Computational Chemistry Laboratory

Luke Nambi Mohanam, Sree Ganesh Balasubramani, Filipp Furche

Winter 2024
Version 4.0

Acknowledgments

Some chapters are based on the Chem 137 Lab manual, which was co-authored by Guo Chen, Matthew M. Agee, Erik P. Almaraz, Ruben Magaña, and Pei-Hsuan Yang.

This material is based upon work supported by the National Science Foundation under CHE-1464828, CHE-1800431, and CHE-2102568.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Chapter 2

Vibrational Spectroscopy of Diimine

2.1 Introduction

Diimine, N_2H_2, the simplest azo compound, occurs in a *cis* and *trans* form. Experimentally, diimine has been characterized by vibrational spectroscopy in solid matrices. This experiment aims to show how DFT calculations may be used to assign experimental spectra and distinguish quantitatively between the two isomers.

2.2 Background

Review the point group symmetry of *cis*- and *trans*-diimine, as well as selection rules for infrared and Raman spectra.

Assignment P 2.1. Point Group Symmetry of Diimine Isomers

1. Draw the Lewis structures of *cis*-, *trans*- and *iso*-diimine in their electronic ground states, and determine the molecular point group symmetry. (3)
2. Does any of the isomers have a permanent dipole moment? What does this imply for rotational spectroscopy? (2)
3. What energetic ordering of the isomers do you expect? Why? (2)

2.3 Experiment

Assignment 2.1. Vibrational spectra of Diimine (22+4*)

Can *cis*- and *trans*-diimine distinguished by vibrational spectroscopy?

1. Obtain suitable initial structures for the two isomers of diimine, by specifying the molecular point group and entering the coordinates of symmetry-unique atoms. (2)

2. Optimize the structure of the two isomers using the TPSS functional with m4 grids and polarized triple-ζ valence basis sets with polarization functions (def2-TZVP). Utilize the RI-*J* approximation ($rij) to speed up the calculations. The SCF convergence threshold ($scfconv) should be set to 7. (6)

3. Tabulate the bond distances and compare them to the CCSD(T) results from Ref. [1].(2)

4. Perform an analytical force constant calculation and tabulate the resulting normal modes, their harmonic frequencies, as well as computed IR and Raman intensities. Plot the IR and Raman spectra with a reasonable line broadening and overlay them for the *cis* and *trans* isomers to show the differences (4).

5. Compare the predicted spectra with the experimental spectra in Ref. [1]. (4*)

6. Write a laboratory report containing a brief introduction motivating the experiment, a computational details section, a results and discussions section and a conclusion section. The research question should be answered clearly in the conclusion (8)

Chapter 3

Structures and Spectroscopy of Water-Halide Clusters

3.1 Introduction

The physical and chemical properties of water-halide solutions depend strongly on dimensionality: Compared to the bulk solution, certain ions may have higher concentrations at or close to the surface at interfaces and in other confined environments [2]. This has important consequences for the structure and properties of low-dimensional ion solutions. For example, the selective enrichment of Br^- at the surface of sea-salt aerosol particles explains the disproportionately large role of bromide ions in tropospheric halogen chemistry [3].

Cluster models of water-halide solutions have a well-defined chemical composition and can be studied by high-resolution spectroscopy and electronic structure calculations. Charged clusters may be generated and separated using mass spectrometry, and characterized by photoelectron spectroscopy as well as UV- and IR-spectrocopy. The key to highly resolved UV- and IR spectra of cluster ions in the gas phase is action spectroscopy: One or several weakly bound inert gas atoms such as Ar are attached to the cluster ions, and the absorption cross section is determined by monitoring dissociation of the weakly bound complex [4].

The goal of this experiment is the determinantion of minimum structures and relative energies of $Cl^-\cdot(H_2O)_3$ clusters. The thus obtained structures will be validated by comparison of computed IR spectra to experimental spectra.

Additional background reading: Schwabl [5]; Tuckerman [6]; Allen and Tildesley [7]; Wilson, Decius, Cross [8]

3.2 Statement of the Model

Our experiment is based on the following key assumptions:

- The spectroscopic properties of the water-halide clusters may be described by a classical grand canonical ensemble.
- All low-lying cluster structures are sampled by *Ab initio* molecular dynamics simulations (AIMD) with sudden quenching.
- The nonrelativistic Born-Oppenheimer ground state provides an adequate approximation for the electronic properties.
- The minimum structures, their relative energies, and nuclear vibrational spectra are well described by hybrid DFT.
- The harmonic oscillator approximation is accurate enough to distinguish IR spectra of different cluster structures.

3.3 Background

3.3.1 Ensembles

Within the canonical ensemble, the high-temperature limit of the statistical operator is

$$\hat{W} = \frac{e^{-\beta \hat{H}}}{Z}, \tag{3.1}$$

where $\beta = 1/(kT)$, k is the Boltzmann constant, T is the temperature, $\hat{H}$ is the molecular Hamiltonian, and

$$Z = \langle e^{-\beta \hat{H}} \rangle \tag{3.2}$$

is the canonical partition function; brakets are used to denote the trace operation.

Assignment P 3.1. Ensembles

1. Define the terms microcanonical and canonical ensemble. What observables are conserved in these ensembles? (2)
2. Describe qualitatively the principle underlying the derivation of the statistical operator in Eq. (3.1). (3)
3. Why is it reasonable to use the high temperature limit for water-halide clusters? (1)

3.3.2 Global Minimum Search by AIMD

In AIMD, the nuclei are approximated by classical particles moving on the (ground state) Born-Oppenheimer potential energy surface obtained from *ab initio* calculations. The classical equations of motion are integrated using time propagation methods, e.g., the leapfrog Verlet algorithm: The nuclear momenta $\mathbf{P}$ and positions $\mathbf{Q}$ at the $i + 1/2$-th and i-th time step are computed from

$$\mathbf{P}(t_{i+1/2}) = \mathbf{P}(t_{i-1/2}) + \mathbf{F}(t_i)\delta t, \tag{3.3a}$$

$$\mathbf{Q}(t_{i+1}) = \mathbf{Q}(t_i) + \mathbf{M}^{-1}\mathbf{P}(t_{i+1/2}), \tag{3.3b}$$

for fixed time-step δt, $t_{i+1} = t_i + \delta t, \quad i = 0, 1, 2, \ldots$. The initial positions of the nuclei at t_0 and their initial momenta are specified as input. $\mathbf{M}$ is the mass tensor, which is diagonal in Cartesian coordinates. The force on the classical nuclei is obtained from the nuclear gradient of the electronic ground state potential energy surface $E(\mathbf{Q})$,

$$E(\mathbf{Q}) = -\nabla_{\mathbf{Q}} E(\mathbf{Q}) = -\frac{dE(\mathbf{Q})}{d\mathbf{Q}}. \tag{3.4}$$

The nuclear potential energy surface and its gradient are only required locally at a given position $\mathbf{Q}$.

For systems with relatively few degrees of freedom that relax on similar time-scales, free dynamics with sudden quenching (FDSQ) is an effective technique to search for the global potential energy minimum. In FDSQ, several nuclear trajectories are started from an initial structure with randomly assigned initial momenta. In free dynamics, no external fields are present, and the total energy remains constant. The initial kinetic energy of the nuclei is often expressed in terms of a pseudo-temperature $T = 2/(3N_n)E^{\text{kin}}$, where N_n is the number of nuclei and E^{kin} denotes the nuclear kinetic energy, and should be chosen large enough to overcome isomerization barriers quickly, but small enough to prevent undesirable side-reactions such as fragmentation. The nuclear potential energy of an FDSQ trajectory is used to detect local minima, which are located by "sudden quenching", i.e., setting the nuclear kinetic energ to zero. In practice, one often performs geometry optimization starting from the local minima determined by the free dynamics simulation.

There is no guarantee that finite-time FDSQ simulations will detect the global minimum and/or other low-energy minima, but the likelihood increases with simulation time.

Assignment P 3.2. Global Minimum Search by AIMD

1. Why is it not strictly correct to refer to a nuclear "temperature" in microcanonical "free dynamics" simulations? (1)
2. In what ways does geometry optimization mimick "sudden quenching", and how does it differ?(1)

3.3.3 Gibbs Free Energies

Under experimental conditions, different isomers of water-halide clusters are in chemical equilibrium. At finite temperature and constant pressure p, the composition of a mixture of different cluster isomers

is determined by the Gibbs free energy (chemical potential) of each species. The Gibbs free energy is related to the canonical partition function by

$$G = -kT \ln Z + pV, \tag{3.5}$$

where V denotes the volume. Within the rigid rotor – harmonic oscillator (RRHO) approximation, the canonical partition function separates into a translational (trans), rotational (rot), and a vibrational (vib) part,

$$Z = Z^{\text{trans}} Z^{\text{rot}} Z^{\text{vib}}. \tag{3.6}$$

The translational partition function is

$$Z^{\text{trans}} = V(2\pi MkT)^{3/2}, \tag{3.7}$$

where M is the molecular mass. Within the rigid rotor approximation, the rotational partition function in the high-temperature approximation for a non-linear molecule is

$$Z^{\text{rot}} = \frac{1}{\sigma}\sqrt{\frac{\pi(kT)^3}{|\mathbf{I}|}}, \tag{3.8}$$

where $\mathbf{I}$ is the inertial tensor and σ is a symmetry factor which is 2 for molecules with an inversion center and 1 for molecules without. The vibrational partition function within the Harmonic oscillator approximation is

$$Z^{\text{vib}} = \frac{1}{2}\Pi_{n=1}^{N^{\text{vib}}} \operatorname{csch}(\beta\Omega_n/2), \tag{3.9}$$

where Ω_n denotes the harmonic vibrational frequencies, and N^{vib} is the number of vibrational degrees of freedom.

