

University of Mississippi

eGrove

Honors Theses

Honors College (Sally McDonnell Barksdale
Honors College)

Spring 5-8-2022

Computational Electronic Spectroscopy Predictions for Astrochemical Ice Analogues and Data Analysis for Experimental Spectra

Austin Wallace

Follow this and additional works at: https://egrove.olemiss.edu/hon_thesis

Recommended Citation

Wallace, Austin, "Computational Electronic Spectroscopy Predictions for Astrochemical Ice Analogues and Data Analysis for Experimental Spectra" (2022). *Honors Theses*. 2549.
https://egrove.olemiss.edu/hon_thesis/2549

This Undergraduate Thesis is brought to you for free and open access by the Honors College (Sally McDonnell Barksdale Honors College) at eGrove. It has been accepted for inclusion in Honors Theses by an authorized administrator of eGrove. For more information, please contact egrove@olemiss.edu.

Computational Electronic Spectroscopy Predictions for
Astrochemical Ice Analogues and Data Analysis for
Experimental Spectra

Austin Wallace

May 2022

Approved By

Advisor: Ryan C. Fortenberry, Ph.D.

Reader: Jared H. Delcamp, Ph.D.

Reader: Steven Davis, Ph.D.

Acknowledgements

Foremost, I want to thank Dr. Ryan Fortenberry for his guidance on my research projects over the past four years. From the very beginning, Dr. Fortenberry was dedicated to my success through tediously teaching me how to use the command line and run quantum calculations. When I started coming up with ideas to pursue within my research projects, Dr. Fortenberry encouraged me to explore those ideas while providing invaluable feedback for refinement. Much of my success and desire to pursue research as a career has come from the positive research experiences fostered by his commitment to undergraduate student success.

Another important faculty member to me has been Dr. Jared Delcamp for taking time out of his busy schedule to talk to me about our collaborative research projects. The encouragement and guidance for refining computational tools and filtering of theoretical dye properties has been crucial for the success of these projects.

My undergraduate academic experience would have been largely incomplete without my college comrades – Rachel Hammond and Michael Valencia. These two exceptional students have inspired me to strive for success and have challenged me to become a better student, researcher, and friend. Additionally, I want to acknowledge my family for their commitment to my success through their ongoing support for my academic endeavors.

Lastly, I want to thank the Barry Goldwater Scholarship & Excellence in Education Foundation. From earning this scholarship, I was able to dedicate more time to research during my senior year to work more on my research projects and develop additional skills.

Abstract

Theoretical chemistry aims to acquire wavefunctions for systems to calculate physical observables to help guide experimentation or provide rationale for particular observations. While quantum chemistry has expanded well beyond Hartree-Fock theory to produce more accurate predictions or faster computational times, this theory has formed the foundation of the field through providing a means to solve the electron-electron repulsion term as simply an average field and to acquire converged molecular orbitals through the self-consistent-field method. When attempting to describe large molecular systems that can populate many microstates, the Boltzmann distribution provides a means for determining the relative abundances of each microstate in the system at a given temperature allowing quantum chemistry predictions to be expanded to larger molecular systems to better match experiment. Through the usage of density functional theory in Chapter 4, carbonic acid clusters are explored to find the lowest in energy and the electronic structure computed to generate theoretical data to compare with experiment. In Chapters 5 and 6, amorphous clusters are generated and electronic excited states computed to construct UV spectra through combining quantum chemistry and the Boltzmann distribution to compare with experiment. Lastly, Chapter 3 is a bit different from the rest of the thesis due to being a data analysis tool for experimentalist that computes the λ_{onset} for absorption and emission spectra. Overall, this thesis primarily focuses on electronic spectroscopy through detailing the generation of spectra for specific molecular systems theoretically or processing experimental spectra to provide a standardized method for approximating the $E_{(0-0)}$ optical energy gap.

Contents

1	Relevant Quantum Mechanics	1
1.1	Background	1
1.1.1	Antisymmetry and Orbitals	3
1.2	Hartree-Fock Theory	6
1.2.1	Solving Hartree-Fock Equations	8
1.2.2	Simplified Procedure	11
1.2.3	Hartree-Fock Final Considerations	11
2	Relevant Statistical Mechanics	13
2.1	Discrete Mathematics and Statistics Background	13
2.2	Boltzmann Distribution	15
3	Accurate Determination of the Onset Wavelength (λ_{onset}) in Optical Spectroscopy	19
3.1	Introduction	19
3.2	The “Onset” Program	22
3.3	Analysis Approach	27
3.4	Results and Discussion	27
3.5	Conclusions	31
4	Linear and Helical Carbonic Acid Clusters	33
4.1	Introduction	33
4.2	Methods & Computational Details	35
4.3	Results & Discussion	36

4.3.1	Structures	36
4.3.2	Excitation Energies	41
4.4	Conclusions	51
5	Computational UV Spectra for Amorphous Solids of Small Molecules	53
5.1	Introduction	53
5.2	Computational Methods & Approach	55
5.3	Results and Discussion	57
5.3.1	Ammonia	57
5.3.2	Carbon Dioxide	60
5.3.3	Water	61
5.4	Conclusions	64
6	Theoretical Characterization of Solid, Amorphous Carbonic Acid in the UV	65
6.1	Introduction	65
6.2	Computational Methods	67
6.3	Results and Discussion	68
6.4	Conclusions	71
7	Final Conclusions	72

List of Figures

3.1	Three examples where absorption and emission intercepts ($E_{(0-0)}$) are non-obtainable. Case 1: The material is non-emissive. Case 2: The emission is beyond detector or photomultiplier tube limits for the instrument available. Case 3: Intersystem crossing (ISC) occurs resulting in the excited state energy lowering and being non-informative relative to the absorption energy.	20
3.2	Onset determination of λ_{onset} for a thin film of C5 with a wide input range. .	23
3.3	The starting input screen on the top and the output graph on the bottom . .	25
3.4	The modified input screen on the top and the output graph on the bottom .	26
3.5	An illustration of the two methods, 10% absorption peak height ($\lambda_{\text{onset } 10\%}$) and line of best fit (λ_{onset}), commonly used in the literature to approximate $E_{(0-0)}$ values when $E_{(0-0)}$ values cannot be obtained.	29
3.6	Onset determination of λ_{onset} emission for $\text{Ru}(\text{bpy})_3\text{Cl}_2$	30
4.1	The monomer isomers with the relative energy in eV increasing from left to right	36
4.2	The dimer isomers explored in this work with the relative energy in eV increasing from left to right	37
4.3	The trimer isomers with the relative energy in eV	37
4.4	The stacked ribbon and spiral geometries	39
4.5	The relative energy (eV) between ribbon and spiral structures according to $\omega\text{B97XD}/6\text{-}31\text{+G}^*$	40
4.6	The binding energies for the spiral and ribbon with adding dimers	40
4.7	The 1B_u excitation energy for ribbons at 8.3 eV	44

4.8	The associated oscillator strength for the 1B_u excited state for ribbons at 8.3 eV	45
4.9	The molecular orbitals for the dimer and tetramer 1B_u excitation around 8.3 eV is depicted here.	45
4.10	The 2B_u excitation energy for ribbons at 9.6 eV	46
4.11	The associated oscillator strength for the 2B_u excited state for ribbons at 9.6 eV	46
4.12	The molecular orbitals for the dimer and tetramer 2B_u excitation around 9.6 eV is depicted here.	47
4.13	The 1B excitation energy for spirals at 7.1 eV	48
4.14	The associated oscillator strength for the 1B excited state for spirals at 7.1 eV	48
4.15	The molecular orbitals for the tetramer and octamer spiral structure 1B excitation is shown here.	49
4.16	The 1B excitation energy for spirals at 8.1 eV	49
4.17	The associated oscillator strength for the 1B excited state for spirals at 8.1 eV to 8.2 eV	50
4.18	The molecular orbitals for the tetramer and octamer spiral structure 1B excitation is shown here.	50
5.1	A set of four subfigures.	58
5.2	A set of four subfigures.	60
5.3	A set of four subfigures.	62
6.1	(a) Computed VUV spectra for the carbonic acid ribbon octamer is compared to experimental data for β -H ₂ CO ₃ from Ref. 84. (b) The ribbon octamer is depicted.	69
6.2	A set of four subfigures.	69
6.3	A set of four subfigures.	70

List of Tables

3.1	Comparing the Previously Attributed and Computed λ_{onset} Values (in nm).	28
4.1	Excited State Data (eV) for the <i>syn-syn</i> Monomer	42
4.2	Excited State Data (eV) for the Dimer	42
4.3	Excited State Data (eV) for the Ribbon Tetramer	43
5.1	Tabulated format of the ammonia spectra displayed in Figure 5.1. Experimental data comes from work produced by Kaiser et al.[105]. All computational spectra are normalized to the maximum peak from CAM-B3LYP/6-311++G(2d,2p).	59
5.2	Tabulated format of the carbon dioxide spectra displayed in Figure 5.2. Experimental data comes from work produced by Kaiser et al.[105]	61
5.3	Tabulated format of the water spectra displayed in Figure 5.3ab. Experimental data comes from work produced by Kaiser et al.[105]	63

Chapter 1

Relevant Quantum Mechanics

Note: This chapter was written referencing Attila Szabo and Neil S. Ostlund's *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. [1]

1.1 Background

One of the most fundamental equations for quantum mechanics is the Schrödinger equation (Equation 1.1) for its ability to provide information about the energy of a quantum system. After acquiring the molecular wave function, Ψ , it can be used to calculate the energy of the system or physical observables such as the dipole moment or spectroscopic information like rotational, vibrational, and electronically excited states.

$$\hat{H}\Psi = E\Psi \tag{1.1}$$

The Hamiltonian operator (\hat{H}) for a system containing nuclei and more than one electron is shown in Equation 1.2 where the lower case letters i, j represent indistinguishable electrons and capital letters A,B represent nuclei. The total number of electrons is N, while the total number of nuclei is represented as M. The position vectors R_{AB} and r_i are used to determine the distance vectors such as the distance vector between electrons and nuclei are described as $r_{iA} = |r_{iA}| = |r_i - R_A|$:

$$\hat{H} = -\sum_{i=0}^N \frac{1}{2} \nabla_i^2 - \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}}. \quad (1.2)$$

The Hamiltonian operator contains several terms based on the the electrons and nuclei present along with their interactions with one another. The first and second terms are the kinetic energy operator for the electrons and kinetic energy operator for the nuclei, respectively. Meanwhile, the third term accounts for the Coloumbic attraction existing between the electrons and nuclei. Finally, the last two terms are for the electron-electron repulsion and nuclei-nuclei repulsion. The ∇_i^2 and ∇_A^2 Laplacian operators are for differentiating the i^{th} electron and A^{th} nucleus with respect to their coordinates. However, due to nuclei containing much more mass and moving much slower than electrons, the Born-Oppenheimer approximation can be employed to treat the second term as zero and fifth term as a constant. These constants can be added to the operator eigenvalues of the other terms since they will not interfere with the operator eigenfunctions.

Therefore, the Hamiltonian simplifies to Equation 1.3, which is referred to as the electronic Hamiltonian:

$$\hat{H} = -\sum_{i=0}^N \frac{1}{2} \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}}. \quad (1.3)$$

Using this form in the Schrödinger equation, the electronic Schrödinger equation is born. The wave function then becomes the electronic wave function Φ_{elec} , which determines the motion of the electrons by explicitly depending on the electronic coordinates and parametrically depending on the nuclear coordinates. Essentially, this means that the electronic wave function is a different function for every unique arrangement of nuclear coordinates:

$$\hat{H}_{elec} \Phi_{elec} = \epsilon_{elec} \Phi_{elec}. \quad (1.4)$$

The primary difficulty in solving the Schrödinger comes from the electronic portion—particularly from effectively describing the electron-electron repulsion due to the classic many body problem in physics; however, after acquiring the electronic portion under the constraint of fixed nuclei, the electronic coordinates can be replaced by their expectation

values to solve for the nuclear Hamiltonian (Equation 1.5):

$$\hat{H} = - \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 + \epsilon_{tot}(\{R_A\}) \quad (1.5)$$

. Solving the resulting nuclear Schrödinger equation (Equation 1.6) will provide descriptions for the vibrations, rotations, and translations of a molecule.

$$\hat{H}_{nucl} \Phi_{nucl} = \epsilon_{nucl} \Phi_{nucl} \quad (1.6)$$

Combining together the electronic and nuclear wavefunctions awards the total wave function Equation 1.7 that can be used to acquire electronic, vibrational, rotational, and translational energy levels for a molecular system.

$$\Phi(\{r_i\}; \{R_A\}) = \Phi_{elec}(\{r_i\}; \{R_A\}) \Phi_{nucl}(\{R_A\}) \quad (1.7)$$

1.1.1 Antisymmetry and Orbitals

Due to the four principle quantum numbers— n , l , m_l , and m_s —every electron can be described uniquely. The m_s term defines the spin of an electron and must be considered within the mathematical construction to fully describe the electrons. The constraint is that there must be two spin functions that are orthonormal to each other such that the following equations are true:

$$\int d\omega \alpha^*(\omega) \alpha(\omega) = \int d\omega \beta^*(\omega) \beta(\omega) = 1, \quad (1.8)$$

$$\langle \alpha | \alpha \rangle = \langle \beta | \beta \rangle = 1, \quad (1.9)$$

$$\int d\omega \alpha^*(\omega) \beta(\omega) = \int d\omega \alpha^*(\omega) \beta(\omega) = 0, \text{ and} \quad (1.10)$$

$$\langle \alpha | \beta \rangle = \langle \beta | \alpha \rangle = 0. \quad (1.11)$$

Thus, the electrons are defined by three spatial coordinates, \mathbf{r} , and a spin coordinate, ω , leading to an electron being described as $\mathbf{x} = \{\mathbf{r}, \omega\}$. Hence, a wave function for an N electron system can be written as a function of electron coordinates: $\Phi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$. To include this into Slater determinants, the interchange of any two electron coordinates must

be antisymmetric (Equation 1.12):

$$\Phi(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_j, \dots, \mathbf{x}_N) = -\Phi(\mathbf{x}_1, \dots, \mathbf{x}_j, \dots, \mathbf{x}_i, \dots, \mathbf{x}_N). \quad (1.12)$$

Now that electrons can be uniquely described, the wave function for an electron can be defined. When the wave function is for a single electron, it is called an orbital. The spatial orbital, $\Psi(\mathbf{r})$, is a function of \mathbf{r} such that its spatial distribution is described with the probability of finding the electron in a volume element is $|\Psi(\mathbf{r})|^2 d\mathbf{r}$. One assumes that the spatial molecular orbitals form an orthonormal set and that if a complete set of spatial orbitals is complete, then any function could be produced through a linear combination of these orbitals (Equation 1.13). Each a_i term represents the coefficient designating the contribution of each spatial orbital to a particular molecular orbital. While an infinite set of spatial orbitals would need to exist for this to be truly complete, computational reality prevents one from accomplishing this. Hence, one will not be able to span the complete space but can increase this description by incorporating more spatial orbitals:

$$f(\mathbf{r}) = \sum_{i=1}^{\infty} a_i \Psi_i(\mathbf{r}). \quad (1.13)$$

In order to incorporate spin into the orbital, the spin orbital $\chi(\mathbf{x})$ is defined by Equation 1.14 where there are K spatial orbitals.

$$\left. \begin{aligned} \chi_{2i-1}(\mathbf{x}) &= \Psi_i(\mathbf{r})\alpha\omega \\ \chi_{2i}(\mathbf{x}) &= \Psi_i(\mathbf{r})\beta\omega \end{aligned} \right\} i = 1, 2, \dots, k \quad (1.14)$$

Spin orbitals are also orthonormal just in case the spatial orbitals are orthonormal.

$$\int d\mathbf{x} \chi_i^*(\mathbf{x}) \chi_j(\mathbf{x}) = \langle \chi_i | \chi_j \rangle = \delta_{ij} \quad (1.15)$$

When expanding to a noninteracting many electron system, the Hamiltonian is simply a summation of single electron kinetic and potential energies without repulsion. Thus, the eigenvalue, E , will simply be a sum of the spin orbital energies and is referred to as the

Hartree product. This is an uncorrelated system where the probabilities of electrons existing in an orbital centered around their position vector is unaffected by the other electrons' existence in spaces nearby. Clearly, this is unsatisfactory, requiring the addition of correlation to the system with many methods expanding on the simple approach provided by Hartree-Fock theory by implementing correlation through more laborious ways for more accuracy.

To include antisymmetry into the wave function, Hartree products are combined through a linear combination of the products with Equation 1.16.

$$\Psi(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{\sqrt{2}}(\chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_2) - \chi_i(\mathbf{x}_2)\chi_j(\mathbf{x}_1)) \quad (1.16)$$

Thus, the antisymmetric wave function (Slater determinant) for a system with N electrons will be described by Equation 1.17. The electron change by rows and the spin orbitals are changed by columns. Thus, interchanging two rows is the same as swapping two electrons because it will change the sign of the determinant; hence, antisymmetry is preserved. This causes the Slater determinant to inherently have exchange correlation because two electrons with parallel spins are correlated; however, this is still considered uncorrelated due to opposite spin electrons not being correlated.

$$\Phi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{bmatrix} \chi_i(\mathbf{x}_1) & \chi_j(\mathbf{x}_1) & \cdots & \chi_k(\mathbf{x}_1) \\ \chi_i(\mathbf{x}_2) & \chi_j(\mathbf{x}_2) & \cdots & \chi_k(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_i(\mathbf{x}_N) & \chi_j(\mathbf{x}_N) & \cdots & \chi_k(\mathbf{x}_N) \end{bmatrix} \quad (1.17)$$

A normalized Slater determinant can be represented by Equation 1.18.

$$\Phi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = |\chi_i(\mathbf{x}_1), \chi_j(\mathbf{x}_2), \dots, \chi_k(\mathbf{x}_N)\rangle \quad (1.18)$$

1.2 Hartree-Fock Theory

The primary objective of Hartree-Fock theory is to solve the non-relativistic time independent electronic Schrödinger equation to acquire the electronic wavefunction from approximating electron-electron repulsion as an average field on each electron. Ultimately, this is accomplished through using a single Slater determinant wave function to perform variational optimization of the electronic energy expectation value under the constraint of spin orthonormality.

Thus, one can start with the definition of the single Slater determinant as Equation 1.19 for an N-electron system:

$$|\Psi_0\rangle = |\chi_1\chi_2\cdots\chi_a\chi_a\cdots\chi_N\rangle. \quad (1.19)$$

From the variational principle, the best spin orbitals will be the ones that produce the lowest possible energy from Equation 1.20, where only the electronic Hamiltonian is considered due to the Born-Oppenheimer approximation Section 1.1:

$$E_0 = \langle\Psi_0|\hat{H}|\Psi_0\rangle = \sum_a \langle a|h|a\rangle = \frac{1}{2} \sum_{ab} \langle ab||ab\rangle \quad (1.20)$$

$$= \sum_a \langle a|h|a\rangle + \frac{1}{2} \sum_{ab} [aa|bb] - [ab|ba]. \quad (1.21)$$

The variational portion comes from different selections of spin orbitals. By systematically varying the spin orbitals under the constraint that they remain orthonormal, the energy can be decreased overall. The equation for finding these best spin orbitals comes from the Hartree-Fock integro-differential equation (Equation 1.22), where the spin orbital energy, χ_a , is ϵ_a and the summation terms are for the electron-electron interactions. The $h(1)$ term is defined by 1.23 as the kinetic and potential energy for electron-nuclei attraction:

$$h(1)\chi_a(1) + \sum_{b \neq a} \left[\int dx_2 |\chi_b(2)|^2 \frac{1}{r_{12}} \right] \chi_a(1) - \sum_{b \neq a} \left[\int dx_2 \chi_b^*(2) \chi_a(2) \frac{1}{r_{12}} \right] \chi_b(1) = \epsilon_a \chi_a(1), \text{ where} \quad (1.22)$$

$$h(1) = -\frac{1}{2}\nabla_1^2 - \sum_A \frac{Z_A}{r_{1A}}. \quad (1.23)$$

The first summation term in 1.22 is the Coulomb term using the Hartree product wave functions, and the second summation term is the exchange term. The Coulomb term replaces the two-electron potential with a one-electron potential through averaging the two-electron operator's ($\frac{1}{r_{12}}$) interaction between electron 1 and electron 2 over all space and spin of \mathbf{x}_2 for electron 2. From the summation, the total averaged potential acting on electron 1 from all the other electrons is acquired. This is denoted by the Coulomb operator:

$$J_b(1) = \int dx_2 |\chi_b(2)|^2 \frac{1}{r_{12}}. \quad (1.24)$$

The exchange term incorporates the antisymmetric property of the single determinant and has an exchange operator that interchanges electron 1 and electron 2 on the right hand side of the two-electron operator. Using these operators, the Hartree-Fock equation can be written as an eigenvalue equation (Equation 1.25):

$$[h(1) + \sum_{b \neq a} J_b(1) - \sum_{b \neq a} K_b(1)]\chi_a(1) = \epsilon_a \chi_a(1). \quad (1.25)$$

Thus, the expectation values for Coulomb and exchange potentials for χ_a will be the Coulomb and exchange integrals.

$$\langle \chi_a(1) | J_b(1) | \chi_a(1) \rangle = \int dx_1 dx_2 \chi_a^*(1) \chi_a^*(1) \frac{1}{r_{12}} \chi_b^*(2) \chi_b^*(2) = [aa|bb] \quad (1.26)$$

$$\langle \chi_a(1) | K_b(1) | \chi_a(1) \rangle = \int dx_1 dx_2 \chi_a^*(1) \chi_b^*(1) \frac{1}{r_{12}} \chi_b^*(2) \chi_a^*(2) = [ab|ba] \quad (1.27)$$

The summations cannot have a equal to b since that would have an electron feeling the presence of an electron in its own spin orbital, which violates the Pauli exclusion principle. This restricted summation in Equation 1.25 causes Equation 1.28,

$$[J_b(1) - K_b(1)]\chi_a(1) = 0, \quad (1.28)$$

producing the Fock operator (f) that can be defined as Equation 1.29:

$$f(1) = h(1) + \sum_b J_b(1) - \sum_b K_b(1). \quad (1.29)$$

This enables one to consider the Fock operator $f(1)$ as the sum of the core-Hamiltonian operator $h(1)$ and the Hartree-Fock potential $\nu^{HF} = \sum_b J_b(1) - \sum_b K_b(1)$. Therefore, the Hartree-Fock equation can be simplified to Equation 1.30. While this produces exact solutions to the integro-differential equation, this approach is constrained to only atoms because molecules require a set of basis functions described by a basis set for the spin orbital expansion where an exact solution only can be acquired when the basis functions approach infinity:

$$f |\chi_a\rangle = \epsilon_a |\chi_a\rangle. \quad (1.30)$$

1.2.1 Solving Hartree-Fock Equations

To provide a method for applying Hartree-Fock Theory to molecules, several adaptations to the steps above must be made to include basis functions. Before acquiring the end result, several parts must be discussed.

