

Ambimodality and Asymmetry in [4+2] Cycloadditions: a Computational and Experimental Study



Alexander Makaveev
St John's College

A thesis submitted for the Honour School of Chemistry
Part II 2023

Acknowledgements

It is only right that I begin by thanking my family, and in particular my parents Ekaterina and Nikolay. Without their unreserved love no endeavour of mine could be successful.

I would like to thank the wonderful educators who taught me at Sofia High School of Mathematics, and especially Petya Todorova, Kiril Martinov, Kalina Tsoneva, Boris Tolev, Ivaylo Ushagelov, Ani Bozhikova, Olga Tuncheva, Tanya Nikolova, and Zornitsa Dimitrova.

Many thanks to Professors Fernanda Duarte and Edward Anderson for giving me the opportunity to work on this exciting and challenging project, and for their support and mentorship throughout the year.

I also thank Tristan for his very patient supervision; Nils, for designing the project, and many useful discussions; Ayan, Kenny, and Jisook, for helping me survive lab work; Tomasz, Aleksy, Shoubhik, and Martin, for teaching me how computers work; Valentin, Tim, Valdas, and Josh, for sharing the joys and struggles of the part II year with me; and last but not least, everyone in the FD and EAA groups for their kindness and the fun we had together.

Finally, I would like to thank Matilde for all she has done for me throughout the year without realising.

Summary

We present a computational and experimental investigation into the mechanism of an apparent inverse electron demand $[4 + 2]$ cycloaddition followed by a SO_2 cheletropic extrusion in a cascade manner. The reaction is proposed to proceed *via* an ambimodal transition state, followed by a minimum energy path bifurcation allowing for the formation of both a normal and an inverse electron demand adduct. This transformation was reported in a study by the Duarte and Anderson groups, and used for the collective synthesis of a family of sesquiterpenes.

This work employs computational and physical organic chemistry methods to further study the mechanism of this cycloaddition. We report quantum mechanical computations that support the ambimodal hypothesis and give new information on the Cope rearrangement that connects the two adducts. *Ab initio* molecular dynamics simulations confirm the bipericyclic nature of the transition state, and provide novel insight into the selectivity and asynchronicity of the reaction. Finally, we report the results from a $^{13}\text{C}/^{12}\text{C}$ kinetic isotope effect experiment, which support the ambimodality hypothesis by ruling out alternative mechanisms, such as a stepwise addition or a regular $[4 + 2]$ cycloaddition.

Contents

List of Figures	vi
List of Tables	viii
List of Abbreviations	ix
1 Introduction	1
1.1 Potential Energy Surfaces	1
1.2 Statistical Theories of Chemical Reactivity	3
1.3 Non-Statistical Behaviour: PES Bifurcation	5
1.4 Cycloadditions with Ambimodal Transition States	9
1.5 Aims and Strategy	12
2 Theory and Methods	14
2.1 Static QM Methods	14
2.1.1 Hartree-Fock Theory	15
2.1.2 Accounting for Correlation: Configuration Interaction and Møller-Plesset Perturbation	16
2.1.3 Density Functional Theory	17
2.1.4 Basis Sets	20
2.1.5 Solvation Models	21
2.2 Dynamic Methods	21
2.3 ^{13}C NMR for KIE studies	23
3 Computational Investigation of the Ambimodal Cycloaddition	25
3.1 Choosing the Level of Theory	25
3.2 Energy Profile of the System of Interest	28
3.3 <i>Ab Initio</i> Molecular Dynamics of the Ambimodal Cycloaddition	30
3.3.1 Evaluation of the Level of Theory	30
3.3.2 Product Ratios and PES Connectivity	32
3.3.3 Bond Formation Dynamics: Asynchronicity and Asymmetry	34
3.4 KIE Predictions	35
3.5 Computational Details	38

4	Experimental KIE Results	39
4.1	Overview of Reactions	40
4.2	KIE Measurements	40
4.3	Experimental Details	45
4.3.1	Synthesis and Characterisation	45
4.3.2	NMR Analysis for KIE Determination	47
5	Conclusions and Outlook	48
5.1	Conclusion	48
5.2	Future Work	49
	Bibliography	51
	Appendices	
A	More on QM	60
A.1	Obtaining the HF Energy	60
A.2	Treatment of Fermionic Correlation in HF	61
A.3	More on DFT and Jacob's Ladder	61
B	KIE Results	65
C	NMR Spectra	67

List of Figures

1.1	Potential energy surfaces (PESs) for systems of different dimensions. Left: 1D PED for a diatomic system. Centre: A 3D PES. Right: Contour plot for the same PES.	2
1.2	Schematic reaction profile with depicted activation barrier ΔG^\ddagger and free energy change $\Delta_r G$	5
1.3	Contour plot of a PES for a schematic bifurcating reaction (yellow: high energy, purple: low energy). Indicated are the MEP (black), ambimodal transition state (TS1), the valley-ridge inflection point (VRI), and the path connecting intermediates IM1 and IM2 through TS2 (yellow).	7
1.4	Some early examples of organic reactions with ambimodal TSs. (A) Electrocyclic ring opening of a cyclopropyl ring. The 1,2-sigmatropic shift and the disrotatory ring opening occur sequentially without an intermediate. Rotation in either direction gives a different product. (B) Cope rearrangement of 1,2,6-heptatriene bifurcates to give both the expected polyene and a much more unstable diradical species. (C) Singlet oxygen $^1\text{O}_2$ ene peroxidation bifurcates to give either enantiomer. (D) Caramella's findings of a bispericyclic TS for the dimerisation of cyclopentadiene.	8
1.5	Proposed ambimodal mechanistic pathway for the reaction of interest. The two DA adducts IM1 (IED) and IM2 (NED) are shown; only IM1 can extrude SO_2 to give the observed product. Adapted from ref [1].	10
1.6	Geometry of bipericyclic TS reported by Duarte and Anderson with key bond lengths indicated. Adapted from ref [1].	12
3.1	The ambimodal cycloaddition step of interest, giving both the normal electron demand adduct IM2 , and the inverse electron demand IM1 , from the same bipericyclic TS.	26
3.2	Ambimodal TS structure with key bond distances labeled.	27
3.3	Reaction energy profile for the system of interest at CPCM(Toluene)-SCS-RI-MP2/de2-TZVPP//CPCM(Toluene)- ω B97X-D3(BJ)/def2-TZVPP.	28

3.4	Basis set benchmark study for AIMD at ω B97X-D3(BJ) using the geometries of IM1 , IM2 , and TS1 . Note SVP means def2-SVP, +C stands for a def2-TZVP expansion on the 6 carbons active in the ambimodal TS1 , +het stands for a def2-TZVP expansion on all heteroatoms, and +het+C is a combination of the latter two. Left: heavy-atom RMSD of geometries obtained using the various basis sets relative to the ones obtained with def2-TZVPP for IM1 , IM2 , and TS1 . Right: computational expense measured as time required for a single point computation on the TS1 structure using 4 cores. .	31
3.5	Product distributions from MD runs at ω B97X-D3(BJ)/def2-SVP initiated at each of the three ambimodal TS structures (see text in sections 3.1 and 3.3). Note the final geometries of all 10 MA trajectories gave IM1 or IM2 upon static optimisation.	32
3.6	Bond formation as a function of time from the ω B97X-D3BJ/def2-SVP AIMD simulations that resulted in a DA adduct. Labelling and colour-coding of bonds shown on TS1 structure on the right.	34
3.7	Schematic TS structure with relevant atoms and bonds labelled. . .	36
4.1	General reaction scheme for the preparation of thiophene dioxide 1 and the cycloaddition step.	40
4.2	Structures of the final product (top) and the ambimodal TS1 (bottom) with atom labels for reference.	42
4.3	Structure of TS1 with atom labels for reference.	43

List of Tables

2.1	Jacob’s ladder of DFT functionals, with a summary of the relevant mathematical terms, and some notable examples. The arrow on the left signifies the general increase in chemical accuracy and computational expense going up the ladder. See Appendix A.3 for further details.	
	*Including range separated hybrids.	20
3.1	Bond distances in Å for the ambimodal TS structure (Fig. 3.2 on the left) at different levels of theory. Computations performed with the def2-TZVPP basis set and CPCM(toluene) correction. The quantity $r_3 - r_2$ is given as a measure of asynchronicity.	27
3.2	Summary of computed KIEs for key carbon atoms in the ambimodal cycloaddition and Cope rearrangement TSs at CPCM(toluene)- ω B97X-D3(BJ). Fig. 3.7 for atom labels.	36
4.1	Summary of average KIE measured.	
	See figure on the right for atom labels.	42
4.2	Comparison between computed (for ambimodal cycloaddition TS1 and Cope rearrangement TS2) and experimental KIEs. See Tables 3.2 and 4.1 for full information.	43
A.1	Jacob’s ladder of DFT functionals, with a summary of the relevant mathematical terms, and some notable examples. The arrow on the left signifies the general increase in chemical accuracy and computational expense going up the ladder.	
	*Including range separated hybrids.	63
B.1	Details from KIE measurements and statistical analysis.	66

List of Abbreviations

PES	Potential energy surface.
TS	Transition state.
RRK(M)	Rice-Ramsperger-Kassel(-Marcus) theory.
TST	Transition state theory.
ZPE	Zero-point energy.
IVR	Intramolecular vibrational relaxation.
IM	Intermediate.
VRI	Valley-ridge inflection.
MEP	Minimum energy path
DA	Diels-Alder
NED	Normal electron demand.
IED	Inverse electron demand.
QM	Quantum mechanics.
KIE	Kinetic isotope effect.
MD	Molecular dynamics.
AIMD	<i>Ab initio</i> molecular dynamics.
RMSD	Root mean square deviation.
HF	Hartree-Fock
CI	Configuration interaction.
MP2	Second-order Møller-Plesset perturbation
SCS	Spin component scaled (relating to MP2)
DFT	Density functional theory
IRC	Intrinsic reaction coordinate
NEB	Nudged elastic band.
LDA	Local density approximation.

GGA	Generalised gradient approximation.
CPCM	Conductor-like polarisable continuum model
NMR	Nuclear magnetic resonance
DMF	Dimethylformamide
TFAA	Trifluoroacetic anhydride
r.t.	Room temperature

1

Introduction

Contents

1.1	Potential Energy Surfaces	1
1.2	Statistical Theories of Chemical Reactivity	3
1.3	Non-Statistical Behaviour: PES Bifurcation	5
1.4	Cycloadditions with Ambimodal Transition States . .	9
1.5	Aims and Strategy	12

The present study is an investigation into a $[4 + 2]$ cycloaddition reaction reported by our two research groups, which is thought to proceed *via* an unusual bipericyclic mechanism.[1] This chapter outlines the theoretical background that underpins the study. We explain the potential shortcomings of traditional methods such as Transition State Theory[2], and introduce alternative approaches to conceptualising chemical reactivity.

1.1 Potential Energy Surfaces

A potential energy surface (PES) describes the energy of a system as a function of a set of variables. In chemistry, these are usually the Cartesian coordinates, which

can be external (defined with respect to some absolute reference frame) or internal (describing the atomic positions relative to each other, e.g. bond distances and angles). Under the Born-Oppenheimer approximation, which states that nuclei move much more slowly than electrons, the PES maps a geometric configuration of atoms to a single energy value.

A key aspect of a PES is its dimensionality, i.e. how many variables determine the energy value (Fig. 1.1). For a reaction involving only two atoms, there is a single chemically relevant coordinate – the interatomic distance. The PES collapses into a one-dimensional potential energy curve as it is a monovariate function.

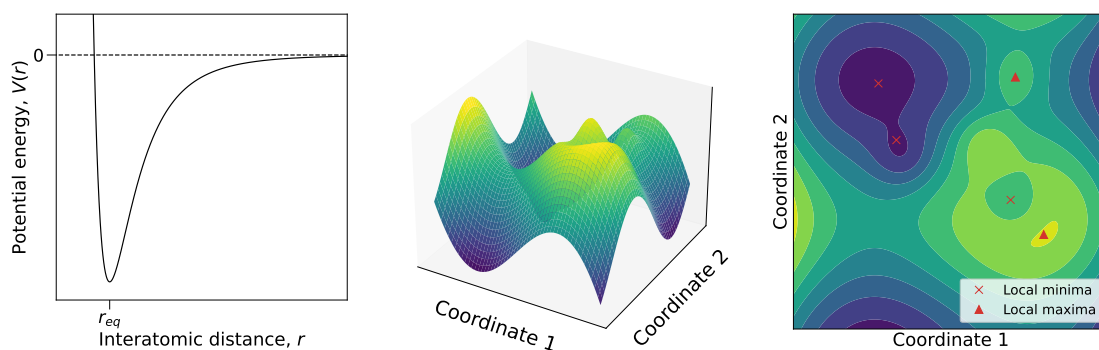


Figure 1.1: Potential energy surfaces (PESs) for systems of different dimensions. Left: 1D PED for a diatomic system. Centre: A 3D PES. Right: Contour plot for the same PES.

However for systems with $N > 2$ atoms, the dimensionality of the PES increases as the number of chemically relevant degrees of freedom, $3N - 6$. This means that for almost any system of interest the PES cannot be visualised in 3D. For this reason, the PES is often represented in terms of fewer coordinates, chosen to describe the process of interest well. This dimensionality reduction inevitably leads to loss of information.

On the PES, intermediates are true minima and have strictly positive curvatures (partial second derivatives) in all dimensions; TSs are first-order saddle points and have negative curvature in exactly one dimension (one negative partial second derivative).

Central to exploring reaction mechanisms is the Minimum Energy Path (MEP) associated with a process of interest. This is the path with the lowest energy barrier (going through the lowest-energy TS) of those that connect two points A and B on the PES. In most cases, the MEP is a good approximation for the average trajectory for a reaction $A \rightarrow B$.

1.2 Statistical Theories of Chemical Reactivity

Chemists have relied on statistical models of chemical reactivity to rationalise and predict mechanisms, rates, and selectivities. An early formulation of such a theory is the Arrhenius equation,[3] which was empirically derived:

$$k = A \cdot e^{-\frac{E_a}{k_B T}}, \quad (1.1)$$

where k is the rate constant, A is a pre-exponential factor, E_a is the activation energy, k_B is the Boltzmann constant, and T is the temperature. The Boltzmann factor $e^{-\frac{E_a}{k_B T}}$ expresses the probability that a molecule has energy $E \geq E_a$, assuming a Boltzmann distribution of energies. The pre-exponential factor A accounts for other contributions, such as the spacial orientation of reactants.

A notable later model was the Rice-Ramsperger-Kassel (RRK) theory for unimolecular gas reactions, developed in the 1920s.[4] It treats the reaction rate as a function of the molecular energy ϵ , and assumes it depends on the probability that the energy distribution is such that a certain critical energy ϵ_c is localised in a particular degree of freedom – which makes RRK a statistical theory.

In the 1930s, Eyring developed Transition State Theory (TST, also known as “activated complex theory”).[2] TST assumes a rapid pre-equilibrium between the reactants and the activated complex, and applies statistical mechanics to express the equilibrium concentration of species in terms of their molecular partition

functions. The theory gives a more proper quantum mechanical description of reactions by accounting for zero point energies (ZPEs). The general form of the Eyring equation is:

$$k = \frac{\kappa k_B T}{h} e^{-\frac{\Delta G^\ddagger}{k_B T}}, \quad (1.2)$$

where κ is a transmission coefficient describing the probability that a visit to the TS results in formation of the product (usually $\kappa \approx 1$), h is Plank's constant, and ΔG^\ddagger is the Gibbs free energy of activation.

Ultimately Marcus, who also modelled electron transfer reactions,[5] generalised the work into the so-called Rice-Ramsperger-Kassel-Marcus (RRKM) theory in the 1950s,[6, 7, 8, 9] which has been the dominant conceptual framework chemists use to gain a physical understanding of chemical reactivity.

RRKM/TST relies on a number of assumptions: (i) there is a pre-equilibrium formed between the activated complex C^\ddagger and the reactants A and B ; (ii) there exists a low-frequency vibrational mode responsible for converting the activated complex to the product; (iii) this vibrational mode is slow enough to allow for intramolecular vibrational relaxation (IVR) to be rapid on the reaction timescale.

A simple single-step reaction can be used to illustrate the basic ideas behind RRKM/TST. The potential energy as function of the reaction coordinate is shown in Fig 1.2. For the reaction to occur, the reactant cluster R needs to possess enough energy to go over the activation barrier ΔG^\ddagger and reach the transition state (TS), which can ultimately collapse to the product cluster P. Therefore, the rate of the reaction will depend on the probability that a reactant cluster R possesses energy no less than the activation barrier, which is the essence of the theory's statistical nature.

Ultimately, within the framework of the statistical theory of chemical reactivity, all properties of the reaction can be derived from the energies of stationary points

and state parameters by applying Boltzmann statistics.

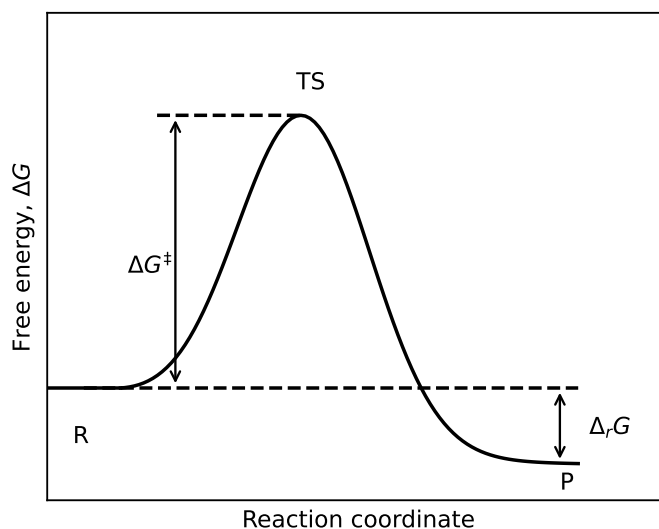


Figure 1.2: Schematic reaction profile with depicted activation barrier ΔG^\ddagger and free energy change $\Delta_r G$.

1.3 Non-Statistical Behaviour: PES Bifurcation

The RRKM/TST conceptual framework has proven to be a reliable way to rationalise reactivity and selectivity, and most examples of quantitative predictions of rates, turn-over frequencies, selectivities etc. rely on it.[10] Nevertheless, this theory is based on a series of assumptions that do not necessarily hold for all systems.

One of these assumptions that is particularly relevant to the current study is that IVR is rapid on the reaction timescale. This means that chemical species will redistribute the energy stored in their vibrational modes to give a Boltzmann distribution before they react further. This leads to a loss of information, meaning that the state of the species is effectively independent of the path that was taken to arrive at it.