Assignment P 3.3. Thermal Corrections

1. What state variables does G depend on? (1)
2. What is N^{vib} for a nonlinear molecule? (1)
3. Explain why electronic contributions to the partition function, Eq. (3.6), may be neglected for this experiment. (2)

3.3.4 Infrared Spectra

The harmonic vibrational frequencies Ω_n and the normal modes $\mathbf{X}$ of a polyatomic molecule are the solutions of the generalized eigenvalue problem

$$\mathbf{VX} = \mathbf{MX\Omega}^2, \tag{3.10}$$

where $\mathbf{\Omega}$ is diagonal with diagonal elements Ω_n.

$$V_{ij} = \frac{\partial^2 E(\mathbf{Q})}{\partial \mathbf{Q}_i \mathbf{Q}_j} \tag{3.11}$$

is the force or spring constant matrix, also called potential energy Hessian. It requires the calculation of all second derivatives of the potential energy with respect to nuclear displacements. The normal modes $\mathbf{X}$ are normalized according to

$$\mathbf{X}^\dagger \mathbf{M} \mathbf{X} = \mathbf{1}. \tag{3.12}$$

The infrared spectrum of a molecular cluster can be crudely interpreted using the harmonic approximation: The main peaks are given by the fundamental vibrational frequencies, and the peak intensity is proportional to the square of the vibrational transition dipole moment

$$\frac{1}{\mu_n \Omega_n} \left| \frac{d\mathbf{D}}{d\mathbf{X}_n} \right|^2, \tag{3.13}$$

which involves the normal mode derivative of the dipole moment $\mathbf{D}$ and the reduced mass μ_n.

Assignment P 3.4. Infrared Spectra

1. What are the main featuresyou expect to see in the infrared spectra of a small water-halide cluster? Why?(4)
2. The main computational cost in computing IR spectra goes into the calculation of the force constant matrix $\mathbf{V}$. With this information, suggest a resource-efficient method for investigating isotope substitution effects computationally. (2)

3.4 Experiment

Assignment 3.1. Structure of $(H_2O)_3Cl^-$

Compute the structures of stable isomers of $(H_2O)_3Cl^-$ and assign the global minimum structure by comparison to the measured IR [9] spectrum and the vertical detachment energy [10].

1. Perform a global minimum seach using AIMD simulations and sudden quenching. A useful functional/basis set combination for this purpose is PBE0/SVP. The initial nuclear kinetic energy should be approximately 10 times higher than a relevant barrier height in your search. It is very important to choose the initial conditions carefully to sample the potential energy surface effectively. Perform 10-20 optimizations starting in the vicinity of local minima on your MD trajectory over at least 1-2 ps. (6)
2. Refine the structures and relative energies of the 4-5 lowest energy isomers using def2-TZVP basis sets and Grimme's D3 dispersion correction. Compute harmonic vibrational spectra to ensure the resulting structures are minima.(4)

3. Tabulate the energies at 0K (relative to the global minimum energy, including zero-point vibrational corrections), and the free energies of the most stable isomers at 298 K within the RRHO approximation. (4)

4. Compute the vertical detachment energies of the thus obtained clusters and compare to experiment [10].(4)

5. Compare the computed harmonic vibrational spectra to the one from Ref. [9].(4)

6. Write a laboratory report containing a brief introduction motivating the experiment, a computational details section, a results section, and a conclusion section. The conclusion should contain a clear answer to the question what is the global minimum structure of $(H_2O)_3Cl^-$, and the answer should be supported by the results. (8)

Chapter 4

Fluoresence Spectroscopy of Luciferin Dyes

4.1 Introduction

Luciferin is the chromophore responsible for firefly luminescence.[11] Fireflies produce the luciferase enzyme which, in the presence of oxygen and ATP, decarboxylate luciferin to generate an oxyluciferin species in the excited state (see Ref. [12] for further background and references). The excited state rapidly decays to the ground state via the spontaneous emisson of a photon.

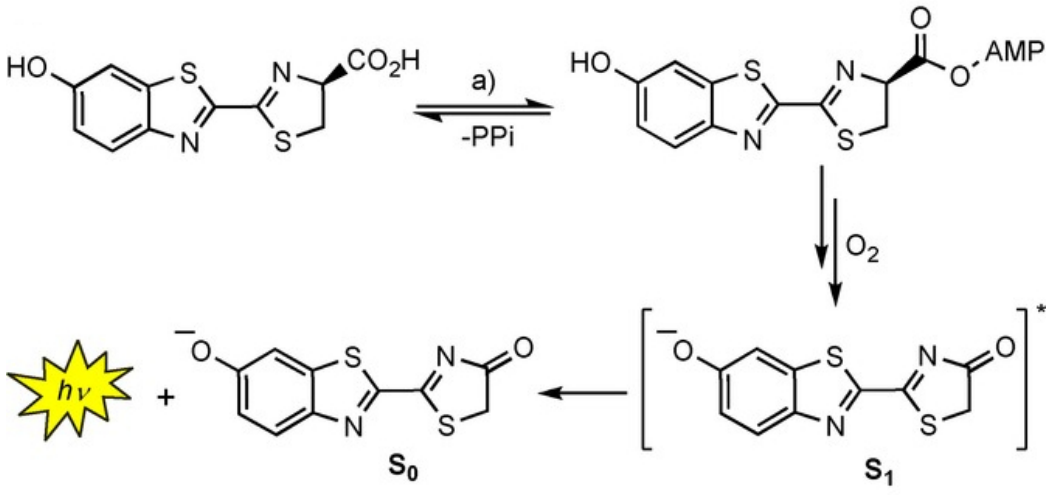

Figure 4.1: General scheme of luciferin luminescence [12]

The luciferase-induced luminescence of luciferin is the basis for bioluminescence imaging (BLI) using luciferin dyes: By attaching luciferase labels to, e.g., antibodies, specific cells can be labeled, and imaged by administration of luciferase. (a useful review of the technique is found in Ref. [13])

Importantly, luciferase not only reacts with wild type luciferin, but also a large number of synthetically functionalized luciferins. [14] Modern CCD detectors can resolve single photons emitted from luciferins, and fluorescence microscopy allows for high spatial resolution. The emission wavelength conveys additional information about the environment. For example, the luciferin emission wavelength shifts from approximately 562 nm to 620 nm as pH decreases from 7.5 to 5. [15] It has been proposed that a single species, the keto-(-1) form of oxy-luciferin, is responsible for both emissions while still bound to the luciferase enzyme. [15]

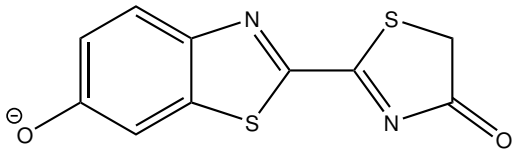

Figure 4.2: Keto-(-1) anion of oxyluciferin

Different explanations for the pH dependence of oxyluciferin emission exist: The pH change could result in conformational changes of luciferase, which in turn could affect the emission wavelength [15]. Alternatively, the excited-state properties of oxyluciferin could be directly affected by pH changes [16, 17].

The aim of this experiment is to simulate the fluorescence emission spectrum of the (neutral) keto tautomer of oxyluciferin, which is obtained by protonation of the keto-(-1) form under acidic conditions.

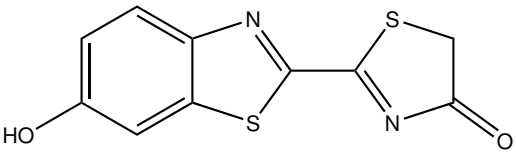

Figure 4.3: Keto tautomer of oxyluciferin

Additional background reading: Schwabl [5]; Levine [18]; Craig and Thirunamachandran [19]; Mukamel [20]

4.2 Background

4.2.1 The Born-Oppenheimer and Franck-Condon Approximations

Assignment P 4.1. Born-Oppenheimer (BO) Approximation

1. Review and state the key assumptions underlying the BO approximation. (1)
2. Why is the BO approximation not capable of describing internal conversion? (1)

The Franck-Condon (FC) approximation (sometimes called Franck-Condon principle) states that excitations and de-excitations between different BO states are "vertical", i.e., they can be considered to take place for fixed nuclear positions. [21, 22]

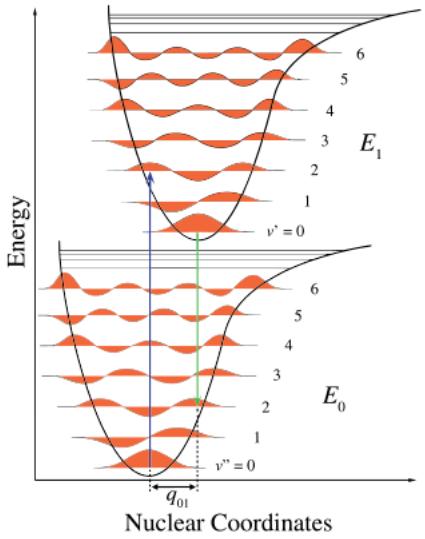

Figure 4.4: Schematic diagram representing the vertical excitation from ground electronic level to excited electronic level(blue line) and the spontaneous emission from the excited to ground state(green line) within the BO approximation[23]

Assignment P 4.2. Franck-Condon Principle

1. Show that, within the Franck-Condon approximation, the transition dipole moment between two molecular states is given by the product of the electronic transition dipole moment times the overlap of the nuclear wavefunctions ("Franck-Condon factor"). (2)

2. Consider a dipole-allowed optical excitation of a diatomic molecule. Assuming a slight increase in intermolecular distance upon excitation, sketch the vibrational fine structure of an electronic excitation band within the FC approximation. (2) *Hint:* Fig. 4.4

3. Now consider vertical deexcitation from the excited state minimum and sketch the resulting fine structure of the fluorescence spectrum in the same figure as above (2).

4.2.2 Emission Intensities

The rate of spontaneous one-photon emission from an excited state is given by the Einstein A coefficient (cgs units),

$$A_{21} = \frac{8\pi^2\Omega_{12}^2}{c^3}\frac{g_1}{g_2}f_{12}, \tag{4.1}$$

where c is the speed of light, f_{12} is the oscillator strength of the emission and Ω_{12} is the excitation energy. g_1 is the degeneracy of the ground state and g_2 is the degeneracy of the excited state.[24, 19]).

Assignment P 4.3. Einstein Coefficients

1. Explain why only the Einstein A coefficient is relevant for the oxyluciferin emission intensity. (1)

2. How is the A coefficent related to the radiative lifetime of excited oxyluciferin? (1)

4.3 Experiment

Assignment 4.1. Fine Structure of Oxyluciferin Fluorescence

Simulate the fluorescence emission spectrum of the keto tautomer of oxyluciferin in the gas phase, and compare it to the experimental spectrum.

