *The Journal of Chemical Physics* **2015**, *143*, 144301.

[23] Yu, X.; Leitner, D. M. Vibrational Energy Transfer and Heat Conduction in a Protein. *The Journal of Physical Chemistry B* **2003**, *107*, 1698–1707.

[24] Kuzmin, M.; Letokhov, V.; Stuchebryukhov, A. Threshold energy dependence of the intramolecular vibrational relaxation rate tor an isolated polyatomic molecule. *Sov. Phys. JETP* **1986**, *63*, 264–271.

[25] Barone, V. Vibrational zero-point energies and thermodynamic functions beyond the harmonic approximation. *The Journal of Chemical Physics* **2004**, *120*, 3059–3065.

[26] Piccardo, M.; Bloino, J.; Barone, V. Generalized vibrational perturbation theory for rotovibrational energies of linear, symmetric and asymmetric tops: Theory, approximations, and automated approaches to deal with medium-to-large molecular systems. *International Journal of Quantum Chemistry* **2015**, *115*, 948–982.

[27] Barnes, L.; Schindler, B.; Compagnon, I.; Allouche, A.-R. iGVPT2 : an interface to computational chemistry packages for anharmonic corrections to vibrational frequencies. **2017**, arXiv:1704.02144.

[28] Samanta, D.; Rana, A.; Schmittel, M. Nonstatistical Dynamics in the Thermal Garratt–Braverman/[1,5]-H Shift of One Ene–diallene: An Experimental and Computational Study. *The Journal of Organic Chemistry* **2014**, *79*, 8435–8439.

[29] Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions. *Physical Chemistry Chemical Physics* **2017**, *19*, 32184–32215.

[30] Mardirossian, N.; Head-Gordon, M. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Molecular Physics* **2017**, *115*, 2315–2372.

[31] Feng, Z.; Guo, W.; Kong, W.-Y.; Chen, D.; Wang, S.; Tantillo, D. J. Analogies between photochemical reactions and ground-state post-transition-state bifurcations shed light on dynamical origins of selectivity. *Nature Chemistry* **2024**, *16*, 615–623.

[32] Baiardi, A.; Stein, C. J.; Barone, V.; Reiher, M. Vibrational Density Matrix Renormalization Group. *Journal of Chemical Theory and Computation* **2017**, *13*, 3764–3777.

[33] Glaser, N.; Baiardi, A.; Reiher, M. Flexible DMRG-Based Framework for Anharmonic Vibrational Calculations. *Journal of Chemical Theory and Computation* **2023**, *19*, 9329–9343.

[34] Baiardi, A.; Reiher, M. Large-Scale Quantum Dynamics with Matrix Product States. *Journal of Chemical Theory and Computation* **2019**, *15*, 3481–3498.

[35] Benten, R. S. V.; Liu, Y.; Abel, B. Dynamical consequences of symmetry breaking in benzene and difluorobenzene. *Journal of Chemical Physics* **2010**, *133*.

[36] Glaser, N.; Baiardi, A.; Lieberherr, A. Z.; Reiher, M. Vibrational Entanglement through the Lens of Quantum Information Measures. *The Journal of Physical Chemistry Letters* **2024**, *15*, 6958–6965.

[37] Reiher, M.; Wiebe, N.; Svore, K. M.; Wecker, D.; Troyer, M. Elucidating reaction mechanisms on quantum computers. *Proceedings of the National Academy of Sciences* **2017**, *114*, 7555–7560.

[38] Sparrow, C.; Martín-López, E.; Maraviglia, N.; Neville, A.; Harrold, C.; Carolan, J.; Joglekar, Y. N.; Hashimoto, T.; Matsuda, N.; O'Brien, J. L.; Tew, D. P.; Laing, A. Simulating the vibrational quantum dynamics of molecules using photonics. *Nature* **2018**, *557*, 660–667.

[39] Zhang, C.; Wolynes, P. G.; Gruebele, M. Quantum information scrambling in molecules. *Physical Review A* **2022**, *105*, 033322.

[40] Kapil, V. et al. i-PI 2.0: A universal force engine for advanced molecular simulations. *Computer Physics Communications* **2019**, *236*, 214–223.

[41] Suleimanov, Y. V.; Aoiz, F. J.; Guo, H. Chemical Reaction Rate Coefficients from Ring Polymer Molecular Dynamics: Theory and Practical Applications. *The Journal of Physical Chemistry A* **2016**, *120*, 8488–8502.

[42] Marsalek, O.; Markland, T. E. Ab initio molecular dynamics with nuclear quantum effects at classical cost: Ring polymer contraction for density functional theory. *The Journal of Chemical Physics* **2016**, *144*.

[43] Rossi, M.; Ceriotti, M.; Manolopoulos, D. E. How to remove the spurious resonances from ring polymer molecular dynamics. *The Journal of Chemical Physics* **2014**, *140*.

[44] Rather, S. R.; Bezdek, M. J.; Chirik, P. J.; Scholes, G. D. Dinitrogen Coupling to a Terpyridine-Molybdenum Chromophore Is Switched on by Fermi Resonance. *Chem* **2019**, *5*, 402–416.

[45] Lindoy, L. P.; Mandal, A.; Reichman, D. R. Quantum dynamical effects of vibrational strong coupling in chemical reactivity. *Nature Communications* **2023**, *14*, 2733.

[46] Mondal, S.; Keshavamurthy, S. Cavity induced modulation of intramolecular vibrational energy flow pathways. *The Journal of Chemical Physics* **2024**, *161*.

[47] O'Neill, R. T.; Boulatov, R. The many flavours of mechanochemistry and its plausible conceptual underpinnings. *Nature Reviews Chemistry* **2021**, *5*, 148–167.

[48] Sun, Y.; Kevlishvili, I.; Kouznetsova, T. B.; Burke, Z. P.; Craig, S. L.; Kulik, H. J.; Moore, J. S. The tension-activated carbon–carbon bond. *Chem* **2024**, *10*, 3055–3066.

[49] Wang, L.-P.; Titov, A.; McGibbon, R.; Liu, F.; Pande, V. S.; Martínez, T. J. Discovering chemistry with an ab initio nanoreactor. *Nature Chemistry* **2014**, *6*, 1044–1048.

[50] Pieri, E.; Lahana, D.; Chang, A. M.; Aldaz, C. R.; Thompson, K. C.; Martínez, T. J. The non-adiabatic nanoreactor: towards the automated discovery of photochemistry. *Chemical Science* **2021**, *12*, 7294–7307.

[51] Zhang, L.; Pios, S. V.; Martyka, M.; Ge, F.; Hou, Y.-F.; Chen, Y.; Chen, L.; Jankowska, J.; Barbatti, M.; Dral, P. O. MLatom Software Ecosystem for Surface Hopping Dynamics in Python with Quantum Mechanical and Machine Learning Methods. *Journal of Chemical Theory and Computation* **2024**, *20*, 5043–5057.