First, an objective for a trial wave function is that the first variation in E will result in no change in E . The linear variational trial wave function can be described as $|\Phi\rangle = \sum_{i=1}^N c_i |\Psi_i\rangle$. While, in most cases since $E[\Phi]$ is a functional of Φ , a small change in Φ will result in a first variation in E as shown in Equation 1.31:

$$E[\Phi + \delta\Phi] = E[\Phi] + \delta E + \dots \quad (1.31)$$

The variational method will search for some Φ such that $E[\Phi]$ is stationary with respect to variation in Φ . In other words, the δE term equals zero. To do this, the energy must be minimized through the use of Lagrange's method of undetermined multipliers.

Through the Roothaan equations, the spatial integro-differential equation can be converted to a set of algebraic expressions for solving with matrix math. This enables one to define the electronic wave function for electron, i , as a set of K known basis functions ex-

panded as molecular orbitals. Thus, as the basis set approaches infinity, the description of the exact molecular orbitals is described completely; however, this is unattainable in practice due to the infinite amount of memory and computational operations required to model such a system.

$$\Psi_i = \sum_{\mu=1}^K C_{\mu i} \phi_{\mu} \quad (1.32)$$

Regardless, for a given basis set, determining molecular orbitals relies on calculating the expansion coefficients $C_{\mu i}$ through substituting Equation 1.32 into the Hartree-Fock equation to acquire Equation 1.33. The indexing for the molecular orbitals is performed using ν :

$$f(1) \sum_{\nu} C_{\nu i} \phi_{\nu}(1) = \epsilon_i \sum_{\nu} C_{\nu i} \phi_{\nu}(1). \quad (1.33)$$

Multiplying by ϕ_{μ}^* on the left hand side and integrating produces Equation 1.34 where the Overlap matrix S and Fock matrix F can be defined to update the Hartree-Fock equation to be Equation 1.37:

$$f(1) \sum_{\nu} C_{\nu i} \phi_{\nu}(1) = \epsilon_i \sum_{\nu} C_{\nu i} \phi_{\nu}(1), \quad (1.34)$$

$$S_{\mu\nu} = \int dr_1 \phi_{\mu}^*(1) \phi_{\nu}(1), \quad (1.35)$$

$$F_{\mu\nu} = \int dr_1 \phi_{\mu}^*(1) f(1) \phi_{\nu}(1), \text{ and} \quad (1.36)$$

$$\sum_{\nu} F_{\mu\nu} C_{\nu i} = \epsilon_i \sum_{\nu} S_{\mu\nu} C_{\nu i}. \quad (1.37)$$

This can be written as the following matrix equation, where C is the expansion of atomic orbital to molecular orbital combination coefficients $C_{\mu i}$ and the diagonal matrix ϵ are the orbital energies:

$$\mathbf{FC} = \mathbf{SC}\epsilon. \quad (1.38)$$

Because the probability of finding an electron in a volume element dr at point \mathbf{r} is $|\Psi_a(\mathbf{r})|^2 dr$, the charge density is the probability distribution function $|\Psi_a(\mathbf{r})|^2$. Assuming a closed-shell molecule with an occupied molecular orbital containing two electrons, the charge density can be described by Equation 1.39. The $P_{\mu\nu}$ term is called the charge-

density bond-order matrix and is

$$p(r) = 2 \sum_a^{N/2} \Psi_a^*(r) \Psi_a(r) \quad (1.39)$$

$$= \sum_{\mu\nu} [2 \sum_a^{N/2} C_{\mu a} C_{\nu a}^*] \Phi_\mu(r) \Phi_\nu^*(r) \quad (1.40)$$

$$= \sum_{\mu\nu} P_{\mu\nu} \Phi_\mu(r) \Phi_\nu^*(r). \quad (1.41)$$

This density matrix is included in the two electron portion of the Fock matrix along with a set of two electron integrals (Equation 1.45) as can be seen from Equation 1.42:

$$F_{\mu\nu} = H_{\mu\nu}^{core} + \sum_a^{N/2} \sum_{\lambda\sigma} C_{\lambda\sigma} C_{\nu a}^* [2(\mu\nu|\sigma\lambda) - (\mu\lambda|\sigma\nu)] \quad (1.42)$$

$$= H_{\mu\nu}^{core} + \sum_{\lambda\sigma} P_{\lambda\sigma} [(\mu\nu|\sigma\lambda) - \frac{1}{2}(\mu\lambda|\sigma\nu)] \quad (1.43)$$

$$= H_{\mu\nu}^{core} + G_{\mu\nu}, \text{ where} \quad (1.44)$$

$$(\mu\nu|\lambda\sigma) = \int dr_1 dr_2 \phi_\mu^*(1) \phi_\nu(1) \frac{1}{r_{12}} \phi_\lambda^*(2) \phi_\sigma(2). \quad (1.45)$$

The $G_{\mu\nu}$ is the part of the total Hartree-Fock procedure that takes the longest due to the N^4 scaling. Therefore, for large molecular systems, other methods have been developed to alleviate this computational cost. A common approach is to use density functional theory (DFT) to alleviate these costs by changing the scaling to N^3 in many cases because only three indices are used.

The term $P_{\mu\nu}$ is used in the SCF procedure by describing the position of the electrons through the molecular orbitals. After making an initial guess at the density matrix through using the core Hamiltonian matrix, $H_{\mu\nu}^{core} = \int dr_1 \phi_\mu^*(1) h(1) \phi_\nu(1)$, the generation of new molecular orbitals from the iterative process will update the density, making it closer to the true density. Once this process reaches self-consistency for the density, and by extension the Fock matrix, the procedure is terminated.

The last necessary piece to describe the complete Hartree-Fock procedure requires a definition of the orthogonalization of basis functions because at the start; the basis functions

are only normalized and not orthogonal to each other. Thus a transformation (Equation 1.46) must be used to transform the set of functions into an orthonormal set:

$$\mathbf{X} \equiv \mathbf{S}^{-1/2} = \mathbf{U}\mathbf{S}^{-1/2}\mathbf{U}^\dagger. \quad (1.46)$$

Applying this concept to the Roothaan equations, the coefficient matrix C can be transformed through $C = \mathbf{X}C'$ and inserted to make equation 1.48, where $\mathbf{F}' = \mathbf{X}^\dagger\mathbf{F}\mathbf{X}$:

$$FXC = SXC'\epsilon \quad (1.47)$$

$$F'C' = C'\epsilon. \quad (1.48)$$

1.2.2 Simplified Procedure

The Hartree-Fock procedure first requires a set of nuclear coordinates \mathbf{R}_A , atomic numbers Z_A , number of electrons N , and basis set ϕ_μ . Then the $S_{\mu\nu}$, $H_{\mu\nu}^{core}$, and $(\mu\nu|\lambda\sigma)$ integrals must be evaluated. Next the diagonalization of \mathbf{S} must be obtained after transforming it to be orthonormal.

The following steps are iteratively repeated until self-consistency is reached. A guess of the density matrix \mathbf{P} is made to calculate \mathbf{G} from \mathbf{P} and $(\mu\nu|\lambda\sigma)$. \mathbf{F} is formed from adding \mathbf{G} to \mathbf{H}^{core} . Then the Fock matrix is transformed for diagonalization to acquire new molecular orbital coefficients (C') and orbital energies ϵ . These orbitals are transformed back into C to form a new density matrix \mathbf{P} .

After convergence is reached, the final matrices are used to calculate expectation values for electronic properties.

1.2.3 Hartree-Fock Final Considerations

Effectively, this approach changes the many body problem into a single electron problem where electron-electron repulsion is simply treated as an average field. However, since the Fock operator depends on its eigenfunctions, the equation is nonlinear rendering a solution only through iteration. Hence, the process of solving Equation 1.22 is called the self-consistent-field (SCF) method because after taking an initial guess at the molecular

orbitals, calculating the average field and solving the eigenvalue equation (1.22) provides a new set of molecular orbitals. Then, the process is continued until convergence is reached between the fields. However, the definition of convergence is truly up to the user.

After acquiring this state, the solution to the Hartree-Fock eigenvalue problem is a set of orthonormal molecular orbitals with orbital energies, where the first N orbitals are considered occupied and all the rest are unoccupied or virtual orbitals. To achieve lower energy expectation values to approach the Hartree-Fock limit, more spatial basis functions can be added. In other words, a larger basis set provides a higher degree of expansion for the spin orbitals to achieve more accurate results. However, another option for more accuracy is to include more determinants than a single determinant for the ground state. The total number of combinations can be determined by the binomial coefficient (Equation 1.49).

$$\binom{2k}{N} = \frac{2K!}{N!(2K - N)!} \quad (1.49)$$

These determinants can be used to classify other possible determinants for the ground state from representing approximate excited states or used in a linear combination with the current ground state determinant to produce more accurate results.

Chapter 2

Relevant Statistical Mechanics

Note: This chapter was written referencing Thomas Egel and Philip Reid's *Thermodynamics, Statistical Thermodynamics, and Kinetics 4e.* [2]

While quantum mechanics focuses on describing behavior of microscopic particles, statistical mechanics examines the macroscopic behavior of a collection of many particles. Fundamentally, statistical mechanics uses distributions of microstates–specific energy arrangements for particular energy levels–to predict bulk properties. The Boltzmann distribution is among the most common distributions observed and can be used to combine together information about distinct microstates.

2.1 Discrete Mathematics and Statistics Background

Probability within statistics utilizes random variables that have values that can change throughout an experiment. These variables can either be discrete–only specific value out of a list of values–or continuous–a value between two defined bounds. The discrete variables will have a probability distribution function using summations to find probabilities. While quantum mechanics teaches that the universe fundamentally operates on discrete values with Planck's constant being the smallest possible subdividing unit, certain mathematical constructions like wave functions lean towards the need for continuous variables. These variables increase by infinitesimally small steps relative to the range of the bounds, permitting the usage of integration on probability densities for evaluating probabilities instead of

summations.

The simplest definition of probability is a specific outcome divided by the total number of possible outcomes. The summation of all these probabilities for a given system must equal 1:

$$p_1 + p_2 + \dots + p_N = \sum_{i=1}^N = 1. \quad (2.1)$$

Outcomes can be described by either a list or a set of values. If the values are contained in a list, then, order matters with each unique list being called a permutation. This is also called a microstate when applied to an arrangement of energy. To determine the total number of permutations of objects, Equation 2.2 can be used, where n is the total permutations of objects and j represents the length of a subset of objects.

$$P(n, j) = \frac{n!}{(n-j)!} \quad (2.2)$$

On the other hand, if arrangements are comprised of sets independent of order, then each outcome is a configuration. The total number of configurations can be computed through Equation 2.3, where there is a subset of j objects from a total of n objects.

$$C(n, j) = \frac{P(n, j)}{j!} = \frac{n!}{j!(n-j)!} \quad (2.3)$$

Since factorials scale quickly rendering the number of permutations and configurations difficult to manage, Stirling's approximation can be used to simplify a natural log surrounding a factorial as in Equation 2.4:

$$\ln(N!) \approx N \ln(N) - N. \quad (2.4)$$

Finally through acquiring all the probabilities, a distribution function can be constructed. One of the most common distribution functions is the Gaussian distribution (Equation 2.5) due to the widespread applicability across disciplines. The broadening of the distribution comes from the variance (σ^2), which is the square of the average deviation subtracted from

the mean of the distribution:

$$P(X)dX = \frac{e^{-\frac{(X-\delta)^2}{2\sigma^2}}}{(2\pi\sigma^2)^{\frac{1}{2}}}. \quad (2.5)$$

Using these mathematical tools, a generalized distribution function can be produced to describe macroscopic bulk behavior from a collection of individual particles.

2.2 Boltzmann Distribution

Throughout statistical mechanics, one of the primary objectives is to determine how energy is distributed across a system's numerous energy levels. To construct the Boltzmann distribution, the weight of a configuration and the dominant configuration must first be defined. The weight of a configuration (W) is the total number of microstates at a particular configuration of energy (Equation 2.6), where a_n are the occupation numbers that specify how many microstates exist at a given energy level.

$$W = \frac{N!}{a_0!a_1!\dots a_n!} = \frac{N!}{\prod_n a_n!} \quad (2.6)$$

The dominant configuration is the configuration with the largest weight, meaning that it will produce the largest probability for observing it in Equation 2.7:

$$P_i = \frac{W_i}{W_1 + W_2 + \dots + W_N} = \frac{W_i}{\sum_{j=1}^N W_j}. \quad (2.7)$$

When a system is sufficiently large, the dominant configuration will be overwhelmingly the single contribution to the observed behavior. Thus, the primary objective is to find the dominant configuration located at the peak of the probability distribution. In essence, this occurs when weight as a function of configurational index (χ) equals zero (Equation 2.8).

$$\frac{d\ln(W)}{d\chi} = 0 \quad (2.8)$$

In order to accomplish this goal of building the energy distribution of the dominant configuration, or Boltzmann distribution, the natural log of the weight while expanding it

with Stirling's approximation from Equation 2.4 is necessary to start:

$$\ln(W) = \ln\left(\frac{N!}{\prod_n a_n!}\right) = \ln(N!) - \ln\left(\prod_n a_n!\right) = N\ln(N) - \sum_n a_n \ln(a_n). \quad (2.9)$$

Since the number of molecules in a given energy level n is described by a_n , the occupation number provides a configurational index. The differentiation of $\ln(W)$ with respect to a_n produces the following equation and steps:

$$\frac{d\ln(W)}{da_n} = \frac{dN}{da_n} \ln(N) + N \frac{d\ln(N)}{da_n} - \frac{d}{da_n} \sum_n (a_n \ln(a_n)) \quad (2.10)$$

$$\frac{d\ln(W)}{da_n} = \ln(N) + N\left(\frac{1}{N}\right) - (\ln(a_n) + 1) = \ln\left(\frac{N}{a_n}\right) = -\ln\left(\frac{a_n}{N}\right). \quad (2.11)$$

Provided that $-\ln\left(\frac{a_n}{N}\right) = 0$, then the search is complete for the dominant configuration; however, the occupation numbers are dependent on each other, meaning that the a loss or gain in energy for one molecule comes at a gain or loss in energy from another molecule in the system. To ensure this, the number of objects and energy in the system are constant, where da_n is the change in occupation number and $d\epsilon_n$ is a given energy level.

$$\sum_n da_n = 0 \quad (2.12)$$

$$\sum_n \epsilon_n da_n = 0 \quad (2.13)$$

Injecting these constants into Equation 2.11 along with Langrange multipliers to maximize the function through the Langrange method of undetermined multipliers to produce

the following:

$$d\ln(W) = 0 = \sum_n -\ln\left(\frac{a_n}{N}\right)da_n + \alpha \sum_n da_n - \beta \sum_n \epsilon_n da_n \quad (2.14)$$

$$0 = \sum_n \left(-\ln\left(\frac{a_n}{N}\right) + \alpha - \beta\epsilon_n\right)da_n \quad (2.15)$$

$$0 = -\ln\left(\frac{a_n}{N}\right) + \alpha - \beta\epsilon_n \quad (2.16)$$

$$\ln\left(\frac{a_n}{N}\right) = \alpha - \beta\epsilon_n \quad (2.17)$$

$$a_n = Ne^\alpha e^{-\beta\epsilon_n}. \quad (2.18)$$

The α term is defined by summing over all energy levels.

$$N = \sum_n a_n = Ne^\alpha \sum_n e^{-\beta\epsilon_n} \quad (2.19)$$

$$1 = e^\alpha \sum_n e^{-\beta\epsilon_n} \quad (2.20)$$

$$e^\alpha = \frac{1}{\sum_n e^{-\beta\epsilon_n}} \quad (2.21)$$

The denominator of Equation 2.21 is referred to as the partition function that serves as a sum of all terms related to the probability of the energy level of interest. The g_n term designates the degeneracy of the energy level:

$$q = \sum_n g_n e^{-\beta\epsilon_n}. \quad (2.22)$$

Thus, the Boltzmann distribution can use this total partition function to find the probability of a particular state existing over the partition function. Equation 2.23 shows a form of the distribution after substituting in $\frac{1}{k_B T}$ for β :

$$p_n = \frac{g_n e^{-\beta\epsilon_n}}{q}. \quad (2.23)$$

Ultimately, the Boltzmann distribution is the most likely distribution of energy, providing a foundational mathematical construct for statistical mechanics. A common application is to use Equation 2.24 to find Boltzmann factors, $\frac{p_n}{p_m}$, between the lowest energy state, p_m ,

and any other desired state in order to effectively to construct the distribution. Due to the dependence on temperature, the Boltzmann factors will increase in size for higher energy states as the temperature increases since more energy is present in the system.

$$\frac{p_n}{p_m} = \frac{\frac{g_n e^{-\beta \epsilon_n}}{\mathcal{Q}}}{\frac{g_m e^{-\beta \epsilon_m}}{\mathcal{Q}}} = \frac{g_n}{g_m} e^{-\beta(\epsilon_n - \epsilon_m)} = \frac{g_n}{g_m} e^{-\frac{(\epsilon_n - \epsilon_m)}{k_B T}} \quad (2.24)$$

Therefore, this implementation of the Boltzmann distribution provides a basis for combining together unique physical observables from different energetic states across various temperatures; however, in every case, the main contributor will always be the lowest energy state. Thus, this statistical mechanics tool can be used to piece together information about individual quantum mechanical systems to generate a prediction that better matches observation for different temperatures.

Chapter 3

Accurate Determination of the Onset Wavelength (λ_{onset}) in Optical Spectroscopy

Note: This chapter contains unmodified text and figures from the following journal article and has been reproduced with permission.

A. M. Wallace et al., “Accurate determination of the onset wavelength (λ_{onset}) in optical spectroscopy”, *Journal of Quantitative Spectroscopy and Radiative Transfer* **265**, 107544 (2021)

3.1 Introduction

The onset wavelength (λ_{onset}) is a convenient heuristic for determining underlying physical properties of a compound based solely on optical spectral data.[4] The value of λ_{onset} is often defined as the x-intercept of a tangent line (or line of best fit) on the inflection point for the lowest energy absorption transition of a compound regardless of whether the lowest energy transition is the highest intensity peak in the spectrum or not.[5] The onset wavelength is meant to be an approximation of the $E_{(0-0)}$ optical energy gap interpreted as the intersection of the absorption and emission curves for singlet state emitting materials.[6] Non-emissive (case 1) and low energy absorbing materials that weakly emit beyond

Si-detector or photomultiplier tube limits (case 2) inherently require estimation of the optical energy gap as shown in Figure 3.1. Similarly, triplet emitters with facile intersystem crossing require the use of a tangent line fit on the high energy side of the emission curve to approximate excited state energy levels since singlet state absorption-triplet state emission curve crossing is non-informative, case 3 from Figure 3.1.

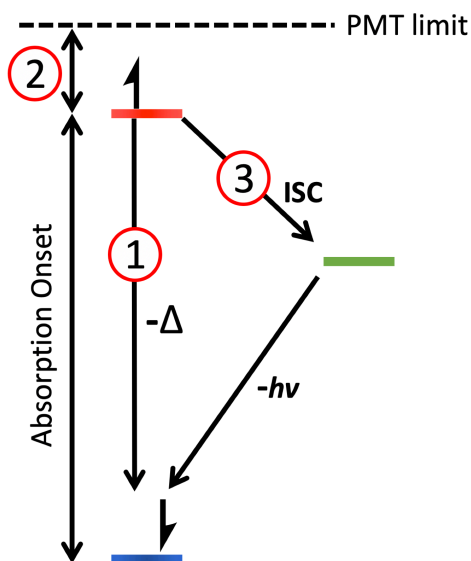


Figure 3.1: Three examples where absorption and emission intercepts ($E_{(0-0)}$) are non-obtainable. Case 1: The material is non-emissive. Case 2: The emission is beyond detector or photomultiplier tube limits for the instrument available. Case 3: Intersystem crossing (ISC) occurs resulting in the excited state energy lowering and being non-informative relative to the absorption energy.

Beyond absorption and emission spectroscopy, several additional types of data require the use of onset value determination via lines of best fit typically manually drawn onto the data samples. Examples with this analysis published in the literature include articles on thermalgravimetric analysis (TGA),^[7] differential scanning calorimetry (DSC),^[8] Tauc plot analysis,^[9, 10] X-ray/ultraviolet photoemission spectroscopy (XPS/UPS),^[11–15] and inverse photoemission spectroscopy (IPES).^[14–16]

Moreover, the onset wavelength is used to quantitatively predict an approximate optical gap of the compound^[17] by utilizing the equation:

$$E_g^{\text{opt}} = \frac{hc}{\lambda_{\text{onset}}} \quad (3.1)$$

The product of h and c , Planck's constant and the speed of light, respectively, is typically approximated to 1240 nm. As a result, λ_{onset} is a straightforward value determined from observed properties that do not require more detailed analysis beyond an optical absorption characterization.[18] Materials applications for such a property range from solar energy production to light emitting diodes to photocatalysis among many other uses for materials utilizing the emission or absorption of photons.