1. Perform a ground state geometry optimization the molecules using DFT (PBE0/TZVP). Confirm that the obtained structure corresponds to a local minimum by computing the harmonic vibrational frequencies. (4)

2. Compute the first 3 vertical singlet excitation energies of the molecules using TDDFT, and verify that S_1 is the optically bright state. Assign the state by analyzing the orbital contributions to the transition density (4).

3. Perform an excited state geometry optimization of the S_1 state of the keto form using TDDFT. Confirm that the converged structures correspond to a local minima by computing numerical vibrational frequencies. (4)

4. Determine the S_1 vertical emission energies, for the keto form, and compare them to the maximum of the experimental fluorescence spectrum from Refs. [15, 17]. (2)

5. Determine the Einstein *A* coefficient for spontaneous emission for the keto form. (2)

6. Using the `radless` tool, simulate the fine structure of the emission spectrum, and compare your result to the literature [16]. (8)

7. Write a lab report containing a brief introduction motivating the experiment, a computational details section, a results section and a conclusion section. (8)

Chapter 5

Bimetallic Lanthanide and Actinide Metallocene Hydrides

5.1 Introduction

Metal-metal multiple bonds in transition metal complexes such as $[Re_2Cl_8]^{2-}$ have been known for a long time [25]. Traditionally, the valence electrons lanthanide and actinide cations are thought to mainly have f-character, based on the orbital energies of gaseous atoms. Thus, the expectation for these valence electrons is that they do not participate in bonding due to the limited radial extent of these f-orbitals, thus metal-metal bonding is unlikely.

Recently, Ln^{2+} complexes with novel $4f^n5d^1$ ground state electronic configurations were synthesized, in contrast to the traditional $4f^{n+1}$ ground state electronic configuration. [26] This was possible as ligand field splitting lowers the energy of the d-orbitals relative to the f-orbitals. Besides being highly colored, these novel electronic configurations suggest the possibility of forming of bimetallic lanthanide and actinide complexes with d-d bonding.

The rare earth metals Sc and Y are chemically similar to the lanthanides, even though they do not have occupied f-orbitals. A DFT study suggested the possibility that the reduction of the rare-earth $[Cp'_2Ln(\mu-H)]_2$ complex can lead to stable metal-metal bonding [27, 28]. Since then, the possibility of Ln-Ln bonding in organometallic complexes has beenn experimentally confirmed [29].

Assignment P 5.1. Lanthanide Electronic Configurations

1. Discuss the relative energies of the $4f$ versus $5d$ atomic orbtails for Sm atom. Qualitatively draw the radial distribution functions for these orbitals. How does the energetic separation of the $4f$ and $5d$ orbitals depend on the oxidation state? (2)

2. Predict the splitting of the $5d$ shell in a trigonal planar ligand field (i.e. D_{3h} symmetry). (2)

5.2 Background

5.2.1 Localized Orbitals

Orbital localization is a unitary transformation of occupied orbitals yielding a set of orthonormal orbitals that are maximally localized by some criterion. For example, Edmiston-Ruedenberg localization[30] maximizes the Coulomb self-energy, whereas Boys[31] and Pipek-Mezey[32] localization maximize orbital dipole moments.

Hund's localization condition can be used to determine whether "perfect" localization of all occupied orbitals is possible.

Assignment P 5.2. Orbital Localization

How does orbital localization change the following quantities? (3)

1. The total energy of a molecule
2. The molecular structure
3. Orbital energies

Assignment P 5.3. Orbital Localization in Hypovalent Compounds

What do you expect to see when applying orbital localization to a hypovalent compound such as cyclopropylium cation? (1)

5.3 Experiment

Assignment 5.1. Multicenter Bonds in B_2H_6

1. Perform a DFT calculation on the B_2H_6 molecule in D_{2h} symmetry using def2-SVP basis sets for all atoms and the TPSSh functional. Plot and assign the resulting canonical KS MOs. (3)
2. Lower the symmetry to C_1 and perform a Boys localization. Plot and assign the resulting MOs. (3)

Assignment 5.2. Reductive chemistry of $[(Cp)_2Y(\mu - H)]_2$

This experiment investigate whether there is Y-Y bonding in the recently synthesized bimetallic rare-earth complex $[(Cp)_2Y(\mu - H)]_2$ and its monoanion.

1. Perform ground state structure optimizations for $[(Cp)_2Y(\mu-H)]_2$ and its monoanion complex using DFT (TPSSh-D3/SVP) in THF solution (COSMO). *Note:* Try different spin states. (8)

2. Tabulate the distance between the metal centers in the complex, and compare the result to X-ray data for the neutral compound [28]. (2)

3. Do you expect the neutral complex and/or the anion to be EPR active? Why (not)? (2)

4. Visualize the frontier orbitals of the monoanion and discuss whether they suggest metal-metal bonding. (4)

5. Propose a minimal molecular orbital model of the electronic structure of the Yμ-H_2Y unit and compare to the Bμ-H_2B unit in diborane. (4)

6. Write a lab report containing a brief introduction motivating the experiment, a computational details section, a results section, and a conclusions section. (8)

Chapter 6

Stereoselective Addition of Tertiary Carbon Radicals

6.1 Introduction

Carbon radicals are playing an increasingly important role in organic chemical synthesis [33]. For example, radical coupling and addition reactions are viable routes for the formation of C-C, C-N, C-S and C-X (X: halogens) bonds [34, 35]. Despite their reputation, radical reactions can be very selective, especially when radicals are stabilized sterically and/or electronically. From transition state theory for these irreversible, kinetically driven reaction mechanisms, the difference in the activation energies between the transition states controls the selectivity of the reaction.

A particularly important application of the stereo-selectivity is carbon-carbon bond formation reactions leading to asymmetric quaternary carbon centers, which occur in many natural products and whose synthesis are notoriously difficult to achieve [36].

However, the selectivity of that rate-determining step is difficult to predict using simple reactivity theories, particularly if the selectivity depends on a subtle balance of noncovalent and steric interactions. Calculations of the transition state and the activation energy, including such noncovalent interactions, are thus valuable for their insight into the relative activation energies of transition states(and hence the selectivity of the reaction). The importance of noncovalent interactions also results in a non-trivial reaction coordinate, hence a calculation to optimize a transition state is more cost-effective than calculations to optimize a possible reaction pathway.

The goal of this experiment is investigate the stereoselectivity of tertiary carbon radical addition reactions from calculations of the rate-determining transition states of different reaction pathways.

6.2 Background

6.2.1 Transition state theory

Transition state theory (TST) is used to qualitatively explain reaction rates for elementary chemical reactions[37]. A general bimolecular reaction between reactants A and B can be represented within simplified TST as

$$[A + B] \overset{k_1}{\rightleftharpoons} X^{\neq} \overset{k_2}{\longrightarrow} C \tag{6.1}$$

where $X^{\neq}$ is the activated complex or the transition state and C is the product, $K^{\neq}$ is the (quasi-)equilibrium constant between the reactants and $X^{\neq}$, k_1 is the rate constant for the formation of $X^{\neq}$ and k_2 is the rate constant for the formation of the products.

The basic assumptions of classical transition state theory are:

- The reactants and the activated complex are in quasi-equilibrium with each other.
- There exists a hypersurface in the potential energy manifold (or the phase space) that divides the space into a reactant region and a product region.
- Trajectories passing through this “dividing surface" going towards the product region and originating from the reactants region will be thermalized or captured in a product state before reaching this surface again. (Often simplified as “Trajectories passing through the dividing surface do not return to the dividing surface.")

Assignment P 6.1. Limitations of transition state theory

1. Is the activated complex equivalent to the “dividing surface"? (2)
2. Give an example of a chemical reactions where TST fails, and explain why. (2)

Assignment P 6.2. Eyring Equation

According to the Eyring equation, the rate constant k_1 can be computed as [38],

$$k_1 = \left(\frac{k_B T}{h}\right) \bar{K}^{\neq} \tag{6.2}$$

where k_B is the Boltzmann constant and h is the Planck constant.

1. Derive an expression for k_1 in terms of the change in Gibbs free energy of activation, $\Delta G^{\neq}$, for the quasi-equilibrium reaction in Eq. 6.1. (2)
2. Consider two competing reactions of the same reactants passing through two different (diastereomeric) transition states. Derive an expression for the relative yield of the two resulting products in terms of the difference in Gibbs free energies of activation, sometimes called $\Delta\Delta G^{\neq}$, under the assumption of kinetic control. (2)

6.2.2 Barton Esters and Radical Reactions

In 1983 Barton and coworkers[39] demonstrated a method to efficiently generate carbon radicals by decarboxylation of carboxylic acid esters ("Barton esters"). These radicals can then electrophilically attack alkenes, forming a new C−*C* bond. (The resulting radical is quenched with a hydrogen donor to yield the product.)

The products of such radical coupling reactions have varying degrees of stereoselectivity. [34] For example, for the Barton reaction in Fig. 6.2.2, the experimentally observed relative syn:anti yield is 1:3.5[40].

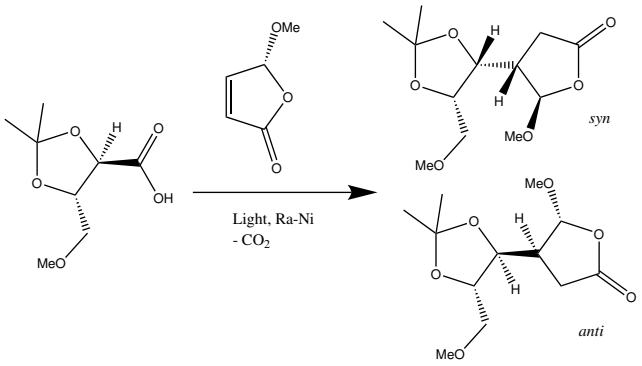

Figure 6.1: The Barton reaction investigated in this experiment [40].