This approximation is however not true for all systems.[11, 12, 13] Breakdowns of this assumption are at the heart of non-statistical phenomena. Systems in which reactivity is determined by non-statistical factors can be separated into three

major groups: those with non-equilibrium dynamic populations (as opposed to a Boltzmann distribution); those wherein reactivity is governed by the shape of the PES (as opposed to only the energies of stationary points); and those where quantum tunneling becomes relevant.

When a reaction pathway goes through a reactive intermediate – i.e., a shallow minimum on the PES, it can exhibit non-statistical behaviour due to an unequilibrated dynamic population. As only molecules with a certain degree of rovibrational excitation can pass through a given TS, the ones that arrive at the intermediate IM will initially possess a non-Boltzmann internal distribution of energies. If the IM is particularly reactive, it is possible for it to undergo further transformations before IVR can take place fully, allowing dynamics to affect reactivity as opposed to solely statistics. Manifestations of this phenomenon that have been observed practically include major products being formed by a higher energy barrier than minor, or a symmetric IM giving an unequal product distribution.[12, 14, 15]

Another type of non-statistical behaviour is bifurcation of the MEP. This is a phenomenon defined by a single TS structure leading to two distinct products. In the general case, the MEP on the PES of the reaction passes through a common TS, usually called ambimodal (TS1 on Fig. 1.3). The MEP then goes through a valley-ridge inflection (VRI) point, where the bifurcation occurs – the valley transforms into a dynamically unstable ridge and the MEP branches into two channels, which ultimately gives two intermediates (local minima) IM1 and IM2 (Fig. 1.3). These are also connected by a second barrier TS2, which lies on the pre-bifurcation MEP trajectory.

Usually, the ambimodal TS1 and the VRI occupy a relatively flat region of the PES, sometimes called a caldera. It is often the case that multiple stationary points (first order saddles or shallow minima) can be found in the caldera.[16, 17]

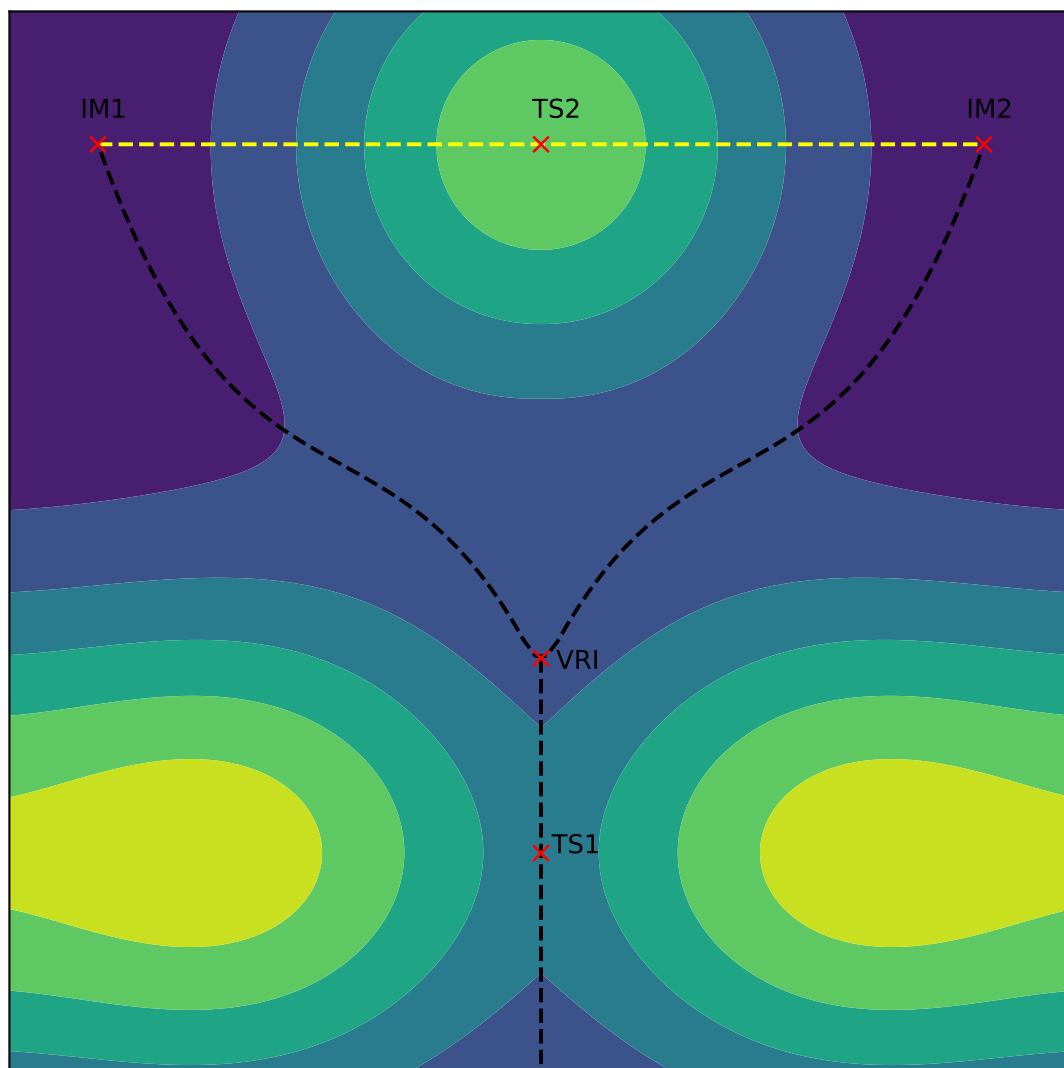


Figure 1.3: Contour plot of a PES for a schematic bifurcating reaction (yellow: high energy, purple: low energy). Indicated are the MEP (black), ambimodal transition state (TS1), the valley-ridge inflection point (VRI), and the path connecting intermediates IM1 and IM2 through TS2 (yellow).

In the recent decades, there has been a considerable and increasing number of reported ambimodal reactions. Examples include unimolecular isomerisations (e.g. methoxy radical[18], cyclooctatriene[19, 20]), substitutions (e.g. aldehyde radical anion and alkyl halide[21]), and importantly numerous pericyclic reactions of all four categories.[11]

Some of the earliest studied ambimodal reactions were the electrocyclic ring

openings of cyclopropylidene to give allene[22], and of cyclopropylcarbinyl carbocation (where the H-1,2-sigmatropic shift and the disrotatory ring opening occur sequentially without an intermediate, Fig. 1.4A).[23] An example of a bifurcating sigmatropic rearrangement is the Cope rearrangement of 1,2,6-heptatriene to give both the concerted product and a diradical by-product from a single ambimodal TS (Fig. 1.4B).[24] Ene reactions, e.g. the singlet-oxygen ene peroxidation, have also been demonstrated to proceed *via* an ambimodal TS (Fig. 1.4C).[11, 25]

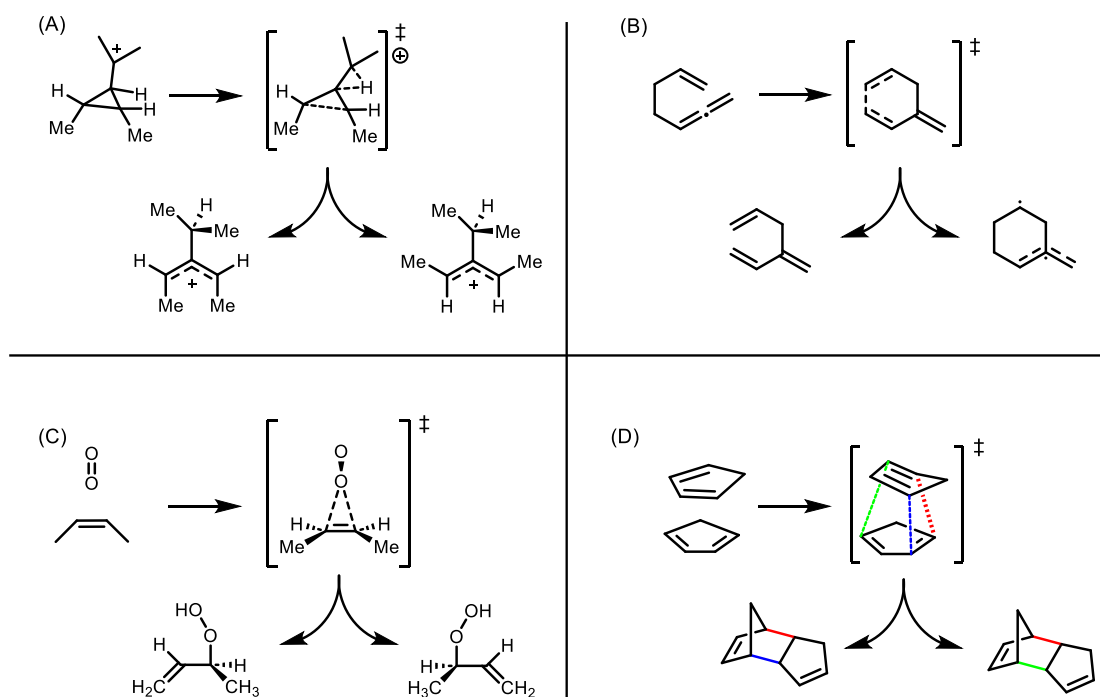


Figure 1.4: Some early examples of organic reactions with ambimodal TSs. (A) Electrocyclic ring opening of a cyclopropyl ring. The 1,2-sigmatropic shift and the disrotatory ring opening occur sequentially without an intermediate. Rotation in either direction gives a different product. (B) Cope rearrangement of 1,2,6-heptatriene bifurcates to give both the expected polyene and a much more unstable diradical species. (C) Singlet oxygen $^1\text{O}_2$ ene peroxidation bifurcates to give either enantiomer. (D) Caramella's findings of a bispericyclic TS for the dimerisation of cyclopentadiene.

1.4 Cycloadditions with Ambimodal Transition States

Perhaps the most diverse and chemically relevant class of pericyclic reactions with bifurcating PESs are cycloadditions. Not surprisingly, Diels-Alder (DA) and related reactions have been studied extensively.[26, 27, 28] Even the simplest DA addition between *s-cis*-butadiene and ethene has been theoretically shown to have a MEP that bifurcates to give either half-chair conformation.[29, 30]

Of greater significance is Caramella’s work showing that the DA dimerisation of cyclopentadiene goes through a highly asynchronous symmetric ambimodal TS.[31] This can be thought of as a mix between a $[4 + 2]$ and $[2 + 4]$ cycloaddition, featuring one short bond that is common for both modes, and two longer ones (Fig. 1.4D). In this case, the latter are equal in length due to the symmetry of the system. By convention, such TSs are termed bispericyclic in the case of dimerisations, where the two cycloaddition modes are equivalent, and bipericyclic otherwise. The two adducts are connected by a Cope rearrangement TS as expected (*vide supra*).

Caramella’s discovery was of historic importance as it inspired many more investigations into ambimodal reactions.[32, 33, 34] It soon became apparent that these are not just a theoretical curiosity, but a phenomenon with relevance to real organic chemistry.[11] For example, it has been shown that bispericyclic TSs are involved in complex cycloaddition reactions employed in natural product (bio)synthesis.[35, 36, 37]

Ambimodal TSs are usually described by their asynchronicity and asymmetry. The asynchronicity refers to the difference in bond formation (quantified usually in terms of bond length and bond order) between the bond that is common for both cycloaddition modes active in the TS, and the other, secondary bonds. Asymmetry

refers to the difference between the secondary bonds. While bispericyclic TSs are by definition fully symmetric, they can be quite asynchronous – this is the case for Caramella’s cyclopentadiene dimerisation (*vide supra*). The term bipericyclic implies some degree of asymmetry in the TS. The asymmetry in bipericyclic TSs has been linked to the product ratio.[38]

The subject of investigation of the current study is another example of a synthetically useful DA cycloaddition that is thought to be ambimodal. In 2022, the Duarte and Anderson groups reported a collective synthesis of illudalane sesquiterpenes.[1] A key step in the synthesis was a cascade inverse electron demand (IED) [4 + 2] DA cycloaddition between thiophene *S,S*-dioxides and furans. This exploits the favourable and virtually irreversible SO₂ extrusion to afford in a single synthetic step a highly functionalised *de novo* benzene ring. Further transformations on the structure then gave the natural products of interest (Fig 1.5).

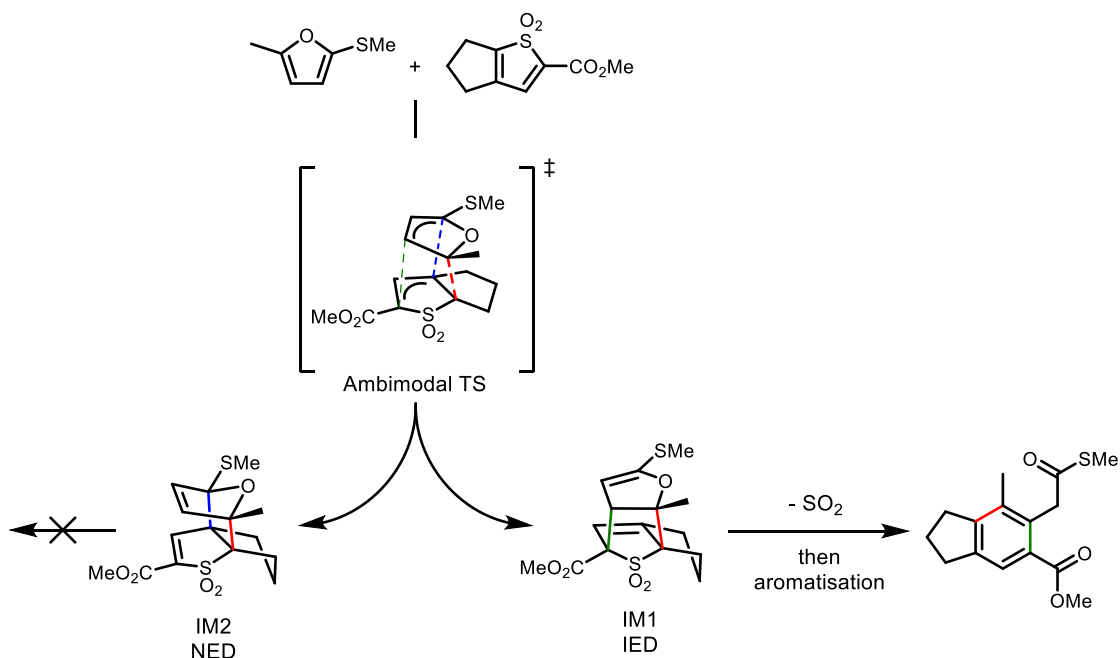


Figure 1.5: Proposed ambimodal mechanistic pathway for the reaction of interest. The two DA adducts IM1 (IED) and IM2 (NED) are shown; only IM1 can extrude SO₂ to give the observed product. Adapted from ref [1].

There is a multitude of possible mechanism for the transformation. These include:

- A concerted inverse electron demand (IED) [4 + 2] DA cycloaddition (in which the furan acts as a dienophile, and the thiophene dioxide as a diene) to give **IM1** directly;
- A concerted normal electron demand (NED) [2 + 4] addition (the furan acts as a diene, and the thiophene dioxide as a dienophile), followed by a Cope rearrangement in a separate step to give **IM1**, which is the adduct that can extrude SO₂;
- A stepwise Michael addition between the electron-rich furan and the electron-deficient thiophene dioxide, followed by an intramolecular second step to give the adduct.

The authors' computational study of the cycloaddition step suggested the activation barrier for the Michael addition is too high. However, a TS for neither of the two formally possible cycloadditions was found. Instead, a highly asynchronous and asymmetric bipericyclic TS structure was reported (Fig 1.6). The authors suggested the reaction proceeds via an ambimodal mechanism, wherein this single TS can lead to both the normal and inverse electron demand products.

Analysis of the computed PES suggested the NED IM2 adduct is formed preferentially from the ambimodal TS. Under the reaction conditions (100 °C, toluene) the only observed product forms *via* the IED intermediate after SO₂ extrusion and aromatisation. The authors postulated that IM2 must therefore be able to revert to the starting materials, presumably through the same ambimodal TS, so that IM1 can be formed. All in all, an unusual mechanistic pathway was proposed for the synthetically useful cycloaddition step, but no experimental evidence was gathered to support it.

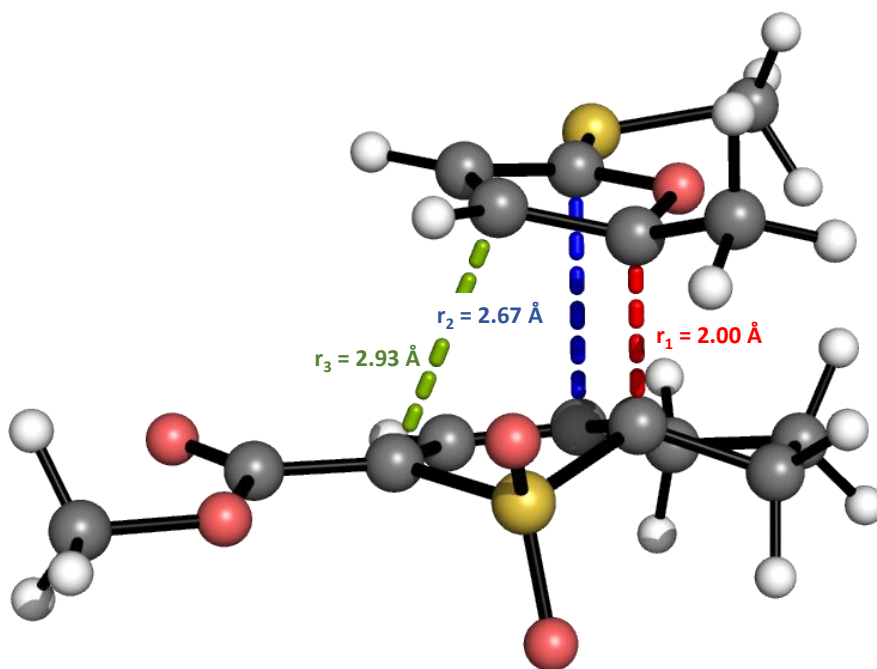


Figure 1.6: Geometry of bipericyclic TS reported by Duarte and Anderson with key bond lengths indicated. Adapted from ref [1].

1.5 Aims and Strategy

The main objective of the current study is to gain further insight into the mechanism of the reaction reported by Duarte, Anderson, and co-workers. Further evidence is needed to confirm the reaction proceeds *via* an ambimodal TS.

To this end, we intend to perform more refined computational studies into the geometries and relative energies of key species (e.g. ambimodal TS, IM1, IM2), the shape of the PES, and how these are connected to the observed reactivity.

To circumvent the limitations of static QM descriptions (*vide supra*), we intend to also employ dynamic methods, such as *ab initio* molecular dynamics (AIMD), to obtain a fuller description of the connectivity between species on the PES and the periselectivity of the ambimodal TS.