As applicable as utilizing λ_{onset} to determine E_g^{opt} is, a generally accepted, rigorous means of determining this value has not been determined. Again, λ_{onset} is most often determined manually. In some cases, the spectrum is projected onto a screen, and a meterstick is stretched from the perceived inflection point to the x-axis in order to determine this intercept. Even when straight lines are applied to an absorption or emission curve visually on a computer screen, proper selection of the inflection point or creation of an approximate tangent line is only accurate within a certain window. Additionally, λ_{onset} is often determined by simply observing where the curve nears zero on the x-axis. Likely the largest error introduction is with this approach. As a result, the determination of λ_{onset} is fraught with human errors and is reminiscent of cutting out paper readouts of NMR spectra and weighing the actual paper to get relative areas for comparison.[19] While such practices were state-of-the-art in the middle of the last century, computer analysis of standard spectral output has revolutionized NMR spectroscopy through numerical integration of the peak areas.[19, 20] Similarly, computationally determining an optical spectrum inflection point and tangent line is straightforward with modern tools (even though it is currently not standard practice) and should make finding λ_{onset} not only more accurate but also more convenient and consistently approximated within the research fields that utilize this metric.

Additionally, the λ_{onset} and its closely-related optical energy gap are challenging to obtain in some instances, especially when working with weakly emissive compounds. This could be due to the intrinsic optical properties for the compound of interest or even due to the spectral region where the lowest energy transition occurs. Projector and meterstick (or digital line and computer screen) analysis breaks down quickly when the peaks are short and wide making the inflection point difficult to isolate.[5] Furthermore, "eyeballing" λ_{onset} is easily skewed when moving into the UV region where the difference of a few nm could

be on-the-order of 0.5 eV or more.[21] Numerical analysis can find the proper inflection point and corresponding x-intercept regardless of the peak height, width, or region of the spectrum.

Consequently, a more robust method is needed in the field of organic/organometallic optical materials research. The tool utilized must interface easily with existing experimental data outputs, work on the sub-second time scale, and be user-friendly enough so that it can have ready uptake by practitioners in the field who are already inundated with dozens of tools. Hence, this paper will present such a program (called “Onset”) and comment on the attribution of recently reported λ_{onset} values from the literature compared with results computed by Onset. This program should reduce the errors (or at least systematize results) for λ_{onset} attribution by introducing a new, user-friendly computational paradigm for analysis.

3.2 The “Onset” Program

The “Onset” program operates on a local host and runs a Python-Flask application through an executable file. The program takes a zeroed baseline dataset in the form of a .csv file and calculates the λ_{onset} after the longest wavelength/lowest energy peak or shoulder in the data. Manual input ranges enable users to choose different ranges from the dataset, and the x-direction feature permits the selection of either side of the local maximum for finding an λ_{onset} line. After selecting the units, clicking the “upload file” button will display a graph with the data and the functional form of the line used to calculate the λ_{onset} . Additionally, this page displays the polynomial fit used for calculating the derivatives and allows users to save a figure with user-specified modifications to the axis, ranges, λ_{onset} line existence, filename, and title.

The program extracts the rows and columns of the data from the file and finds the maximum within the given range. In order to ensure that the plotted lines match the data in a future step, the y-values are all divided by the maximum y-value normalizing the functional ordinate. The program then runs the algorithm several times in the background shifting the input range up to form several λ_{onset} lines from different polynomial fittings. The fourth degree polynomial fitting slightly changes the onset value; therefore, several

shifted ranges are used to form a more accurate and reproducible output. The onset lines are determined through calculating the numerical first and second derivatives over the shifted input ranges to find the inflection point.

After the algorithm plots several lines (which are not presented to the user), the lines are evaluated based on their slope's steepness and how well the line fits the original dataset within 15 data points from the inflection point. The matching is determined through a list of booleans produced by `numpy.isclose()`, which uses the evaluation $absolute(a - b) \leq (atol + rtol * absolute(b))$ where $rtol=1e-06$ and $atol=1e-02$. The total number of Trues are summed to be the matches. The line with the highest value from the equation $value = (matches) - 1500 * (slope)$ is used as the final λ_{onset} line. Due to normalizing the maximum absorption data point to 1.0, the evaluation will work regardless of units provided in the input dataset. If the original units are desired for display, then the option can be selected under the "Y Values" dropdown in the input page.

The first onset value: **855.2 nm**

The input interval was ['820', '920'] and the used interval was [836.0, 856.0]

$$y = -0.0073412029596511275x + 6.2780584143078535$$

RMS Error for the Polynomial Fit: 0.00035673374800042394

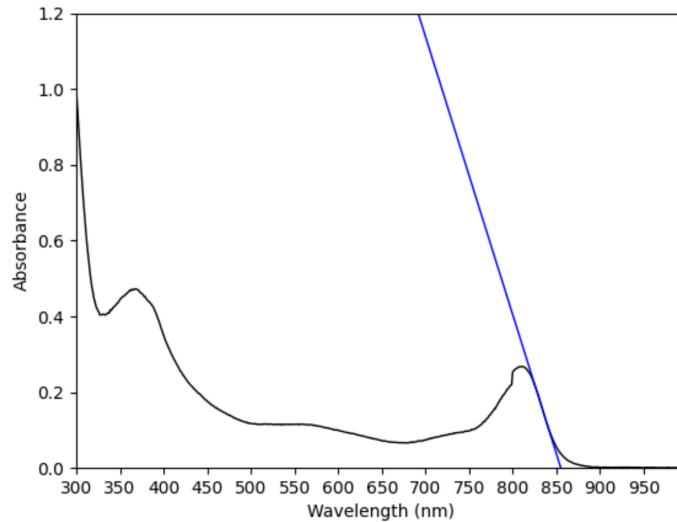


Figure 3.2: Onset determination of λ_{onset} for a thin film of C5 with a wide input range.

The equation of the line is acquired through plotting a line that passes through the

inflection point with the slope determined by the first derivative at that point and the y-intercept coming from the original function at that point. Then, the line is translated to pass through the inflection point, and where the onset value is reported as the x-intercept of the final line generated.

Although the program can find an onset after the longest wavelength peak without manual input ranges, the program should be run again with manual input ranges to ensure that the desired peak is selected and that the polynomial fitting finds the appropriate inflection point. For example, Figure 3.2 displays the the output graph for the molecule C5 with a phosphorus hexafluoride couterion in the solid state.[18] When the program is run with no manual ranges - the “Minimum X Value” and “Maximum X Value” both are “0.0” - it finds the maximum in the dataset and the next inflection point. In this example, the initial estimate over the entire range of the data places the λ_{onset} at 851 nm. However, a more sensible, yet still large, range (800 nm to 900 nm) encompasses the entire low-energy side of the curve produces a minimally shifted λ_{onset} of 855 nm as shown in Figure 3.2. Further refinements to a smaller range do not change this onset wavelength. The units are also variable; nm, cm^{-1} , and eV are currently available.

Consequently, after a first attempt, the user should define a manual input range that starts before the local maximum of the desired peak without including a higher maximum before the peak so as to isolate the actual curve feature of interest. Next, the minimum should be a point after the local minimum. This provides a large enough range for the program to make several polynomial fits in order to ensure the best λ_{onset} line is selected and that it contains the true inflection point in the data. Using AP14 as an example in Figure 3.3, the initial guess without a data range does not produce the proper inflection point and the corresponding λ_{onset} value is not near a sensible range. However, it permits the user to see the graph and what interval the program selected for the input range. The input page in the top of Figure 3.4 shows the selected input range for this dataset and the output shows the correct λ_{onset} line through the appropriate inflection point. Therefore, the process should be to first run the program without an input range and then run it again with a more defined range for a particular peak.

Upload new File

For the input X Values, use wavelength integers in the form "200"

If you would like to use the maximum value on the graph, then use the input range of "0.0"

Choose File AP14.csv

Minimum X Value 0.0 Maximum X Value 0.0

X unit : Y unit
nm abs

X direction
increasing

Y Values
normalize

Upload

Note: Any saved figures will appear in the directory above the program

The first onset value: **884.1 nm**

The input interval was [0.0, 0.0] and the used interval was [415.0, 435.0]

$$y = -0.0020533648748823907x + 1.8152919872385218$$

RMS Error for the Polynomial Fit: 5.627062307286492e-05

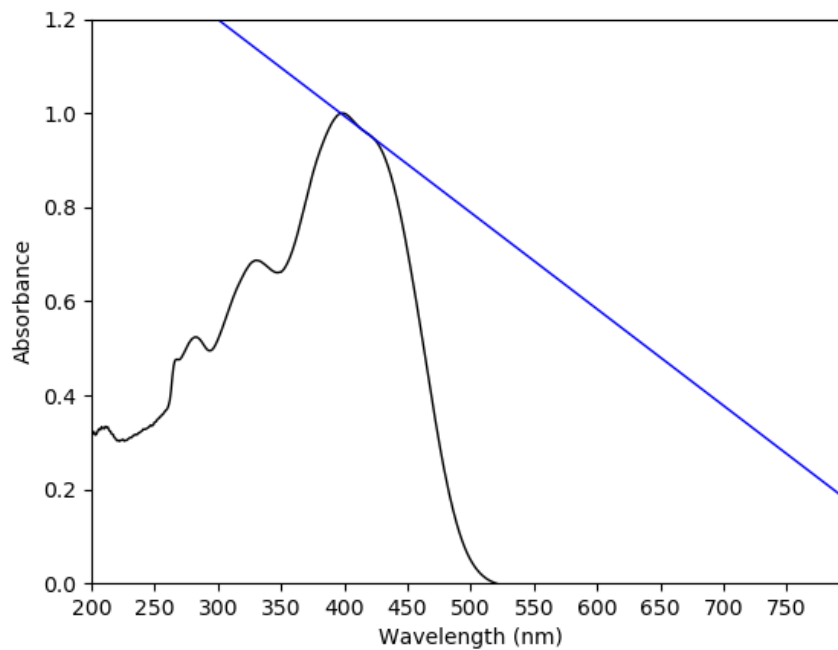


Figure 3.3: The starting input screen on the top and the output graph on the bottom

Upload new File

For the input X Values, use wavelength integers in the form "200"

If you would like to use the maximum value on the graph, then use the input range of "0.0"

Choose File AP14.csv

Minimum X Value 440 Maximum X Value 540

X unit : Y unit
nm abs

X direction
increasing

Y Values
normalize

Upload

Note: Any saved figures will appear in the directory above the program

The first onset value: 492.3 nm

The input interval was ['440', '540'] and the used interval was [456.0, 476.0]

$$y = -0.017006097407829657x + 8.372881562514642$$

RMS Error for the Polynomial Fit: 8.834229594842812e-05

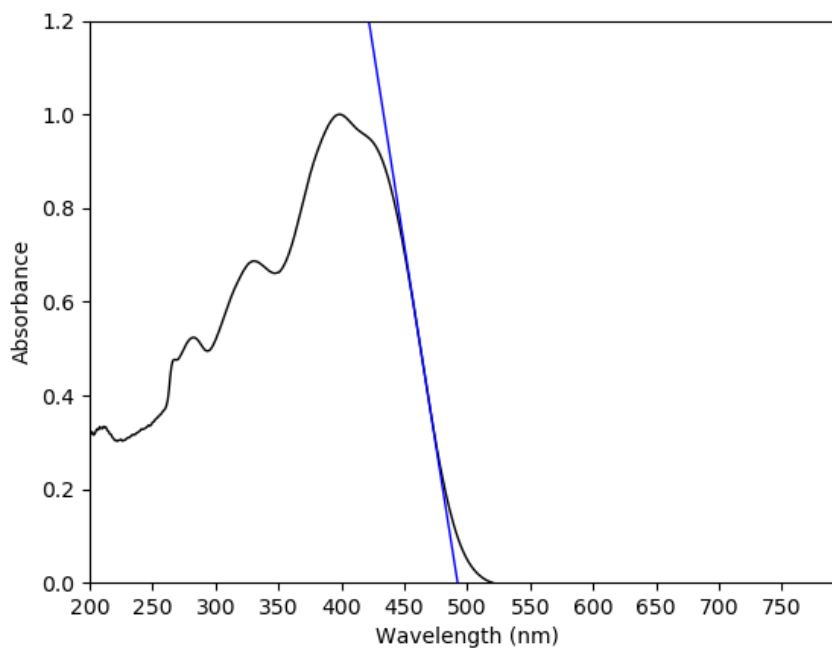


Figure 3.4: The modified input screen on the top and the output graph on the bottom

3.3 Analysis Approach

In order to assess the proficiency for the “Onset” program, 38 previously attributed λ_{onset} values and their associated spectra have been taken from the literature and from our group. The “Onset” computed λ_{onset} values are determined by utilizing the above approach where an initial run produces an estimated x-intercept for consideration before the range containing the inflection point is limited by the user. This results in much better behaved tangent lines as the example of AP14 has highlighted in Figures 3.3 and 3.4.

The differences in the previously attributed and the presently computed λ_{onset} values are reported in this work. The raw differences are averaged as well as the absolute values of the differences where the latter gives the mean absolute error or MAE. Percent differences are also computed as the difference in the numerator and the previously attributed λ_{onset} as the denominator. This is interpreted to show in how much error the manual attributions may result.

3.4 Results and Discussion

Table 3.1 reports a list of 38 different λ_{onset} values taken from the recent literature as produced in the Delcamp group and the corresponding Onset-computed values from this present work. The entries selected include organic dyes with high energy charge transfer bands ($\lambda_{\text{onset}} \leq 500$ nm), organic dyes with low energy charge transfer bands ($\lambda_{\text{onset}} \geq 600$ nm), organic cyanine-type absorption spectra ($750 \text{ nm} \leq \lambda_{\text{onset}} \leq 1020$ nm), and metal-to-ligand charge transfer transition metal complexes. A variety of spectral shapes are represented here with some broad charge transfer transitions and sharp cyanine type transitions. Spectra with and without shoulders as the low energy transition are purposefully selected to probe the capabilities of the Onset program. The MAE for the difference is 3 nm. The raw average is -2 nm. The percent errors, which are agnostic to the region of the spectrum involved, have averages of 0.55% and -0.24%, respective of absolute and raw datasets. The previously, manually-attributed λ_{onset} values are typically reported in 5 nm increments, and the errors, regardless of raw or absolute, are on the order of such an increment. While

Table 3.1: Comparing the Previously Attributed and Computed λ_{onset} Values (in nm).

Dye	Reported	Onset Value	10% Height	$E_{(0-0)}$	Diff. ^a	Percent Diff. (%)	Reference
C5 w/ TFSI in DCM	875	872	876	850	-3	-0.34%	18
C5 w/ BARF in DCM	875	871	876	847	-4	-0.46%	18
C5 w/ Cl in DCM	870	869	875	840	-1	-0.11%	18
C5 w/ NO ₃ in DCM	870	871	878	846	1	0.11%	18
C5 w/ PF ₆ in DCM	870	870	876	848	0	0.00%	18
C5 w/ TPB in DCM	875	870	876	851	-5	-0.57%	18
PhIndzCl1 in 1:1 MeCN/Hi ₂ O	695	699	702	669	4	0.58%	22
PhIndzC3 in DMSO	740	739	744	720	-1	-0.14%	22
PhIndzC5 in MeCN	855	854	859	786	-1	-0.12%	22
IndzOMe-C5 in DMSO	885	880	884	853	-5	-0.56%	22
PhCN-C5 in DCM	880	879	883	848	-1	-0.11%	22
PhOMe-C5 in DMSO	870	867	872	830	-3	-0.34%	22
PB1 in DCM	650	648	656	585	-2	-0.31%	23
AP11 in DCM	470	470	464	476	0	0.00%	24
AP14 in DCM	495	496	493	493	1	0.20%	24
AP16 in DCM	475	477	473	446	2	0.42%	24
AP17 in DCM	455	453	452	441	-2	-0.44%	24
bisCF ₃ PhIndzSQ in toluene	745	744	751	735	-1	-0.13%	25
bistBuOH PhIndzSQ in toluene	740	736	744	726	-4	-0.54%	25
CF ₃ PhIndzSQ in toluene	750	749	762	738	-1	-0.13%	25
CNPhIndzSQ in toluene	755	752	764	734	-3	-0.40%	25
MesPhIndzSQ in toluene	745	743	754	725	-2	-0.27%	25
NaphIndzSQ in toluene	755	753	767	740	-2	-0.26%	25
OH PhIndzSQ in DMSO	755	752	770	726	-3	-0.40%	25
OMe PhIndzSQ in toluene	750	748	768	731	-2	-0.27%	25
PhIndzSQ in toluene	750	748	764	731	-2	-0.27%	25
PyrIndzSQ in toluene	750	750	761	736	0	0.00%	25
RhIndz in DCM	1040	1033	1044	1004	-7	-0.67%	26
Mn(bpy)(CO) ₃ Br in MeCN	480	491	482	non-emissive	11	2.29%	27
Mn(bpy) ₂ Br ₂ in MeCN	455	451	465	non-emissive	-4	-0.80%	27
ICG in MeOH	830	829	835	804	-1	-0.12%	28
SO3C5 in MeOH	850	848	858	805	-2	0.24%	28
SO3SQ in MeOH	735	731	744	710	-4	-0.54%	28
W(pyNHC)(CO) ₄ in MeCN	450	446	540	non-emissive	-4	0.89%	29
AP25 in DCM	765	761	770	N/A ^b	-4	-0.52%	30
Ru(bpy) ₃ Cl ₂ in MeCN ^c	550	551	549	triplet emitter	1	0.18%	31
Ir(ppy) ₃ in MeCN (410 nm) ^{c,d}	480	480	477	triplet emitter	0	0.00%	31
Ir(ppy) ₃ in MeCN (490 nm) ^{c,d}	490 ^e	461	N/A ^f	triplet emitter	-29	-5.92%	31
Raw Average					-2	-0.24	
Absolute Average					3	0.55	

^a Difference taken between the Onset value and the reported value.

^b Beyond photomultiplier tube ~850 nm detection limit.

^c Data is reported from the emission curve on the high energy side.

^d The emission maxima and low energy onset shifts with a change in excitation energy.

^e Rough estimation due to minimal curve shape observed on the high energy side of the emission due to overlap with the excitation signal.

^f Not observable due to overlap with the excitation signal.

the overall averages do not produce a large error, on average, the range of error deviates from the average to 11 nm in $\text{Mn}(\text{bpy})_2\text{Br}_2$ to -29 nm in $\text{Ir}(\text{ppy})_3$, or 2.29% to -5.92%, respectively, in our set.

Some λ_{onset} attributions appear to be more straight-forward than others. For instance, C5 with PF_6 [18] is identical between the attributed experimental and the computer-analyzed λ_{onset} results. However, the raw average of the difference between the previously attributed and the Onset λ_{onset} is negative, with 27 of the 38 differences in Table 3.1 resulting in negative values. This indicates that the observer is most often subconsciously inclined to artificially red-shift the attributed λ_{onset} consequently producing a smaller E_g^{opt} and underestimating the optical energy gap. Such a difference could lead to difficult and time-consuming synthesis of boundary pushing materials that may not have needed the property for the desired application. This would result, then, in a significant inaccuracy for many optoelectronics applications.

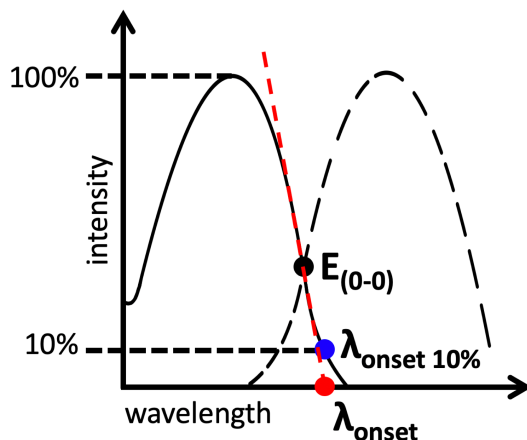


Figure 3.5: An illustration of the two methods, 10% absorption peak height ($\lambda_{\text{onset } 10\%}$) and line of best fit (λ_{onset}), commonly used in the literature to approximate $E_{(0-0)}$ values when $E_{(0-0)}$ values cannot be obtained.

Whether the human determination of the inflection point itself or “eyeballing” of the slope of the tangent line (or a combination of the factors) is the cause for this common error is unclear. The 38 entries in our dataset have, in large part, been synthesized with aims of creating longer wavelength absorption/emission properties. This objective could also be shifting the results by a subconscious desire from the observer to create materials with λ_{onset} values more toward the red. In any case, the remedy for this seemingly systematic

error is to utilize a non-biased computational analysis to determine the λ_{onset} value in the future.

$E_{(0-0)}$ is generally the preferred method of determining optical energy gap values when measurable. However, when these values can not be obtained, alternative optical energy gap estimations are used with line of best fit and 10% absorbance peak height measurements being the most commonly employed (Figure 3.5). Many researchers prefer the 10% peak height method due to the lack of ambiguity or human error in obtaining the value. However, with computerized approaches, this ambiguity can be eliminated. We are unaware of a systematic study comparing the two methods without potential human bias. For all of the cases with experimental $E_{(0-0)}$ data available (31 dyes), the average Onset and 10% height values are compared. On average, the Onset values are within 19 nm (0.05 eV) of the $E_{(0-0)}$ values. In all cases except one, the Onset values are red-shifted relative to the $E_{(0-0)}$ values. Comparatively, the use of 10% half height values results in a less correlated value to $E_{(0-0)}$ than Onset with red-shifts on average of 31 nm (0.07 eV). Notably, the 10% height analysis and the human estimated line of best fit method give substantially different values (40 nm or a 0.18 eV difference) when comparing values on curves with overlapping features such as in the case of $\text{W}(\text{pyNHC})(\text{CO})_4$.

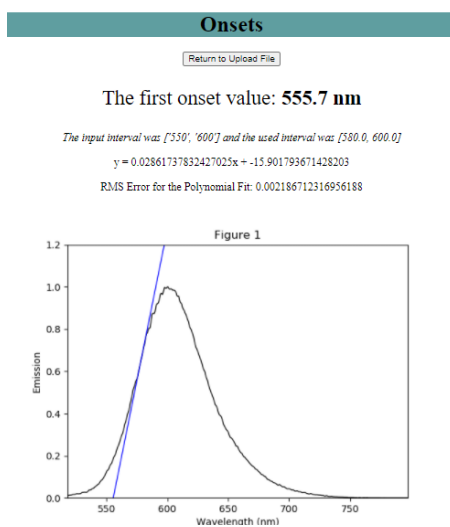


Figure 3.6: Onset determination of λ_{onset} emission for $\text{Ru}(\text{bpy})_3\text{Cl}_2$.