Assignment P 6.3. Lewis structures and their limitations

1. Draw the Lewis structure of the reacting radical for the reaction in Fig. 6.2.2. (1)
2. Draw structures for the transition states leading to the syn and the anti products. (2)
3. Why is the syn/anti selectivity difficult to predict without a calculation?

6.3 Experiment

Assignment 6.1. Stereoselectivity of Radical Addition

Investigate the stereoselectivity of the Barton reaction in Fig. 6.2.2 using transition state theory. Consider only the rate-determining radical addition step.

1. Optimize the ground state structures for reasonable reactant complexes using DFT (TPSSh-D4/def2-SVP/dichloromethane solvent using COSMO). Verify that these structures are local

minima on the potential energy surface. *Hint:* Initial structures in Tables D.6 and D.7 can be used. (6)

2. Optimize the structures of the products in the same way, and verify that they are local minima. (4)

3. Using a localized internal bond stretching coordinate along the nascent C-C bond as approximate reaction coordinate, perform a "potential energy scan", i.e., constrained geometry optimizations for values of the reaction coordinate interpolating between the reactants and the products. Plot the resulting energy profile and graphically locate the approximate location of the transition state for both the anti and the syn pathways. (8)

4. Perform a force constant calculation using the thus obtained approximate transition state structure to ensure that there is exactly one imaginary vibrational frequency. Then start an *unconstrained* transition state optimization from this structure, using your force constant matrix as input Hessian. Perform a final force constant calculation on the converged transition state to ensure that the resulting structure is at a saddle point. (6)

5. Using the above results, compute the free energies of activation for the anti and syn mechanisms, and compute the expected theoretical syn:anti ratio. Compare to the experimental result of 1:3.5 [40]. (4)

6. Write a lab report containing a brief introduction motivating the experiment, a computational details section, a results section, and a conclusions section. (8)

Chapter 7

Photocatalytic Reactivity of TiO_2 Clusters

7.1 Introduction

Photocatalysts transform electronic excitation into chemical energy. Heterogeneous photocatalysis has many important applications, from artificial photosynthesis and solar fuel generation to UV activated water purification.[41] Titania (TiO_2) is a prototype semiconductor photocatalyst that is abundant, inexpensive, and nontoxic, although pure TiO_2 photocatalysts have poor quantum efficiency.[41] In this experiment, you will explore radiationless deactivation channels of a small model for a TiO_2 photocatalyst using nonadiabatic molecular dynamics.

Additional background reading: Muuronen et al.[42]

7.2 Background

Nonadiabatic molecular dynamics (NAMD) simulations capture effects beyond the Born-Oppenheimer (BO) approximation. In the semi-classical approximation, an electronic wavefunction, $|\Phi(t);\mathbf{R}\rangle$, is propagated together with nuclear positions ($\mathbf{R}$) and velocities. The electronic wavefunction may be expanded in the basis of time-independent BO eigenstates,

$$|\Phi(t);\mathbf{R}\rangle = \sum_k C_k(t)|\phi_k;\mathbf{R}\rangle, \tag{7.1}$$

where the electronic wavefunction and the BO states ϕ_k, depend parametrically on the nuclear position, $\mathbf{R}$. The electronic wavefunction evolves according to

$$\mathrm{i}\dot{\mathbf{C}} = (\mathbf{H} - \mathrm{i}\mathbf{W})\mathbf{C}, \tag{7.2}$$

where

$$W_{nm} = \langle \phi_n; \mathbf{R} | \frac{\mathrm{d}}{\mathrm{d}\mathbf{R}} | \phi_m; \mathbf{R} \rangle \cdot \dot{\mathbf{R}} \tag{7.3}$$

is the nonadiabatic coupling and

$$d_{nm} = \langle \phi_n; R | \frac{\mathrm{d}}{\mathrm{d}R} | \phi_m; R \rangle \tag{7.4}$$

is the derivative coupling that represents coupling between two electronic states induced by nuclear motion.

Assignment P 7.1. What is the quantum mechanical nuclear momentum and kinetic energy operators in the nuclear position representation?

Assignment P 7.2. Derive (7.2) from the time-dependent electronic Schrödinger equation.

Within the fewest switches surface hopping (FSSH) algorithm proposed by Tully[43], the nonadiabatic couplings give rise to sudden switches or hops between BO surfaces used to propagate the nuclei. The probablility of a hop between two adiabatic states i to j is given by

$$P_{i \to j} = \frac{dt}{a_{ii}} \left(2Im(a_{ji} V_{ij}) - 2Re(a_{ji} \dot{\mathbf{R}}.d_{ij}) \right) \tag{7.5}$$

where $a_{ij} = c_i^* c_j$ and $v_{ij} = \langle \phi_i | H | \phi_j \rangle \delta_{ij}$.

Assignment P 7.3. How can energy conservation be maintained when a nuclear hop takes place between two adiabatic states?

Assignment P 7.4. During the surface hop what is the direction of the change in the nuclear velocity? (Hint: The nuclear force can be computed as $\langle \phi_i | \nabla_{\mathbf{R}} H | \phi_j \rangle$)

Assignment 7.1. Excited state lifetime of $Ti(OH)_4$[44] using Fewest Switches Surface Hopping[43]

1. Run a ground state AIMD simulation of $Ti(OH)_4$ (using PBE0/SVP). Choose a reasonable initial temperature and run the simulation for few hundreds of femtoseconds. Collect 30 structures from this trajectory spaced out at equal time intervals. This will be the ensemble to be used in the following NAMD simulations.

2. Set up an NAMD simulation for each of the 30 structures starting at the S_1 state Franck-Condon geometry with random nuclear velocities. Determine the nonradiative excited state lifetime by log-linear regression of the excited state population

$$\rho_{11}^{\mathrm{av}}(t) = \frac{1}{N_{\mathrm{ensemble}}} \sum_{k \in \mathrm{ensemble}} n_k(t)$$

$$n_k(t) = \begin{cases} 1 & \text{if excited state is active in trajectory } k \text{ at time } t \\ 0 & \text{if ground state is active in trajectory } k \text{ at time } t \end{cases} .$$

3. Plot the S_1 and S_0 potential energy as a function of simulation time for representative trajectories.

4. According to your simulations, which nuclear degree(s) of freedom are mainly responsible for the excited state deactivation?

5. Write a lab report containing a brief introduction motivating the experiment, a computational details section, a results section, and a conclusions section. The conclusions section should answer the question above that is supported by your results.

Appendix A

Good Scientific Computing Practices

A.1 Principles

- Organization: Clarity about goals, scientific priorities
- Accountability: Keeping records
- Transparency: Methods and approach must be scientifically motivated and justified
- Efficiency: Achieving the scientific goals with the smallest necessary effort of (human and computational) resources
- Replicability and reproducibility: Results must be numerically replicable and repeatable by someone else under different conditions
- Critical awareness: Pro-actively checking for systematic errors, bias, human error; willingness to submit results to audit and criticism by others
- Reusability: Making results and code available to others

A.2 Project Design

- Scientific goals need to be as clearly and specifically stated as possible before computation and/or coding starts: What, why, how
- Scientific goals must inform the methodology and project set-up, not vice versa
- Hyotheses should be stated as crisply as possible
- Before the start of code development projects, working equations and/or pseudocode shold be completely written down

- The literature should be carefully reviewed to avoid duplication and giving due credit to others
- The project should be broken down in tasks (to-do list) with a rough time-line

A.3 Data Generation

- Create a project directory in an easily accessible location in your backed-up home directory, e.g., `projects/ln_compounds` and add the following files/directories:
  - A `README` file briefly describing the purpose of the project (text format)
  - `/data` directory for storage of raw data
  - `/doc` directory for notes, reports, and manuscript files
  - `/results` directory for spreadsheets and data extraction
  - `/src` directory for scripts, programs, source code
- Calculations are always performed on local scratch or work disks.
- Choose a directory structure that reflects the methodology. For example: `/work/phi /ln_compounds/thcpp3/tpss/svp`
- Don't overwrite raw data. Instead, make new directories for revised calculations. Use descriptive names, e.g., "`grid_m3`, `grid_3`, ... rather than "`run1`, `run2`, ...".
- Once the calculations have finished, copy or sync the results to your `/data` directory:
  - At the minimum, all input and the main output files need to be copied.
  - The output must contain the version of the code. Additional information on the platform/compiler options used is desirable. Commands used to run applicatins, run scripts (e.g. for batch jobs, queueing systems) need to be saved as well.
  - Intermediates should not be saved, particularly when they are in binary format and consume lots of disk space.
- Everything (except for raw data) should be version contolled. Simplest version control method is using a `CHANGELOG` file.

A.4 Data Analysis

- Transfer data from output to a spreadsheet by
  - copy and paste,
  - existing scripts,

  - custom scripts. If you use the latter, scripts must have a version and need to be tested and archived.

- Spreadsheets need to be well commented. Use tabs/sections to structure data. Each table needs a caption defining all data shown, column and row names must be descriptive and unique.
- Spreadsheet computing is error prone. Each formula needs to be accompanied by a descriptive comment and tested. Avoid doing complicated computations using a spreadsheet.
- Spreadsheet formats: .csv, xls (Google Sheets, OpenOffice Calc), Emacs org-mode
- Use specialized software for further analysis and/or plotting if necessary. Examples: gnuplot, R, jupyter notebook.
- Separate data generation and analysis work cycles
- The data need to be analyzed in a meaningful way: Give answers to the scientific questions, test hypotheses

A.5 Publication

- Timely, effective publication of results is critical for accountability and reuseability
- Computational manuscripts need to include a "computational details" section describing computational methods
- Coordinates of optimized structures, computed spectra and other key observables should be supplied as supporting information
- Additional data (e.g. MD trajectories) should be deposited in electronic repositories such as UC's escholarship and linked to the original publication.
- Published code repositories should be citeable by DOI
- Style [Strunk and White]: Concise and clear, determined by logical structure of the paper
- Graphical excellence [Tufte]:
  - Well-designed presentation of relevant data: Substance, statistics, design
  - Complex ideas communicated with clarity, precision, efficiency (least ink, smallest space, shortest time, maximum information)
  - Nearly always multivariate
  - Requires telling the truth about the data

A.6 Version Control

A.6.1 Basic Concepts

- All important data of a project are gathered in a repository (repo)
- A single version of project is defined at all times by the master branch or trunk
- Contributors may locally pull or check out versions of the master branch, make changes and additions, and then push or check in to the master branch
- If the version of the master branch has changed, the new changes need to be merged into the local version before pushing

A.6.2 Version Control Using Git

Git was started by Linus Torvalds for version control of the Linux kernel. It is a widely used version control system, and often combined with online repositories provided through services such as GitHub or Bitbucket. A number of useful online tutorials [45, 46] are available.