To probe the ambimodality hypothesis experimentally, we plan to perform natural abundance ^{13}C kinetic isotope effect (KIE) studies on the cycloaddition

step. These are expected to provide information about the degree of bond formation in the TS.

2

Theory and Methods

Contents

2.1	Static QM Methods	14
2.2	Dynamic Methods	21
2.3	¹³C NMR for KIE studies	23

2.1 Static QM Methods

Quantum mechanics (QM) is at the centre of many theoretical models for the rationalisation and prediction of chemical phenomena. At the heart of the theory lies the Schrödinger equation, which in its time-independent form reads:

$$\hat{H}\Psi = E\Psi, \quad (2.1)$$

where Ψ is the wavefunction, which contains all the information about the system, \hat{H} is the Hamiltonian operator, and E is the energy of the system. In atomic units, the electronic Hamiltonian takes the general form

$$\hat{H} = -\frac{1}{2} \sum_{i=1}^n \nabla_i^2 - \sum_{i=1}^n \sum_{I=1}^N \frac{Z_I}{r_{iI}} + \sum_{i<j}^n \frac{1}{r_{ij}}, \quad (2.2)$$

where n , N , and Z_I are the number of electrons, nuclei, and nuclear charges respectively. The first term describes the electronic kinetic energy, the second – the electron-nucleus attractions, and the third – the electron-electron repulsions. For systems with multiple electrons, this last term makes the Schrödinger equation analytically insoluble. This is a central issue in QM and has been approached in various ways, some of which are discussed below.

Static QM methods deal with single molecular geometries and try to approximate their energy. There are many approaches towards this end and some of these are outlined in this section.

2.1.1 Hartree-Fock Theory

Hartree-Fock (HF) theory is a conceptually simple yet useful theory for approximating the electronic energy of molecules. For a closed-shell system, the HF wavefunction is an antisymmetrised product of orthonormal one-electron spin-orbitals. This symmetry requirement stems from the Pauli exclusion principle, and is equivalently accounted for by the use of a so-called Slater determinant.

The main objective of the method is to estimate the HF energy. Decomposing the Hamiltonian in (2.2) into operators acting on one and on two electrons, and considering the Pauli antisymmetry requirement, we arrive at the HF energy:

$$E_{\text{HF}} = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \sum_i \langle \phi_i | \hat{h}_i | \phi_i \rangle + \frac{1}{2} \sum_{ij} \langle \phi_i \phi_j | r_{12}^{-1} | \phi_i \phi_j \rangle - \frac{1}{2} \sum_{ij} \langle \phi_i \phi_j | r_{12}^{-1} | \phi_j \phi_i \rangle. \quad (2.3)$$

The first term is the isolated energy contribution of each spin-orbital. The second is the energy of the Coulombic electron-electron repulsion, which is treated classically within the mean field approximation. The third term accounts for the purely QM phenomenon of electronic exchange. See Appendix A.1 for further details.

The expression for the energy can now be used to find a good approximation for

the HF wavefunction using the variational principle, which ensures the approximate energy is never an underestimate of the true ground state energy. This is done by iteratively solving a self-consistent set of equations until convergence is achieved. This is known as the self-consistent field (SCF) method.

HF theory is relatively simple and provides an important approximation for the ground state Hamiltonian and wavefunction. However, its description of the correlated motion of electrons is limited. While HF accounts correctly for Fermionic correlation, which has its origin in the Pauli exclusion principle (details in Appendix A.2), it treats Coulombic repulsion in the classical mean-field approximation. This fails to fully describe the effect that electron-electron repulsions have on the correlated motion of electrons, and on the probability distribution function.

Different approaches to accounting for correlation have given rise to a plethora of more refined electronic structure theories and methods, many of which use HF as a starting point. Some of these are outlined in the remainder of the section.

2.1.2 Accounting for Correlation: Configuration Interaction and Møller-Plesset Perturbation

The correlated motion of electrons arising from Coulomb repulsion beyond the mean-field approximation gives rise to some correlation energy, which is defined as the difference between the HF energy and true ground state energy:

$$E_{\text{corr}} = E_{\Psi} - E_{\text{HF}}. \quad (2.4)$$

One approach to account for correlation more fully is to describe wavefunctions as linear combinations of Slater determinants, as opposed to a single one as is done in HF theory. This approach is termed Configuration Interaction (CI) theory. In CI, electron-electron repulsion is described by mixing in some excited state

configurations, which allows for more radially diffuse (e.g. $1s^1 2s^1$) and/or angularly varying (e.g. $1s^1 2p_x^1$) densities.

A fundamentally different approach for treating electronic correlation is to apply perturbation theory, wherein a reference (unperturbed) Hamiltonian \hat{H}_0 is selected such that its eigenstates are known and it is a good approximation to the true Hamiltonian \hat{H} . The perturbation is defined as $\hat{H}_1 = \hat{H} - \hat{H}_0$. This then allows for expressions for the n^{th} order perturbations on the wavefunction Ψ_n and the energy E_n .

The application of perturbation theory to the electronic structure problem wherein the Fock operator is taken as the unperturbed Hamiltonian \hat{H}_0 , the HF wavefunction is Ψ_0 , and the electron-electron repulsion is treated as the perturbation, is called Møller-Plesset perturbation method.[39]

The theory is usually applied to the second order of perturbation, and is termed MP2 in such cases. Going beyond second order does not necessarily increase accuracy, and may lead to convergence issues. MP2 is an efficient method for the treatment of electron correlation, and gives comparable accuracy to DFT. It is a pair theory, i.e. the contribution to the correlation of each electron pair is taken into account separately, which leads to an unfavourable scaling with system size.

2.1.3 Density Functional Theory

So far we have considered methods involving searching for the ground state wavefunction to obtain the energy. Density Functional Theory (DFT) is a fundamentally different approach in that the search for the ground state energy is performed in terms of the electron density without knowledge of the wavefunction.[40]

The treatment of the system with regards only to the electron density is made possible by the two Hohenberg-Kohn theorems.[41] The first one states that the

potential $v(\mathbf{r})$ can be inferred up to a constant solely from the ground state density ρ . This implies that both the ground state wavefunction and the ground state energy can be expressed as functionals of the density, i.e. $\Psi \equiv \Psi[\rho]$ and $E \equiv E[\rho]$. The second theorem states that there exists an energy functional that maps any density to a single energy.

For the purposes of DFT, the Hamiltonian is usually broken down into a universal n -electron part \hat{F} , containing the electronic kinetic energy \hat{T} , and the electron-electron repulsions \hat{V}_{ee} :

$$\hat{F} = \hat{T} + \hat{V}_{ee} = \sum_i^n -\frac{1}{2}\nabla_i^2 + \sum_{i<j}^n \frac{1}{r_{ij}}, \quad (2.5)$$

and a system-dependent potential $v(\mathbf{r})$:

$$\hat{V} = v(\mathbf{r}) = -\sum_{iI}^{nN} \frac{Z_I}{r_{iI}}, \quad (2.6)$$

for a system with n electrons and N nuclei with charges Z_I . The universal part of the Hamiltonian is a functional of the density, $\hat{F} \equiv F[\rho]$. The ground state energy becomes

$$E_0 = \min_{\Psi} \langle \Psi | \hat{F} + \hat{V} | \Psi \rangle = \min_{\rho} \left\{ F[\rho] + \int d\mathbf{r} v(\mathbf{r}) \rho(\mathbf{r}) \right\}. \quad (2.7)$$

The main objective of DFT is to find $F[\rho]$, as this allows variational searches for the ground state energy E_0 using only the density ρ . However, this proves hard as the kinetic contribution to $F[\rho]$ is large and very hard to approximate.

Kohn-Sham theory provides an effective solution to this problem. Within it, a fictitious system of n non-interacting electrons is introduced. The universal functional is then decomposed into

$$F[\rho] = T_s[\rho] + E_H[\rho] + E_{xc}[\rho]. \quad (2.8)$$

$T_s[\rho]$ is the non-interacting kinetic energy functional, which extracts a large fraction of the kinetic energy from the system by acting on the system of non-interacting

electrons introduced above. $E_H[\rho]$ is the Hartree energy functional, which accounts for the classical electrostatic repulsion for the charge distribution of the system; it can be computed exactly, but contains a non-physical self-interaction (an electron repelling itself). $E_{xc}[\rho]$ is the exchange-correlation functional, which contains the non-classical contributions to exchange and correlation, and a correction for the self-interaction.

In principle, DFT is exact, provided the universal functional $F[\rho]$ is known. However, the exchange correlation functional $E_{xc}[\rho]$ cannot be computed exactly and has to be approximated. Thus, the practical simplifications provided by Kohn-Sham DFT lead to the breakdown of the exact nature of the theory.

Kohn-Sham DFT is made practically useful by the availability of good density functional approximations which allow the cheap evaluation of functionals and potentials. Different approaches to approximating $E_{xc}[\rho]$ have given rise to a plethora of DFT methods, colloquially known as the functional zoo. These functionals can be grouped into separate levels depending on their level of sophistication, and are often depicted as rungs on the so-called Jacob’s ladder (Table 2.1).[42] Generally speaking, both computational expense and chemical accuracy increase going up the ladder, but it is not always true that a more sophisticated functional gives more accurate results.[43] For example, hybrid functionals have been shown to be less accurate than methods lower on the ladder for metallic systems.[44] See Appendix A.3 for further details.

The choice of functional depends on the nature of the system (e.g. types of atoms and interactions studied), and is often a balance between accuracy and computational cost. Determining an appropriate level of theory often requires a benchmarking study. In this work, we considered the M06-2X and ω B97X hybrid functionals.

As density functional approximations do not account for London dispersion

Functional	Terms	Examples
↑ Double-Hybrid GGA	$\rho, \nabla\rho, E_x^{HF}, E_c^{MP2}$	B2-PLYP
Hybrid GGA*	$\rho, \nabla\rho, E_x^{HF}, (E_x^{lr, HF})^*$	B3LYP, M06(-2X), ω B97X*
Meta GGA	$\rho, \nabla\rho, \nabla^2\rho, \tau$	TPSS, SCAN
GGA	$\rho, \nabla\rho$	LYP
LDA	ρ	PW92

Table 2.1: Jacob’s ladder of DFT functionals, with a summary of the relevant mathematical terms, and some notable examples. The arrow on the left signifies the general increase in chemical accuracy and computational expense going up the ladder. See Appendix A.3 for further details.

*Including range separated hybrids.

interactions (long-range correlation caused by coupled instantaneous dipoles in distant parts of the molecule),[45] semi-empirical corrections to the dispersion are often added to the DFT result. In this work, we tested the common examples D3-Zero and D3(BJ).[46, 47]

2.1.4 Basis Sets

The numerical evaluation of any property of interest necessitates that molecular orbitals can be described with an appropriate mathematical function. For molecular systems, there are two common choices: Slater-type orbitals (STOs) and Gaussian-type orbitals (GTOs), which take the general forms

$$\phi_{\text{STO}} = N r^{n-1} e^{-\zeta r} Y_{l,m}(\theta, \phi), \quad (2.9)$$

and

$$\phi_{\text{GTO}} = N x^i y^j z^k e^{-\alpha r^2}, \quad (2.10)$$

respectively.

While STOs represent hydrogenic wavefunctions better, especially at and around the cusp ($r \rightarrow 0$), GTOs have an analytical form which makes them easier to

process computationally. The compromise between the accuracy of STOs and the convenience of GTOs is realised by describing orbitals as linear combinations of GTOs.[48] The choice of basis set is an important balance between computational cost and accuracy, and usually requires benchmarking. In this work, we considered the def2-SVP, def2-TZVP, and def2-TZVPP split valence basis sets.[49]

2.1.5 Solvation Models

Most chemical processes occur in solution, therefore making it crucial to account for solvation effects during modelling. The solvation models used in computational chemistry can broadly be divided into explicit and implicit. In the former, the solvent molecules are included explicitly, along with the solute-solvent interactions. As multiple solvent molecules are necessary for an accurate description, this increases the computational cost by enlarging the system size and the number of degrees of freedom significantly.

Implicit solvation models are a simplification as they treat the whole solvent as a uniform medium described by certain parameters.[50] A commonly utilised method is the Conducting Polarizable Continuum Model (CPCM), in which the solute is described as a cavity using its van der Waals surface, and the solvent as a polarisable continuum interacting with the cavity.[51, 52, 53, 54] The solvent and cavity polarise each other, and this interaction is computed iteratively until convergence. This is a relatively cheap model, but inaccurate for systems where the explicit solute-solvent interactions matter, e.g. for polar protic solvents.[55, 56]

2.2 Dynamic Methods

Static QM computations rely on a series of assumptions that are an incomplete description of chemical reality. Firstly, the interest is focused on local minima

and first order saddles, despite the fact that at chemically relevant temperatures molecules are constantly moving, and properties are determined by ensembles of microstates rather than by isolated structures. Secondly, vibrations are treated as harmonic, which is a bad approximation when the bond is close to dissociation. This presents an issue for computing entropic contributions. Further to that, hindered rotations are treated as low-frequency vibrations, which can also be a source of error.

Molecular dynamics (MD) can partially overcome these challenges by giving insight into the motion of molecules in time resulting from the forces acting on each atom. The evolution of atomic positions is treated within Hamiltonian mechanics, and necessitates the computation of the forces acting on the atoms in the system. In classical MD, these are computed with the help of parametrised forcefields which treat bonds as harmonic oscillators. In *ab initio* MD (AIMD), forces are computed at the QM level. This approach is much more expensive, and has only become viable for large systems in the recent decades.

One major advantage AIMD has over classical MD is that it accounts for the anharmonicity of chemical bonds. This allows for the explicit and time-dependent description of bond-breaking and forming events, which in turn enables the computation of properties such as the asynchronicity of bond formations.

Once the forces are known, momenta can be easily obtained from Newton's second law, and the displacement can be computed by integrating over time. In the general case, however, analytical integration is not possible. MD provides a solution to this problem by considering the change that occurs after a finite time-step Δt , which allows for the numerical approximation of the displacement integral. Generally, a smaller Δt is expected to give more accurate results. AIMD is implemented in an iterative manner, wherein for each geometry along the trajectory, the forces are computed at the chosen level of theory; the forces are then used

to compute the atomic displacements; finally, applying the displacements gives the new geometry, and the loop is closed.

To obtain statistically and chemically meaningful results, multiple AIMD runs are performed with a stochastic distribution of initial conditions. This is done during a process termed sampling of the ensemble, where a distinct microstate is chosen at the beginning of each AIMD trajectory. In a canonical ensemble, where the number of particles N , the volume V , and the temperature T are kept constant, a thermostat is used to adjust the velocity of atoms and allow for approximately constant T . An example is the Nosé-Hoover chain (NHC), which was used in this work.[57, 58, 59, 60]

2.3 ^{13}C NMR for KIE studies

Kinetic experiments, and in particular Kinetic Isotope Effect (KIE) studies are frequently involved in investigating the mechanisms of reactions as they indirectly provide information for experimentally unobservable species such as TSs, and can provide a measure for activation barriers.[61, 62, 63, 64, 65] The Kinetic Isotope Effect (KIE) is a physical chemical phenomenon wherein the reaction rates differ for different isotopes of the same element. The KIE can be defined as the ratio of the rate constants associated with either isotopic substitution:

$$\text{KIE} = \frac{k_1}{k_2} \quad (2.11)$$

The investigation of a C-C bond formation/breaking often necessitates the measurement of $^{13}\text{C}/^{12}\text{C}$ KIE. A commonly used procedure for measuring multiple C atom KIEs at natural abundance was described by Singleton.[66] The method relies on the fractional enrichment in the faster-reacting isotope of the products, and in the slower-reacting isotope of the starting material during the course of the reaction. The analysis is performed either on the recovered starting material at

high conversion, or on the obtained product at low conversion. The enrichment in a certain isotope is then determined by quantitative ^{13}C NMR. The KIE is given by

$$\text{KIE} = \frac{\ln(1 - F)}{\ln[1 - F(R/R_0)]} \quad (2.12)$$

for analysis on product obtained at low conversion. F is the conversion, R is the relative peak integral for the sample, and R_0 is the relative peak integral from the natural abundance standard.

Two problems can be identified in this approach. Firstly, the mass ratio between the two isotopes is relatively small, and the natural abundance of the NMR active ^{13}C is around 1%, which makes the KIE inherently small and harder to measure accurately. To account for that, multiple experiments are usually performed, and multiple spectra obtained from each sample, so that a statistically significant averaging can be achieved. The second issue is that it is always the minor component of the reaction mixture that is examined - this implies the reaction needs to be scalable enough so that enough material can be retrieved for analysis.

Despite these issues, the method is advantageous in that it operates at natural abundance, i.e. it does not require synthetic isotopic enrichment, it is principally general for any element, and can be implemented using routine NMR instruments. Many organic reactions, including a variety of cycloadditions, have been studied by the Singleton method.[67, 68, 69]

3

Computational Investigation of the Ambimodal Cycloaddition

Contents

3.1	Choosing the Level of Theory	25
3.2	Energy Profile of the System of Interest	28
3.3	<i>Ab Initio</i> Molecular Dynamics of the Ambimodal Cycloaddition	30
3.4	KIE Predictions	35
3.5	Computational Details	38

3.1 Choosing the Level of Theory

We began our investigation by attempting to reproduce the geometries reported by Duarte and Anderson for key species (**IM1**, **IM2**, ambimodal TS) in the bipericyclic cycloaddition of interest at the M06-2X/def2-TZVPP level of theory (Figs 3.1 and 3.2). We used the same level of theory, but a more recent version of the software ORCA, version 5.0 as opposed to 4.2 employed the original study. We also used a finer integration grid as functionals from the M06 suite are known

to be sensitive to that.[70]

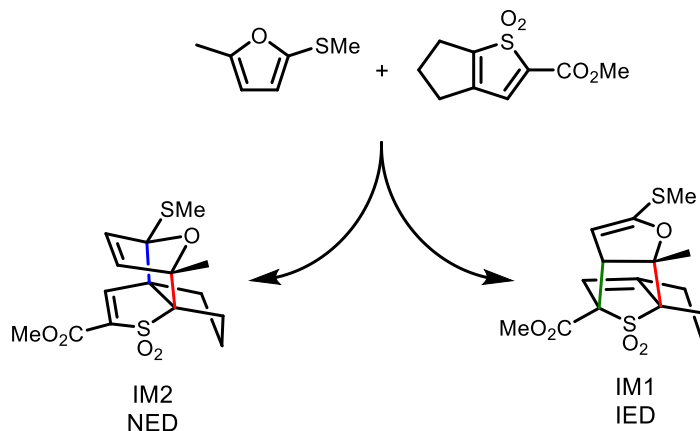


Figure 3.1: The ambimodal cycloaddition step of interest, giving both the normal electron demand adduct **IM2**, and the inverse electron demand **IM1**, from the same bipericyclic TS.