Finally, the applicability of the Onset program toward emission curve λ_{onset} estimations on the high energy side of the curve are shown in Figure 6. This analysis is important since

non-singlet state emitters often have the first triplet excited state energy level (T_1) values estimated from the emission curve rather than the absorption curve as was used in all of the prior examples. The λ_{onset} of the emission curve is an applicable method of estimating the energetic distance from the T_1 to the ground state energy level (S_0) for compounds exhibiting intersystem crossings (Figure 1, case 3). Ir(ppy)₃ provides a key example of the utility of the Onset program. Ir(ppy)₃, when excited at 490 nm, results in an emission λ_{onset} estimate at 461 nm via the Onset program. The higher energy emission than excitation observed by the compound is due to the photoexcitation of an electron in a higher S_0 vibrational energy level followed by an emission to a lower S_0 vibrational energy level. As with any compound that participates in intersystem crossing, the absorption and emission transitions are from different spin states which does not allow for the use of E(0—0) values if the curves cross. Thus, the use of the Onset program is attractive for estimating the T_1 energy level. When finding the emission curve λ_{onset} , the high energy side is used which Onset is amenable to with a simple selection click for x-axis direction shift set to “decreasing.” Figure 3.6 shows the emission curve obtained with Ru(bpy)₃Cl₂ with a human estimated onset value of 550 nm. The Onset estimation is in close agreement with this value at 551 nm.

3.5 Conclusions

Previous λ_{onset} values are in error from accurately computed x-intercepts of lines tangent to the inflection point of the lowest energy λ_{max} by an average of -2 nm. While, this is not a tremendous error, this present analysis shows that for the sample set, human error tends to lean towards longer wavelength values artificially decreasing the associated E_g^{opt} and optical energy gaps by as much as 5%. The percent errors between the human-derived and computed λ_{onset} are typically within 1%, but future analysis could benefit from a more robust and, most notably, more consistent approach for determining the onset value, especially for applications to shorter wavelengths and higher energies into the UV region like those for Mn(bpy)(CO)₃Br, Ir(ppy)₃, and Mn(bpy)₂Br₂ reported herein.

Finally, this program makes use of standard .csv files, is built on an HTML interface,

is available for the both the Windows and Unix-based operating systems, and is freely downloadable via GitHub. This provides easy access and usage to experimental groups who are determining the spectra of novel, synthesized compounds with application to next-generation optical materials. Most importantly, this software has the potential to unify how λ_{onset} values are obtained since very often these are only estimated by simply looking (without even human-drawn tangents), a point of discussion about how the value is obtained since methods can vary widely between research teams.

Chapter 4

Linear and Helical Carbonic Acid Clusters

Note: This chapter contains unmodified text and figures from the following journal article and has been reproduced with permission.

A. M. Wallace and R. C. Fortenberry, “Linear and helical carbonic acid clusters”, *The Journal of Physical Chemistry A* **125**, PMID: 34029078, 4589–4597 (2021)

4.1 Introduction

Carbonic acid is a simple organic molecule examined for numerous applications, but even such a simple, well-studied system can still have surprising properties. From its association with ocean acidification and CO₂ transport in blood, carbonic acid is prevalent on earth. However, it also plays roles in extraterrestrial environments, such as potentially on the surface of Mars with implications for both organic and inorganic chemistry [33]. Within the vacuum of space, carbonic acid may exist as an ice especially since both CO₂ and water ices (as well as their mixture) are well-known TODO: add citations to bib [34–37], but it has yet

to be directly observed beyond the Earth.

Experimentally, trace amounts of carbonic acid have been produced through the irradiation of H₂O and CO₂ ice (1:1) after slow warming from 20 K and identified through

infrared spectroscopy at 215 K simulating various astrochemical conditions[38]. Solid carbonic acid has also been produced through acid-base reactions between HBr and KHCO_3 under vacuum at low-temperature[39]. Ultimately, carbonic acid has been shown to form readily through ionizing radiation and vacuum-UV light on $\text{H}_2\text{O}+\text{CO}_2$ ice[40]. As noted by Peeters and coworkers, both H_2O and CO_2 have been found in the outer Solar System on satellites orbiting Jupiter, Saturn, Uranus and Neptune[39]. These environments have radiation which could enable the formation of carbonic acid. However, detection of carbonic acid has yet remained elusive implying that any possible observations of this material might require it to exist in a solid, ice phase or as small clusters in the gas phase.

If it exists in space, solid carbonic acid is almost certainly in some largely, amorphous solid phase under interstellar or interplanetary conditions. However, smaller clusters may nucleate in motifs similar to bulk solids closer to STP. Carbonic acid was thought to reside in one of two solid phases, α and β , with the latter having received much more analytical scrutiny. The β phase was thought to produce macroscopic tendrils when observed via microscopy whereas the α phase produced clumps of material [41–43]. However, recent work with more advanced experimental techniques has shown that α -carbonic acid is actually the monomethyl ester of carbonic acid and is not a polymorph negating the need for any further analysis of the α phase [44]. Even so, efforts to describe β -carbonic acid computationally have been successful in correlating with experiment for infrared and Raman spectra of a linear octamer [45]. Attempts to construct a crystal structure through molecular dynamic methods suggest that the crystals with “sheet-like hydrogen bonding topologies” are among the most stable[46].

However, experimental characterization for nucleating solids from clusters of carbonic acid may require other spectroscopies that are more sensitive to structure than infrared or Raman techniques. Therefore, exploration into the ultraviolet and visible spectra of small clusters of carbonic acid may provide more options for detecting carbonic acid and its isomers in outer Solar System environments and beyond. The present work will provide structural and electronic spectral characterization for associations of carbonic acid clusters.

4.2 Methods & Computational Details

The present study focuses on computing the lowest energy structures for n -mers of carbonic acid for $n = 1 - 6$ with selected higher n values. Additionally, excited states for the carbonic acid monomer, dimer, and some higher clusters of carbonic acid will also be explored. The geometry optimizations, energies, and harmonic zero-point vibrational energy corrections for the carbonic acid molecules are determined primarily by the ω B97XD method with its long-range functionality [47, 48] and the 6-31+G* basis set [49–52] both computed through Gaussian16 [53]. This method and basis set are standard for computing clusters of molecules where intermolecular forces between molecules are necessary, like that for carbonic acid here, and ω B97XD captures both short-range and long-range interactions more accurately than B3LYP [54].

After optimizing the geometries, excited states are computed using time dependent density functional theory (TD-DFT) [55, 56] with the B3LYP method [57–59] and 6-311G* [60] basis set, again, through Gaussian09. A more accurate calculation for excited states is accomplished for some clusters as noted in the discussion through equation-of-motion coupled-cluster theory at the singles and doubles level (EOM-CCSD) [61–63] run through Molpro [64] and the aug-cc-pVDZ basis set [65] [66]. Although B3LYP is perceived not to be as systematically accurate as EOM-CCSD, B3LYP provides a less computationally expensive calculation for the larger carbonic acid systems. The present purpose is to determine excitation behavior changes as the structures become larger. Therefore, EOM-CCSD is used as a benchmark calculation on the monomer, dimer, and tetramer structures in order to indicate a subsequent energy value shift on the energies generated by the TD-DFT/B3LYP computations. The excited state calculations provide spectral information for the isomers and increasing cluster sizes. The Kohn-Sham orbitals and molecular orbitals are plotted in the figures present in the next section, respectively, for the TD-DFT and EOM-CCSD calculations with the 6-311G* and aug-cc-pVDZ basis sets, also respectively. These orbitals are investigated in order to examine the behavior of the excitations within a series of increasing monomer units with the aim to establish a link between small carbonic acid systems and extrapolating to approximate bulk behavior.

4.3 Results & Discussion

4.3.1 Structures

The relative carbonic acid monomer energies are shown on Figure 1 with the most stable structure depicted on the far left. The *syn-syn* carbonic acid monomer (0.00 eV) is the lowest energy structure followed by the *syn-anti* (+0.08 eV, in line with previous computational results [41, 67]) and then the *anti-anti* (+0.49 eV), according to the ω B97XD/6-31+G*. The *syn-syn* molecules have H-O-C *syn* angles that are 108.9° with the C-O, C=O, and O-H bonds being 1.333 Å, 1.209 Å, and 0.965 Å, respectively. The *syn-anti* molecule has one H-O-C *syn* angle at 107.4° with the associated C-O bond being 1.351 Å. The other H-O-C *trans* angle is 110° and has C-O and C=O bond distances of 1.333 Å, and 1.199 Å. The *anti-anti*, has two H-O-C *down* angles at 114.476° with the C-O, C=O, and O-H bond lengths of 1.352 Å, 1.192 Å, and 0.962 Å. These relative energies vary by less than 0.02 eV compared with QCISD(T)/6-311++G** results from previous research by Zapata-Escobar et al.[68] implying that the present approach is fitting for analysis of larger clusters.

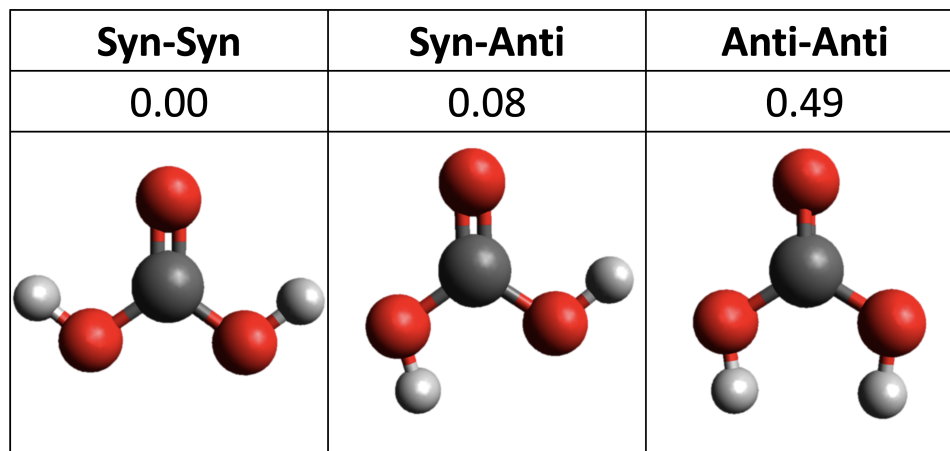


Figure 4.1: The monomer isomers with the relative energy in eV increasing from left to right

For the dimers, the two *syn-syn* molecules engaging in hydrogen bonding have the lowest relative energy also in line with previous work [67]. This strong connection is due to two hydrogen bonds forming a hexagonal ring while the outside hydrogens remain in the *syn* position. Additionally, the molecules both lie flat on the same xy-plane due to the trigo-

nal planar sp^2 hybridized carbon in the center of each molecule. One can compare the two *syn-anti* molecules - D3 and D4 from Figure 4.2 - to show that hydrogen bonding with one H-O-C *syn* hydrogen is more energetically favorable than having both H-O-C *anti* hydrogens engaged in making the two hydrogen bonds. These observations support the *syn* position as the lowest energy form for the hydrogen and ketone. Ultimately, the structures containing the *anti-anti* motif are all relatively high in energy.

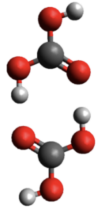
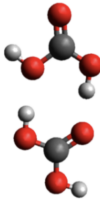
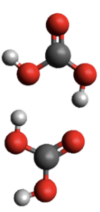
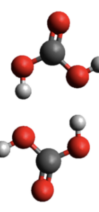
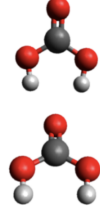
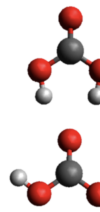
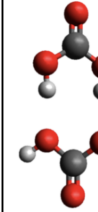
D1	D2	D3	D4	D5	D6	D7
0.0	0.33	0.39	0.60	1.21	0.80	0.89
						

Figure 4.2: The dimer isomers explored in this work with the relative energy in eV increasing from left to right

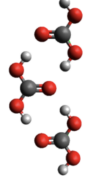
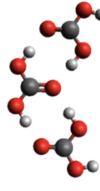
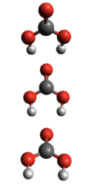
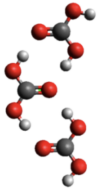
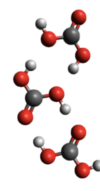
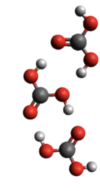
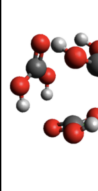
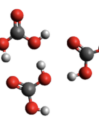
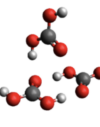
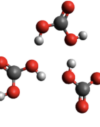
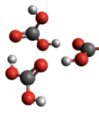
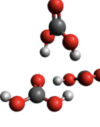
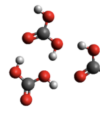
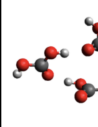
Tri1	Tri2	Tri3	Tri4	Tri5	Tri6	Tri7
0.0	0.10	1.66	0.05	0.45	0.16	1.13
						
Tri8	Tri9	Tri10	Tri11	Tri12	Tri13	Tri14
0.75	0.43	1.11	0.49	0.89	0.63	0.57
						

Figure 4.3: The trimer isomers with the relative energy in eV

The lowest relative energy for the trimer is a *syn-syn* molecule adding onto the D1 structure in Figure 4.2, which is the beginning of a ribbon structure. This ribbon structure

is extended with more *syn-syn* molecules. At three molecules, three-dimensional isomers begin to arise; however, these are all higher in energy than the ribbon structure. Comparing the energies of Tri1, Tri2, and Tri3 in Figure 3, the cost of a *syn-anti* molecule at the end of the ribbon is about 0.5 eV. When the *anti-anti* trimer completes a circuit as in Tri7, it becomes more stable compared to Tri3, albeit higher in intermolecular angular strain, since the hydrogens are all engaging in hydrogen bonding. This example highlights a pattern that when more hydrogens are unable to engage in hydrogen bonding, the relative energy increases significantly. Therefore, although the three-dimensional structures are higher in relative energy; these structures might become more favorable in a solid structure where the non-interacting hydrogens on the outside no longer cause an increase in relative energy from not engaging in hydrogen bonding.

An exception to the *syn-syn* molecule being lower in energy than the *syn-anti* is found when comparing Tri11 and Tri14. These molecules engage in a similar pattern and exist in nearly the same space when overlapped. However, the Tri11 structure is about 0.75 eV lower in energy. In any case, as the ribbon motif is maintained going to larger clusters (i.e. longer ribbons) of carbonic acid dimers, the structures maintain C_{2h} symmetry.

Further exploration of larger systems with more *syn-anti* and *anti-anti* molecules shows that the relative energies increase when the system is not comprised solely of the *syn-syn* molecules. The supplemental information contains more isomers and relative energy figures for the tetramer and larger. From the dimer to dodecamer, the *syn-syn* forms are the lowest relative energy, with the ensuing ribbon structure being the lowest isomeric form, until stacking two hexamer ribbons. The stacking of ribbons can be seen in Figure 4 and is the three-dimensional form of the ribbon structure. When stacking the ribbons, the tips of a ribbon bend toward the tips of the other ribbon, for example the 12 Molecules Ribbon image in Figure 4.4. The oxygens are puckering in an attempt to hydrogen bond with hydrogens present that lack a hydrogen bond interaction, and this is likely only an artifact of the present computations. The centers of the ribbons keep a distance of about 3.0 Å.

Another viable, three-dimensional *syn-syn* isomer is the spiral. The structure contains the most favorable *syn-syn* dimer, like the ribbon; however, instead of extending in one dimension, dimers engage in hydrogen bonding from the side. The spiral requires a minimum

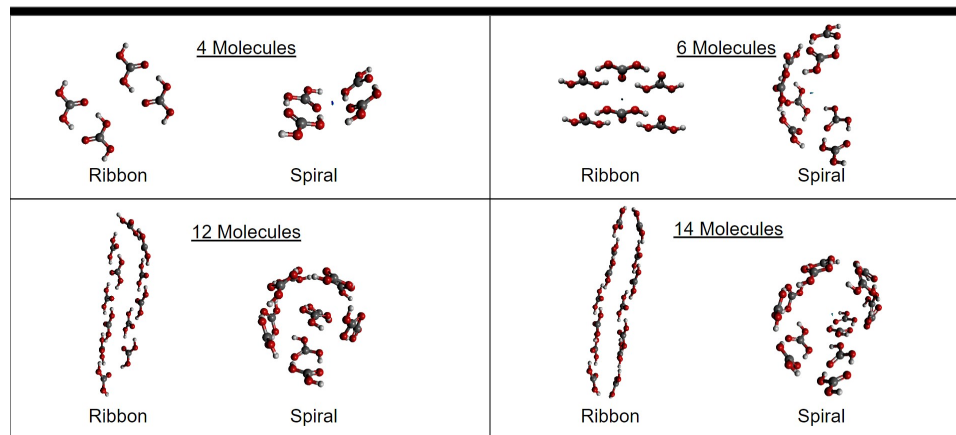


Figure 4.4: The stacked ribbon and spiral geometries

of four carbonic acid molecules in a slipped position with one rotated out of the plane by 40.4° as shown in Figure 4. It maintains C_2 symmetry even as more dimers are added to the cluster in this motif. The spiral structure is inspired by Candidate V from previous work by Reddy et al.[46]; however, the geometry optimizations in the present study on such a molecular motif settled upon the ribbon structure. One complete spiral of the helix requires 7 dimers - 14 carbonic acid molecules.

These two main isomeric structures - the stacked ribbon and spiral - comprised of only the *syn-syn* molecules and extending in three-dimensions are compared in relative energies and binding energies in Figures 4.5 and 4.6. The ribbon and the spiral structures both have the strong hexagonal ring between adjacent *syn-syn* molecules. Hence, both have the lowest energy dimer formation within them, and the structures deviate from each other by the location where more dimers are added. Additionally, both the stacked ribbon and the spiral attempt to minimize the amount of hydrogens not engaging in hydrogen bonding. The ribbon is an extension of *syn-syn* monomers from the D1 structure above in Figure 4.2. Stacking these ribbons produces the structure as is seen in Figure 4, which is a three-dimensional chain. When stacking the ribbons, the tips of a ribbon bend toward the tips of the other ribbon, for example the 12 Molecules Ribbon image in Figure 4.4. The oxygens are puckering in an attempt to hydrogen bond with hydrogens present that lack a hydrogen bond interaction, and this is likely only an artifact of the present computations. The centers of the ribbons keep a distance of about 3.0 \AA .

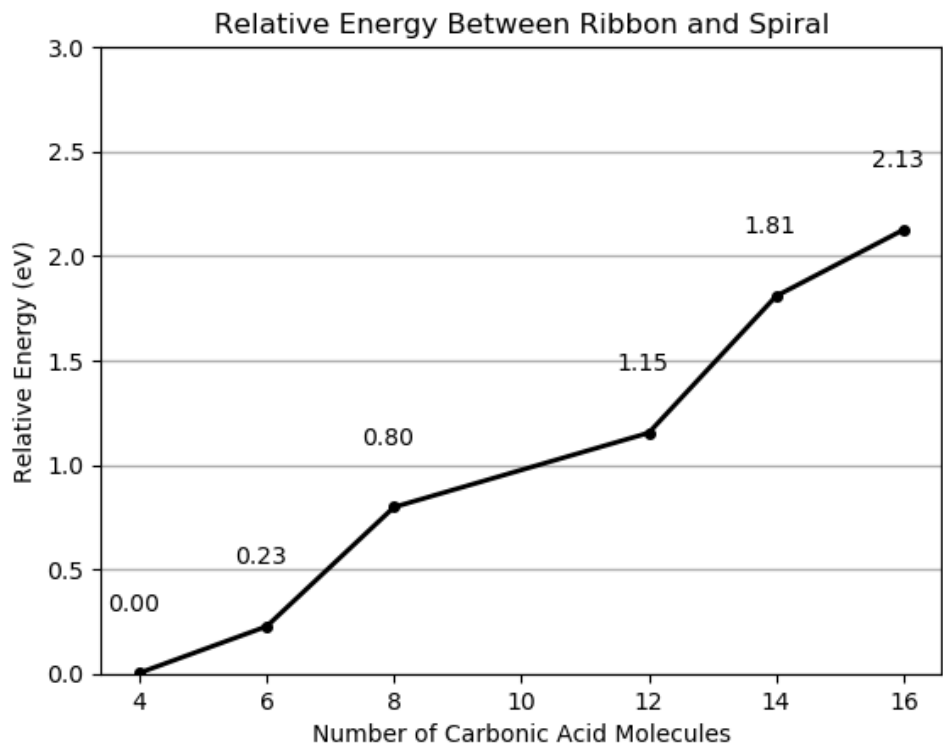


Figure 4.5: The relative energy (eV) between ribbon and spiral structures according to ω B97XD/6-31+G*

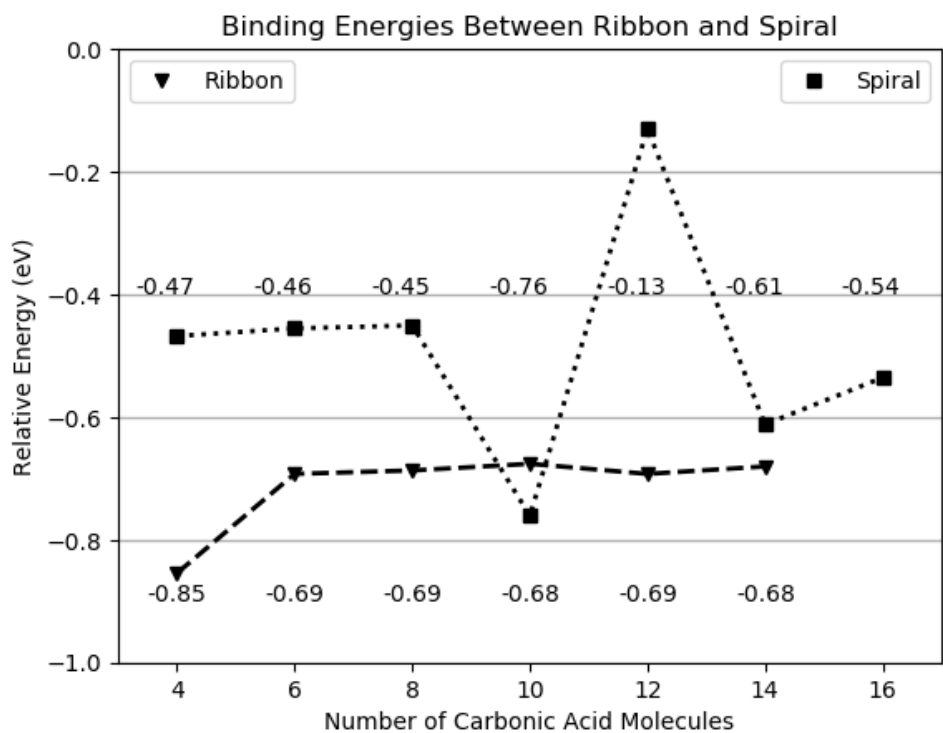


Figure 4.6: The binding energies for the spiral and ribbon with adding dimers

These two main isomeric structures - the stacked ribbon and spiral - comprising of only the *syn-syn* molecules and extending in three-dimensions are compared in relative energies and binding energies in Figures 5 and 6. The ribbon and the spiral structures both have the strong hexagonal ring between adjacent *syn-syn* molecules. Hence, both have the lowest energy dimer formation within them, and the structures deviate from each other by the location where more dimers are added. Additionally, both the stacked ribbon and the spiral attempt to minimize the amount of hydrogens not engaging in hydrogen bonding. Comparing the relative energies of these two structures in Figure 5 shows the ribbon structure as the more favorable form, since the consistent positive energy increase indicates favorability toward the stacked ribbon structure as the clusters become larger. Therefore, the binding energies for the ribbon structure are expected to be lower than that of the spiral.