Basic git workflow:

- Creating the initial repo copy in the working directory:

  ```
  git clone <remote repo>
  ```

  The repo may either be specified as URL or ssh location.

- Adding new files:

  ```
  git add <files>
  ```

  Directories will be added recursively.

- Display local changes against the origin:

  ```
  git diff <files>
  ```

- Undoing a change:

  ```
  git checkout <file>
  ```

  will restore the file to the origin version,

  ```
  git checkout -p
  ```

  will let you revert changes interactively in batches.

- Status of the local repo relative to origin:

  ```
  git status
  ```

- View commit logs:

  ```
  git log
  ```

- Committing changes (local):

  ```
  git commit -a
  ```

- Communicating with the remote server:

  ```
  git pull
  ```

  to get the latest changes,

  ```
  git push
  ```

  to send your local commits to the repo.

Commit messages should start with a short single line summary written in the imperative, followed by a blank line and a more detailed description.

A.7 Typesetting Documents with LaTeX

A.7.1 Visual vs. Logical Design

LaTeXwas developed as a front-end for Donald Knuth's TeXtypesetting system in the 1980s. The key idea of LaTeXis to empower authors to focus on the structure and contents of a document rather than distracting them with formatting. Thus, LaTeXis text-based and does not support WYSIWYG.

A.7.2 LaTeXDocument and Data Structure

LaTeXdocuments consist of

1. a preamble that defines the document class, loads additional packages, changes defaults, and adds local definitions,
2. the main text document.

Important classes:

- article: Papers, lab reports, general purpose documents.
- report: Longer papers and small books, theses.
- book: Large books.
- letter: (Formal) letters
- beamer: Presentations
- minimal: Plain class for standalone figures etc.

A LaTeXdocument is a plain ASCII file with `.tex` extension. It must be compiled using `latex` to generate typset documents that may be printed or viewed:

- `latex document.tex` will generate the device independent `document.dvi`, which may be viewed (e.g. using `xdvi`) or converted to postscript (`dvips`) or pdf (`dvipdf`).
- Documents including bitmap graphics such as `.jpg` or `.png` may be directly converted to pdf using `pdflatex document.tex`.
- Documents containing references or bibliographies need to be compiled at least twice to resolve all dependencies.

A.7.3 LaTeXBasics

LaTeXhas three modes:

- Paragraph mode (default): Text is typeset in paragraphs and pages. Text is structured in words, separated by one or more spaces, and paragraphs, separated by one or more lines.
- LR mode: Text is typeset as a single line
- Math mode: All letters are assumed to be mathematical symbols, spaces are ignored.

LaTeXhas a number of environments to accomplish special tasks. They are enclosed by `\begin{myenvironment}` and `\end{myenvironment}`.

- Paragraph formatting environments: `center` or `\centering` (single paragraph), `flushleft` or `\raggedright`, `flushright` or `\raggedleft`, `verbatim` or `\verb`, `quote`, `theorem`
- List type environments: `list`, `itemize`, `enumerate`, `description`, `thebibliography`
- Equation-type environments: `equation`, `eqnarray`, `multline` (AMS), `align` (AMS)
- Table type environments: `tabular`, `tabbing`

- Float type environments: `figure`, `table`, `float` (float package)
- Array type environments (math mode): `array`, `matrix` (AMS), `pmatrix` (AMS), `bmatrix` (AMS), `vmatrix` (AMS)
- Some other environments: `figure`, `minipage`, `titlepage`, `verse`

Lengths, spaces, boxes:

1. If more or less space is needed, first try rubber lengths `\hfill`, `\vfill`, then `\bigskip` etc., then `\vspace{2in}` etc.
2. `\parbox` generates paragraph box, `\mbox` generates textbox
3. Try to change individual lengths and parameters as little as possible

Additional packages:

1. Include only necessary packages, avoid outdated or bloated ones
2. See LaTeXcompanions for documentation and best practices

A.7.4 Typesetting a Paper in LaTeX

1. Choose the `article` document class
2. Structure the document using `\section`, `\subsection` etc.
3. Insert tables and figures
4. Write methods section, followed by results, introduction, conclusions
5. Insert references using `\cite`
6. Write abstract
7. Proofread and re-write
8. Address any remaining formatting issues (overfull hboxes, float positions) at the very end

Appendix B

UNIX Basics

B.1 UNIX Philosophy

Characteristic UNIX features:

- Time-sharing multi-user multi-tasking environment
- Portable to many platforms
- Hierarchical file system, devices and streams are files
- Pipe-based command line interface
- Modular structure (kernel-shell model)

D. McIlroy (1978):

1. Make each program do one thing well. To do a new job, build afresh rather than complicate old programs by adding new "features".
2. Expect the output of every program to become the input to another, as yet unknown, program. Don't clutter output with extraneous information. Avoid stringently columnar or binary input formats. Don't insist on interactive input.
3. Design and build software, even operating systems, to be tried early, ideally within weeks. Don't hesitate to throw away the clumsy parts and rebuild them.
4. Use tools in preference to unskilled help to lighten a programming task, even if you have to detour to build the tools and expect to throw some of them out after you've finished using them.

B.2 Your UNIX User Account

You will receive a username (typically your UCINetID) and a password at the beginning of this course. The first thing to do in the lab is to log on a workstation and to change your password. Your new password should have at least 8 characters, include at least two numbers and two special characters. Your password should not contain your user name or any other words or personal information that can be easily found out (e.g. doing an internet search on your name), but you should be able to remember it. UNIX passwords are case-sensitive. To change your password, open up a shell terminal (clicking on the terminal icon on your application panel) and type:

```
passwd
```

You are personally responsible for your account including damage caused by abuse, so make sure your password is safe and do not share it with anyone. Please note that while you are free to use all installed software on the Modeling Facility workstations, it is illegal to copy or download installed proprietary software, even for your personal use.

Any computational project should be run on a local disc to avoid excessive network traffic. The default location is the so-called work disk. Create this directory and access it by doing the following:

```
mkdir work
cd work
```

Every UNIX system has at least one superuser (traditionally called root user). By definition, root can do anything.

B.3 Basic UNIX Shell Commands

Table B.1 lists some basic shell (bash) commands. For a more complete introduction see, e.g., the O'Reilly "in a Nutshell" books on UNIX, bash, and other topics.

The default shell you will be using is the Bourne-again shell (bash). bash has a number of very useful command-line features listed in Table B.2 you should familiarize yourself with.

B.4 The Vi and Emacs Editors

Vi is a basic UNIX text editor that is available in almost any UNIX environment. There are two basic modes in Vi, the command mode and the append mode. Table B.3 summarizes useful vi commands.

A more user-friendly editor is Emacs (and its X-windows version, Xemacs). Emacs has many options and modes, and Table B.4 displays only only a small number of commands you need to get started with Emacs. You can access Emacs tutorials by clicking on "Help".

Command	Description
pwd	Print working directory
ls	List all files in the current directory
ls -l	List all files in current directory with more details
cp file1 file2	Copy “file1" to “file2"
cp -r dir1 dir2	Recursively copy directory “dir1" to directory “dir2"
mkdir chemistry	Create directory “chemistry"
rmdir chemistry	Remove directory “chemistry"
cd chemistry/	Change working directory to “chemistry”
cd ..	Go to parent directory (..)
rm file1	Remove file1. Warning: once it’s gone, it’s gone
mv file1 file2	Rename “file1" to “file2". Warning: if file 2 already exists it will be lost
mv * chemistry/	Move all files into the directory chemistry
find * -name *.pdf -print	Find all .pdf files in all directories and their subdirectories
cat file1	Print contents of file1 on screen
less file1	Pager that allows you to scroll through file1 pressing <space>
grep “string" file1	Search file1 for “string” and print matching lines
which openoffice	Shows location of executable “openoffice”
man (command)	Shows manual page for “command”
bzip2 file, bunzip2 file.bz2	Compresses/Uncompresses ’file’
tar -vcjf dir.tar.bz2 *	Copy all files into compressed tape archive "dir.tar.bz2"
ps auwx	Print list of all running processes and their process IDs (PIDs)
kill <PID>	Kill (terminate) process
vi file1	Creates or opens file1 in the vi text editor
emacs file1	Creates or opens file1 in the emacs editor

Table B.1: Basic shell commands

Command	Description
<UP>, <DOWN>	Scroll through command history
<CTRL>-r	Search command history
<TAB>	Automatic completion
<CTRL>-c	Abort present command
<CTRL>-z	Stop present command
<CTRL>-a, <CTRL>-e	Jump to beginning/end of line
<CTRL>-d, <ALT>-d	Delete character/word
ls -l | less	Combine or pipe two commands (pipe is "|")
dscf > dscf.out	run program dscf and direct output to "dscf.out"
dscf > dscf.out 2>&1 or dscf &> dscf.out	run program dscf, direct error output to standard output, and standard output to "dscf.out"
dscf 2>&1 | tee dscf.out	run program dscf, direct stderr to stdout, and stdout both to screen and "dscf.out"
dscf &> /dev/null	run program dscf, discard any output
dscf >> dscf.out	run program dscf and append the output to "dscf.out"
define < define.input	run program define and redirects interactive input from file "define.input"
dscf &	Run program dscf in the background
nohup jobex &	Run program jobex in the background and continue running it even if the present shell is terminated
bg <PID>	Run (stopped) forground process in the background
fg <PID>	Run (stopped) background process in the foreground

Table B.2: Command-line features of bash

Command	Description
a	Enter append mode
i, I	Insert text before cursor position (goes into append mode)
<Esc>	Leave append mode
o,O	Open new line below/before the current line in append mode
x	Delete character under cursor
dd	Delete whole line
D	Delete line from cursos position to the end of the line
r	Replace character under cursor
R	Enter replace mode
:w (file1)	Save current contents (in file1)
:wq	Save file and exit vi
:q!	Exit without saving changes

Table B.3: Table: Basic Vi commands.