Our attempt to reproduce the ambimodal TS failed - we instead obtained a TS structure that corresponded to the step following a potential Michael addition. Although simply attempting to re-optimize a TS structure is admittedly not a sufficiently rigorous technique, this discrepancy led us to doubt the reliability of the M06-2X functional for geometry optimisation.

To go beyond DFT, we chose to re-optimize the obtained geometries using spin-component-scaled (SCS) MP2, which has been shown to describe the asynchronicity and asymmetry in pericyclic reactions better than pure MP2.[39, 71, 72] This time, we were able to reproduce the ambimodal TS within a small deviation in the geometry relative to the reported one (differences in key bond lengths within 0.1 Å, Table 3.1).

The computational expense of MP2 clashed with our intention to run a statistically significant number of AIMD simulations (in the order of hundreds, section 3.3). This motivated us to search for a DFT functional more appropriate than M06-2X. We examined the range-separated correlation hybrid functionals ω B97 and ω B97X[73], and considered various empirical dispersion corrections. Among

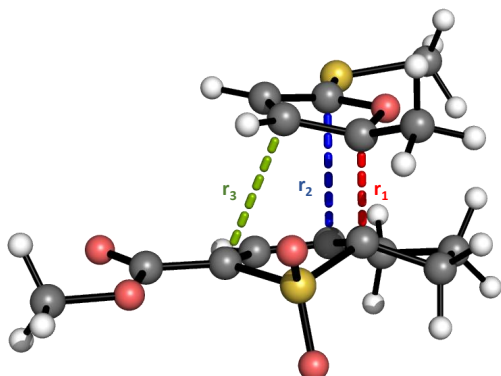


Figure 3.2: Ambimodal TS structure with key bond distances labeled.

	Duarte & Anderson M06-2X	This work	
		SCS- RI-MP2	ω B97X- D3(BJ)
r_1	2.00	1.95	2.02
r_2	2.67	2.67	2.75
r_3	2.93	2.84	2.87
$r_3 - r_2$	0.16	0.17	0.12

Table 3.1: Bond distances in Å for the ambimodal TS structure (Fig. 3.2 on the left) at different levels of theory. Computations performed with the def2-TZVPP basis set and CPCM(toluene) correction. The quantity $r_3 - r_2$ is given as a measure of asynchronicity.

those, and based on our ability to obtain appropriate true minima and first order saddles, we chose ω B97X with a D3 empirical dispersion under Becke-Johnson damping: ω B97X-D3(BJ). [46, 47]

The ambimodal TS structure at this level of theory is characterised by a similar asynchronicity ($r_1 \ll r_2, r_3$), but a reduced asymmetry when compared to the structure reported at M06-2X ($r_3 - r_2$ goes from 0.16 Å at M06-2X to 0.12 Å at ω B97X-D3(BJ), Table 3.1). This change in asymmetry hints towards the method-dependence of the TS structure, which has been shown to be significant for asynchronous DA reactions.[72]

In addition to **TS1**, we also explored TS structures from Nudged Elastic Band (NEB) computations connecting either intermediate to the reactant complex: **TS1'** from **IM1** and **TS1''** from **IM2**. The three TSs are almost identical in structure, with heavy-atom root-mean-square-deviations (RMSD) not exceeding 0.02 Å.

The findings reported in this section led us to believe that the ω B97X-D3(BJ) DFT functional provided a reasonable balance between accuracy and computational

cost for the study of this system.

3.2 Energy Profile of the System of Interest

Gibbs free energy profiles for the potential reaction pathways were computed at the SCS-RI-MP2/def2-TZVPP// ω B97X-D3(BJ)/def2-TZVPP level of theory with a CPCM(toluene) solvation correction. These pathways include the Michael addition through **MATS** giving **MA**, the bipericyclic addition **TS1** giving adducts **IM1** and **IM2**, the Cope rearrangement **TS2** that connects the latter two, and the SO_2 extrusion giving the product **P** (Fig. 3.3, computational details in section 3.5). All structures were confirmed as either true minima or first-order saddles accordingly.

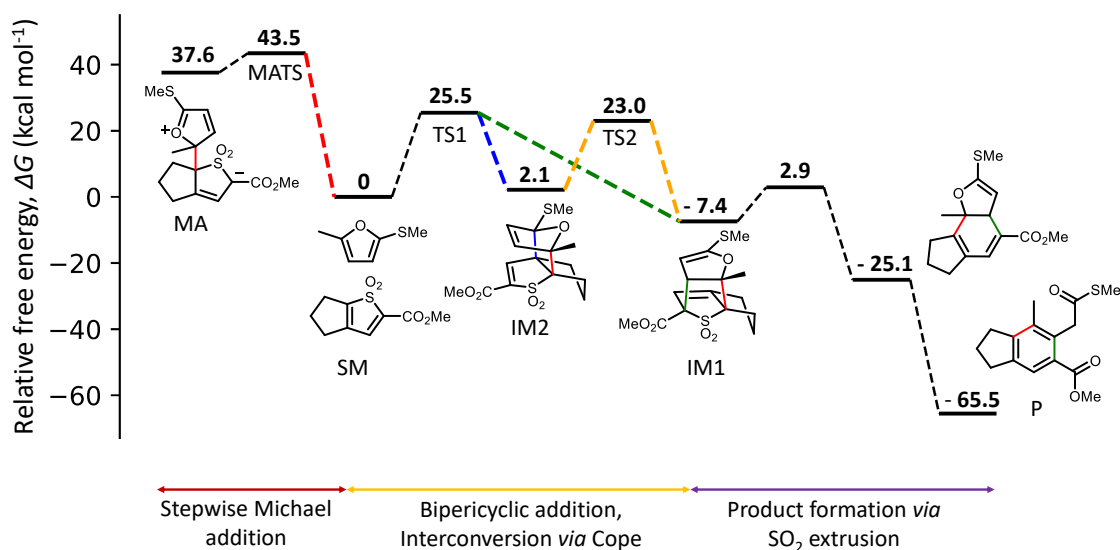


Figure 3.3: Reaction energy profile for the system of interest at CPCM(Toluene)-SCS-RI-MP2/de2-TZVPP//CPCM(Toluene)- ω B97X-D3(BJ)/def2-TZVPP.

IRC computations starting from the TSs for the Michael addition (**MATS**) and the Cope rearrangement (**TS2**) confirmed the proposed connectivity. As an IRC cannot give both products for a bifurcating system, the ambimodal nature of **TS1** can be neither confirmed nor denied by this otherwise robust method. The IRC initiated at **TS1** gave the reactants in one direction, and **IM2** in the

other. We concluded that since IRC operates in a steepest descend algorithm, the branch of the MEP that forms **IM2** after the bifurcation is steeper than the one forming **IM1** (further discussion in Section 3.3).

For the most part, the obtained profile qualitatively reproduces the geometries and relative energies reported by Duarte and Anderson. The discrepancies in the key activation barriers for the Michael addition and the bipericyclic pathway are within 2.5 kcal/mol, which can be rationalised by the use of a different level of theory. Importantly, our results are in agreement with the original study in that the stepwise Michael addition is unlikely to be operating under the reaction conditions given the 18 kcal mol⁻¹ difference between **TS1** and **TS2** (section 1.4).

One notable difference is the significantly lower barrier for the Cope rearrangement that we obtained. In the original study, the Cope TS was higher in energy than the ambimodal one by around 7 kcal mol⁻¹. In the current work the order is reversed, and **TS2** is more stable than **TS1** by 2.5 kcal mol⁻¹. Moreover, a substantial difference in geometries is observed: the RMSD is 1.65 Å. At the ω B97X-D3(BJ)/def2-TZVPP level of theory and a temperature of 373 K, the main mode of interconversion of **IM2** to **IM1** appears to be the Cope rearrangement through **TS2**, as opposed to the retroaddition through **TS1** as suggested previously. Although this finding poses an interesting question on the method-sensitivity of studying ambimodal cycloadditions, it is of secondary importance as any attainable mode of conversion of **IM2** to **IM1** is sufficient for the rationalisation of experimental observations.

The reaction profile obtained here suggests that once formed, **IM1** readily and irreversibly extrudes SO₂ with a relatively small activation barrier of $\Delta G^\ddagger = 10.3$ kcal mol⁻¹, in line with experimental observations for this system and the general knowledge on SO₂ extrusions in cascade cycloadditions.[74, 75, 76] The

extrusion is also highly exergonic, $\Delta_r G = -19.6 \text{ kcal mol}^{-1}$.

The rearrangement that follows extrusion is not modelled here but is postulated to not have a significant barrier as it is an aromatisation *via* proton transfer. The relatively stable aromatic final product **P** acts as a thermodynamic sink for the reaction, and is the only compound observed in meaningful quantities experimentally.

All in all, the energy profile constructed is in agreement with the previous computations, except for the more stable Cope **TS2** identified in this work, and can be used to explain experimental observations.

3.3 *Ab Initio* Molecular Dynamics of the Ambimodal Cycloaddition

As we could not assert the ambimodal nature of **TS1** using only static QM methods such as IRC, we set out to perform downhill AIMD simulations initiated at the ambimodal TS to prove its connection to the two DA adducts **IM1** and **IM2** on the PES of the system.

3.3.1 Evaluation of the Level of Theory

We initially set out to run AIMD simulations at the more expensive and more accurate CPCM(Toluene)-SCS-RI-MP2 level. To reduce the computational cost of the simulations, we used a smaller basis set: def2-SVP with a def2-TZVP expansion on all heteroatoms and the six carbons active in **TS1**.

At this level of theory, we performed 100 AIMD simulations, each spanning 100 fs in 200 steps. Due to randomisation issues wherein some trajectories were initialised with the same sampling random seed (details in section 3.5), our 100 runs gave only around 40 distinct trajectories. Out of these, only one appeared to give the Michael adduct, and all the rest resulted in the experimentally unobserved

NED **IM2**. The strong bias for **IM2** can be explained by the large asymmetry in the TS1 obtained at the SCS-RI-MP2 level ($r_3-r_2 = 0.17$ Å, Table 3.1).

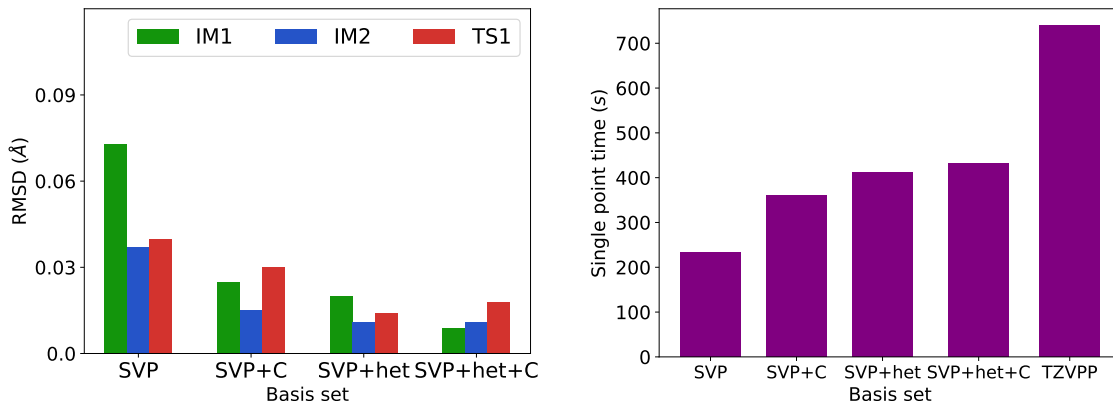


Figure 3.4: Basis set benchmark study for AIMD at ω B97X-D3(BJ) using the geometries of **IM1**, **IM2**, and **TS1**. Note SVP means def2-SVP, +C stands for a def2-TZVP expansion on the 6 carbons active in the ambimodal **TS1**, +het stands for a def2-TZVP expansion on all heteroatoms, and +het+C is a combination of the latter two. Left: heavy-atom RMSD of geometries obtained using the various basis sets relative to the ones obtained with def2-TZVPP for **IM1**, **IM2**, and **TS1**. Right: computational expense measured as time required for a single point computation on the **TS1** structure using 4 cores.

These unsatisfactory results, together with the computational expense of the MP2 method for this rather large system, prompted us to change our strategy. We turned to the level we had used for the construction of the energy profile: ω B97X-D3(BJ). To choose an appropriate basis set, we performed a small benchmark study by re-optimising the structures of **IM1**, **IM2**, and **TS1** using def2-SVP and a def2-TZVP expansion on important atoms in the system (the six carbons active in the TS, and heteroatoms, Fig. 3.4), and considering how these affect the geometries. We concluded that the gain in accuracy from the expansion of the basis set was not sufficient to warrant the increase in computational time. Using def2-SVP for all atoms presented a good balance between accuracy and expense and allowed us to run a greater number of simulations (a total of 300, *vide infra*).

3.3.2 Product Ratios and PES Connectivity

Having established an acceptable level of theory, we set out to perform a statistically significant number of AIMD simulations. Initiating at each of the three TS structures obtained previously (**TS1**, **TS1'**, **TS1''**, section 3.1), we performed 100 explicitly randomised downhill AIMD runs, each spanning 175 fs in 350 steps, mimicking the reaction conditions (373 K, toluene solvent, details in section 3.5).

The distributions of species (final geometries) obtained from the total of 300 geometries are presented in Fig. 3.5. All runs resulted in the formation of either intermediate (**IM1** or **IM2**), the reactant complex, or very rarely the Michael adduct (**MA**).

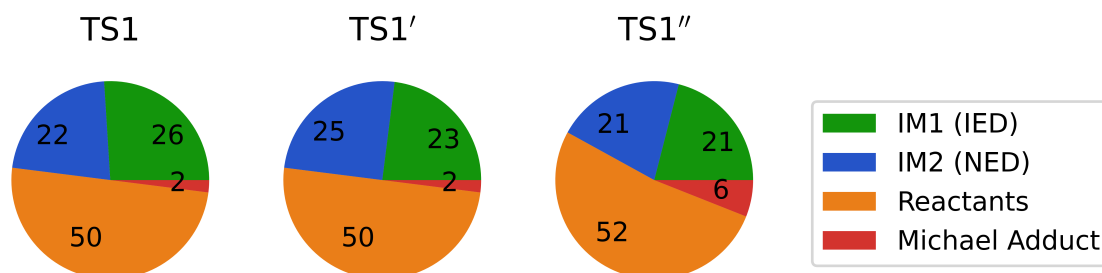


Figure 3.5: Product distributions from MD runs at ω B97X-D3(BJ)/def2-SVP initiated at each of the three ambimodal TS structures (see text in sections 3.1 and 3.3). Note the final geometries of all 10 **MA** trajectories gave **IM1** or **IM2** upon static optimisation.

The very similar product distributions for **TS1**, **TS1'**, and **TS1''**, alongside with the negligible RMSDs mentioned above, lead us to postulate that the three TS structures either correspond to the same chemical species (and the differences are due to optimisation convergence criteria), or to three species which occupy the same small region of a relatively flat part of the system PES (see discussion on multiple TSs on the caldera in section 1.3 and refs [16, 17]).

As the Michael addition was found to proceed through a high activation barrier (section 3.2), we decided to further investigate the total of 10 (out of 300) trajectories

that result in the Michael adduct **MA**. Static optimisations of the final geometry at CPCM(Toluene)- ω B97X-D3(BJ)/def2-TZVPP all resulted in **IM1**, which lead us to believe these trajectories were in fact en route to form **IM1**, but the formation of the second bond was pronouncedly retarded and was not achieved in the 175 fs window.

Our results suggest that the structures we considered as corresponding to a bipericyclic TS are indeed each connected to both the intermediates **IM1** (IED) and **IM2** (NED) and the reactants on the PES at this level of theory. We interpret this finding as evidence in favour of the proposed ambimodal mechanistic pathway.

The obtained product ratio between **IM1** and **IM2** of nearly 1:1 (Fig. 3.5) is a significant departure from the ValleyRidge[77, 78] analysis performed in the original study, which predicted a 99% selectivity for **IM2** at the M06-2X level. It is also not in accordance with the significant asymmetry in the ambimodal **TS1**. Houk and co-workers established a linear relationship between TS geometries and MD product distributions[38], which would predict a ratio of approximately 1:3.4 based on the bond difference of $r_3 - r_2 = 0.12 \text{ \AA}$ found in **TS1** at ω B97X-D3(BJ).

We should note that the product formed in each individual MD trajectory will depend on the momenta assigned to the atoms during the initial sampling of the NVT canonical ensemble (sections 2.2 and 3.5). In this work, we initiated AIMD trajectories from the ambimodal **TS1** geometry. This means any dynamic information about the path taken to arrive at the TS structure is lost (see discussion in section 1.3). This may have biased our results in an unrealistic way, e.g. increased the proportion of the MD trajectories that result in the reactant complex (roughly half in our results, Fig. 3.5). Even if the local shape of the PES around the ambimodal TS is such that it favours formation of the reactants, in reality the trajectories that started from the reactant cluster and reached the TS are more likely to be associative rather than dissociative.

At any rate, the main objective of our AIMD simulations was to assert that **TS1** is connected to both **IM1** and **IM2** on the PES, and it was achieved. Furthermore, any comments on the selectivity of the reaction currently remain entirely in the realm of theoretical discussion as **IM2** has not been experimentally observed, and the practical determination of the product ratio is not possible given the interconvertibility of **IM2** and **IM1**, and the irreversible SO₂ extrusion.

3.3.3 Bond Formation Dynamics: Asynchronicity and Asymmetry

Finally, we examined the bond formation dynamics by considering the change of interatomic distances (averaged over runs) with time for the runs terminating in **IM1** and **IM2** respectively (Fig. 3.6). Both the asynchronicity and asymmetry in the TS structure are reflected in the dynamics. In either case, the formation of the bond common to both pericyclic modes (r_1 , which is also the shortest in **TS1**, Table 3.1) is much quicker than of the other (r_2 and r_3 respectively).