From Figure 6, the stacked ribbon and spiral binding energies behave differently. The ribbon's binding energy approaches a constant value around -0.75 eV for every dimer added past 6 total carbonic acid molecules in the cluster. The spiral has two regions with a slight trend upwards separated by a discrepancy at the decamer and dodecamer. This discrepancy is the energy manifestation of the spiral helix being one dimer away from making a complete circulation. Ultimately, the stacked ribbon structure is more favorable in energy than the spiral structure as the cluster grows in size because the ribbon's average binding energy is -0.71 eV, while the spiral's average binding energy is -0.70 eV. Consequently, the ribbon will decrease in energy at a faster rate than the spiral when extending the size of the cluster.

4.3.2 Excitation Energies

Ribbon Excitation Energies

The electronic excitations for carbonic acid clusters of both ribbon and spiral motifs are calculated and their molecular orbitals plotted to aid in characterizing how the excitation energies shift as the cluster sizes increase towards the bulk. Additionally, the energies between the EOM-CCSD/aug-cc-pVDZ and the B3LYP/6-311G(d,p) excited state calculations are compared through tables and spectral graphs of the ribbon monomer, dimer, tetramer, and spiral tetramer resulting in a mean absolute difference of 0.18 eV.

Table 4.1: Excited State Data (eV) for the *syn-syn* Monomer

Excited State	EOM-CCSD/aug-cc-pVDZ		B3LYP/6-311G(d,p)		Energy Difference
	Exc. Energy	f	Exc. Energy	f	
1 1A_1	8.49	0.017	8.42	0.004	0.07
2 1A_1	9.28	0.222	9.70	0.165	-0.42
1 1B_1	8.15	0.001	8.19	0.001	-0.04
2 1B_1	10.10	0.053	9.99	0.016	0.11
1 1B_2	7.54	0.043	7.54	0.038	0.00
2 1B_2	9.72	0.001	10.05	0.161	-0.33
1 1A_2	7.22	0.000	7.04	0.000	0.18

Table 4.2: Excited State Data (eV) for the Dimer

Excited State	EOM-CCSD/aug-cc-pVDZ		B3LYP/6-311G(d,p)		Energy Difference
	Exc. Energy	f	Exc. Energy	f	
1 1A_g	8.06	0.000	8.23	0.000	-0.17
2 1A_g	8.74	0.000	8.75	0.000	-0.01
1 1A_u	7.58	0.001	8.35	0.000	-0.78
2 1A_u	8.46	0.001	8.69	0.001	-0.23
1 1B_u	8.08	0.086	8.27	0.065	-0.19
2 1B_u	9.11	0.303	9.57	0.322	-0.46
1 1B_g	7.59	0.000	7.34	0.000	0.24

The *syn-syn* monomer has an excited state at 9.28 eV with the highest oscillator strength (f) of 0.22. This f value is much larger than the other noticeable excitations at 10.10 eV with an f of 0.053 and at 7.54 eV with an f of 0.043. The B3LYP/6-311G(d,p) and EOM-CCSD/aug-cc-pVDZ calculations vary by about ± 0.5 eV with the EOM-CCSD calculation likely providing a more accurate number [69]. For instance, EOM-CCSD computes the second 1 1A_1 to be 9.28 eV, while B3LYP is 9.70 eV with a -0.42 eV difference.

Next, Table 2 shows the excited states for the most stable dimer from Figure 4.2. The most noticeable excited states are both 1B_u with the EOM-CCSD energies of 8.08 eV and 9.11 eV with the respective oscillator strengths of 0.086 and 0.303. These values differ from the B3LYP/6-311G(d,p) energies by less than 0.5 eV with the lower energy value at 8.27 eV and the higher energy point at 9.57 eV for the TD-DFT computations. The relatively large f value on the 1B_u 9.11 eV excitation likely produces a detectable signal.

In Table 3, the excited states for the lowest energy form of the ribbon tetramer are displayed. Once again the EOM-CCSD/aug-cc-pVDZ values for two of the 1B_u excitations are around the same energies as the TD-DFT excited states, although f decreases. The lower energy value for the EOM-CCSD calculation reports excitations at 8.14 eV and 8.76

Table 4.3: Excited State Data (eV) for the Ribbon Tetramer

Excited State	EOM-CCSD/aug-cc-pVDZ		B3LYP/6-311G(d,p)		Energy Difference
	Exc. Energy	f	Exc. Energy	f	
1^1A_g	8.14	0.000	8.31	0.000	-0.17
2^1A_g	8.66	0.000	8.77	0.000	-0.11
1^1A_u	7.58	0.001	7.33	0.001	0.25
2^1A_u	7.93	0.000	8.26	0.000	-0.33
1^1B_u	8.14	0.094	8.31	0.067	-0.17
2^1B_u	8.76	0.031	9.51	0.073	-0.75
1^1B_g	7.58	0.000	7.33	0.000	0.25

eV with the oscillator strengths of 0.093 and 0.031, respectively. These relate to the B3LYP energies of 8.31 eV and 9.51 eV. From these points, a trend is starting to emerge with the increase of the ribbon's size having negligible effects on the excitation energies. Figure S3 depicts this data visually on an absorption plot. Since the systems start to become too costly to run EOM-CCSD/aug-cc-pVDZ, the electronically excited states of the larger ribbon structures are calculated with only B3LYP/6-311G(d,p).

In Figure 4.7, the 1^1B_u excitation energy for the ribbon clusters around 8.3 eV are displayed. Excluding the monomer, the excitation energies are approaching an asymptote at 8.37 eV, while the oscillator strength stays approximately the same after 7 carbonic acid molecules are in the cluster as shown in Figure 4.8. However, in order to ensure that these excitations are the same excitation throughout the polymers, the orbitals are plotted in Figure 4.9. Comparing the highest contributors of the dimer to the ribbon tetramer excitations in Figure 4.9, the orbitals are what would be expected for similar excitations in more distributed systems. The most notable excited state for the dimer is the 1^1B_u state at 8.27 eV and is comprised of molecular orbital (MO) contributions from an in-plane π orbital to a σ^* MO with a character of 0.52. These orbitals are the HOMO-1 and the LUMO+1. This behavior is mimicked in the ribbon tetramer 1^1B_u excitation at 8.31 eV from the HOMO -1 and the LUMO+4; however, the in-plane π orbital is centralized on the ending two carbonic acid molecules. This has a 0.49 character of the excited state making it the largest contributor. Looking at the second-highest contributors, the dimer and ribbon tetramer have another in-plane π to a σ^* excitation. However, the in-plane π orbital is distributed across all of the molecules in the dimer and tetramer alike. This contribution for dimer and ribbon tetramer is 0.49 and 0.46, respectively. Overall, the energy difference between the dimer

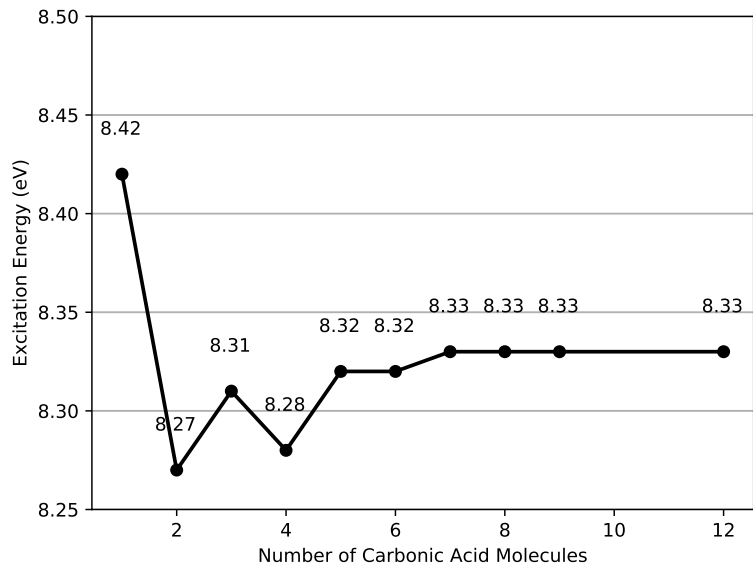


Figure 4.7: The 1B_u excitation energy for ribbons at 8.3 eV

and tetramer excitations is 0.04 eV and increases from the dimer to tetramer, but these are the same type of excitation.

Next, the tetramer ribbon and the hexamer ribbon excitation around 8.3 eV are compared. The tetramer excitations described above are similar to the two, corresponding 1B_u excitations from the hexamer, with excitation energy for the hexamer at 8.32 eV, a 0.01 eV increase. The contributions are 0.48 for the HOMO-3 to LUMO+5 and 0.46 from the HOMO-2 to the LUMO+6. These contributions are in very close agreement with the tetramer. After analyzing the orbitals for the octamer and dodecamer as well, Figure 4.7 is depicting the same type of excitation.

The Figures 4.10-4.12 depict a similar analysis to the excited state described above; however, the 1B_u excitation around 9.6 eV for ribbons is investigated. From Figure 4.10, extrapolation of the points appears to give an asymptotic convergence to around 9.70 eV as the number of carbonic acids approaches the bulk. 4.10, when extrapolated with a natural logarithmic function - excluding the monomer and trimer, shows that the excitation energy will top out around 9.70 eV when the number of carbonic acid molecules becomes large. Additionally, the oscillator strength seems to be dampening towards a constant value between 0.2 and 0.3 as the system becomes larger.

The molecular orbitals are displayed in Figure 4.12 to determine if the excitations are

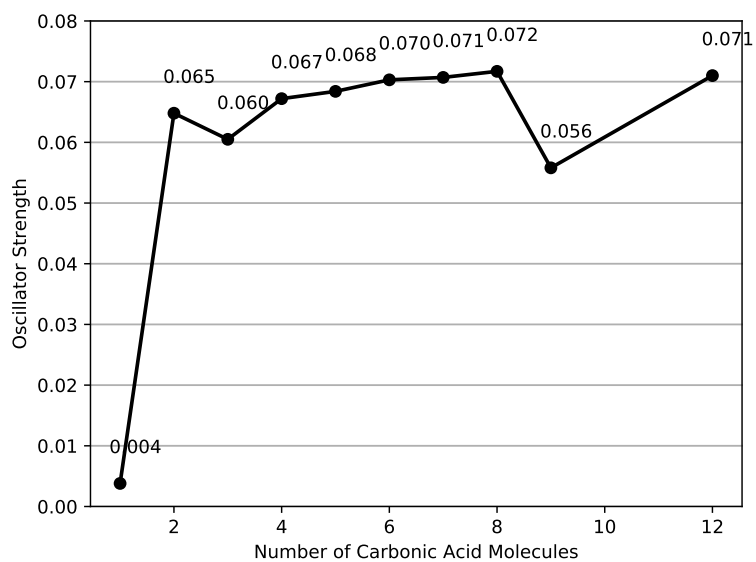


Figure 4.8: The associated oscillator strength for the 1B_u excited state for ribbons at 8.3 eV

Ribbon B_u Excitation 8.3 eV

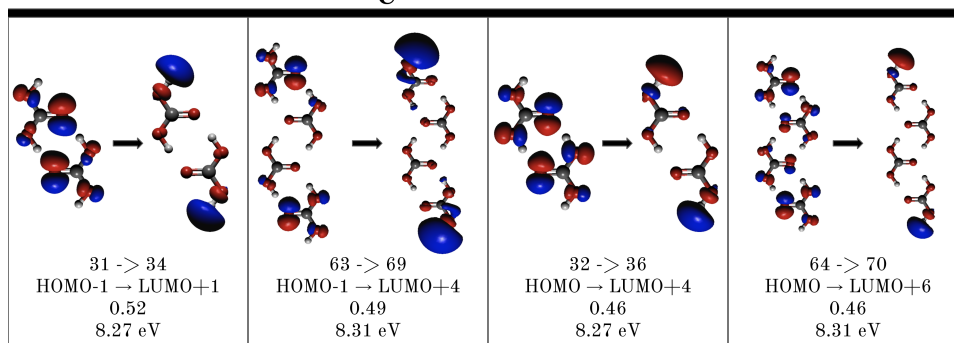


Figure 4.9: The molecular orbitals for the dimer and tetramer 1B_u excitation around 8.3 eV is depicted here.

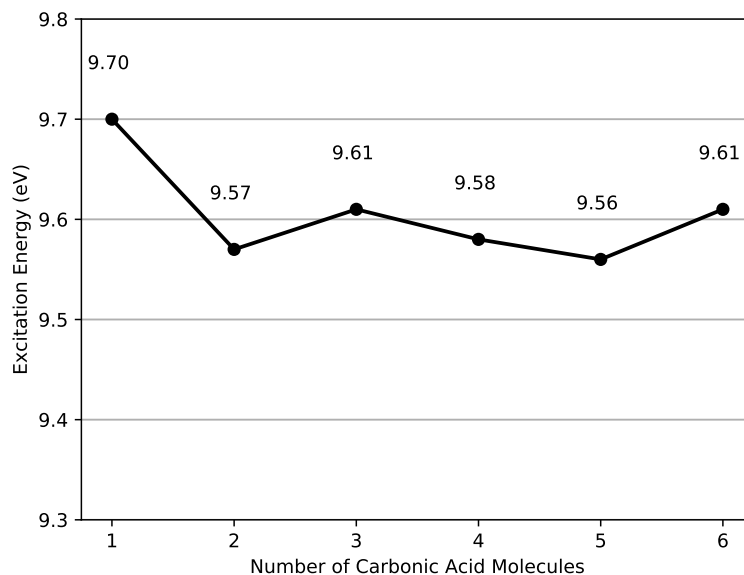


Figure 4.10: The 2B_u excitation energy for ribbons at 9.6 eV

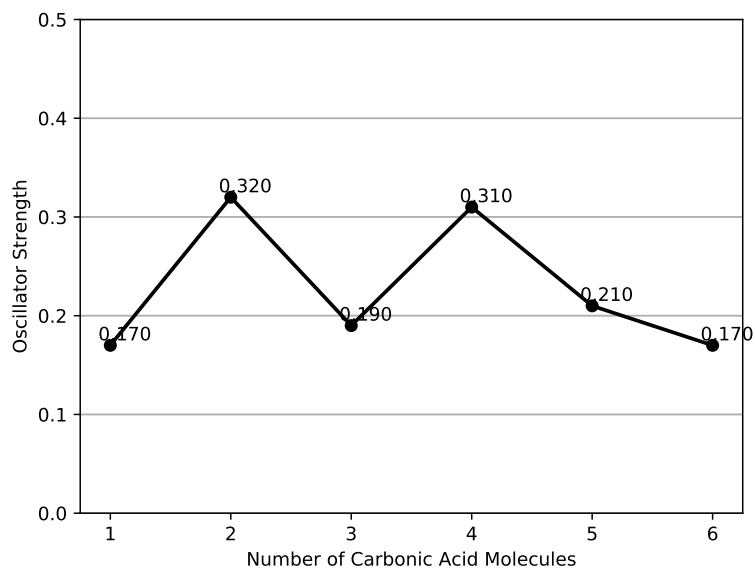


Figure 4.11: The associated oscillator strength for the 2B_u excited state for ribbons at 9.6 eV

Ribbon B_u Excitation 9.6 eV

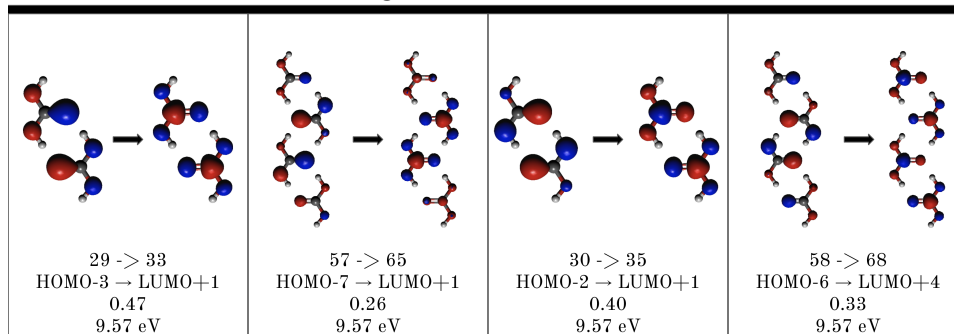


Figure 4.12: The molecular orbitals for the dimer and tetramer 2B_u excitation around 9.6 eV is depicted here.

similar. For the dimer 1B_u excitation, the energy is 9.57 eV and is an out-of-plane, in-phase π to out-of-plane, in phase π^* excitation. The dimer excited state is the HOMO-3 to LUMO+1. the HOMO-3 to LUMO+1? This description matches the highest contributor of the ribbon tetramer's 1B_u excitation at 9.57 eV from HOMO-7 to LUMO+1. The contributions from the dimer and tetramer are 0.47 and 0.26. The second highest contributors are out-of-phase, out-of-plane π to out-of-phase, out-of-plane π^* 1B_u excitation. These contributions are 0.40 and 0.33. Therefore, these two excitations are of the same type.

Moving on to the larger systems for comparison, the hexamer has a 1B_u excitation energy of 9.61 eV. The tetramer HOMO-7 to LUMO+1 orbitals involved in the excitation match that of the hexamer HOMO-11 to LUMO+1 orbitals and has a contribution of 0.13. Meanwhile the tetramer HOMO-6 to LUMO+3 excitation matches the largest contributor of the hexamer of orbital 86 to 100 - HOMO-10 to LUMO+3 - with the contribution of 0.24. For this 1B_u excitation, the contributions from more orbitals increases with the carbonic acid molecules present. Therefore, although these contributions seem small, they are still the largest contributors to the excitation. Overall, these 1B_u excited states are around the same energy with the addition of many carbonic acid molecules.

Spiral Excitation Energies

Next, the spiral tetramer structure has a 1B excited state with an energy and oscillator strength of 7.36 eV and 0.001 for EOM-CCSD/aug-cc-pVDZ and 7.13 eV and 0.001 for

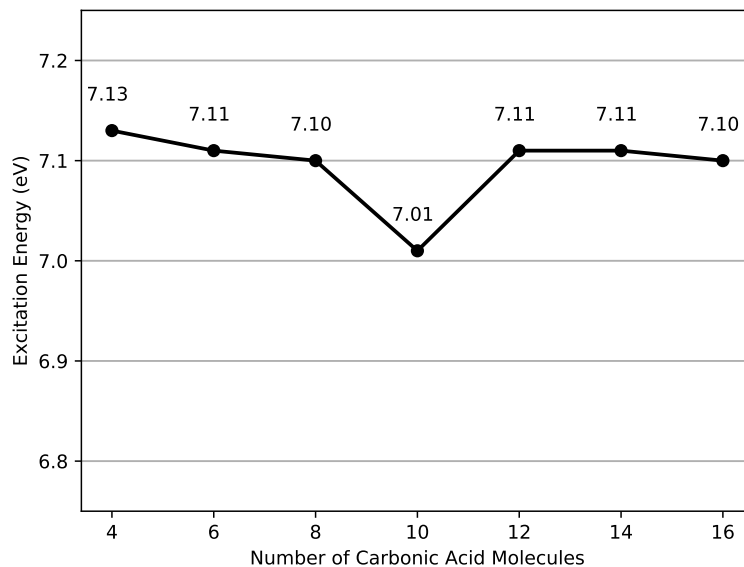


Figure 4.13: The 1B excitation energy for spirals at 7.1 eV

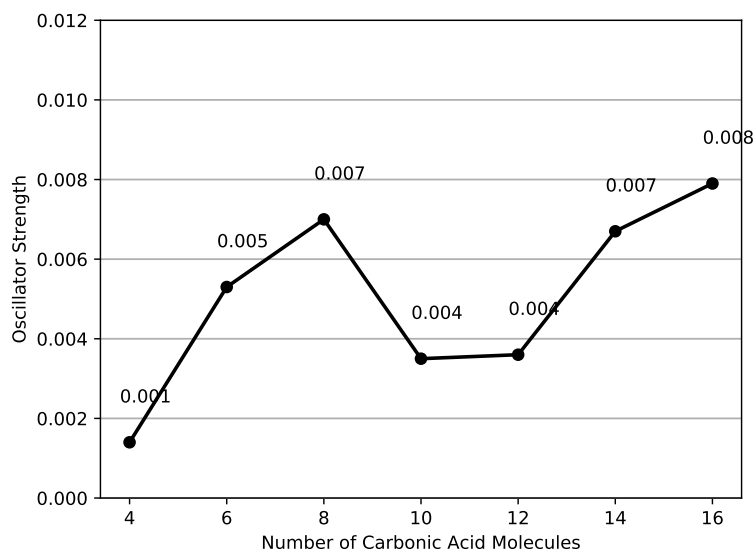


Figure 4.14: The associated oscillator strength for the 1B excited state for spirals at 7.1 eV

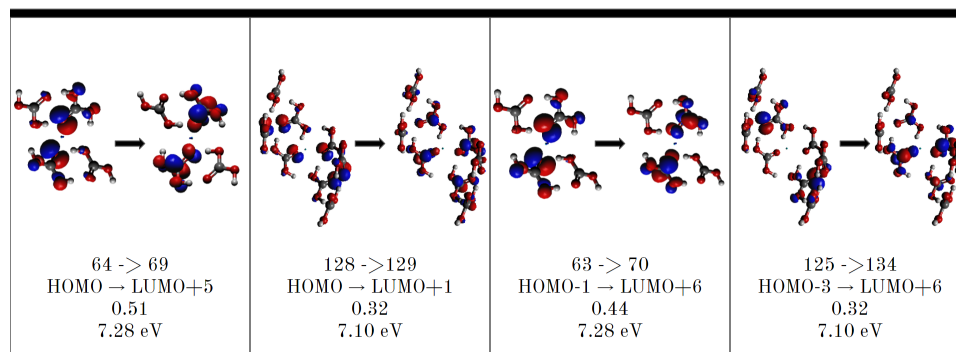


Figure 4.15: The molecular orbitals for the tetramer and octamer spiral structure 1B excitation is shown here.