Command	Description
<CTRL>-x <CTRL>-a	Save file and exit emacs
<CTRL>-x <CTRL>-c	Exit emacs
<CTRL>-u	Undo
<CTRL>-k	Delete line from cursor on to the end of line
<CTRL>-<space>	Mark beginning of a selected region
<CTRL>-y	Paste region

Table B.4: Useful key-combinations for the emacs editor

There are a number of other text editors available. If you are used to Microsoft Word, try OpenOffice, which has word compatibility.

Another useful package is Open Babel [**babel**] which is able to read, write and convert over ninety chemical file formats. This is particularly useful if the user wishes to open a certain file with different packages. For example, users could use Open Babel when they wish convert a Protein Data Bank file (.pdb, standard format used by RCSB) that can normally be viewed with the VMD package to a XMOL molecule model file (XYZ, graphical coodinates) that can be opened with Avogadro. The users could use Open Babel to convert the PDB file to an XYZ file through using the terminal and then opening the files with the respective packages.

B.5 UNIX File Systems

Unix keeps all files (including directories) using what is called a file system. After logging into your account, you will be placed in your home directory. Here you can create files, directories, and subdirectories. Your home directory is “mounted" on all Modeling Facility workstations and physically resides on the Modeling Facility Server. Your home directory (and all subdirectories) are backed-up in regular intervals.

Important UNIX system directories:

/bin Generic binaries

/boot Location of startup files and kernel

/dev System devices (drives, RAM, tty, loop devices)

/etc System configuration

/home User home directories

/lib System (dynamic) libraries

/media Removable file systems (e.g. usb)

/mnt Temporary file systems mounted after boot

/opt Third-party programs

/proc Information about processes, system

/sbin Binaries for system administration

/root Home directory of root user

/tmp Scratch directory

/usr User-shared programs and files

/var Directory for various files, such as logs, mail, print spooling files

Disk allocation is very important especially when dealing with I/O intensive calculations. The command “df -h" will print out the system hard disk space usage. This shows how much space is available and how much is being used in your main locations. The command “du -h" will list the disk usage in the current directory.

Any computational project should be run on a local disc to avoid excessive network traffic. The default location is the so-called work disk. Create this directory and access it by doing the following:

```
mkdir work
cd work
```

Note that scratch disks are not backed up and may be lost in case of a hardware problem.

Appendix C

Introduction to High Performance Computing (HPC)

C.1 HPC Basics

HPC, as opposed to general purpose computing, uses special hard- and software to achieve high performance in scientific and technical applications.

Important hardware performance measure for scientific applications: Floating point operations per second (FLOPS) for theoretical peak performance,

$$\text{FLOPS} = \#\,\text{CPUs} \times \frac{\#\,\text{cores}}{\text{CPU}} \times \frac{\#\,\text{cycles}}{\text{s}} \times \frac{\#\,\text{FLOPS}}{\text{cycle}}. \tag{C.1}$$

However, performance in actual applications is often drastically lower than peak performance:

- I/O or memory bound algorithms
- Inefficient use of parallelism and vectorization

Hardware	FLOPS
Pentium P5, 75 MHz	8 M
Core i7 930	2 G
Xeon Haswell 3.8 GHz	500 G
Xeon Skylake	~ 1-3 T
Nvidia Tesla P100	~ 5 T
Top supercomputers	~ 100 p

Table C.1: Approximate FLOPS counts

Energy efficiency is becoming increasingly important for HPC. Modern hardware can reach efficiencies of $\sim$ 10 GFLOPS per Watt.

C.2 Parallelization

C.2.1 Efficiency Measures

Parallel speedup (ratio):

$$S(N) = \frac{T(N)}{T(1)}, \tag{C.2}$$

where $T(N)$ is the wall time observed for parallel execution with N workers.

Parallel efficiency:

$$E(N) = \frac{S(N)}{N} \tag{C.3}$$

Amdahl's Law: If f denotes the fraction of execution time spent in parallelized sections of code, then

$$S < \frac{1}{f + \frac{1-f}{N}} < \frac{1}{f}. \tag{C.4}$$

High parallel efficiency requires algorithms that

- balance load well between workers,
- minimize (communication) overhead and maximize cache efficiency.

C.2.2 Parallel Computing Paradigms

Shared-memory parallel computing: See Fig. C.1

- Advantage: Requires relatively minor modifications of serial code
- Disadvantage: Scalability can be limited

Distributed memory parallel computing: See Fig. C.2

- Advantage: Massive parallelism possible
- Disadvantage: Extensive code modification or specialized codes/algorithms needed; performance may depend on interconnect speed

Hierarchical parallelization: Combine distributed (coarse grain) and shared-memory (fine grain) parallelism

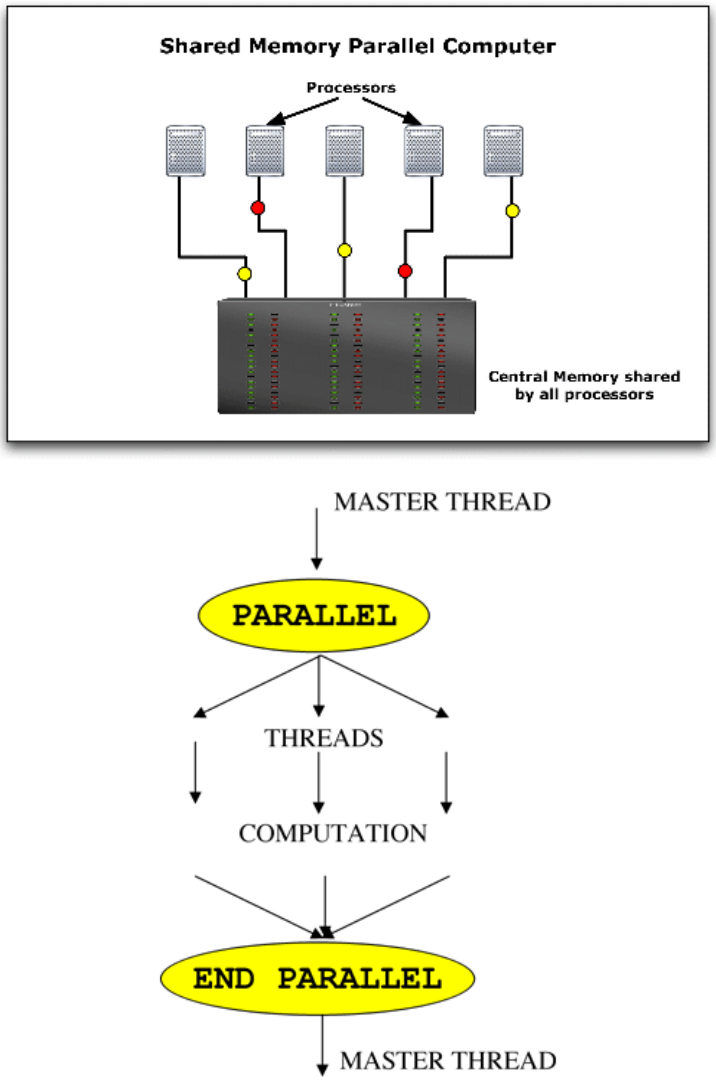

Figure C.1: Shared-memory parallel computing

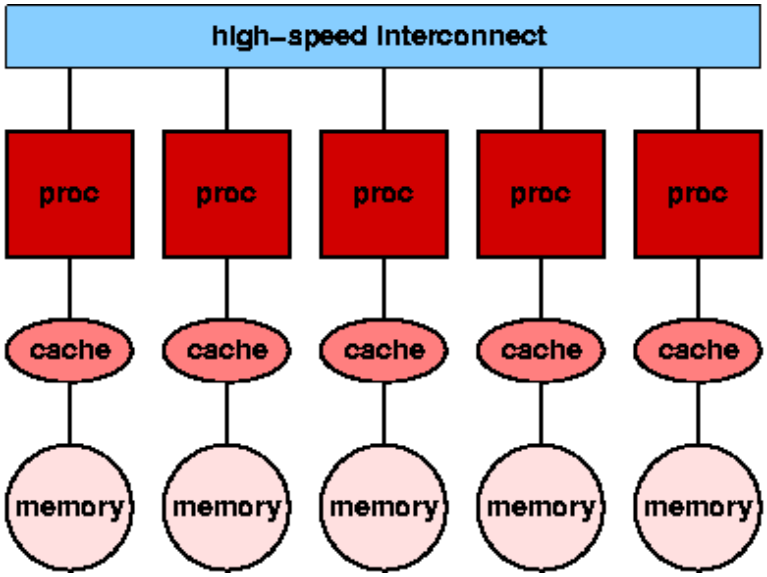

Figure C.2: Distributed memory parallel computing

Appendix D

Input Structures

Table D.1: Initial geometry of 5′-BrLuc

21			
O	-6.7891639	3.5566622	-0.0032265
C	-5.7123659	2.9455578	-0.0020496
C	-4.4263046	3.6087163	-0.0048395
C	-3.2568236	2.8885796	-0.0034073
C	-3.2398164	1.4495888	0.0009261
C	-4.4786221	0.7552349	0.0037855
C	-5.6411938	1.4692353	0.0023466
S	-1.6061438	3.4546740	-0.0061729
C	-1.0610640	1.7704864	-0.0014388
N	-2.0314378	0.8797553	0.0018765
C	0.3135098	1.4331837	-0.0013169
S	0.7500242	-0.2931721	0.0036376
N	1.2894669	2.3021346	-0.0046127
C	2.5322110	1.7265635	-0.0037090
C	2.4881321	0.1859963	0.0010092
H	3.0086956	-0.1874802	0.8932456
H	3.0074937	-0.1929026	-0.8896401
O	3.5923193	2.3188898	-0.0062399
H	-4.4424913	4.6989112	-0.0081132
Br	-7.2975886	0.5503267	0.0061108
H	-4.4737861	-0.3346915	0.0070588