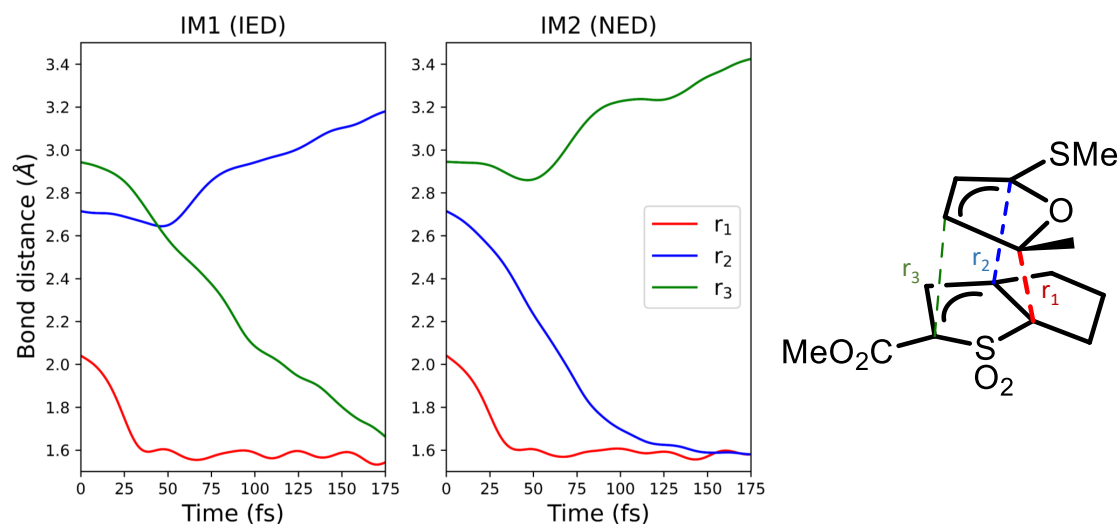


Figure 3.6: Bond formation as a function of time from the ω B97X-D3BJ/def2-SVP AIMD simulations that resulted in a DA adduct. Labelling and colour-coding of bonds shown on **TS1** structure on the right.

A comparison between the formations of the second bonds in the two adducts makes apparent that the asynchronicity is much more pronounced for the formation of **IM1** (IED) than for **IM2** (NED), with the second bond being formed around 60 fs later on average. This is consistent with our assertion that the branch of the bifurcating MEP that gives **IM2** is steeper (see IRC discussion in section 3.2). It also reflects the asymmetry of the TS structure as expressed in the bond difference $r_3 - r_2 = 0.12 \text{ \AA}$ (Table 3.1).

The results reported so far in this section demonstrate the inherent complexity of ambimodal systems. While the structure of **TS1** would suggest that **IM1** is formed preferentially, our results show an almost even distribution between the two adducts (Fig. 3.5). The asynchronicity in the **TS1** geometry is however reflected in the bond formation dynamics (Fig. 3.6). As stated above, the product distribution obtained from AIMD depends on the initial sampling, while the rates of bond formation depend on the shape of the PES and the steepness of the MEP branches. A further point of interest is that the less steep branch leads to the more thermodynamically stable **IM1** (Fig. 3.2 for relative energies).

Overall, our AIMD simulations have provided a more complete and holistic description that further supports the ambimodal nature of the system under study.

3.4 KIE Predictions

To explore the bipericyclic nature of the reaction further, we turned our attention to studying the kinetic isotope effect (KIE) on key atoms in the system both computationally and experimentally (Chapter 4).

The energetic profile obtained (section 3.2 and Fig. 3.3) suggests that **TS1** is the rate-determining step. However, the Cope rearrangement **TS2** still plays an important role in the reaction system, interconverting **IM2** to **IM1**. We concluded

that both TSs are likely to influence the isotopic enrichment in the product, and set out to compute the theoretical KIE values associated with each of them. These predictions were made by simply computing the activation barrier ΔG^\ddagger at different isotopic substitutions (details in section 3.5) at the CPCM(toluene)- ω B97X-D3(BJ) level. The results are summarised in Table 3.2. Note that the KIE for the Cope rearrangement was computed for the direction **IM2** \rightarrow **IM1**.

Atom	Ambimodal cycloaddition KIE	Cope rearrangement KIE
C1	1.028	1.004
C2	1.008	1.018
C4	1.006	1.013
C5	1.004	1.005
C6	1.041	1.008
C13	1.001	1.024

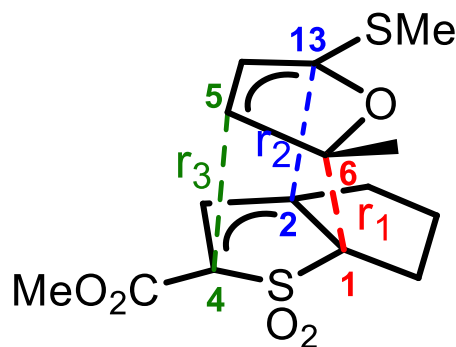


Table 3.2: Summary of computed KIEs for key carbon atoms in the ambimodal cycloaddition and Cope rearrangement TSs at CPCM(toluene)- ω B97X-D3(BJ). Fig. 3.7 for atom labels.

Figure 3.7: Schematic TS structure with relevant atoms and bonds labelled.

The KIEs computed for the ambimodal **TS1** reflect the structure of the TS and the imaginary frequency vibrational mode corresponding to it. The largest KIEs are on the atoms forming the shortest bond r_1 between C1 and C6 (Fig. 3.7).

In the Cope rearrangement **TS2**, the largest KIE values are obtained for C13 and C2 forming r_2 . C4 and C5 forming r_3 are also exhibiting KIEs, but these are broadly comparable with the ones on C1 and C6 albeit r_1 is not formally forming or breaking in the rearrangement. This is consistent with the observation that r_2 is the most active in the imaginary vibrational mode of **TS2**. Further to that, the IRC computation initiated at **TS2** gave **IM2** after fewer steps than for **IM1**. From

this it can be concluded that the path for formation of **IM2** is steeper than for **IM1** when starting from either **TS1** or **TS2** (section 3.3.3). We have therefore indirectly obtained information about the shape of the PES. Although **IM1** is the more thermodynamically stable adduct, **IM2** lies on the steeper side of the ridge on the PES of the bifurcating system.

The predictions of the KIEs reported in this section were obtained subject to a number of approximations and simplifications. Firstly, the simple computational protocol employed relies on activation barriers as described in transition state theory (TST); however, TST is not fully reliable for systems with bifurcating MEPs as already discussed in Chapter 1. Secondly, we have assumed KIE to be a purely QM phenomenon, driven by changes in ZPE and shapes of potential curves upon isotopic substitution. However, another type of KIE, which has its origin in Newton’s second law of motion and the difference in mass between isotopes, has been proposed as relevant for bifurcating reactions.[79]

Finally, it ought to be considered that the reaction proceeds *via* a complex mechanism, wherein the bipericyclic addition **TS1** is reversible, and the Cope rearrangement **TS2** plays a tentative role. These make it hard to predict the effective KIE that would be experimentally observed from the computations.

To conclude, we have obtained computational predictions for the KIEs at key atoms in the system of interest. These are in general agreement with the obtained TS structures and IRC computations. Although obtained using a number of approximations, they provide a qualitative reference frame for the experimental results reported in Chapter 4.

3.5 Computational Details

All QM computations were performed using the ORCA 5.0.3 suite of programs.[80, 81] The initial attempts to reproduce the reported geometries were performed using the M06-2X functional [82] and a def2-TZVPP basis set[49] and applying a CPCM solvation model (toluene).[51, 52, 53, 54]

The reaction profile was constructed from geometries obtained using the ω B97X functional with D3 empirical dispersion with Becke-Johnson damping[46, 47] and the def2-TZVPP basis set. Single point corrections were obtained at the SCS-RI-MP2/def2-TZVPP level with a def2-TZVPP/C auxiliary basis set.[39, 71, 83, 84] Free energy corrections were applied *via* the Grimme method[85] (vibrations computed at the CPCM(Toluene)- ω B97X-D3BJ/def2-TZVPP level of theory, 373 K, 1M standard state) within the Otherm program.[86] Nudged elastic band (NEB) computations were performed as implemented in ORCA.[87, 88, 89] KIE were computed by generating vibrations for isotopically substituted molecules from hessian matrices using the `orca_vib` miniprogram interfaced by a modified version of Otherm.

AIMD simulations were performed in the Orca MD package by Brehm.[90] Trajectories were initiated from the optimised TS geometries. 350 steps of 0.5 fs were requested. An NHC thermostat was used to sample the NVT ensemble under a Maxwell-Boltzmann distribution at 373 K. This process requires a random seed, which is taken as the system time unless specified. The MP2 simulations were not explicitly randomised and so those of them initiated at the same time gave identical trajectories. The ω B97X-D3(BJ) simulations were explicitly randomised using the "Randomize X" flag in the ORCA MD module, where X was replaced by the sequential number of the run.

4

Experimental KIE Results

Contents

4.1	Overview of Reactions	40
4.2	KIE Measurements	40
4.3	Experimental Details	45

The aim of the experimental part of this study was to gain insight into the bond formation processes taking place during the ambimodal cycloaddition. The chosen strategy was to measure the KIE at various carbon atoms using Singleton's method (section 2.3). This would provide information on the degree of bond formation and infer the nature of the TS(s) the system goes through to yield the product. The ultimate aim was to be able to compare and combine theoretical computations and experimental results to judge the viability of the proposed ambimodal mechanistic route.

4.1 Overview of Reactions

To this end, we set out to synthesise thiophene dioxide **1** using the literature procedure.[1, 91] The reaction sequence (shown in Fig. 4.1A) was performed in batches to yield just under 1 gram of the thiophene dioxide **1**.

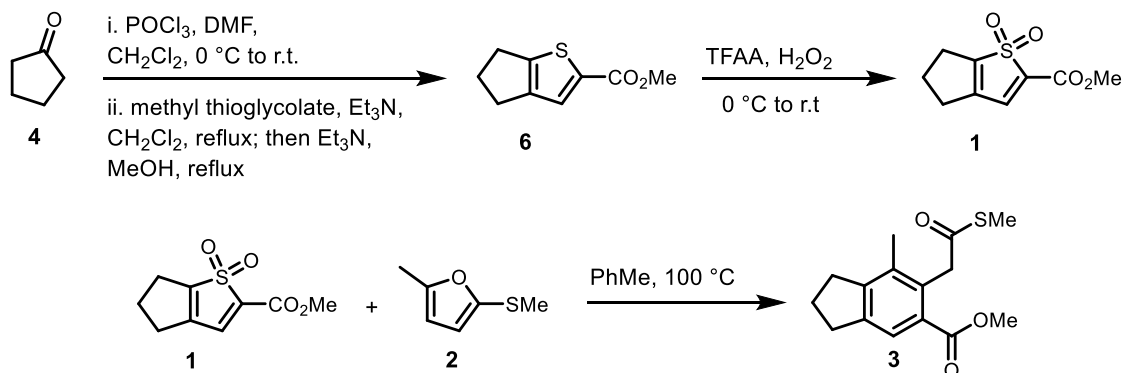


Figure 4.1: General reaction scheme for the preparation of thiophene dioxide **1** and the cycloaddition step.

Commercially available cyclopentanone **4** is treated under Vilsmeier formylation-type conditions[92] to give aldehyde intermediate **5** (see section 4.3.1), which is then further reacted with methyl thioglycolate in the presence of base to give thiophene **6**. Oxidation with hydrogen peroxide in trifluoroacetic anhydride yields target thiophene dioxide **1**. [93]

4.2 KIE Measurements

Kinetic isotope effects for various carbon atoms throughout the system of interest were determined by Singleton’s natural abundance ¹³C NMR procedure described in Section 2.3. We chose to analyse the product at low conversions rather than the recovered starting materials at high conversion. This was motivated by the difficulty in purification for the thiophene dioxide, and the considerable excess in which the furan was added.

Conversion was determined by ^1H NMR spectroscopy using 1,3,5- trimethoxybenzene (TMB) as internal standard. Subsequent ^{13}C NMR spectroscopy on the isolated product allowed the determination of isotopic enrichment by careful integration of the relevant peaks.

To ensure identical conditions for different runs, one stock solution of thiophene dioxide and TMB in dichloromethane, and another of furan in toluene, were prepared. These were used for setting up the reactions for the KIE study. A reaction performed to the fullest attainable conversion (70 h reaction time, 50% yield; details in section 4.3.1) gave a sample for the product **3** which was used as a standard mimicking natural abundance throughout the study.

Guided by the reaction time of 20 hours previously reported[1], we chose to run the reaction for 1 hour at the same temperature of 100 °C. Crude ^1H NMR showed a negligible amount of thiophene dioxide left in the reaction mixture, and we determined this reaction time to be unsuitably long.

Consequently, we performed the reaction at the same conditions but with a 10 minute reaction time. This time, NMR conversion was 24%. As we did not want to drive conversion too low and thus limit our capability to isolate sufficient product, we determined this reaction time was suitable for our aims. We performed another experiment at identical conditions and for 10 minutes, and determined 11% conversion.

The products from both reaction mixtures were purified and subsequently analysed in NMR. Three spectra were collected for each sample to allow for a statistical averaging. Peaks were integrated using a strict and uniform procedure, and the integraion for the peak corresponding to atom C8 from the aliphatic backbone was set as 1000 to be used as a reference (details in section 4.3.2). The KIE at C5 could not determined due to persistent impurity peaks obscuring the integration.

The results from the KIE study (together with their standard deviations) are summarised in Table 4.1. The KIEs obtained are not particularly reliable due to the large errors. An inversion of the sense of the effect is within a standard deviation for each atom. Furthermore, C8 was chosen for the internal reference as the aliphatic backbone was assumed inert in the cycloaddition, however average KIEs of 1.019 ± 0.044 and 0.997 ± 0.044 were obtained for the other atoms on the backbone, C7 and C9 respectively. While some secondary effect could be at work at these two atoms, this results overall undermines the reliability of the measurements. The methyl carbons C17 and C19, which would also be assumed inert, exhibit some effect as well (0.993 ± 0.048 and 1.007 ± 0.045 respectively). Another problem is the KIEs computed from the two experiments are not in good agreement (Appendix B), even though KIEs calculated by Singleton’s method should not vary with conversion (section 2.3).

Atom	KIE	Atom	KIE
C1	1.008 ± 0.044	C8	1 (assumed)
C2	1.034 ± 0.044	C7	1.019 ± 0.044
C3	1.015 ± 0.040	C9	0.997 ± 0.044
C4	1.014 ± 0.043	C17	0.993 ± 0.048
C6	0.992 ± 0.045	C19	1.007 ± 0.045
C13	1.013 ± 0.042		

Table 4.1: Summary of average KIE measured. See figure on the right for atom labels.

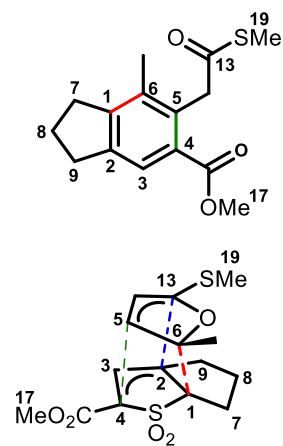


Figure 4.2: Structures of the final product (top) and the ambimodal **TS1** (bottom) with atom labels for reference.

The substantial error in KIEs is dominated by the integration error, introduced mainly from the differences between the spectra obtained for the standard (full conversion) sample. A significant source of error is also the assumed 10% uncertainty

in the conversion measurement brought about by the non-quantitative crude ^1H NMR spectra.

A comparison between computed and experimentally obtained KIE values for the relevant atoms is presented in Table 4.2. The values are not in good agreement for some atoms, which is to be expected in light of the limitations discussed here and in section 3.4.

	Computed		Experimental
Atom	TS1	TS2	
C1	1.028	1.004	1.008 ± 0.044
C2	1.008	1.018	1.034 ± 0.044
C4	1.006	1.013	1.014 ± 0.043
C6	1.041	1.008	0.992 ± 0.045
C13	1.001	1.024	1.013 ± 0.042

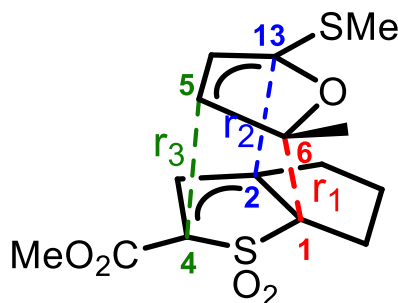


Table 4.2: Comparison between computed (for ambimodal cycloaddition **TS1** and Cope rearrangement **TS2**) and experimental KIEs. See Tables 3.2 and 4.1 for full information.

Figure 4.3: Structure of **TS1** with atom labels for reference.

A small effect is observed at C1 going against the computational predictions for **TS1**. This may allude to the TS having the red bond formed to a larger extent in reality than in our modeling: a late TS should give a smaller KIE. This may however simply be the result of the simple TST approach taken to compute the KIEs not being sophisticated enough to reflect the reality of system.

The most dramatic discrepancy between computation and experiment is at C6. This is also the only atom in the aromatic core that exhibits an experimental effect smaller than 1. This could be the result of a balance between the normal primary KIE due to a change in the ZPE on the one hand, and a secondary effect attributed to the $sp^2 \rightarrow sp^3$ rehybridisation this atom undergoes. That said, care

should be taken when making conclusions about the KIEs measured at C6 and C13, as these belong to the furan which was added in a great stoichiometric excess. This scenario is not described well within Singleton’s method. Further to that, the measurement on C6 is the one with the largest error.

One major problem with the investigated reaction is that it does not go to completion. Even though the yield for the cycloaddition step previously reported[1] (67%) exceeds the one in this work (50 %), the transformation is still nowhere near quantitative. Our observations have lead us to believe that the thiophene dioxide is depleted through some side reactions. Presumably, these reactions may have KIEs themselves, which would obscure the measurements performed on the cycloaddition. This complexity is not accounted for in the simple Singleton model, and cannot be resolved if the side reactions are unknown.

Despite the problems with the reliability of this data, some rationalisations can be drawn from it. A potential stepwise mechanism, where a Michael addition is the first and rate-determining step, would exhibit significant effects only at C1 and C6. The effects at C2 and C4 can therefore be interpreted as evidence against that mechanistic pathway.

The effect at C2 being larger than at C4 is in accordance with the computational predictions for both TSs. Importantly, it suggests the reaction does not proceed *via* a simple [4+2] IED addition, and that the NED mode plays an important role. The effect at C13 is also in agreement with that. This result further bolsters the ambimodal hypothesis, and our suggestion that **IM2** converts to **IM1** primarily through the Cope rearrangement pathway as opposed to *via* the retroaddition originally proposed by Duarte and Anderson.

Overall, the experimental KIE evidence argues against a stepwise Michael addition mechanism and a simple IED cycloaddition. Although it cannot prove

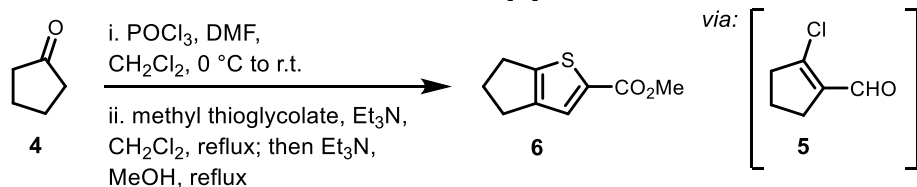
it conclusively, it is in qualitative agreement with the proposed ambimodal cycloaddition mechanism.