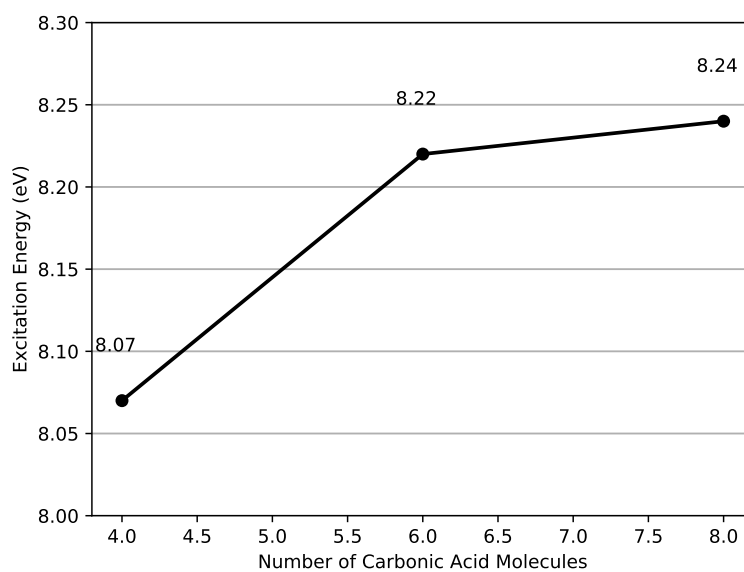


Figure 4.16: The 1B excitation energy for spirals at 8.1 eV

B3LYP/6-311G(D,p). These energies differ by 0.23 eV. This is the only spiral oligomer that has both EOM-CCSD/aug-cc-pVDZ and B3LYP/6-311G(d,p) excitation energies due to costs of the EOM-CCSD calculation. The plotted molecular orbitals both have a large contribution from an in-plane π to out-of-plane π^* excitation.

For the spiral excitations, the Figures 4.13 and 4.14 plot the excitation energy and f of the 1B excitation around 7.1 eV. Except for the dip at 10 molecules, the spiral has an excitation that decreases slowly to a value around 7.1 eV. The f starts very weak, but increases slightly as the structure becomes larger.

The molecular orbitals in Figure 4.15 show that these excitations are the same as the

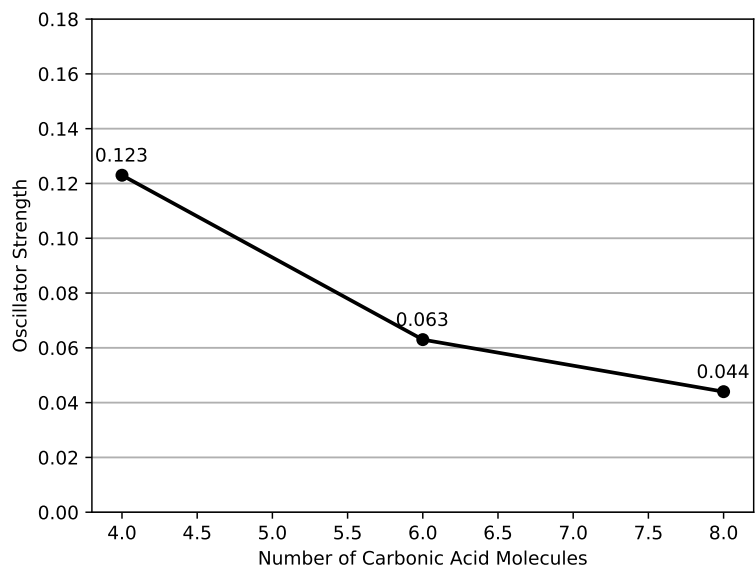


Figure 4.17: The associated oscillator strength for the 1B excited state for spirals at 8.1 eV to 8.2 eV

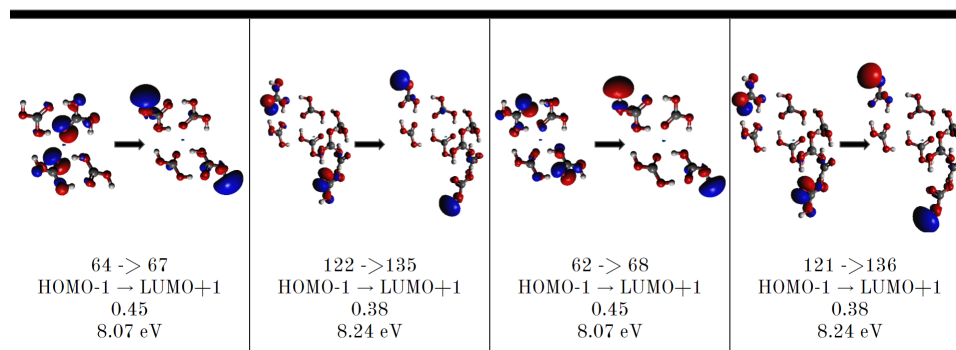


Figure 4.18: The molecular orbitals for the tetramer and octamer spiral structure 1B excitation is shown here.

structure increases. The major contributors are, again, both in-plane π to out-of-plane π^* excitation. For the tetramer, the orbital excitations are from the HOMO to LUMO+4 with a 0.51 contribution and HOMO-1 to LUMO+5 with a contribution of 0.44. The larger contributor's excited state is from the HOMO to the LUMO+4, and the second highest contributor is the HOMO-1 to the LUMO+5. For the octamer, the major contributors are the HOMO to the LUMO and the HOMO-4 to LUMO+6 both with the respective excitation characters of 0.32.

Finally, the last notable excitation in the spiral is around 8.1 to 8.2 eV. Figure 4.16 shows an upward trend in energy as the structure increases in size; meanwhile the f decreases as shown in Figure 4.17. From Figure 4.18, the molecular orbitals that are largely responsible for this excitation both start in an in-plane π orbital and end in a σ^* orbital; however, in both instances the phases are different. The tetramer larger contributor is from orbital 64 to 67 - HOMO to the LUMO+2 - with 0.45, while the second contributor is from orbital 62 to 68 - HOMO-2 to LUMO+3 - with 0.45 as well. Within the octamer, the larger contributor that lines up with the tetramer is from orbital 122 to 135 - HOMO-6 to LUMO+6 - with 0.38 and the second contributor is from orbital 121 to 136 - HOMO-7 to LUMO+7 - with 0.38.

4.4 Conclusions

A new motif for the early clustering of carbonic acid incorporates the lowest energy dimer into a spiral or helical pattern. While this structural behavior is not the most energetically favorable when compared to the ribbon (or linear) motif, this novel arrangement of carbonic acid molecules costs roughly only about 0.2 eV of energy for each dimer unit added. Furthermore, the helical oligomerization is chiral and would induce optical activity in any incident light interacting with such material. This chirality combined with the importance and ubiquity of the constituent water and carbon dioxide molecules implies that the helical spiral motif of carbonic acid may have implications for astrobiology and origins of life studies, especially in evaporation studies of photoprocessed astrophysical ices. Laboratory differentiation of these motifs is likely straightforward with the helical spiral

structure absorbing at lower UV energies (< 8.1 eV) and the linear, ribbon form absorbing at higher energies (> 8.1 eV). Ultimately, these data should assist in detecting small clusters of carbonic acid as they begin to nucleate in various extended systems.

Chapter 5

Computational UV Spectra for Amorphous Solids of Small Molecules

Note: This chapter contains unmodified text and figures from the following journal article and has been reproduced with permission.

A. M. Wallace and R. C. Fortenberry, “Computational uv spectra for amorphous solids of small molecules”, *Physical Chemistry Chemical Physics*, 24413–24420 (2021)

5.1 Introduction

In cold interstellar regions and protoplanetary disks, amorphous solids exist due to the lack of energy required to produce crystalline solids. While specific molecular concentrations vary depending on the astrophysical environment, H₂O, H₂CO, N₂, CO, O₂, CO₂, H₂O₂, CH₄, and NH₃ are the primary constituents of grain mantles[71–75] that exist as amorphous solids in low temperatures. These amorphous solids can act as a surface upon which molecules accumulate and reaction pathways are accelerated. At temperatures around 10 K, most molecules—except H₂ and He—that collide with these ice analogs will stick to the surface[76]. More specifically, the surface of amorphous water has nanopores that have strong binding sites which allow adsorption of molecules, such as CO, permitting additional surface chemistry.[77, 78] Due to increased flexibility, amorphous solids can behave as superior catalysts compared to their corresponding crystalline solid form[77]. Consequently,

interstellar amorphous solids provide an environment for bringing molecules together and, subsequently, increasing their reactivity.

Additionally, amorphous solids can exist as ices and provide the material for forming larger molecules independently. Ultraviolet photolysis of ice analogs consisting of H₂O, CH₃, NH₃ and CO ultimately produces H₂CO₃, CO₂, CH₄, HCO and more complex molecules in simulated interstellar environments[79–82]. These products and other similar small molecules containing carbon atoms are likely precursors for forming larger organic and biologically relevant molecules in the interstellar medium (ISM)[76, 79, 83–89] Some of these compounds include glycine, alanine, and serine[90] with reaction pathways explored computationally[91]. Furthermore, previous theoretical and laboratory work shows that methane, ethylene, and acetylene ices can undergo radical reactions to produce larger alkanes[92]. Therefore, better ice analog characterization can lead to increased understanding of extraterrestrial environments in which organic residues form.

While many computational approaches exist that describe amorphous solids with molecular dynamics and machine learning[93–96], few attempt to use higher-levels of theory, such as density functional theory (DFT), for describing amorphous solids due to the high computational costs. Attempts at combining DFT and machine learning have produced favorable results for describing *ta*-C surfaces[97]; however, a gap exists in the literature regarding the description of electronic spectra of small molecule amorphous solids potentially present in the ISM. Other computational descriptions of ices have utilized B3LYP/6-31+G** with up to 12 water molecules and implicit solvent effects to describe water as an ice environment for simulating interactions between ammonia and formaldehyde[98]. Chen and Woon conclude that increasing the number of water molecules in the clusters seems to produce infrared modes that start converging in value with observations[98]. Other approaches with DFT investigate interactions with water ices with good agreement to experimental results[99, 100]. Thus, DFT appears predictive for interstellar-like clusters in the infrared region.

Previous research on ices in interstellar regions primarily focuses on vibrational spectroscopy in the infrared region[101–104] with less investigation into electronic spectroscopy within the ultraviolet light region[105, 106]. Infrared spectroscopy of amorphous CO₂

ices has shown that IR can detect mixtures of CO₂ with other small molecules and discern between ice and gas phase through attention to minor details[107, 108]. While water in the gas phase has its first vertical excitation at about 7.5 eV, the amorphous solid is blueshifted significantly due to less favorable interactions with adjacent water molecules from the excited state's smaller dipole moment.[106] Due to ionization limits, each ice has a specific upper-bound energy for the observable electronically excited states before the energy changes the ice itself. As such, electronic excited states confined to a region below the ionization limit provides a region for comparing computation and experiment.

In order to provide a computational description of the amorphous solids of ammonia, carbon dioxide, and water, the present work describes a method for generating these solids computationally and comparing the computed electronic spectroscopic data with experiment in the literature.[105] The generation of molecular structures uses randomization, and the optimizations and electronically excited states are calculated using DFT. After generating clusters of each of these molecules, the use of DFT for molecular optimizations and electronic spectroscopy in the current work aims to provide electronic spectral characterization for small molecule amorphous solids with application to elucidating the behavior of interstellar ices both in the laboratory and potentially even in astrophysical environments.

5.2 Computational Methods & Approach

Amorphous solids are computationally generated through a randomization program written in Python3. The program relies on two main parameters: number of clusters and number of molecules per cluster. After building the clusters and running optimization and electronic excitation calculations, spectroscopic data is extracted and weighted according to a Boltzmann distribution of the clusters' energies. The final output is a normalized spectrum based on combining all the clusters' data.

Prior to using the program, the desired molecule is optimized as a monomer. In the present work, ammonia, water, and carbon dioxide monomers are optimized with ω B97XD/6-31G(d) through Gaussian16 [109–111]. The optimized molecular geometry and cluster parameters, such as number of molecules in the system, number of clusters, and

size of the box, are then input into the program. Additionally, differing geometries can be used with specific ratios to construct mixtures. Regardless of the parameters, each molecule in the cluster starts at the origin and undergoes randomized rotations and displacements according to values generated by the Mersenne Twister series[112], which acts as a pseudorandom number generator. The random rotations occur in three planes. Then, the molecules are displaced by a randomly generated three-dimensional vector. In order to ensure that the molecular geometries do not overlap in the cluster—causing errors in the optimization calculations—the program checks the distance between each of the monomers in the system and the newly added monomer. If the distance is too small, the molecule undergoes the randomization process again. This cycle continues until the system contains the specified number of molecules. Upon completion, the program constructs Gaussian16 geometry optimization input files with Cartesian coordinates and frozen internal coordinates of bond length and bond angles within the individual molecules themselves.

The process described above is conducted for as many clusters as the user specifies. Next, each optimization calculation runs with ω B97-XD/6-31G(d) until the constrained molecular internal coordinates cause the calculation to fail to converge; however, this optimizes the distance between the molecules providing a better guess as to the preferred arrangement of the molecules. The last molecular geometry is extracted and placed into another Gaussian16 input file without freezing internal molecular coordinates in order to calculate the optimized geometry and harmonic frequency zero-point energy for the cluster. This two step process accelerates the optimization process and avoids nearly all imaginary frequencies. Then, the optimized geometries undergo time-dependent density functional theory (TD-DFT) electronic excitation calculations. Notably, the current work uses the optimized structures as a reference geometry for the exploration of different functionals and basis sets. The methods include B3LYP[113], PBE0[114], ω B97-XD[111], CAM-B3LYP[115], and B97D3[116]. The basis sets include 6-311G(d,p) and 6-311++G(2d,2p)[117, 118].

After the optimizations finish, the program uses a Boltzmann distribution from the relative energies to acquire a scaling factor. The scaling factor is used to weight the contributions of each cluster's excitations oscillator strengths, which approximates Beer's Law. Additionally, the temperature parameter is set by the user to match the energy levels of

the environment. For ammonia, water, and carbon dioxide, the temperature is calculated through converting the binding energies of the dimers into units of temperature. These binding energies come from optimizing the dimers with ω B97-XD/6-31G(d) and subtracting the energy from twice the monomer energy. From the Boltzmann distribution, the more stable clusters will contribute more to the overall spectrum than the less stable structures. The program then compiles the excitations into one file for a sub-processed artificial spectrum broadening program that uses the Gaussian line shape procedure with a full width at half maximum height (FWHM) of 2 nm to produce a continuous function from the discretely calculated spectra. Finally, a normalized spectrum (compared to the highest peak) of the oscillator strength is plotted as a function of energy.

In order to explore the capabilities of generating amorphous solid electronic spectra, four datasets are generated comprising of pure water, ammonia, or carbon dioxide. Since water has the fewest electrons of the listed molecules, two datasets of water are analyzed. One dataset comprises of 105 randomized clusters of 32 water molecules, while the other is 30 randomized clusters of 8 water molecules. The datasets for ammonia and carbon dioxide agree in size with the smaller water dataset of 30 randomized clusters of 8 molecules.

5.3 Results and Discussion

5.3.1 Ammonia

A total of 30 octamer clusters of ammonia are generated with their spectra compared with experiment. The temperature for the Boltzmann distribution with ammonia is 1348 K corresponding to the binding energy of the dimerization which implies that all of the binding energy will be thermally released into the amorphous ice. This large assumption ultimately has little effect on the observed results due to the similarities of the excited state properties for the various molecular geometries. Initially, a dataset for several functionals are computed with the 6-311G(d,p) basis set with 25 electronic states to extend over 10 eV as shown in Figure 5.1a. All of the peaks are normalized to the oscillator strength of the maximum intensity peak from the CAM-B3LYP/6-311++G(2d,2p) spectrum to provide a standard normalization factor for comparing relative intensities of the functionals. The

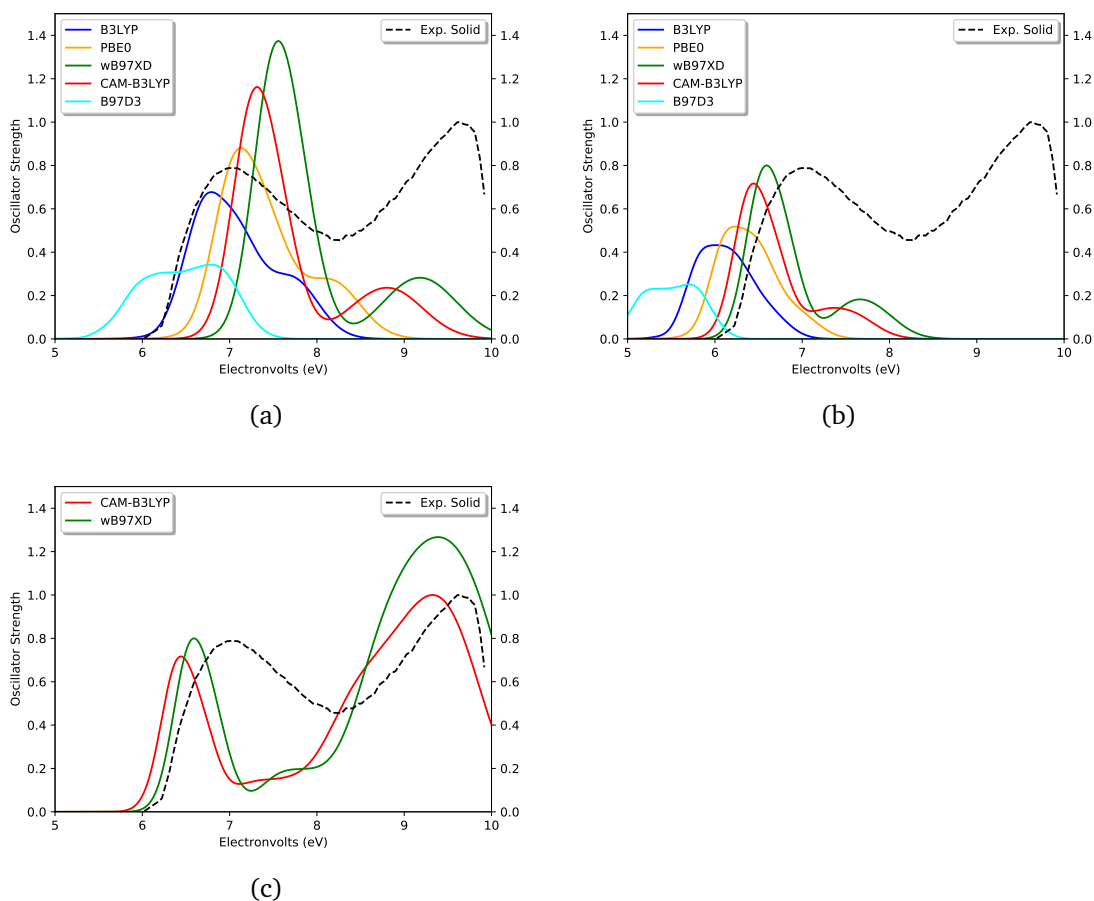


Figure 5.1: VUV spectra for 30 clusters of 8 ammonia molecules normalized to the maximum oscillator strength in cfor consistency, and the experimental plot[105] normalized to itself: a Basis set: 6-311G(d,p) and States: 25; b Basis set: 6-311++G(2d,2p) and States: 25; and, c Basis set: 6-311++G(2d,2p) and States: 125

relative intensities this basis set at describing the first peak shown from work by Kaiser et al.[105] while CAM-B3LYP, ω B97-XD, and PBE0 blueshift.

However, in order to better describe the hydrogen bonding interactions in the amorphous ammonia, diffuse orbitals and a larger basis set are computed. Figure 5.1b utilizes the 6-311++G(2d,2p) basis set with the same functionals showing that all the hybrid functionals are redshifted. Now, it appears as though CAM-B3LYP and ω B97-XD are matching the first peak the best. To extend the spectrum to 10 eV, a spectrum is built with CAM-B3LYP/6-311++G(2d,2p) and ω B97-XD/6-311++G(2d,2p) with 125 electronic states producing Figure 5.1c. The artificial spectrum matches with the experimental spectrum very well qualitatively and even semi-quantitatively. Both peaks of the CAM-B3LYP artificial spec-

Method	Basis Set	Excitation (eV)	Oscillator Strength (Normalized)
B3LYP	6-311G(d,p)	6.79	0.68
PBE0	6-311G(d,p)	7.13	0.88
wB97XD	6-311G(d,p)	7.56	1.37
wB97XD	6-311G(d,p)	9.18	0.28
CAM-B3LYP	6-311G(d,p)	7.32	1.16
CAM-B3LYP	6-311G(d,p)	8.80	0.24
B97D3	6-311G(d,p)	4.79	0.00
B97D3	6-311G(d,p)	6.30	0.31
B97D3	6-311G(d,p)	6.79	0.34
B3LYP	6-311++G(2d,2p)	6.02	0.43
PBE0	6-311++G(2d,2p)	6.23	0.52
wB97XD	6-311++G(2d,2p)	6.59	0.80
wB97XD	6-311++G(2d,2p)	7.66	0.18
wB97XD	6-311++G(2d,2p)	9.39	1.27
CAM-B3LYP	6-311++G(2d,2p)	6.44	0.72
CAM-B3LYP	6-311++G(2d,2p)	7.36	0.14
CAM-B3LYP	6-311++G(2d,2p)	9.32	1.00
B97D3	6-311++G(2d,2p)	5.70	0.25
Exp. Solid		7.00	0.79
Exp. Solid		8.33	0.48
Exp. Solid		9.61	1.00

Table 5.1: Tabulated format of the ammonia spectra displayed in Figure 5.1. Experimental data comes from work produced by Kaiser et al.[105]. All computational spectra are normalized to the maximum peak from CAM-B3LYP/6-311++G(2d,2p).

trum match nicely with regards to relative oscillator strengths to the experimental solid; while ω B97-XD approximates the relative intensities with less accuracy. Therefore, 30 randomized clusters of 8 ammonia molecules with CAM-B3LYP/6-311++G(2d,2p) with 125 electronic excited states appears to effectively describe amorphous ammonia’s UV spectrum.