Table D.2: Initial geometry of 7′-BrLuc

21			
O	-6.8117381	3.4595796	0.0029335
C	-5.7216478	2.8707448	0.0013385
C	-4.4348731	3.5452103	0.0045764
C	-3.2531015	2.8403188	0.0026474
C	-3.2340551	1.4008789	-0.0027133
C	-4.4737326	0.7034700	-0.0060358
C	-5.6422345	1.4048733	-0.0040951
S	-1.6157756	3.4241695	0.0058077
C	-1.0583751	1.7446135	-0.0001276
N	-2.0207710	0.8437951	-0.0040662
C	0.3182677	1.4165633	-0.0005116
S	0.7643403	-0.3082045	-0.0066373
N	1.2897944	2.2900858	0.0033428
C	2.5355001	1.7211731	0.0020153
C	2.4996733	0.1802397	-0.0037318
H	3.0212415	-0.1961605	0.8866843
H	3.0223923	-0.1895294	-0.8962514
O	3.5926833	2.3185410	0.0048960
Br	-4.4145662	5.4330358	0.0116693
H	-4.4573940	-0.3881333	-0.0100862
H	-6.6069278	0.8930852	-0.0065650

Table D.3: Initial geometry of $[(Cp)_2Y(\mu - H)]_2^{-1}$

44

Y	-0.5427143	0.0817929	-1.6460822
H	-1.0985778	-0.5809541	0.3335138
C	-1.7212416	1.7421958	-3.3148775
C	-2.7486827	1.3398723	-2.4073929
C	-2.4444571	1.8927943	-1.1296865
C	-1.2306305	2.6282408	-1.2410282
C	-0.7769303	2.5362584	-2.5938499
C	-0.4081738	-2.5548602	-2.0941625
C	0.9230371	-2.0707268	-2.2375921
C	0.9648327	-1.2173391	-3.3841215
C	-0.3482894	-1.1821481	-3.9468017
C	-1.1977417	-2.0044216	-3.1450914
Y	0.5535175	-0.0985302	1.6459379
H	1.1115166	0.5603398	-0.3346629
C	1.7020687	-1.7728529	3.3218380
C	2.7405148	-1.3779941	2.4236517
C	2.4419669	-1.9253885	1.1423262
C	1.2206848	-2.6501657	1.2420601
C	0.7568287	-2.5570038	2.5913609
C	0.4305645	2.5422882	2.0772683
C	-0.8998999	2.0619090	2.2387010
C	-0.9308409	1.2166511	3.3915877
C	0.3884043	1.1828453	3.9401896
C	1.2306305	1.9977361	3.1234391
H	-0.7689962	-3.2255880	-1.3104301
H	1.7687926	-2.3137241	-1.5932758
H	1.8520416	-0.7174405	-3.7827556
H	-0.6458030	-0.6405961	-4.8477112
H	3.0433503	-1.7973572	0.2390400
H	3.6177846	-0.7766575	2.6788686
H	-1.8144520	0.7212565	3.8036899
H	0.7844147	3.2063141	1.2847078
H	-1.7521230	2.3014709	1.6014969
H	0.6949230	0.6485721	4.8424660
H	-3.6228271	0.7304928	-2.6542300
H	-0.7333050	3.1743924	-0.4384809
H	-1.6804068	1.5062779	-4.3807120
H	-3.0367640	1.7604766	-0.2210742
H	0.7252602	-3.1902598	0.4343169
H	1.6543310	-1.5386160	4.3878072
H	-0.1383002	-3.0352882	2.9986306
H	2.2964228	2.1847539	3.2831427
H	0.1096838	3.0225835	-3.0101081
H	-2.2615946	-2.1911922	-3.3171236

Table D.4: Initial geometry of $[(Cp)_2Y(\mu - H)]_2$

44

Y	-0.5427143	0.0817929	-1.6460822
H	-1.0985778	-0.5809541	0.3335138
C	-1.7212416	1.7421958	-3.3148775
C	-2.7486827	1.3398723	-2.4073929
C	-2.4444571	1.8927943	-1.1296865
C	-1.2306305	2.6282408	-1.2410282
C	-0.7769303	2.5362584	-2.5938499
C	-0.4081738	-2.5548602	-2.0941625
C	0.9230371	-2.0707268	-2.2375921
C	0.9648327	-1.2173391	-3.3841215
C	-0.3482894	-1.1821481	-3.9468017
C	-1.1977417	-2.0044216	-3.1450914
Y	0.5535175	-0.0985302	1.6459379
H	1.1115166	0.5603398	-0.3346629
C	1.7020687	-1.7728529	3.3218380
C	2.7405148	-1.3779941	2.4236517
C	2.4419669	-1.9253885	1.1423262
C	1.2206848	-2.6501657	1.2420601
C	0.7568287	-2.5570038	2.5913609
C	0.4305645	2.5422882	2.0772683
C	-0.8998999	2.0619090	2.2387010
C	-0.9308409	1.2166511	3.3915877
C	0.3884043	1.1828453	3.9401896
C	1.2306305	1.9977361	3.1234391
H	-0.7689962	-3.2255880	-1.3104301
H	1.7687926	-2.3137241	-1.5932758
H	1.8520416	-0.7174405	-3.7827556
H	-0.6458030	-0.6405961	-4.8477112
H	3.0433503	-1.7973572	0.2390400
H	3.6177846	-0.7766575	2.6788686
H	-1.8144520	0.7212565	3.8036899
H	0.7844147	3.2063141	1.2847078
H	-1.7521230	2.3014709	1.6014969
H	0.6949230	0.6485721	4.8424660
H	-3.6228271	0.7304928	-2.6542300
H	-0.7333050	3.1743924	-0.4384809
H	-1.6804068	1.5062779	-4.3807120
H	-3.0367640	1.7604766	-0.2210742
H	0.7252602	-3.1902598	0.4343169
H	1.6543310	-1.5386160	4.3878072
H	-0.1383002	-3.0352882	2.9986306
H	2.2964228	2.1847539	3.2831427
H	0.1096838	3.0225835	-3.0101081
H	-2.2615946	-2.1911922	-3.3171236

Table D.5: Initial ground state geometry of $Ti(OH)_4$

9			
Ti	0	0	0.000176
O	0.408463	-1.412581	1.075608
O	-1.413303	-0.405653	-1.07524
H	-2.007516	-1.169882	-1.004803
O	-0.408463	1.412581	1.075608
H	-1.171167	2.007632	0.994781
O	1.413303	0.405653	-1.07524
H	2.007516	1.169882	-1.004803
H	1.171167	-2.007632	0.994781

Table D.6: Initial ground state geometry of (R)-methoxybutenolide

14			
C	2.2465546	-0.5280731	0.3939748
C	2.4833467	0.7793383	0.3807771
C	2.8706760	1.1753251	-0.9880774
O	2.8471987	0.0614729	-1.7801166
C	2.4459797	-1.0897256	-0.9812463
O	3.4304194	-2.0622708	-0.9758795
C	3.5354108	-2.7687692	-2.2238903
O	3.1646278	2.2631105	-1.4247210
H	2.4305656	1.4843232	1.1954783
H	4.2565591	-3.5658964	-2.0612544
H	3.8876467	-2.1029071	-3.0136463
H	2.5654413	-3.1924438	-2.4998968
H	1.9547916	-1.1506549	1.2269290
H	1.5219119	-1.4612392	-1.4376605

Table D.7: Initial grounsd state geometry of an acetonide radical

23

C	0.0568071	0.0553493	-0.0541491
O	-0.5898570	-0.6142435	-1.0573746
C	-1.1316343	0.3597692	-1.9843611
O	-0.9737061	1.6340939	-1.3537822
C	-0.3968816	1.4798820	-0.0471507
C	-1.4016311	1.8399085	1.0428969
O	-2.4483616	0.8859329	1.0623708
C	-3.4935937	1.2659563	1.9458490
C	-0.3180144	0.3366572	-3.2666729
C	-2.6027129	0.0483762	-2.1831685
H	0.4362818	2.1864382	0.0385355
H	0.7338164	0.5131942	-3.0386862
H	-0.6731800	1.1207475	-3.9367583
H	-0.4238728	-0.6273145	-3.7666713
H	-3.1084686	0.0926026	-1.2203358
H	-2.7238008	-0.9436220	-2.6213226
H	-3.0420486	0.7845959	-2.8576334
H	-0.8818413	1.8650547	2.0113552
H	-1.8019789	2.8431549	0.8439660
H	0.2399402	-0.5352085	0.8339454
H	-3.1309474	1.3356440	2.9785471
H	-3.9194739	2.2336590	1.6541896
H	-4.2618006	0.4961819	1.8830412

Bibliography

[1] Jan M. L. Martin and Peter R. Taylor. "Benchmark ab initio thermochemistry of the isomers of diimide, N2H2, using accurate computed structures and anharmonic force fields". In: *Mol. Phys.* 96.4 (1999), pp. 681–692. DOI: 10.1080/00268979909483004. eprint: https://doi.org/10.1080/00268979909483004. URL: https://doi.org/10.1080/00268979909483004.

[2] P. Jungwirth and D. J. Tobias. "Specific Ion Effects at the Air/Water Interface". In: *Chemical Reviews* 106.4 (2006), pp. 1259–1281. DOI: 10.1021/cr0403741.

[3] B. J. Finlayson-Pitts. "The Tropospheric Chemistry of Sea Salt: A Molecular-Level View of the Chemistry of NaCl and NaBr". In: *Chemical Reviews* 103.12 (2003), pp. 4801–4822. DOI: 10.1021/cr020653t.

[4] W. H. Robertson and M. A. Johnson. "Molecular aspects of halide ion hydration: the cluster approach". In: *Annual Review of Physical Chemistry* 54.1 (2003), pp. 173–213. DOI: 10.1146/annurev.physchem.54.011002.103801.

[5] F. Schwabl. *Statistical Mechanics*. 2nd ed. Springer, 2006. ISBN: 978-3540323433.

[6] *Statistical Mechanics: Theory and Molecular Simulation*. 2nd ed. Oxford Graduate Texts. Oxford University Press, 2023.