4.3 Experimental Details

4.3.1 Synthesis and Characterisation

General synthetic procedures are given below. The reactions were performed in batches and at different scales to give the required amount of thiophene dioxide 1. NMR spectra are collated in Appendix C.

Methyl 5,6-dihydro-4*H*-cyclopenta[*b*]thiophene-2-carboxylate, 6.

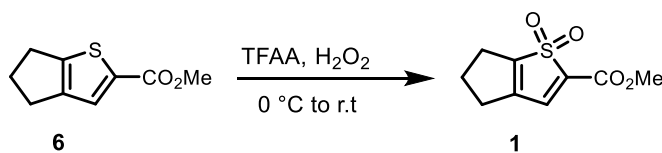


(i) To a stirred solution of anhydrous DMF (7.4 ml, 95 mmol, 1.6 equiv.) at 0 °C was added phosphorus oxychloride (6.6 mL, 71 mmol, 1.2 equiv.) dropwise. The resulting mixture was allowed to warm up to room temperature, and then stirred for 1 h before a solution of cyclopentanone (5 g, 5.26 mL, 59 mmol, 1.0 equiv.) in CH₂Cl₂ (25 mL) was added. The resulting mixture was stirred and heated under reflux for 4 h before it was cooled down and poured onto a mixture of ice, deionised water, and NaOAc (15 g, 181 mmol, 3.1 equiv.). The resulting layers were separated and the aqueous layer was extracted with CH₂Cl₂ (2 × 30 mL). The combined organic layers were washed with water (20 mL), brine (20 mL), dried over Na₂SO₄, filtered, and concentrated *in vacuo* to afford aldehyde **5** as a brown oil (3.26 g, 25 mmol, 42%), which was used directly without further purification.

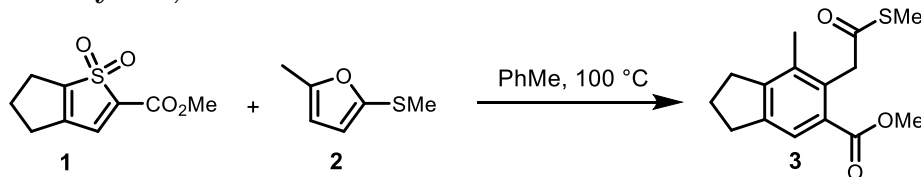
(ii) To a stirred solution of crude aldehyde **5** (3.2 g, 25 mmol, 1.0 equiv.) in CH₂Cl₂ (20 mL) was added methylthioglycolate (2.46 mL, 27.5 mmol, 1.1 equiv.) and Et₃N (13.9 mL, 100 mmol, 4.0 equiv.). The resulting mixture was heated to

reflux and stirred for 16 h before it was cooled to room temperature and concentrated *in vacuo*. To the resulting residue in MeOH (20 mL) at room temperature was added Et₃N (13.9 mL, 100 mmol, 4.0 equiv). The resulting mixture was warmed to reflux and stirred for 16 h before it was cooled to room temperature and concentrated under reduced pressure. Flash column chromatography (silica gel, pentane:Et₂O 20:1 → 6:1) afforded **6** as a pale yellow powder (2.97 g, 16.3 mmol, 28% over two steps). *R_f* = 0.6 (pentane:Et₂O, 5:1). ¹H NMR (400 MHz, CDCl₃): δ 7.51 (s, 1H), 3.84 (s, 3H), 2.92 (t, *J* = 7.2 Hz, 2H), 2.74 (t, *J* = 7.2 Hz, 2H), 2.44 (p, *J* = 7.4 Hz, 2H). ¹³C NMR (101 MHz, CDCl₃): δ 163.1, 151.1, 147.2, 135.5, 128.9, 52.0, 29.5, 29.3, 28.1. Data identical to literature values.[1]

Methyl 5,6-dihydro-4*H*-cyclopenta[*b*]thiophene-2-carboxylate 1,1-dioxide, 1



To a stirred solution of trifluoroacetic anhydride (0.80 mL, 5.5 mmol, 10 equiv.) at 0 °C was added hydrogen peroxide (0.20 mL 30% aq., 2.1 mmol, 3.8 equiv.) dropwise. The mixture was stirred at room temperature for 15 minutes before adding starting material thiophene **6** (100 mg, 0.55 mmol, 1.0 equiv.) at 0 °C. The resulting mixture was stirred for 22 h at r.t. before being reduced *in vacuo*. Flash column chromatography (silica gel, pentane:ethyl acetate, 10:1 → 1:1) afforded thiophene dioxide **1** as a white crystalline solid (25 mg, 0.11 mmol, 21%). ¹H NMR (400 MHz, CDCl₃): δ 7.47 (s, 1H), 3.92 (s, 3H), 2.81-2.74 (m, 2H), 2.69-2.60 (m, 2H), 2.49-2.40 (m, 2H). ¹³C NMR (101 MHz, CDCl₃): δ 158.4, 146.7, 146.1, 139.1, 135.2, 53.0, 29.1, 26.6 (2 C). Data identical to literature values.[1]

Methyl 7-methyl-6-(2-(methylthio)-2-oxoethyl)-2,3-dihydro-1*H*-indene-5-carboxylate, **3**

To a stirred solution of thiophene dioxide **1** (47 mg, 0.21 mmol, 1.0 equiv) in toluene (1.5 mL) was added furan **2** (0.11 mL, 0.88 mmol, 4.0 equiv.). The mixture was stirred at 100 °C for 70 h. Flash column chromatography (silica gel, pentane:Et₂O, 10:1) afforded **3** as a yellow oil (29 mg, 0.10 mmol, 50%). ¹H NMR (600 MHz, CDCl₃): δ 7.69 (s, 1H), 4.34 (s, 2H), 3.86 (s, 3H), 2.96 (t, *J* = 7.5 Hz, 2H), 2.91 (t, *J* = 7.5 Hz, 2H), 2.27 (s, 3H), 2.24 (s, 2H), 2.10 (p, *J* = 7.5 Hz, 2H). ¹³C NMR (151 MHz, CDCl₃): δ 198.1, 169.5, 148.5, 143.2, 134.7, 131.3, 129.1, 124.2, 52.0, 44.5, 33.0, 32.7, 24.6, 16.5, 11.7. Data identical to literature values.[1]

4.3.2 NMR Analysis for KIE Determination

The analysis was performed in accordance with literature procedures.[67, 68, 69] ¹³C NMR spectra of the samples were obtained on a Bruker 600 MHz spectrometer at 150.99 MHz carrier frequency, with 128 scans and an 85 s relaxation delay. The obtained FIDs were analysed in MNova with a zero filling to 2,097,152 points and a Whittaker smoother baseline correction.[94, 95] Three independent spectra were obtained from each sample to achieve a statistically significant averaging of the results.

5

Conclusions and Outlook

Contents

5.1	Conclusion	48
5.2	Future Work	49

5.1 Conclusion

We have presented a study into the mechanism of a bipericyclic inverse electron demand Diels-Alder cycloaddition between a furan and a thiophene dioxide originally reported by Duarte and Anderson.

Our DFT computations suggest the reaction is occurring *via* a highly asynchronous concerted cycloaddition rather than a stepwise Michael addition, and confirm the existence of an ambimodal transition state at the ω B97X-D3(BJ)/def2-TZVPP level. The energy barriers computed suggest a Cope rearrangement likely operates to interconvert the normal to the inverse electron demand adduct, and the existence of an accessible and practically irreversible SO₂ cheletropic extrusion for the latter, which is in agreement with the experimental observations.

The AIMD simulations of the cycloaddition step provide novel information about the system of interest. They confirm the ambimodal TS is connected to both the NED and IED adducts on the PES, which had only been assumed previously. These simulations also give some insight into the dynamics of the bond formation processes in the reaction, and the shape of the PES.

Albeit inconclusive, the $^{13}\text{C}/^{12}\text{C}$ KIE experimental results do not support a stepwise addition and an ordinary $[4 + 2]$ IED cycloaddition, thus hinting towards a bipericyclic mechanistic pathway.

Overall, the aims outlined in the beginning of the thesis have been fulfilled, and new knowledge of the system has been acquired.

5.2 Future Work

One way to refine the computational study of the system would be to perform a more exhaustive benchmarking of the level of theory. Confirming the bipericyclic nature of the cycloaddition step at multiple levels (ideally ones which have been shown in the literature to be appropriate for similar systems) would increase the reliability of the claim of ambimodality. In particular, obtaining meaningful results at the M06-2X level would show the results originally reported by Duarte and Anderson were not accidental. Comparing the performance of multiple DFT functionals and benchmarking against more expensive and reliable methods (e.g. coupled cluster[96]) would provide information about the method-dependence of the system.

The AIMD simulations can be refined by utilising a larger basis set, as the one used in this study (def2-SVP) is unlikely to be accurate enough. Going beyond downhill dynamics and employing enhanced sampling methods would allow for obtaining dynamic trajectories initiated at a minimum, going through a TS, and terminating in another minimum, i.e. describing a full kinetic step.[97, 98] One

approach towards this would be to employ umbrella sampling along the IRC paths of the ambimodal cycloaddition **TS1** and the Cope rearrangement **TS2**.^[99] The final aim would be to obtain dynamic trajectories traversing the whole chemical space from reactants to products.

The part of this work that needs the most refinement is the KIE studies. This could be achieved by performing further experiments until better agreement is achieved between the results obtained from separate samples. Performing a larger number of independent NMR measurements on each sample would aid the statistical averaging and reduce errors. It would be worth investigating the possibility of applying Singleton’s method to recovered starting materials at high conversion as well.

Finally, obtaining any experimental evidence for the formation of the NED **IM2** and its conversion to **IM1** would strongly support the proposed mechanistic pathway. This could be approached with careful NMR observation of the system as the reaction progresses. Alternatively, a study can be devised wherein **IM2** is formed *in situ* through another reaction under the reaction conditions, to confirm whether it would form the same product.

Bibliography

- [1] Park, K. H. K., Frank, N., Duarte, F. & Anderson, E. A. Collective synthesis of illudalane sesquiterpenes via cascade inverse electron demand (4 + 2) cycloadditions of thiophene s,s-dioxides. *Journal of the American Chemical Society* **144**, 10017–10024 (2022).
- [2] Eyring, H. The activated complex in chemical reactions. *The Journal of Chemical Physics* **3**, 107–115 (1935).
- [3] Arrhenius, S. Über die dissociationswärme und den einfluss der temperatur auf den dissociationsgrad der elektrolyte. *Zeitschrift für Physikalische Chemie* **4U**, 96–116 (1889).
- [4] Rice–ramspberger–kassel (RRK) theory. In *The IUPAC Compendium of Chemical Terminology* (International Union of Pure and Applied Chemistry (IUPAC), 2014). URL <https://doi.org/10.1351/goldbook.r05390>.
- [5] Marcus, R. A. On the theory of electron-transfer reactions. VI. unified treatment for homogeneous and electrode reactions. *The Journal of Chemical Physics* **43**, 679–701 (1965).
- [6] Marcus, R. A. & Rice, O. K. The kinetics of the recombination of methyl radicals and iodine atoms. *The Journal of Physical Chemistry* **55**, 894–908 (1951).
- [7] Marcus, R. A. Lifetimes of active molecules. I. *The Journal of Chemical Physics* **20**, 352–354 (1952).
- [8] Marcus, R. A. Lifetimes of active molecules. II. *The Journal of Chemical Physics* **20**, 355–359 (1952).
- [9] Marcus, R. A. Unimolecular dissociations and free radical recombination reactions. *The Journal of Chemical Physics* **20**, 359–364 (1952).
- [10] Truhlar, D. G., Garrett, B. C. & Klippenstein, S. J. Current status of transition-state theory. *The Journal of Physical Chemistry* **100**, 12771–12800 (1996).
- [11] Ess, D. H. *et al.* Bifurcations on potential energy surfaces of organic reactions. *Angewandte Chemie-International Edition* **47**, 7592–7601 (2008).
- [12] Carpenter, B. K. Dynamic behavior of organic reactive intermediates. *Angewandte Chemie International Edition* **37**, 3340–3350 (1998).
- [13] Rehbein, J. & Carpenter, B. K. Do we fully understand what controls chemical selectivity? *Physical Chemistry Chemical Physics* **13**, 20906 (2011).

- [14] Carpenter, B. K. Dynamic matching: The cause of inversion of configuration in the [1, 3] sigmatropic migration? *Journal of the American Chemical Society* **117**, 6336–6344 (1995).
- [15] Carpenter, B. K. Bimodal distribution of lifetimes for an intermediate from a quasiclassical dynamics simulation. *Journal of the American Chemical Society* **118**, 10329–10330 (1996).
- [16] Collins, P., Carpenter, B. K., Ezra, G. S. & Wiggins, S. Nonstatistical dynamics on potentials exhibiting reaction path bifurcations and valley-ridge inflection points. *Journal of Chemical Physics* **139** (2013).
- [17] Collins, P., Kramer, Z. C., Carpenter, B. K., Ezra, G. S. & Wiggins, S. Nonstatistical dynamics on the caldera. *Journal of Chemical Physics* **141** (2014).
- [18] Taketsugu, T., Tajima, N. & Hirao, K. Approaches to bifurcating reaction path. *The Journal of Chemical Physics* **105**, 1933–1939 (1996).
- [19] Wenthold, P. G., Hrovat, D. A., Borden, W. T. & Lineberger, W. C. Transition-state spectroscopy of cyclooctatetraene. *Science* **272**, 1456–1459 (1996).
- [20] Hrovat, D. A. & Borden, W. T. CASSCF calculations find that a D8h geometry is the transition state for double bond shifting in cyclooctatetraene. *Journal of the American Chemical Society* **114**, 5879–5881 (1992).
- [21] Shaik, S., Danovich, D., Sastry, G. N., Ayala, P. Y. & Schlegel, H. B. Dissociative electron transfer, substitution, and borderline mechanisms in reactions of ketyl radical anions. differences and difficulties in their reaction paths. *Journal of the American Chemical Society* **119**, 9237–9245 (1997).
- [22] Valtazanos, P. & Ruedenberg, K. Bifurcations and transition states. *Theoretica Chimica Acta* **69**, 281–307 (1986).
- [23] Nouri, D. H. & Tantillo, D. J. Hiscotropic rearrangements: Hybrids of electrocyclic and sigmatropic reactions. *The Journal of Organic Chemistry* **71**, 3686–3695 (2006).
- [24] Hrovat, D. A., Duncan, J. A. & Borden, W. T. Ab initio and DFT calculations on the cope rearrangement of 1, 2, 6-heptatriene. *Journal of the American Chemical Society* **121**, 169–175 (1998).
- [25] Stephenson, L. M., Grdina, M. J. & Orfanopoulos, M. Mechanism of the ene reaction between singlet oxygen and olefins. *Accounts of Chemical Research* **13**, 419–425 (1980).
- [26] Zou, Y. & Houk, K. N. Mechanisms and dynamics of synthetic and biosynthetic formation of delitschiapyrones: Solvent control of ambimodal periselectivity. *Journal of the American Chemical Society* **143**, 11734–11740 (2021).
- [27] Zhang, H. *et al.* Computational exploration of the mechanism of critical steps in the biomimetic synthesis of preisolactone A, and discovery of new ambimodal (5 + 2)/(4 + 2) cycloadditions. *Journal of the American Chemical Society* **143**, 6601–6608 (2021).

- [28] Jamieson, C. S., Sengupta, A. & Houk, K. N. Cycloadditions of cyclopentadiene and cycloheptatriene with tropones: All *endo*-[6+4] cycloadditions are ambimodal. *Journal of the American Chemical Society* **143**, 3918–3926 (2021).
- [29] Nguyen, M. T., Chandra, A. K., Uchimaru, T. & Sakai, S. Comment on the electronic reorganization in 1, 3-dipolar cycloaddition of fulminic acid to acetylene. *The Journal of Physical Chemistry A* **105**, 10943–10945 (2001).
- [30] Kraka, E., Wu, A. & Cremer, D. Mechanism of the diels-alder reaction studied with the united reaction valley approach: mechanistic differences between symmetry-allowed and symmetry-forbidden reactions. *The Journal of Physical Chemistry A* **107**, 9008–9021 (2003).
- [31] Caramella, P., Quadrelli, P. & Toma, L. An unexpected bispericyclic transition structure leading to 4+2 and 2+4 cycloadducts in the *endo* dimerization of cyclopentadiene. *Journal of the American Chemical Society* **124**, 1130–1131 (2002).
- [32] Martin-Somer, A., Xue, X.-S., Jamieson, C. S., Zou, Y. & Houk, K. Computational design of a tetrapericyclic cycloaddition and the nature of potential energy surfaces with multiple bifurcations. *Journal of the American Chemical Society* **145**, 4221–4230 (2023).
- [33] Zhou, Q., Thøgersen, M. K., Rezayee, N. M., Jørgensen, K. A. & Houk, K. N. Ambimodal bispericyclic [6 + 4]/[4 + 6] transition state competes with diradical pathways in the cycloheptatriene dimerization: Dynamics and experimental characterization of thermal dimers. *Journal of the American Chemical Society* **144**, 22251–22261 (2022).
- [34] Ito, T., Maeda, S. & Harabuchi, Y. Kinetic analysis of a reaction path network including ambimodal transition states: A case study of an intramolecular diels-alder reaction. *Journal of Chemical Theory and Computation* **18**, 1663–1671 (2022).
- [35] Zhang, C. *et al.* Dynamical trajectory study of the transannular [6+4] and ambimodal cycloaddition in the biosynthesis of heronamides. *The Journal of Organic Chemistry* **85**, 9440–9445 (2020).
- [36] Zhang, Z. *et al.* Enzyme-catalyzed inverse-electron demand diels-alder reaction in the biosynthesis of antifungal ilicicolin h. *Journal of the American Chemical Society* **141**, 5659–5663 (2019).
- [37] Gagnepain, J., Castet, F. & Quideau, S. Total synthesis of (+)-aquaticol by biomimetic phenol dearomatization: Double diastereofacial differentiation in the diels-alder dimerization of orthoquinols with a C₂-symmetric transition state. *Angewandte Chemie International Edition* **46**, 1533–1535 (2007).
- [38] Yang, Z. *et al.* Relationships between product ratios in ambimodal pericyclic reactions and bond lengths in transition structures. *Journal of the American Chemical Society* **140**, 3061–3067 (2018).
- [39] Møller, C. & Plesset, M. S. Note on an approximation treatment for many-electron systems. *Physical Review* **46**, 618–622 (1934).