Finally, the timings for the electronic excited states depend heavily on the number of states and the basis set size. The best results from the CAM-B3LYP and ω B97-XD with 6-311++G(2d,2p) and 125 states take an average of 13.20 and 13.56 hours, respectively, on the local high-performance computing cluster. With the 6-311++G(2d,2p) basis set and 25 states, these functionals run for an average of 4.67 and 5.17 hours, a reduction in time cost of roughly one-third. Lastly, the quickest option for these two functionals with 6-311G(d,p) and 25 states required 1.02 and 1.16 hours, which is relatively fast but produces the worst results. Thus, CAM-B3LYP finishes slightly faster than ω B97-XD and provides closer energies

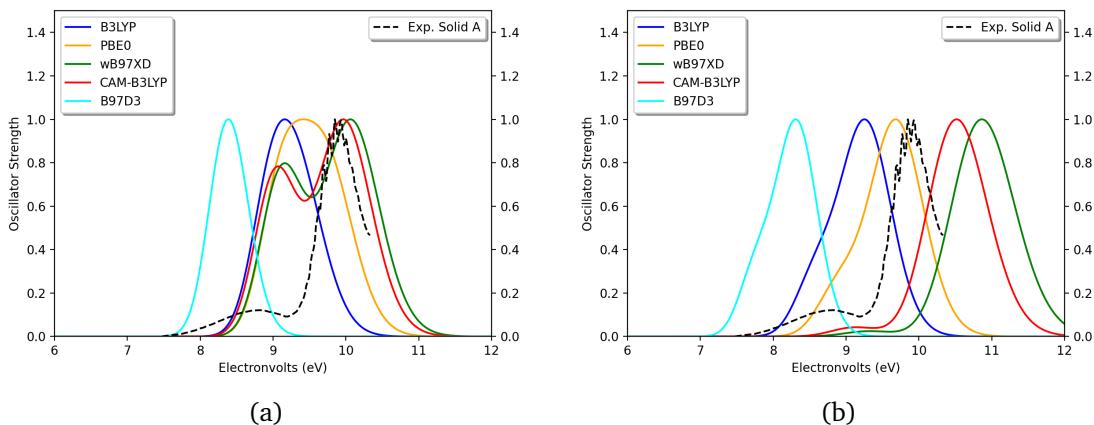


Figure 5.2: VUV spectra for 30 clusters of 8 carbon dioxide molecules normalized to the maximum oscillator strength of each functional, and the experimental solid carbon dioxide plot[105]normalized to itself: a Basis set: 6-311G(d,p) and States: 25; and b Basis set: 6-311++G(2d,2p) and States: 50;

and relative intensities when utilizing larger basis sets and number of states.

5.3.2 Carbon Dioxide

Amorphous carbon dioxide is simulated through the creation of 30 randomized clusters of eight carbon dioxide molecules. Since carbon dioxide interacts with itself much less than ammonia due to the lack of hydrogen bonding, the binding energy from the dimerization yields a lower temperature of 457 K for the Boltzmann distribution. Spectra from several functionals with 6-311G(d,p) are displayed in Figure 5.2a. Once again, the two functionals that match experiment the best are CAM-B3LYP and ω B97-XD; both predict a smaller peak at around 9 eV and a larger peak just over 10 eV. However, the relative oscillator strengths do not match well. B3LYP and B97D3 both redshift the major peak and do not display two peaks with this basis set. Finally, PBE0 has two peaks that are of nearly equal strength. Therefore, only CAM-B3LYP and ω B97-XD appear to perform decently with this basis set.

In order to improve the possible physical representation, the basis set is increased to 6-311++G(2d,2p). While no hydrogen bonding exists in the carbon dioxide clusters, the oxygen and carbon atoms do benefit from extra *d* and *p* orbitals. The effects of the basis set are apparent as shown in Figure 5.2b. In order to cover the energy range of the experiment, 50 electronic states for each functional are calculated. While PBE0 aligns closer in energy

Method	Basis Set	Excitation (eV)	Oscillator Strength (Normalized)
B3LYP	6-311G(d,p)	9.16	1.00
PBE0	6-311G(d,p)	9.42	1.00
ω B97XD	6-311G(d,p)	9.16	0.80
ω B97XD	6-311G(d,p)	10.06	1.00
CAM-B3LYP	6-311G(d,p)	9.07	0.78
CAM-B3LYP	6-311G(d,p)	9.97	1.00
B97D3	6-311G(d,p)	8.39	1.00
B3LYP	6-311++G(2d,2p)	9.25	1.00
PBE0	6-311++G(2d,2p)	9.68	1.00
ω B97XD	6-311++G(2d,2p)	9.33	0.02
ω B97XD	6-311++G(2d,2p)	10.86	1.00
CAM-B3LYP	6-311++G(2d,2p)	9.14	0.04
CAM-B3LYP	6-311++G(2d,2p)	10.52	1.00
B97D3	6-311++G(2d,2p)	8.31	1.00
Exp. Solid		8.83	0.12
Exp. Solid		9.63	0.60
Exp. Solid		9.70	0.79
Exp. Solid		9.77	0.93
Exp. Solid		9.85	1.00
Exp. Solid		9.93	1.00
Exp. Solid		10.00	0.91

Table 5.2: Tabulated format of the carbon dioxide spectra displayed in Figure 5.2. Experimental data comes from work produced by Kaiser et al.[105]

to the experiment with this basis set, the qualitative description has a large shoulder. This shoulder is merged into the major peak due to the artificial spectrum broadener; however, the height for the peak is still too large. For both of the predictive peaks in CAM-B3LYP and ω B97-XD, the energies are blueshifted. However, the curves for CAM-B3LYP and ω B97-XD match qualitatively well with experiment through a small peak separated by a much larger peak over 1 eV away. Thus, the CAM-B3LYP and ω B97-XD functionals describe amorphous carbon dioxide the effectively but not quite the best as they did with ammonia.

5.3.3 Water

Two datasets of water are generated at varying cluster sizes to compare the results with each other. The larger dataset, 105 randomized clusters of 32 water molecules, should better approximate the reality of amorphous water; however, the objective is to determine

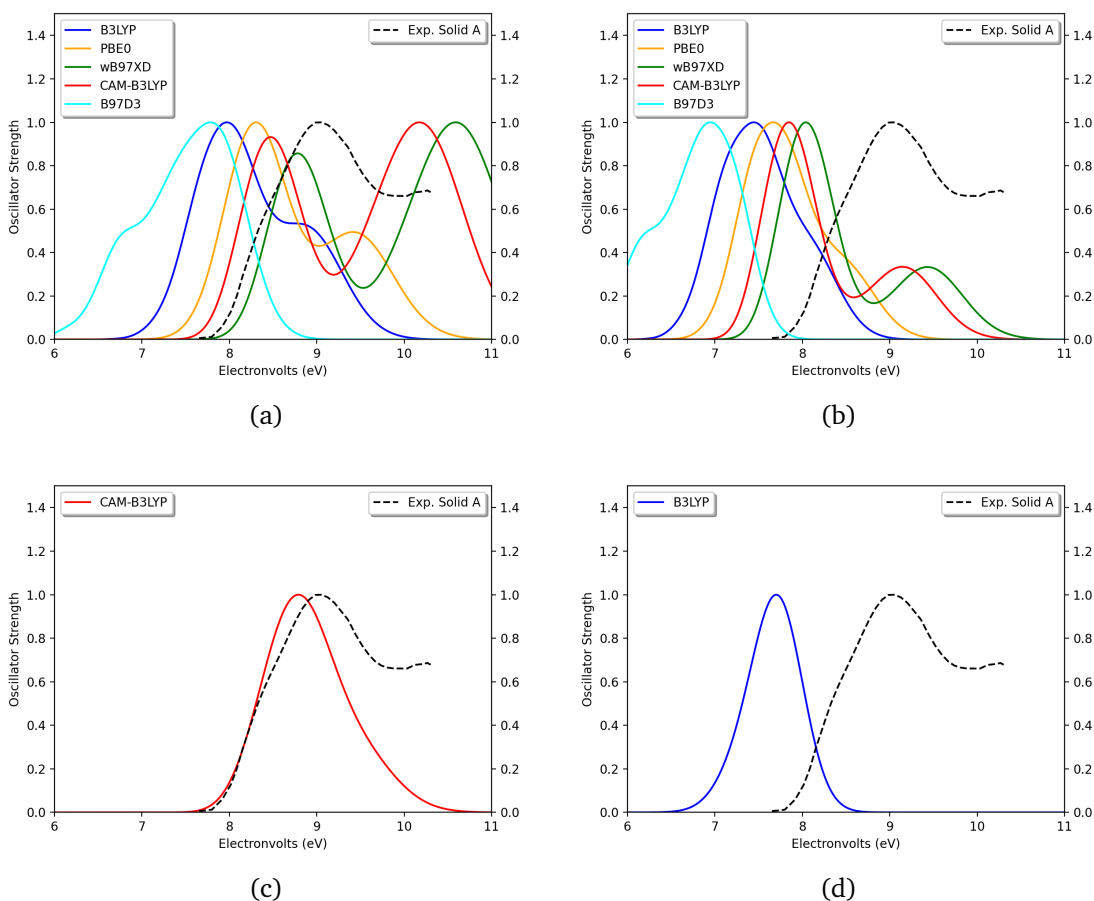


Figure 5.3: VUV spectra for 30 clusters of 8 waters ab and 105 cluster of 32 waters cd compared with experiment[105]. All plots are normalized to themselves: a Basis set: 6-311G(d,p) and States: 25; b Basis set: 6-311++G(2d,2p) and States: 25; c Basis set: 6-311G(d,p) and States: 50; and, d Basis set: 6-311G(d,p) and States: 25

if a smaller dataset, such as 30 randomized clusters of 8 water molecules, can make an equally valid approximation.

First, a dataset of 30 clusters with eight water molecules is generated to investigate with the same functionals listed above. Figure 5.3a displays the 30 clusters of 8 water molecules for several functionals with 6-311G(d,p). From this water dataset, B3LYP and PBE0 appear to perform the best qualitatively for water, while CAM-B3LYP and ω B97-XD predict a larger secondary peak that appears to be beyond the energy range of the experiment. With a larger basis set in Figure 5.3b, the functionals produce peaks more redshifted from experiment. While all of functionals report redshifted excitation energies, CAM-B3LYP and ω B97-XD report the closest value along with a secondary small peak as they did with ammonia. While

Method	Basis Set	Excitation (eV)	Oscillator Strength (Normalized)
B3LYP	6-311G(d,p)	7.98	1.00
B3LYP	6-311G(d,p)	8.76	0.56
PBE0	6-311G(d,p)	8.31	1.00
PBE0	6-311G(d,p)	9.42	0.51
wB97XD	6-311G(d,p)	8.79	0.86
wB97XD	6-311G(d,p)	10.61	1.00
CAM-B3LYP	6-311G(d,p)	8.48	0.96
CAM-B3LYP	6-311G(d,p)	10.19	1.00
B97D3	6-311G(d,p)	6.82	0.49
B97D3	6-311G(d,p)	7.81	1.00
B3LYP	6-311++G(2d,2p)	7.48	1.00
PBE0	6-311++G(2d,2p)	7.68	1.00
wB97XD	6-311++G(2d,2p)	8.05	1.00
wB97XD	6-311++G(2d,2p)	9.43	0.32
CAM-B3LYP	6-311++G(2d,2p)	7.86	1.00
CAM-B3LYP	6-311++G(2d,2p)	9.15	0.32
B97D3	6-311++G(2d,2p)	6.21	0.49
B97D3	6-311++G(2d,2p)	6.97	1.00

Table 5.3: Tabulated format of the water spectra displayed in Figure 5.3ab. Experimental data comes from work produced by Kaiser et al.[105]

these qualitative descriptions are in decent agreement with experiment at the cluster size of 8 water molecules, the energy and oscillator strength differences between the major peaks and experiment could be improved. The inaccuracy could be due to the small molecular weight of the water molecules and the hydrogen bonding not being fully represented.

Upon increasing the number of molecules from 8 to 32 in a cluster and the amount of clusters from 30 to 105, the spectrum redshifts closer to the experimental values as in Figure 5.3c. CAM-B3LYP with 6-311G(d,p) and 50 states matches the experimental spectrum. Alternatively, B3LYP with 6-311G(d,p) and 25 states as in Figure 5.3d displays a redshift, as well, albeit not as strong. Regardless, both of these spectra support the notion that a dataset with more clusters and more molecules in each cluster better represents reality even with a smaller basis set.

To compare timings of the two datasets, the clusters with eight water molecules took an average of 0.51 hours for B3LYP/6-311G(d,p), whereas the 32 water molecule clusters averaged 15.83 hours for the same method, basis set and number of states. Additionally, the

CAM-B3LYP/6-311G(d,p) with 25 states for the smaller clusters takes an average of 2.45 hours, while the much more accurate results with same method and basis set with 50 states takes an average of 31.46 hours per TD-DFT calculation. Clearly, the larger water cluster produces better results with the CAM-B3LYP/6-311G(d,p) and 50 states than the CAM-B3LYP/6-311++G(2d,2p) and 25 states; however, the cost is much higher but nowhere near prohibitively so.

5.4 Conclusions

Ultimately, the methodology implemented herein utilizing a randomization program along with DFT calculations can produce predictive electronic spectrum descriptions for amorphous ices in the ISM or for laboratory analogues based on the present benchmarks for H₂, NH₃, and CO₂. The usage of a randomization procedure for generating arbitrary input geometries and a Boltzmann distribution to weight the excitations yields high qualitative, and even semi-quantitative, agreement with experiment for small molecule amorphous solid electronic spectra. Overall, the best functionals for this application appear to be CAM-B3LYP and ω B97-XD. While increasing the basis set size on ammonia and carbon dioxide produces better results, water requires additional water molecules and larger number of clusters to resemble reality better, even with a smaller basis set. Regardless, the solid correlation with experiment in the UV region provides evidence for these clusters mimicking amorphous solids as ice and that increasing cluster size improves the description. Therefore, the method described in this work effectively characterizes ice analogues through the usage of DFT and should be able to do so for other ices. Mixtures of molecules in the ices are, naturally, a next step and will be explored in future work.

Chapter 6

Theoretical Characterization of Solid, Amorphous Carbonic Acid in the UV

6.1 Introduction

Carbonic acid is believed to exist in various terrestrial environments performing a wide range of functions associated with biology and geology. In a biological context, carbonic acid is believed to serve as a vital component for physiological processes as a weak acid transporting carbon dioxide through the circulatory system[119]. Geologically, carbonic acid is essential in the carbon cycle[120, 121] and is hypothesized to be key to the formation of carbonate minerals as the starting material for the ultimate carbon trioxide moiety [122, 123]. While carbonic acid is likely essential for various chemical processes on Earth, no detection of the molecule exists in extraterrestrial environments, and by extension, other key, potential functions of carbonic acid throughout the universe are still unknown.

Even though carbonic acid has not been detected in the interstellar medium (ISM) nor the Solar System, previous work has demonstrated that carbonic acid may form through the irradiation of H₂O-CO₂ ices with ultraviolet (UV) light matching numerous interstellar environmental conditions.[38, 84, 124, 125]. With the high amounts of H₂O and CO₂ ices throughout the Solar System and ISM, the formation of carbonic acid seems quite likely; still, no detection has been confirmed [74, 126–131]. Therefore, infrared (IR) investiga-

tions into identifying carbonic acid have looked for it in various astronomical sources[132, 133]. However, since carbonic acid has many overlapping vibrational frequencies with other more abundant species in the solid phase,[84] detection in the Solar System through IR spectroscopy is a challenge leading investigations to favor the UV region. To this end, recent laboratory experimental work has produced solid carbonic acid UV spectra for what has been dubbed Experimental Solid A with unknown structure and its counterpart Experimental Solid B, which is believed to correspond to the known β -carbonic acid polymorph[84].

Overall, three polymorphs have been proposed for the solid structures of carbonic acid, but only β -carbonic acid is confirmed to exist[32, 84, 134]. Early research into solid phase carbonic acid had shown support for α -carbonic acid and β -carbonic acid through the formation under acid-base reactions conducted at low temperatures[135]. However, the formation of α -carbonic acid relies upon the usage of a methanolic solution that actually produces the monomethyl ester of carbonic acid ($\text{CH}_3\text{OCO}_2\text{H}$), as shown through comparing matrix-isolation IR spectra of both α -carbonic acid and monomethyl ester.[134, 136–140]. Meanwhile, β -carbonic acid can be produced from the irradiation of $\text{CO}_2:\text{H}_2\text{O}$ ice mixtures raised to the temperature of 220 K.[84, 124]

Previous computational work has demonstrated that the lowest energy isomer of the carbonic acid monomers is the *syn-syn* conformation with the hydrogens both pointing towards the ketone oxygen—a unit of β -carbonic acid [32, 141]. This monomer is followed closely in relative energy by the *syn-anti* conformational isomer at 0.08 eV higher in energy. The *anti-anti* conformational isomer is substantially higher in energy at 0.49 eV higher than the *syn-syn* isomer.[32, 142]. Furthermore, computational and experimental investigations into the β -carbonic acid polymorph support the idea that the structure is comprised of *syn-syn* carbonic acid monomers linked together in a ribbon structure.[32, 141–143] . In the UV region, β -carbonic acid is characterized by a single major peak at 8.92 eV (139nm)[84]. Finally, another carbonic acid solid has appeared in literature through the formation of carbonic acid through a radical reaction between CO and OH in a watery environment at low temperatures (10-40 K)[133, 140, 144]. The defining feature for this form of carbonic acid — Experimental Solid A — comes from a minor peak at approximately 6.2 eV (200nm) on the shoulder of the 8.92 eV feature; however, theoretical work has yet to describe the

carbonic acid structure responsible for this peak in the UV region. The present work aims to achieve this.

Simultaneously, a new procedure leveraging quantum mechanics and statistical mechanics has been able to generate UV spectra for amorphous water, ammonia, and carbon dioxide in agreement with experiment with a mean absolute error of 3.3% for CAM-B3LYP/6-311G(d,p).[70]. Therefore, the present work will utilize this procedure to create amorphous carbonic acid UV spectra to potentially characterize the unknown carbonic acid structure observed in Experimental Solid A [84]. Additionally, the ribbon structure is investigated further with respect to the Experimental Solid B spectrum. These combined efforts will be able to provide novel insights into the UV spectrum of carbonic acid in the solid phase where its presence in extraterrestrial ices may yet be confirmed.

6.2 Computational Methods

Amorphous carbonic acid is simulated by generating 40 randomized clusters.[70] Each cluster consists of 8 carbonic acid molecules comprised of the two lowest energy conformational isomers of the monomer. While 30 clusters are comprised of the lowest energy *syn-syn* carbonic acid monomers, 5 clusters are built from the second lowest conformational isomer, *syn-anti* monomers, and the remaining 5 clusters are a mixture of *syn-syn* and *syn-anti* monomers. Incorporating some of the *syn-anti* monomers provides a more complete sampling of the amorphous structure. Then, each cluster is optimized with ω B97-XD/6-31G(d)[111, 117, 118] with Gaussian16[109] with the energies stored for use in the Boltzmann distribution. Next, the optimized cluster structures are used to calculate the electronically excited states with time-dependent density functional theory (TD-DFT) via the CAM-B3LYP[58, 59, 115, 145] and ω B97-XD[112] methods conjoined to the 6-311G(d,p) basis set[117, 118] for 15 excited states, since 15 states cover the main features observed in the experimental spectra. In order to predict the spectrum of a larger, bulk amorphous solid, each isomer's excited states' oscillator strengths are scaled by a Boltzmann factor. This factor is produced for each cluster from relative energies to mimic the relative abundances of each cluster's contribution in an amorphous solid ice. By approximating all the bind-

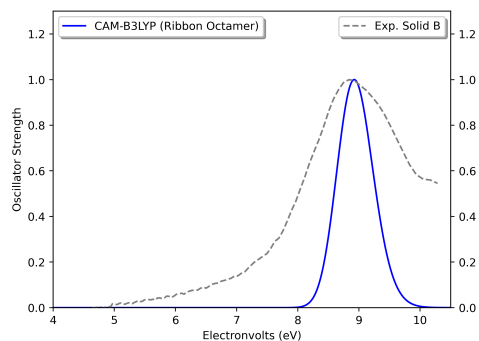
ing energy turning into the thermal energy, the temperature for calculating the Boltzmann factor comes from the binding energy of the lowest energy dimer (0.80 eV)[32], yielding 9259 K. The full set of scaled excitations are input into a Gaussian lineshape procedure with a full width at half maximum height of 2 nm which constructs a normalized artificial UV spectrum. Again, a similar procedure has been shown to reproduce the UV experimental spectrum of amorphous ammonia, carbon dioxide, and water in a semi-quantitative fashion such that unique interpretation of the experimental results based on the theory is possible.

The linear octamer ribbon structure is the extended structure from previous work that appears to correspond with β -carbonic acid[32, 141]. This octamer is optimized with ω B97-XD/6-31G(d) and the electronically excited states are calculated with CAM-B3LYP/6-311G(d,p) with 15 states. Both the simulated, amorphous carbonic acid clusters and ribbon structure results are compared with experimental data on what is reported to be solid carbonic acid from previous work presented in Figure 6 of Ioppolo et al., Ref. 84.