[7] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids*. 2nd ed. Oxford, 2017.

[8] E. B. Wison, J. C. Decius, and P. C. Cross. *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra*. Dover, 1980. ISBN: 987-0-486-63941-3.

[9] P. Ayotte, G. H. Weddle, and M. A. Johnson. "An infrared study of the competition between hydrogen-bond networking and ionic solvation: Halide-dependent distortions of the water trimer in the $X^-\cdot(H_2O)_3$, (X = Cl, Br, I) systems". In: *J. Chem. Phys.* 110 (1999), pp. 7129–7132. DOI: 10.1063/1.478616.

[10] G. Markovich et al. "Photoelectron spectroscopy of Cl^-, Br^-, and I^- solvated in water clusters". In: *J. Chem. Phys.* 101.9344–9353 (1994). DOI: 10.1063/1.467965.

[11] Joao Vieira, Luís Pinto da Silva, and Joaquim C.G. Esteves da Silva. "Advances in the knowledge of light emission by firefly luciferin and oxyluciferin". In: *J. Photochem. Photobio. B: Bio.* 117 (2012), pp. 33–39. ISSN: 1011-1344. DOI: https://doi.org/10.1016/j.jphotobiol.2012.08.017. URL: http://www.sciencedirect.com/science/article/pii/S1011134412001893.

[12] R. C. Steinhardt et al. "Brominated Luciferins Are Versatile Bioluminescent Probes". In: *ChemBioChem* 18 (2017), pp. 96–100. DOI: 10.1002/cbic.201600564.

[13] Miranda A. Paley and Jennifer A. Prescher. "Bioluminescence: a versatile technique for imaging cellular and molecular features". In: *Med. Chem. Commun.* 5 (2014), pp. 255–267. DOI: 10.1039/C3MD00288H. URL: http://dx.doi.org/10.1039/C3MD00288H.

[14] Krysten A. Jones et al. "Orthogonal Luciferase-Luciferin Pairs for Bioluminescence Imaging". In: *J. Am. Chem. Soc.* 139 (2017). PMID: 28106389, pp. 2351–2358. DOI: 10.1021/jacs.6b11737. eprint: https://doi.org/10.1021/jacs.6b11737. URL: https://doi.org/10.1021/jacs.6b11737.

[15] Luís Pinto da Silva and Joaquim C. G. Esteves da Silva. "Computational Investigation of the Effect of pH on the Color of Firefly Bioluminescence by DFT". In: *Chem. Phys. Chem.* 12 (2011), pp. 951–960. DOI: 10.1002/cphc.201000980. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cphc.201000980. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/cphc.201000980.

[16] Miyabi Hiyama et al. "Vibronic Structures in Absorption and Fluorescence Spectra of Firefly Oxyluciferin in Aqueous Solutions". In: *Photochem. Photobio.* 91.4 (2015), pp. 819–827. DOI: 10.1111/php.12463. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/php.12463. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/php.12463.

[17] Avisek Ghose et al. "Emission Properties of Oxyluciferin and Its Derivatives in Water: Revealing the Nature of the Emissive Species in Firefly Bioluminescence". In: *J. Phys. Chem. B* 119 (2015). PMID: 25364813, pp. 2638–2649. DOI: 10.1021/jp508905m. eprint: https://doi.org/10.1021/jp508905m. URL: https://doi.org/10.1021/jp508905m.

[18] *Quantum Chemistry*. 7th ed. Pearson, 2013.

[19] David Parker Craig and T. Thirunamachandran. *Molecular quantum electrodynamics: an introduction to radiation-molecule interactions*. Dover Publication, 2005.

[20] *Principles of Nonlinear Optical Spectroscopy*. Optical and Imaging Sciences. Oxford University Press, 1999.

[21] J. Franck and E. G. Dymond. "Elementary processes of photochemical reactions". In: *Trans. Faraday Soc.* 21 (February 1926), pp. 536–542. DOI: 10.1039/TF9262100536.

[22] Edward Condon. "A Theory of Intensity Distribution in Band Systems". In: *Phys. Rev.* 28 (6 Dec. 1926), pp. 1182–1201. DOI: 10.1103/PhysRev.28.1182.

[23] URL: https://en.wikipedia.org/wiki/Franck%E2%80%93Condon_principle.

[24] Enrico Tapavicza, Filipp Furche, and Dage Sundholm. "Importance of Vibronic Effects in the UV-Vis Spectrum of the 7,7,8,8-Tetracyanoquinodimethane Anion". In: *J. Chem. Theory Comput.* 12 (2016). PMID: 27585186, pp. 5058–5066. DOI: 10.1021/acs.jctc.6b00720. eprint: https://doi.org/10.1021/acs.jctc.6b00720. URL: https://doi.org/10.1021/acs.jctc.6b00720.

[25] F. A. Cotton. "Metal-Metal Bonding in $[Re_2X_8]^{2-}$ Ions and Other Metal Atom Clusters". In: *Inorganic Chemistry* 4.3 (1965), pp. 334–336. DOI: 10.1021/ic50025a016.

[26] W. J. Evans. "Tutorial on the Role of Cyclopentadienyl Ligands in the Discovery of Molecular Complexes of the Rare-Earth and Actinide Metals in New Oxidation States". In: *Organometallics* 35.18 (2016), pp. 3088–3100. DOI: 10.1021/acs.organomet.6b00466.

[27] Ryan R. Langeslay et al. "Expanding Thorium Hydride Chemistry Through Th2+, Including the Synthesis of a Mixed-Valent Th4+/Th3+ Hydride Complex". In: *J. Am. Chem. Soc.* 138.12 (2016), pp. 4036–4045. DOI: 10.1021/jacs.5b11508.

[28] Megan T. Dumas et al. "Synthesis and reductive chemistry of bimetallic and trimetallic rare-earth metallocene hydrides with $(C_5H_4SiMe_3)^{1-}$ ligands". In: *J. Organomet. Chem.* 849-850 (2017), pp. 38–47. DOI: 10.1016/j.jorganchem.2017.05.057.

[29] C. A. Gould et al. "Ultrahard magnetism from mixed-valence dilanthanide complexes with metal-metal bonding". In: *Science* 375 (2022), pp. 198–202. DOI: 10.1126/science.abl5470.

[30] C. Edmiston and K. Ruedenberg. "Localized Atomic and Molecular Orbitals". In: *Rev. Mod. Phys.* 35 (3 1963), pp. 457–464. DOI: 10.1103/RevModPhys.35.457.

[31] J. M. Foster and S. F. Boys. "Canonical Configurational Interaction Procedure". In: *Rev. Mod. Phys.* 32 (2 Apr. 1960), pp. 300–302. DOI: 10.1103/RevModPhys.32.300.

[32] J. Pipek and P. G. Mezey. "A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions". In: *J. Chem. Phys.* 90.9 (1989), pp. 4916–4926. DOI: 10.1063/1.456588.

[33] Samir Z. Zard. "Radicals in Action: A Festival of Radical Transformations". In: *Organic Letters* 19.6 (2017), pp. 1257–1269. DOI: 10.1021/acs.orglett.7b00531.

[34] Maurício F. Saraiva et al. "The Barton ester free-radical reaction: a brief review of applications". In: *Tetrahedron* 65.18 (2009), pp. 3563–3572. DOI: https://doi.org/10.1016/j.tet.2009.01.103.

[35] Derek H. R. Barton, David Crich, and Gerhard Kretzschmar. "The invention of new radical chain reactions. Part 9. Further radical chemistry of thiohydroxamic esters; formation of carbon-carbon bonds". In: *J. Chem. Soc., Perkin Trans. 1* (0 1986), pp. 39–53. DOI: 10.1039/P19860000039.

[36] M. Büschleb et al. "Synthetic Strategies toward Natural Products Containing Contiguous Stereogenic Quaternary Carbon Atoms". In: *Angew. Chem. Int. Ed.* 55 (2016), pp. 4156–4186. DOI: 10.1002/anie.201507549. URL: http://dx.doi.org/10.1002/anie.201507549.

[37] Donald G. Truhlar, Bruce C. Garrett, and Stephen J. Klippenstein. "Current Status of Transition-State Theory". In: *J. Phys. Chem.* 100.31 (1996), pp. 12771–12800. DOI: 10.1021/jp953748q.

[38] Henry Eyring. "The Activated Complex in Chemical Reactions". In: *J. Chem. Phys.* 3.2 (1935), pp. 107–115. DOI: 10.1063/1.1749604.

[39] Derek H. R. Barton, David Crich, and William B. Motherwell. "New and improved methods for the radical decarboxylation of acids". In: *J. Chem. Soc., Chem. Commun.* (17 1983), pp. 939–941. DOI: 10.1039/C39830000939.

[40] Daniel J. Tao et al. "Diastereoselective Coupling of Chiral Acetonide Trisubstituted Radicals with Alkenes". In: *Chem. Eur. J.* 22.26 (2016), pp. 8786–8790. DOI: 10.1002/chem.201601957.

[41] V. Augugliaro et al. "Overview on oxidation mechanisms of organic compounds by TiO_2 in heterogeneous photocatalysis". In: *J. Photochem. Photobiol., C* 13.3 (2012), pp. 224–245. DOI: 10.1016/j.jphotochemrev.2012.04.003.

[42] M. Muuronen et al. "Mechanism of photocatalytic water oxidation on small TiO_2 nanoparticles". In: *Chem. Sci.* 8 (2017), pp. 2179–2183. DOI: 10.1039/C6SC04378J.

[43] J. C. Tully. "Molecular dynamics with electronic transitions". In: *J. Chem. Phys.* 93.2 (1990), p. 1061. DOI: 10.1063/1.459170.

[44] A. Kazaryan, R. van Santen, and E. J. Baerends. "Light-induced water splitting by titanium-tetrahydroxide: a computational study". In: *Phys. Chem. Chem. Phys.* 17 (31 2015), pp. 20308–20321. DOI: 10.1039/C5CP01812A.

[45] *tryGit Online Tutorial*. URL: https://try.github.io (visited on 01/23/2018).

[46] *Atlassian Tutorials*. URL: https://www.atlassian.com/git/tutorials (visited on 01/23/2018).