- [40] Parr, R. G. Density functional theory of atoms and molecules. In *Horizons of Quantum Chemistry*, 5–15 (Springer Netherlands, 1980). URL https://doi.org/10.1007/978-94-009-9027-2_2.
- [41] Hohenberg, P. & Kohn, W. Inhomogeneous electron gas. *Physical Review* **136**, B864–B871 (1964).
- [42] Perdew, J. P. Jacob’s ladder of density functional approximations for the exchange-correlation energy. In *AIP Conference Proceedings* (AIP, 2001). URL <https://doi.org/10.1063/1.1390175>.
- [43] Vega, L., Ruvireta, J., Viñes, F. & Illas, F. Jacob’s ladder as sketched by Escher: Assessing the performance of broadly used density functionals on transition metal surface properties. *Journal of Chemical Theory and Computation* **14**, 395–403 (2017).
- [44] Notario-Estévez, A., Kozlov, S. M., Viñes, F. & Illas, F. Electronic-structure-based material descriptors: (in)dependence on self-interaction and hartree–fock exchange. *Chemical Communications* **51**, 5602–5605 (2015).
- [45] London, F. Zur theorie und systematik der molekularkräfte. *Zeitschrift für Physik* **63**, 245–279 (1930).
- [46] 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. *The Journal of Chemical Physics* **132**, 154104 (2010).
- [47] Grimme, S., Ehrlich, S. & Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry* **32**, 1456–1465 (2011).
- [48] Hehre, W. J., Stewart, R. F. & Pople, J. A. Self-consistent molecular-orbital methods. I. Use of gaussian expansions of slater-type atomic orbitals. *The Journal of Chemical Physics* **51**, 2657–2664 (1969).
- [49] Weigend, F. & Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Physical Chemistry Chemical Physics* **7**, 3297 (2005).
- [50] Tomasi, J., Mennucci, B. & Cammi, R. Quantum mechanical continuum solvation models. *Chemical Reviews* **105**, 2999–3094 (2005).
- [51] Klamt, A. & Schüürmann, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2* 799–805 (1993).
- [52] Andzelm, J., Kölmel, C. & Klamt, A. Incorporation of solvent effects into density functional calculations of molecular energies and geometries. *The Journal of Chemical Physics* **103**, 9312–9320 (1995).
- [53] 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* **102**, 1995–2001 (1998).

- [54] Cossi, M., Rega, N., Scalmani, G. & Barone, V. Energies, structures, and electronic properties of molecules in solution with the c-PCM solvation model. *Journal of Computational Chemistry* **24**, 669–681 (2003).
- [55] Yin, J. *et al.* Overview of the SAMPL5 host–guest challenge: Are we doing better? *Journal of Computer-Aided Molecular Design* **31**, 1–19 (2016).
- [56] da Silva, C. O., Mennucci, B. & Vreven, T. Combining microsolvation and polarizable continuum studies: new insights in the rotation mechanism of amides in water. *The Journal of Physical Chemistry A* **107**, 6630–6637 (2003).
- [57] Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics* **81**, 511–519 (1984).
- [58] Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A* **31**, 1695–1697 (1985).
- [59] Martyna, G. J., Klein, M. L. & Tuckerman, M. Nosé–hoover chains: The canonical ensemble via continuous dynamics. *The Journal of Chemical Physics* **97**, 2635–2643 (1992).
- [60] Martyna, G. J., Tuckerman, M. E., Tobias, D. J. & Klein, M. L. Explicit reversible integrators for extended systems dynamics. *Molecular Physics* **87**, 1117–1157 (1996).
- [61] Gómez-Gallego, M. & Sierra, M. A. Kinetic isotope effects in the study of organometallic reaction mechanisms. *Chemical Reviews* **111**, 4857–4963 (2011).
- [62] Karandashev, K., Xu, Z.-H., Meuwly, M., Vaníček, J. & Richardson, J. O. Kinetic isotope effects and how to describe them. *Structural Dynamics* **4**, 061501 (2017).
- [63] Shao, L. & Hewitt, M. The kinetic isotope effect in the search for deuterated drugs. *Drug News & Perspectives* **23**, 398 (2010).
- [64] Gu, H. & Zhang, S. Advances in kinetic isotope effect measurement techniques for enzyme mechanism study. *Molecules* **18**, 9278–9292 (2013).
- [65] Huang, M. *et al.* Dissecting the mechanisms of a class of chemical glycosylation using primary ¹³C kinetic isotope effects. *Nature Chemistry* **4**, 663–667 (2012).
- [66] Singleton, D. A. & Thomas, A. A. High-precision simultaneous determination of multiple small kinetic isotope effects at natural abundance. *Journal of the American Chemical Society* **117**, 9357–9358 (1995).
- [67] Li, K., Huang, S., Liu, T., Jia, S. & Yan, H. Organocatalytic asymmetric dearomatizing hetero-diels–alder reaction of nonactivated arenes. *Journal of the American Chemical Society* **144**, 7374–7381 (2022).
- [68] Liu, C.-X. *et al.* Explicit mechanism of Rh(I)-catalyzed asymmetric C-H arylation and facile synthesis of planar chiral ferrocenophanes. *Journal of the American Chemical Society* **145**, 4765–4773 (2023).

- [69] Guo, R. *et al.* Photochemical dearomative cycloadditions of quinolines and alkenes: Scope and mechanism studies. *Journal of the American Chemical Society* **144**, 17680–17691 (2022).
- [70] Wheeler, S. E. & Houk, K. N. Integration grid errors for meta-GGA-predicted reaction energies: Origin of grid errors for the M06 suite of functionals. *Journal of Chemical Theory and Computation* **6**, 395–404 (2010).
- [71] Grimme, S. Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *The Journal of Chemical Physics* **118**, 9095–9102 (2003).
- [72] Linder, M. & Brinck, T. On the method-dependence of transition state asynchronicity in Diels–Alder reactions. *Physical Chemistry Chemical Physics* **15**, 5108 (2013).
- [73] Chai, J.-D. & Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *The Journal of Chemical Physics* **128**, 084106 (2008).
- [74] Nájera, C. & Yus, M. Desulfonylation reactions: Recent developments. *Tetrahedron* **55**, 10547–10658 (1999).
- [75] Nicolaou, K. C., Barnette, W. E. & Ma, P. A remarkably simple, highly efficient, and stereoselective synthesis of steroids and other polycyclic systems. Total synthesis of estra-1, 3, 5(10)-trien-17-one *via* intramolecular capture of o-quinodimethanes generated by cheletropic elimination of sulfur dioxide. *The Journal of Organic Chemistry* **45**, 1463–1470 (1980).
- [76] Prilezhaeva, E. N. Sulfones and sulfoxides in the total synthesis of biologically active natural compounds. *Russian Chemical Reviews* **69**, 367–408 (2000).
- [77] Lee, S. & Goodman, J. M. Rapid route-finding for bifurcating organic reactions. *Journal of the American Chemical Society* **142**, 9210–9219 (2020).
- [78] Lee, S. & Goodman, J. M. VRAI-selectivity: calculation of selectivity beyond transition state theory. *Organic Biomolecular Chemistry* **19**, 3940–3947 (2021).
- [79] Kelly, K. K., Hirschi, J. S. & Singleton, D. A. Newtonian kinetic isotope effects. Observation, prediction, and origin of heavy-atom dynamic isotope effects. *Journal of the American Chemical Society* **131**, 8382–8383 (2009).
- [80] Neese, F. The ORCA program system. *WIREs Computational Molecular Science* **2**, 73–78 (2011).
- [81] Neese, F., Wennmohs, F., Becker, U. & Riplinger, C. The ORCA quantum chemistry program package. *The Journal of Chemical Physics* **152**, 224108 (2020).
- [82] Zhao, Y. & Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theoretical Chemistry Accounts* **120**, 215–241 (2007).

- [83] Weigend, F. & Häser, M. RI-MP2: first derivatives and global consistency. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)* **97**, 331–340 (1997).
- [84] Weigend, F., Häser, M., Patzelt, H. & Ahlrichs, R. RI-MP2: optimized auxiliary basis sets and demonstration of efficiency. *Chemical Physics Letters* **294**, 143–152 (1998).
- [85] Grimme, S. Supramolecular binding thermodynamics by dispersion-corrected density functional theory. *Chemistry - A European Journal* **18**, 9955–9964 (2012).
- [86] Young, T. & Duarte, F. Otherm: thermochemical contributions from ORCA calculations. <https://github.com/duartegroup/otherm>.
- [87] Mills, G. & Jónsson, H. Quantum and thermal effects in H₂ dissociative adsorption: Evaluation of free energy barriers in multidimensional quantum systems. *Physical Review Letters* **72**, 1124–1127 (1994).
- [88] Mills, G., Jónsson, H. & Schenter, G. K. Reversible work transition state theory: application to dissociative adsorption of hydrogen. *Surface Science* **324**, 305–337 (1995).
- [89] Ásgeirsson, V. *et al.* Nudged elastic band method for molecular reactions using energy-weighted springs combined with eigenvector following. *Journal of Chemical Theory and Computation* **17**, 4929–4945 (2021).
- [90] Brehm, M. ORCA Molecular Dynamics Module. <https://brehm-research.de/orcamd.php>.
- [91] Wang, X. *et al.* Discovery of potent and selective tricyclic inhibitors of bruton's tyrosine kinase with improved druglike properties. *ACS Medicinal Chemistry Letters* **8**, 608–613 (2017).
- [92] Vilsmeier, A. & Haack, A. Über die einwirkung von halogenphosphor auf alkyl-formanilide. eine neue methode zur darstellung sekundärer und tertiärer ip/i-alkylamino-benzaldehyde. *Berichte der deutschen chemischen Gesellschaft (A and B Series)* **60**, 119–122 (1927).
- [93] Nenajdenko, V. G., Moiseev, A. M. & Balenkova, E. S. A novel method for the oxidation of thiophenes. Synthesis of thiophene 1, 1-dioxides containing electron-withdrawing substituents. *Russian Chemical Bulletin* **53**, 2241–2247 (2004).
- [94] Whittaker, E. T. On a new method of graduation. *Proceedings of the Edinburgh Mathematical Society* **41**, 63–75 (1922).
- [95] Cobas, C. Applications of the whittaker smoother in NMR spectroscopy. *Magnetic Resonance in Chemistry* **56**, 1140–1148 (2018).
- [96] Bartlett, R. J. & Musiał, M. Coupled-cluster theory in quantum chemistry. *Reviews of Modern Physics* **79**, 291–352 (2007).
- [97] Bernardi, R. C., Melo, M. C. & Schulten, K. Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1850**, 872–877 (2015).

- [98] Bartolucci, G., Orioli, S. & Faccioli, P. Transition path theory from biased simulations. *The Journal of Chemical Physics* **149**, 072336 (2018).
- [99] Kästner, J. Umbrella sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **1**, 932–942 (2011).
- [100] Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Physical Review* **140**, A1133–A1138 (1965).
- [101] March, N. The thomas-fermi approximation in quantum mechanics. *Advances in Physics* **6**, 1–101 (1957).
- [102] Meijer, E. J. & Sprik, M. A density-functional study of the intermolecular interactions of benzene. *The Journal of Chemical Physics* **105**, 8684–8689 (1996).
- [103] Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Physical Review Letters* **77**, 3865–3868 (1996).
- [104] Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics* **98**, 5648–5652 (1993).

Appendices



More on QM

A.1 Obtaining the HF Energy

For a closed-shell system, the HF wavefunction is an antisymmetrised product of orthonormal one-electron spin-orbitals. In Dirac notation:

$$|\Psi\rangle = \sqrt{n!} \hat{A} |\phi_1 \phi_2 \dots \phi_n\rangle, \quad (\text{A.1})$$

where

$$\hat{A} = \frac{1}{n!} \sum_{\mu} (-1)^{p_{\mu}} \hat{\pi}_{\mu} \quad (\text{A.2})$$

is the antisymmetric projection operator necessary to ensure the wavefunction satisfies the Pauli exclusion principle, and $\hat{\pi}_{\mu}$ is the permutation operator exchanging a pair of electrons μ . Alternatively, this antisymmetric property is described using a determinant which has spin-orbitals as its elements called a Slater determinant.

The HF energy can be written as the expectation value of the Hamiltonian acting on the HF wavefunction:

$$E_{\text{HF}} = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle}. \quad (\text{A.3})$$

The HF Hamiltonian in (2.2) can be decomposed into operators acting on one and on two electrons:

$$\hat{H} = \hat{h} + \hat{g} = \sum_i \hat{h}_i + \sum_{i < j} \hat{g}_{ij}, \quad (\text{A.4})$$

where

$$\hat{h}_i = \frac{1}{2} \nabla_i^2 - \sum_I \frac{Z_I}{r_{iI}} \quad (\text{A.5})$$

and

$$\hat{g}_{ij} = \frac{1}{r_{ij}}. \quad (\text{A.6})$$

We now arrive at

$$\langle \Psi | \hat{h} | \Psi \rangle = \sum_i \langle \phi_i | \hat{h}_i | \phi_i \rangle \quad (\text{A.7})$$

and

$$\langle \Psi | \hat{g} | \Psi \rangle = \sum_{i < j} \langle \phi_i \phi_j | r_{12}^{-1} (1 - \hat{\pi}_{12}) | \phi_i \phi_j \rangle \quad (\text{A.8})$$

as all the other terms vanish due to the antisymmetriser \hat{A} . Finally, the HF energy is

$$E_{HF} = \sum_i \langle \phi_i | \hat{h}_i | \phi_i \rangle + \frac{1}{2} \sum_{ij} \langle \phi_i \phi_j | r_{12}^{-1} | \phi_i \phi_j \rangle - \frac{1}{2} \sum_{ij} \langle \phi_i \phi_j | r_{12}^{-1} | \phi_j \phi_i \rangle. \quad (\text{A.9})$$

A.2 Treatment of Fermionic Correlation in HF

Two main origins of correlation can be identified - the Fermionic nature of electrons, and Coulombic forces. The former refers to Pauli exclusion principle stating that the wavefunction is asymmetric with respect to exchange of any pair of electrons (as they are Fermions), i.e. $\hat{\pi}_{ij} \Psi = -\Psi$ for any pair of electrons i and j . The latter refers to the repulsion arising between electrons on grounds of them being like charges.

The original Hartree description, whereby the system is treated as a product of spin-orbitals, does not account for either type of correlation. The more refined Hartree-Fock theory constructs wavefunctions with the help of the antisymmetriser operator \hat{A} , and thus allows for correlation arising from the Pauli exclusion principle to be correctly described.

To illustrate this, we can take the HF wavefunction corresponding to one of the triplet states of a two-electron system:

$$\Psi_{HF}^3 = \frac{1}{\sqrt{2}} [\phi_1(\mathbf{r}_1) \phi_2(\mathbf{r}_2) - \phi_2(\mathbf{r}_1) \phi_1(\mathbf{r}_2)] \alpha(1) \alpha(2) \quad (\text{A.10})$$

The density probability function ρ is proportional to $|\Psi|^2$ under the Born interpretation. When $\mathbf{r}_1 = \mathbf{r}_2$, i.e. the electrons occupy the same point in space, Ψ_{HF}^3 vanishes and so does ρ . This impossibility for two electrons of the same spin to occupy the same point in space has its origin in the Pauli exclusion principle and the Fermionic nature of electrons, and gives rise to a region of space with depleted electron density known as a Fermi hole. HF theory correctly describes this phenomenon.

It is noteworthy that the HF description of the singlet state of a two-electron system is completely uncorrelated, as no Fermi hole appears for electrons of opposite spins. This is clearly inconsistent with the Coulombic repulsion between like-charged electrons. The failure of HF theory to account for Coulomb correlation makes it quantitatively unreliable in many cases.

A.3 More on DFT and Jacob's Ladder

For the purposes of DFT, the Hamiltonian is usually broken down into a universal n -electron part, containing the electronic kinetic energy and the electron-electron repulsions:

$$\hat{F} = \hat{T} + \hat{V}_{ee} = \sum_i^n -\frac{1}{2} \nabla_i^2 + \sum_{i < j}^n \frac{1}{r_{ij}}, \quad (\text{A.11})$$

and a system-dependent potential $v(\mathbf{r})$:

$$\hat{V} = v(\mathbf{r}) = - \sum_{iI}^{nN} \frac{Z_I}{r_{iI}}, \quad (\text{A.12})$$

for a system with n electrons and N nuclei with charges Z_I .

The treatment of the system with regards only to the electron density is made possible by the two Hohenberg-Kohn theorems.[41] The first one states that the potential $v(\mathbf{r})$ can be inferred up to a constant solely from the ground state density ρ . This implies that both the ground state wavefunction and the ground state energy can be expressed as functionals of the density, i.e. $\Psi \equiv \Psi[\rho]$ and $E \equiv E[\rho]$. A rationalisation for this is offered by the fact that the density contains all the necessary information of the system: integrating over all space specifies the number of electrons:

$$\int d\mathbf{r} \rho(\mathbf{r}) = n, \quad (\text{A.13})$$

and the cusps in ρ specify the number, positions, and charges of all nuclei.

The second theorem states that there exists an energy functional that maps any density to a single energy. The functional can be variationally optimised such that the energy is greater than or equal to the true ground state energy. The equality holds only for the exact ground state density.

As the Hamiltonian is broken down to $\hat{H} = \hat{T} + \hat{V}$, the ground state energy becomes

$$E_0 = \min_{\Psi} \langle \Psi | \hat{T} + \hat{V} | \Psi \rangle = \min_{\rho} \left\{ F[\rho] + \int d\mathbf{r} v(\mathbf{r}) \rho(\mathbf{r}) \right\}, \quad (\text{A.14})$$

where the universal density functional is variationally optimised over all wavefunctions that integrate to n electrons $\Psi \rightarrow \rho$:

$$F[\rho] = \min_{\Psi \rightarrow \rho} \langle \Psi | \hat{T} | \Psi \rangle. \quad (\text{A.15})$$

A main objective of DFT is to find a good approximation for $F[\rho]$, as this allows variational searches for the ground state energy E_0 using only the density ρ .

Within Kohn-Sham theory, a fictitious system of n non-interacting electrons is introduced. The wavefunction for this system is a single Slater determinant like in HF theory:

$$|\Phi_s\rangle = |\phi_1 \dots \phi_n\rangle. \quad (\text{A.16})$$

This single determinant of n spin-orbitals can give any density ρ that corresponds to n electrons. Furthermore, a significant part of the kinetic energy can be extracted from the non-interacting system. The universal density functional is broken down into:

$$F[\rho] = T_s[\rho] + E_H[\rho] + E_{xc}[\rho], \quad (\text{A.17})$$

where

$$T_s[\rho] = \min_{\Phi_s \rightarrow \rho} \langle \Phi_s | \hat{T} | \Phi_s \rangle \quad (\text{A.18})$$

is the non-interacting kinetic energy functional,

$$E_H[\rho] = \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1) \rho(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2, \quad (\text{A.19})$$

is the Hartree energy functional, which takes into account the classical electrostatic repulsion for the mean field, and E_{xc} is the exchange correlation functional. This

Functional	Terms	Examples
Double-Hybrid GGA	$\rho, \nabla\rho, E_x^{HF}, E_c^{MP2}$	B2-PLYP
Hybrid GGA*	$\rho, \nabla\rho, E_x^{HF}, (E_x^{lr, HF})^*$	B3LYP, PBE0, M06(-2X), ω B97X*
Meta GGA	$\rho, \nabla\rho, \nabla^2\rho, \tau$	TPSS, SCAN
GGA	$\rho, \nabla\rho$	LYP
LDA	ρ	PW92

Table A.1: Jacob’s ladder of DFT functionals, with a summary of the relevant mathematical terms, and some notable examples. The arrow on the left signifies the general increase in chemical accuracy and computational expense going up the ladder.

*Including range separated hybrids.

provides a great simplification, because the largest contributors to the energy can be evaluated: $T_s[\rho]$ is found variationally, and $E_H[\rho]$ can be computed exactly, albeit it includes a non-physical self-interaction.

The only part of $F[\rho]$ that still needs to be addressed is the exchange-correlation functional $E_{xc}[\rho]$.

We now describe the different rungs on Jacob’s ladder in greater detail than in the main text.

In the local density approximation (LDA) the exchange energy is described as a function of the local density $\rho(\mathbf{r})$ and calculated analytically under the infinite uniform electron gas (UEG) model.[100, 101] LDA work well for extended structures with approximately uniform densities, but fail to be accurate for molecular systems.[102]

In the generalised gradient approximation (GGA)[103], the exchange correlation functional depends on the density ρ and on its gradient $\nabla\rho$. The latter is said to provide some semilocal information about the system. A further improvement are meta GGAs, which also take into account the curvature $\nabla^2\rho$ and the non-interacting positive kinetic energy density $\tau(\mathbf{r})$.

The main shortcoming of the GGA approach is the significant self-interacting error, which leads to unrealistically delocalised densities. In hybrid functionals, this issue is partially resolved by mixing in some of the HF exchange energy obtained from Kohn-Sham orbitals $E_x^{HF}[\Phi]$. An example is the M06-2X functional.[82] In double-hybrid functionals, an MP2 correlation E_c^{MP2} is added *a posteriori*. An example is B2-PLYP.[104]

Another type of functional are the range separated hybrids, in which short- and long-range exchange and correlation are treated and scaled separately. An example is the ω B97X functional.[73]

As density functional approximations do not account for London dispersion interactions (long-range correlation caused by coupled instantaneous dipoles in distant parts of the molecule), semi-empirical corrections to the dispersion are often added to the DFT result. Common examples are D, D3-Zero, and D3(BJ).[46, 47]

In the local density approximation (LDA) the exchange energy is described as a function of the local density $\rho(\mathbf{r})$ and calculated analytically under the infinite uniform electron gas (UEG) model.[100, 101] LDA work well for extended structures with approximately uniform densities, but fail to be accurate for molecular systems.[102]

In the generalised gradient approximation (GGA)[103], the exchange correlation functional depends on the density ρ and on its gradient $\nabla\rho$. The latter is said to provide some semilocal information about the system. A further improvement are meta GGAs, which also take into account the curvature $\nabla^2\rho$ and the non-interacting positive kinetic energy density $\tau(\mathbf{r})$.

The main shortcoming of the GGA approach is the significant self-interacting error, which leads to unrealistically delocalised densities. In hybrid functionals, this issue is partially resolved by mixing in some of the HF exchange energy obtained from Kohn-Sham orbitals $E_x^{\text{HF}}[\Phi]$. An example is the M06-2X functional.[82] In double-hybrid functionals, an MP2 correlation E_c^{MP2} is added *a posteriori*. An example is B2-PLYP.[104]

Another type of functional are the range separated hybrids, in which short- and long-range exchange and correlation are treated and scaled separately. An example is the ω B97X functional.[73]

B

KIE Results

Details from the KIE experiment are tabulated on the next page. S_i , $E_{1.i}$, and $E_{2.i}$ stand for the integral from the i^{th} spectrum for the standard, the first, and the second experiment respectively. R/R_0 is the integral ratio with respect to the standard.

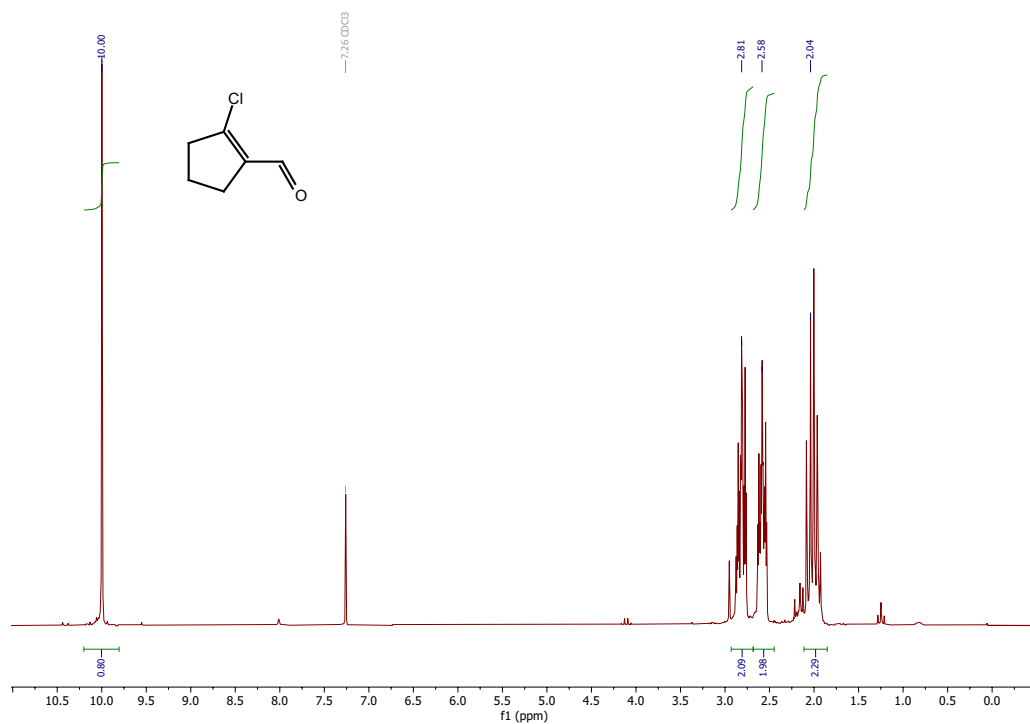
	C1	C2	C3	C4	C6	C13
S ₁	994.46	987.43	914.63	1008.60	994.94	984.77
S ₂	928.60	921.27	874.69	957.76	926.45	933.79
S ₃	948.72	952.48	900.63	966.44	967.64	959.32
S _{avg}	957.26 ± 27.56	953.72 ± 27.03	896.65 ± 16.55	977.60 ± 22.20	963.01 ± 28.15	959.32 ± 20.81
E _{1.1}	941.07	885.32	859.00	963.82	958.54	936.48
E _{1.2}	958.56	930.35	862.75	950.97	966.25	921.71
E _{1.3}	937.85	898.19	874.38	971.84	959.33	949.04
E _{1,avg}	945.83 ± 9.10	904.62 ± 18.94	865.38 ± 6.55	962.21 ± 8.60	961.37 ± 3.46	935.74 ± 11.17
R/R ₀	0.988 ± 0.030	0.949 ± 0.033	0.965 ± 0.019	0.984 ± 0.024	0.998 ± 0.029	0.975 ± 0.024
KIE ₁	1.014 ± 0.041	1.063 ± 0.041	1.042 ± 0.037	1.018 ± 0.039	1.002 ± 0.042	1.029 ± 0.039
E _{2.1}	946.25	950.90	907.89	999.54	990.54	958.14
E _{2.2}	952.76	963.76	884.74	933.24	969.27	945.10
E _{2.3}	968.71	930.21	924.66	972.99	978.91	981.07
E _{2,avg}	955.91 ± 9.44	948.29 ± 13.82	905.76 ± 16.37	968.59 ± 27.25	979.57 ± 8.70	961.44 ± 14.87
R/R ₀	0.999 ± 0.030	0.994 ± 0.032	1.010 ± 0.026	0.991 ± 0.036	1.017 ± 0.031	1.002 ± 0.027
KIE ₂	1.002 ± 0.016	1.006 ± 0.016	0.989 ± 0.016	1.010 ± 0.017	0.982 ± 0.017	0.998 ± 0.016
KIE _{avg}	1.008 ± 0.044	1.034 ± 0.044	1.015 ± 0.040	1.014 ± 0.043	0.992 ± 0.045	1.013 ± 0.042

Table B.1: Details from KIE measurements and statistical analysis.

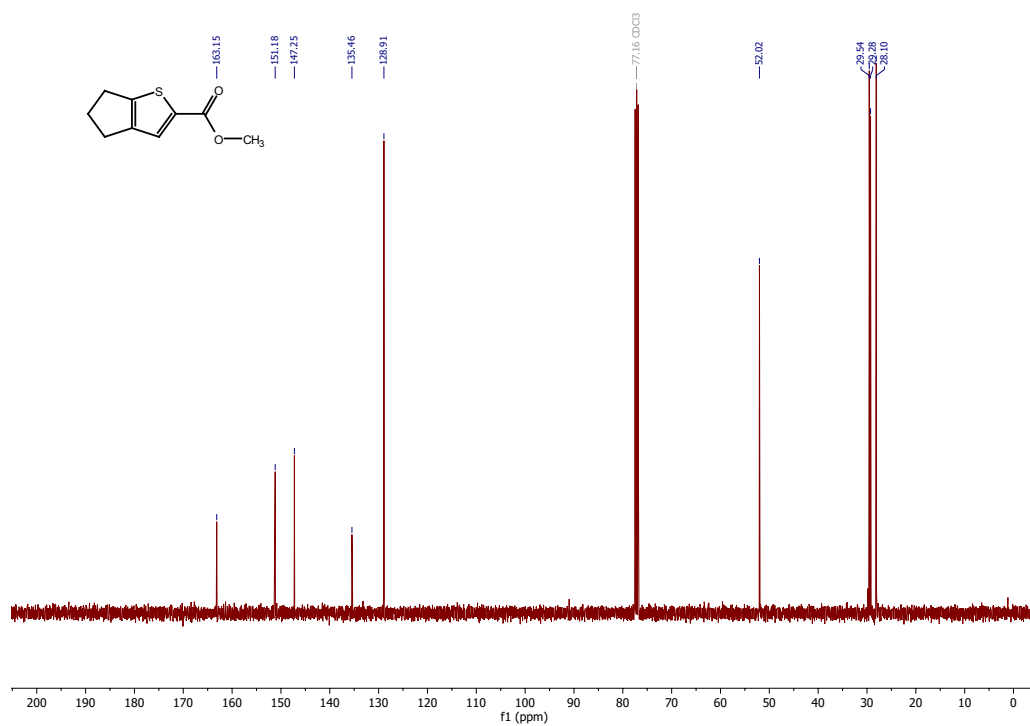
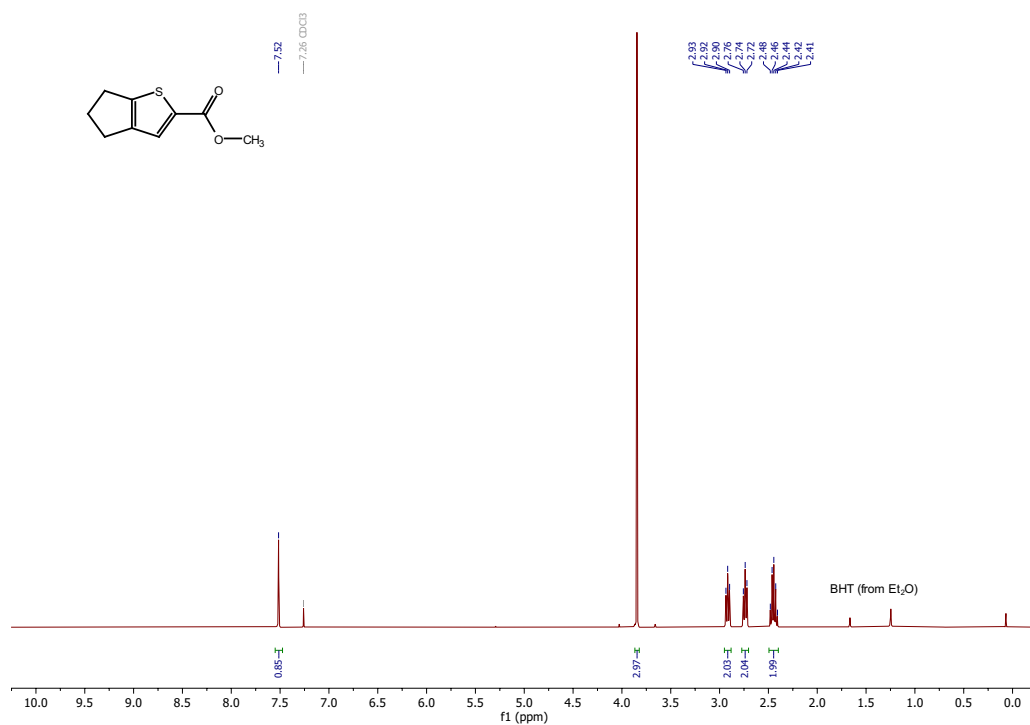
C

NMR Spectra

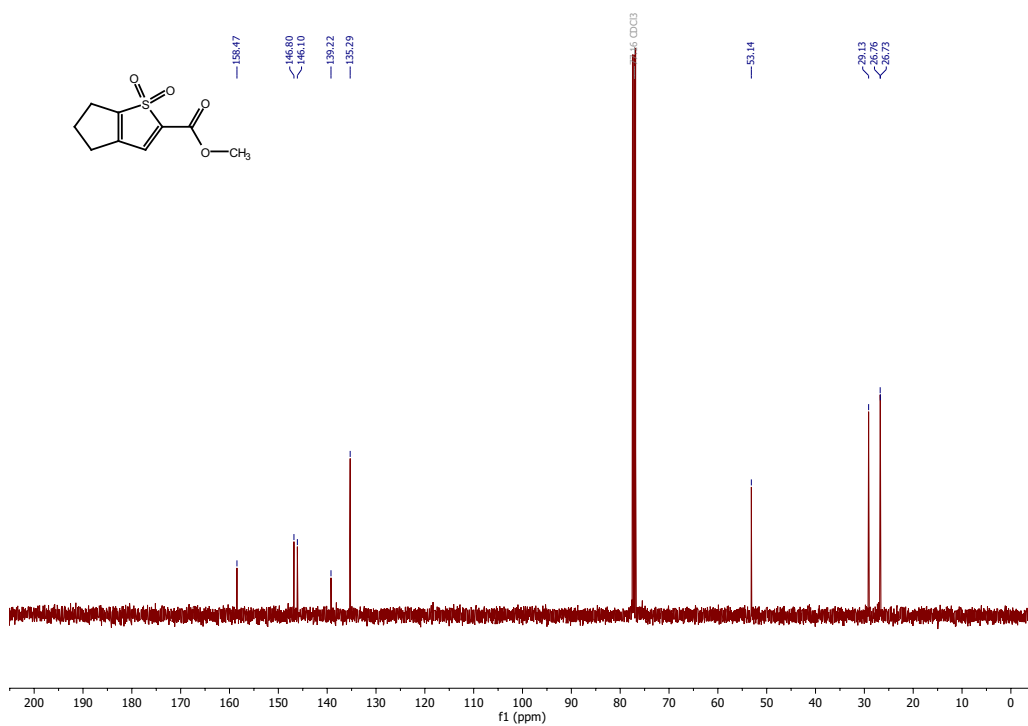
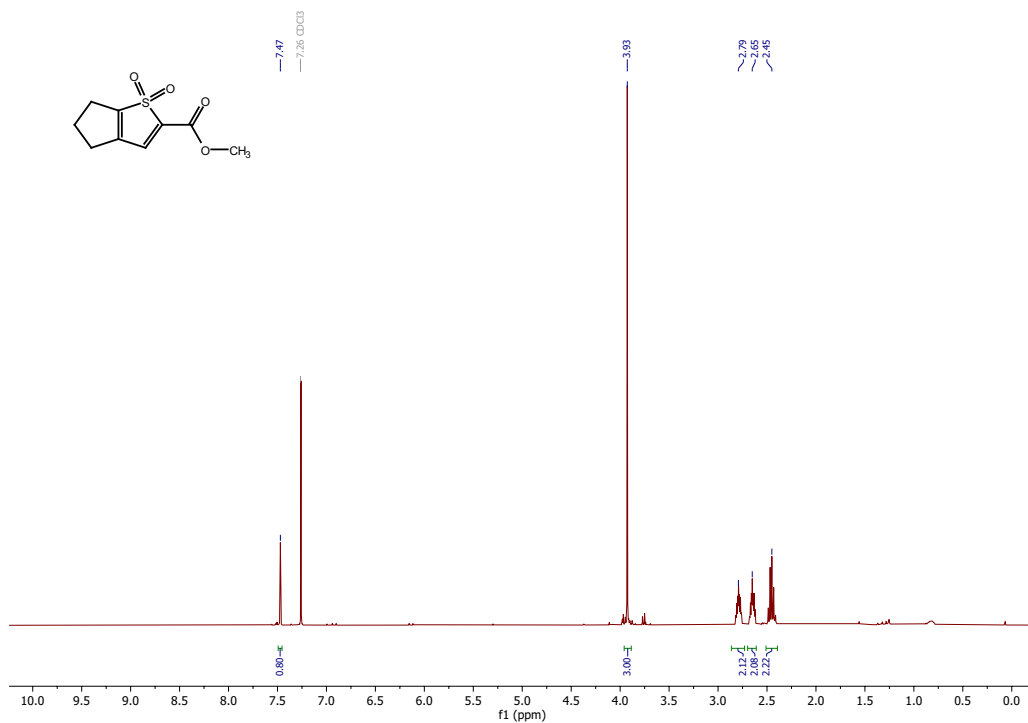
2-chlorocyclopent-1-ene-1-carbaldehyde, 5
 ^1H NMR (CDCl_3 , 200 MHz, crude)



Methyl 5,6-dihydro-4*H*-cyclopenta[*b*]thiophene-2-carboxylate.
 ^1H NMR (CDCl_3 , 400 MHz); ^{13}C NMR (CDCl_3 , 101 MHz)



Methyl 5,6-dihydro-4*H*-cyclopenta[*b*]thiophene-2-carboxylate 1,1-dioxide
 ^1H NMR (CDCl_3 , 400 MHz); ^{13}C NMR (CDCl_3 , 101 MHz)



Methyl 7-methyl-6-(2-(methylthio)-2-oxoethyl)-2,3-dihydro-1*H*-indene-5-carboxylate ^1H NMR (CDCl_3 , 600 MHz); ^{13}C NMR (CDCl_3 , 151 MHz)