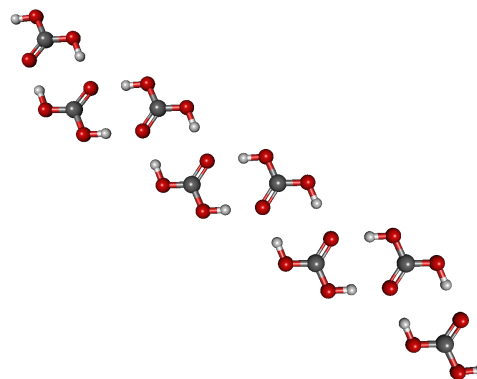
6.3 Results and Discussion

The UV spectrum for the linear octamer ribbon structure is depicted against the experimental solid B in Figure 6.1[84]. The Experimental Solid B plot comes from work done by Ioppolo et al. in which solid carbonic acid is formed at 80 K, annealed to 225 K and cooled down to 80 K. The computed electronically excited states with CAM-B3LYP show high qualitative and even quantitative agreement through having a single major peak at 8.90 eV which is notably close to the experimental value of 8.92 eV (139nm)[32, 141]. This supports previous claims that the ribbon structure Experimental Solid B spectra and, by extension, β -carbonic acid through the additional agreement in the UV region.

The 40 carbonic acid clusters simulating an amorphous solid produce the UV spectrum in Figure 6.2 with two major peaks for both CAM-B3LYP and ω B97XD. The relative energies for these clusters are contained in the supplemental information section. These depictions qualitatively match Experimental Solid A[84]. The Experimental Solid A plot also comes from work done by Ioppolo et al. in which solid carbonic acid is formed at 20 K, annealed to 225 K and cooled down to 20 K. CAM-B3LYP and ω B97XD both depict the largest peak

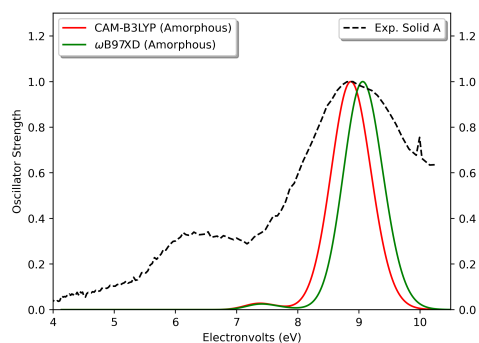


(a)

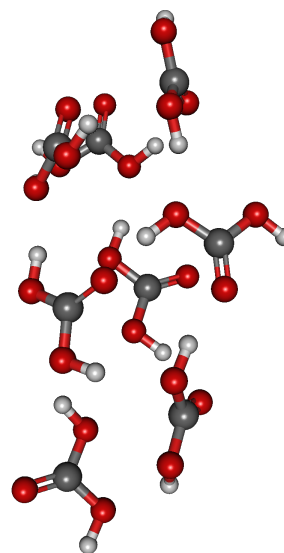


(b)

Figure 6.1: (a) Computed VUV spectra for the carbonic acid ribbon octamer is compared to experimental data for β -H₂CO₃ from Ref. 84. (b) The ribbon octamer is depicted.



(a)



(b)

Figure 6.2: (a) Computed VUV spectra for 40 clusters of 8 carbonic acid molecules is compared with experiment from Ref. 84. (b) The lowest energy cluster is depicted from the 40 used to build the graph.

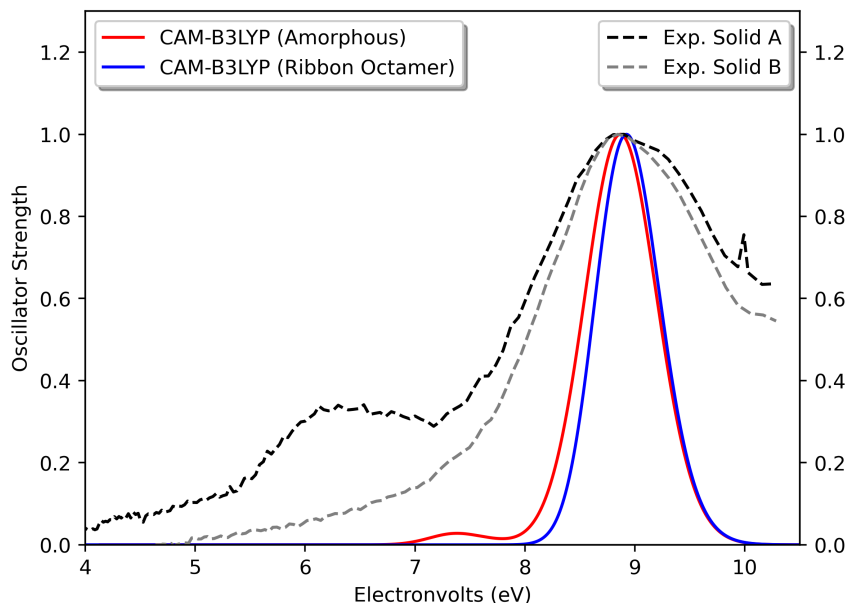


Figure 6.3: Computed VUV spectra of amorphous and ribbon carbonic acid calculated with CAM-B3LYP is compared with Experimental Solid A and B from Ref. 84.

around 9 eV - 8.87 eV and 9.06 eV, respectively - corresponding well with the dominant experimental peak at 8.92 eV[84]. The source of this excitation is from a non-bonding molecular orbital (MO) to a π^* MO in the external carbonic acid molecules. This is the same feature as that present in the linear, ribbon structure discussed above.

The computational amorphous solid spectrum for CAM-B3LYP and ω B97XD both predict a secondary peak, as well, appearing at 7.38 eV and 7.41 eV, respectively, coming from an excitation starting once again in a non-bonding MO promoting the electron into a σ^* MO. This excitation is also present in the ribbon structure, but the symmetry of the ribbon forces this excitation to have a zero oscillator strength. The amorphous simulations have C_1 symmetry which allows these excitations to produce non-zero (albeit still relatively small) oscillator strengths for these lower energy excitations. Only semi-quantitative agreement exists between the amorphous solid and the Experimental Solid A because the computed excitations are approximately 1 eV away. From previous benchmarking on clusters with 8 molecules and CAM-B3LYP/6-311G(d,p), the amorphous spectra predictions for water, ammonia, and carbon dioxide are off from experiment by 0.51 eV, 0.32 eV, and 0.24 eV, re-

spectively[70]. Even so, this excitation with a small oscillator strength appears consistently throughout all of the amorphous clusters and is the closest theoretical carbonic acid cluster spectral feature to match the experimental curve.

Figure 3 brings together the best theoretical data with the experimental data from Ref. 84 to display the unique differences between the two forms of carbonic acid that can be experimentally produced. While both the ribbon and amorphous structures have a major peak at approximately 9 eV, the lack of symmetry in the amorphous clusters allows the non-bonding to σ^* excitation to exhibit some oscillator strength that the C_{2v} and C_{2h} ribbon structures forbid. Experimental Solid A qualitatively, and even semi-quantitatively, matches the CAM-B3LYP spectrum of the 40 clusters of carbonic acid octamers simulating the amorphous solid phase, while Experimental Solid B agrees with the ribbon octamer structure. The CAM-B3LYP curve for the amorphous solid produces slightly better results than the ω B97XD method regarding the maximum peak in line with previous work [70].

6.4 Conclusions

The UV spectrum for carbonic acid appears to have two competing characterizations, split into Experimental Solid A and B. Computationally, the linear octamer ribbon structure matches the Experimental Solid B, further strengthening the claim that the ribbon is responsible for the β -carbonic acid polymorph. Meanwhile, Experimental Solid A is characterized by having two major peaks with the smaller peak arising on the lower energy side, which maps to the computationally predicted spectrum of the amorphous carbonic acid clusters. Therefore, the ribbon structure has further support for matching β -carbonic acid, and Experimental Solid A — the speculated γ -carbonic acid polymorph[84] — appears to correspond with the amorphous form of carbonic acid. With these distinctions, future UV studies simulating carbon-rich Solar System ices can distinguish β -carbonic acid from the amorphous solid state.

Chapter 7

Final Conclusions

Theoretical and computational chemistry has rapidly expanded over the last 100 years to provide mathematical models for predicting physical observables from quantum mechanical principles and to generate efficient algorithms for feasibly performing the resulting, labor-intensive procedures. The basis for modern quantum chemistry comes from using Hartree-Fock theory as a starting point for more accurate methods which, in turn, treat additional correlation. Alternatively, DFT modifies the two-electron portion of the Fock matrix from the Hartree-Fock method to reduce the computational costs dramatically while still treating some amount of correlation. Such a compromise allows for the study of much larger molecular systems still through the lense of quantum mechanics instead of only classical mechanics.

The present work in Chapters 4, 5, and 6 employ DFT to optimize molecular geometries of clusters and TD-DFT with various functionals to calculate electronically excited states for the resulting structures. These computations have led to the discovery of a helical, linear, and amorphous carbonic acid structures along with identifying UV spectral features. Furthermore, from generating randomized clusters and combining quantum mechanical predictions through the statistical mechanical Boltzmann distribution, a new method for generating UV spectra for amorphous solids has been developed, benchmarked, and applied. Lastly, the development of the `ONSET` program in Chapter 3 has provided experimentalists the ability to easily and reliably calculate the λ_{onset} for organic chromophores.

Bibliography

- ¹A. Szabo and N. S. Ostlund, *Modern quantum chemistry introduction to advanced electronic structure theory* (Dover Publications, Inc, 2020).
- ²T. Engel and P. J. Reid, *Physical chemistry: thermodynamics, statistical thermodynamics, and kinetics*, 4th ed. (Pearson Education, Inc, 2018).
- ³A. M. Wallace, C. Curiac, J. H. Delcamp, and R. C. Fortenberry, “Accurate determination of the onset wavelength (λ_{onset}) in optical spectroscopy”, *Journal of Quantitative Spectroscopy and Radiative Transfer* **265**, 107544 (2021).
- ⁴W. S. Choi, S. S. A. Seo, and H. N. Lee, “Optoelectronics: optical properties and electronic structures of complex metal oxides”, in *Epitaxial growth of complex metal oxides*, edited by G. Koster, M. Huijben, and G. Rijnders (Elsevier, Amsterdam, 2015), pp. 331–363.
- ⁵M. P. Cifuentes, M. G. Humphrey, J. E. McGrady, P. J. Smith, R. Stranger, K. S. Murray, and B. Moubaraki, “High nuclearity ruthenium carbonyl cluster chemistry: 5. local density functional, electronic spectroscopy, magnetic susceptibility, and electron paramagnetic resonance studies on (carbido)decaruthenium carbonyl clusters”, *J. Am. Chem. Soc.* **119**, 2647–2655 (1997).
- ⁶J. H. Delcamp, A. Yella, M. K. Nazeeruddina, and M. Grätzel, “Modulating dye $E_{(S^+/S^*)}$ with efficient heterocyclic nitrogen containing acceptors for dscs”, *Chem. Comm.* **48**, 2295–2297 (2012).
- ⁷C. M. García Santander, S. M. Gómez Rueda, N. de Lima da Silva, C. L. de Camargo, T. G. Kieckbusch, and M. R. Wolf Maciel, “Measurements of normal boiling points of

- fatty acid ethyl esters and triacylglycerols by thermogravimetric analysis”, *Fuel* **92**, 158–161 (2012).
- ⁸G. Kolb, S. Scheiber, H. Antrekowitsch, P. Uggowitz, D. Pöschmann, and S. Pogatscher, “Differential scanning calorimetry and thermodynamic predictions—a comparative study of al-zn-mg-cu alloys”, *Metals* **6**, 180 (2016).
- ⁹D. M. Fabian and S. Ardo, “Hybrid organic–inorganic solar cells based on bismuth iodide and 1,6-hexanediammonium dication”, *J. Mater. Chem. A* **4**, 6837–6841 (2016).
- ¹⁰M. P. P. Makula and W. Macyk, “How to correctly determine the band gap energy of modified semiconductor photocatalysts based on uv-vis spectra”, *J. Phys. Chem. Lett.* **9**, 6814–6817 (2018).
- ¹¹P. H. Carey, F. Ren, D. C. Hays, B. P. Gila, S. J. Pearton, S. Jang, and A. Kuramata, “Band offsets in ito/ga₂o₃ heterostructures”, *Appl. Surface Sci.* **422**, 179–183 (2017).
- ¹²R. G. Egdell, J. Rebane, and T. J. Walker, “Competition between initial- and final-state effects in valence- and core-level x-ray photoemission of sb-doped sno₂”, *Phys. Rev. B* **59**, 1792–1799 (1999).
- ¹³T. Zhang, M. A. Hossain, C.-Y. Lee, Y. Zakaria, A. A. Abdallah, and B. Hoex, “Atomic layer deposited zn_xni_{1-x}o: a thermally stable hole selective contact for silicon solar cells”, *Appl. Phys. Lett.* **113**, 262102 (2018).
- ¹⁴C. Arantes, M. Scholz, R. Schmidt, V. Dehm, M. L. M. Rocco, A. Schöll, F. Reinert, and F. Würthner, “Comparative analysis of the energy levels of planar and core-twisted perylene bisimides in solution and solid state by uv/vis, cv, and ups/ipes”, *Appl. Phys. A* **108**, 629–637 (2012).
- ¹⁵S. Y. Liu, J. H. Chang, I. W. Wu, and C. I. Wu, “Alternating current driven organic light emitting diodes using lithium fluoride insulating layers”, *Sci. Rep.* **4**, 7559 (2014).
- ¹⁶M. Terashima, T. Miyayama, T. Shirao, H. W. Mo, Y. Hatae, H. Fujimoto, and K. Watanabe, “The electronic band structure analysis of oled device by means of *in Situ* leips and ups combined with gcib”, *Surf. Interface Anal.*, *in press* DOI: 10.1002/sia.6777 (2020).

- ¹⁷R. E. Benfield, “Magnetic properties and uv-visible spectroscopic studies of metal cluster compounds”, in *Physics and chemistry of metal cluster compounds*, edited by L. J. de Jongh (Kluwer Academic Publishers, Amsterdam, 1994), pp. 249–270.
- ¹⁸J. N. Gayton, S. Autry, R. C. Fortenberry, N. I. Hammer, and J. H. Delcamp, “Counter anion effect on the photophysical properties of emissive indolizine-cyanine dyes in solution and solid state”, *Molecules* **23**, 3051 (2018).
- ¹⁹R. R. Ernst, “Without computers — no modern nmr”, in *Computational aspects of the study of biological macromolecules by nuclear magnetic resonance spectroscopy*, edited by J. C. Hoch (Springer, Berlin, 1991), pp. 1–25.
- ²⁰P. A. Hays and R. A. Thompson, “A processing method enabling the use of peak height for accurate and precise proton nmr quantitation”, *Magn. Res. Chem.* **47**, 819–824 (2009).
- ²¹H. A. Naveen Dharmagunawardhane, W. R. Woerner, Q. Wu, H. Huang, X. Chen, A. Orlov, P. G. Khalifah, and J. B. Parise, “Photocatalytic hydrogen evolution using nanocrystalline gallium oxynitride spinel”, *J. Mater. Chem. A* **2**, 19247–19252 (2014).
- ²²J. Gayton, S. A. Autry, W. Meador, S. R. Parkin, G. A. H. Jr., N. I. Hammer, and J. H. Delcamp, “Indolizine-cyanine dyes: near infrared emissive cyanine dyes with increased stokes shifts”, *J. Org. Chem.* **84**, 687–697 (2019).
- ²³P. Brogdon, F. Giordano, G. A. Punecky, A. Dass, S. M. Zakeeruddin, M. K. Nazeeruddin, M. Grätzel, G. S. Tschumper, and J. H. Delcamp, “A computational and experimental study of thieno[3,4-b]thiophene as a proaromatic π -bridge in dye-sensitized solar cells”, *Chem. Eur. J.* **22**, 694–703 (2015).
- ²⁴R. R. Rodrigues, A. Peddapuram, A. L. Dorris, N. I. Hammer, and J. H. Delcamp, “Thienopyrroledione-based photosensitizers as strong photoinduced oxidants: oxidation of $\text{Fe}(\text{bpy})_3^{2+}$ in a > 1.3 v dye-sensitized solar cell”, *ACS Appl. Energy Mater.* **2**, 5547–5556 (2019).
- ²⁵L. E. McNamara, T. A. Rill, A. J. Huckaba, V. Ganeshraj, J. Gayton, R. A. Nelson, E. A. Sharpe, A. Dass, N. I. Hammer, and J. H. Delcamp, “Indolizine-squaraines: nir fluorescent materials with molecularly engineered stokes shifts”, *Chem. Eur. J.* **23**, 12494 (2017).

- ²⁶C. S. L. Rathnamalala, J. N. Gayton, A. L. Dorris, S. A. Autry, W. Meador, N. I. Hammer, J. H. Delcamp, and C. N. Scott, “Donor–acceptor–donor nir ii emissive rhodindolizine dye synthesized by c–h bond functionalization”, *J. Org. Chem.* **84**, 13186–13193 (2019).
- ²⁷H. Shirley, S. Parkin, and J. H. Delcamp, “Photoinduced generation of a durable thermal proton reduction catalyst with in situ conversion of $\text{mn}(\text{bpy})(\text{co})_3\text{br}$ to $\text{mn}(\text{bpy})_2\text{br}_2$ ”, *Inorg. Chem.* **59**, 11266–11272 (2020).
- ²⁸W. E. Meador, S. A. Autry, R. N. Bessetti, J. N. Gayton, A. S. Flynt, N. I. Hammer, and J. H. Delcamp, “Water-soluble nir absorbing and emitting indolizine cyanine and indolizine squaraine dyes for biological imaging”, *J. Org. Chem.* **85**, 4089–4095 (2020).
- ²⁹A. J. Huckaba, H. Shirley, R. W. Lamb, S. Guertin, S. Autry, H. Cheema, K. Talukdar, T. Jones, J. W. Jurss, A. Dass, N. I. Hammer, R. H. Schmehl, C. E. Webster, and J. H. Delcamp, “A mononuclear tungsten photocatalyst for h_2 production”, *ACS Catal.* **8**, 4838–4847 (2018).
- ³⁰H. Cheema, J. Watson, A. Peddapuram, and J. H. Delcamp, “A 25 ma cm^{-2} dye-sensitized solar cell based on a near-infrared-absorbing organic dye and application of the device in ssm-dscs”, *Chem. Commun.* **56**, 1741–1744 (2020).
- ³¹R. R. Rodrigues, C. Boudreaux, E. T. Papish, and J. H. Delcamp, “Photocatalytic reduction of co_2 to co and formate: do reaction conditions or ruthenium catalysts control product selectivity?”, *ACS Appl. Energy Mater.* **2**, 37–46 (2019).
- ³²A. M. Wallace and R. C. Fortenberry, “Linear and helical carbonic acid clusters”, *The Journal of Physical Chemistry A* **125**, PMID: 34029078, 4589–4597 (2021).
- ³³G. Strazzulla, J. R. Brucato, and M. E. Palumbo, “Carbonic acid on mars?”, *Planetary and Space Science* **44**, 1447–1450 (1996).
- ³⁴W. Zheng and R. I. Kaiser, “On the formation of carbonic acid (h_2co_3) in solar system ices”, *Chemical Physics Letters* **450**, 55–60 (2007).
- ³⁵B. M. Jones, R. I. Kaiser, and G. Strazzulla, “Carbonic acid as a reserve of carbon dioxide on icy moons: the formation of carbon dioxide (co_2) in a polar environment”, *The Astrophysical Journal* **788**, 170 (2014).

- ³⁶S. Radhakrishnan, M. S. Gudipati, W. Sander, and A. Lignell, “Photochemical processes in $\text{CO}_2/\text{H}_2\text{O}$ ice mixtures with trapped pyrene, a model polycyclic aromatic hydrocarbon”, *The Astrophysical Journal* **864**, 151 (2018).
- ³⁷S. A. Sandford, M. Nuevo, P. P. Bera, and T. J. Lee, “Prebiotic astrochemistry and the formation of molecules of astrobiological interest in interstellar clouds and protostellar disks”, *Chemical reviews* **120**, 4616–4659 (2020).
- ³⁸R. K. M.H. Moore, “Infrared and mass spectral studies of proton irradiated $\text{H}_2\text{O} + \text{CO}_2$ ice: evidence for carbonic acid”, *Spectrochimica Acta Part A: Molecular Spectroscopy* **47**, 255–262 (1991).
- ³⁹Z. Peeters, R. Hudson, M. Moore, and A. Lewis, “The formation and stability of carbonic acid on outer solar system bodies”, *Icarus* **210**, 480–487 (2010).
- ⁴⁰S. Ioppolo, Z. Kaňuchová, R. James, A. Dawes, N. Jones, S. Hoffmann, N. Mason, and G. Strazzulla, “Vacuum ultraviolet photoabsorption spectroscopy of space-related ices: 1 keV electron irradiation of nitrogen- and oxygen-rich ices”, *Astronomy & Astrophysics* **641**, A154 (2020).
- ⁴¹J. Bernard, M. Seidl, I. Kohl, K. R. Liedl, E. Mayer, Ó. Gálvez, H. Grothe, and T. Loerting, “Spectroscopic observation of gas-phase carbonic acid isolated in matrix”, (2011).
- ⁴²C. Mitterdorfer, J. Bernard, F. Klauser, K. Winkel, I. Kohl, K. R. Liedl, H. Grothe, E. Mayer, and T. Loerting, “Local structural order in carbonic acid polymorphs: Raman and FT-IR spectroscopy”, *Journal of Raman Spectroscopy* **43**, 108–115 (2012).
- ⁴³J. Bernard, R. G. Huber, K. R. Liedl, H. Grothe, and T. Loerting, “Matrix isolation studies of carbonic acid? the vapor phase above the β -polymorph”, *Journal of the American Chemical Society* **135**, 7732–7737 (2013).
- ⁴⁴E.-M. Köck, J. Bernard, M. Podewitz, D. F. Dinu, R. G. Huber, K. R. Liedl, H. Grothe, E. Bertel, R. Schlögl, and T. Loerting, “Alpha-carbonic acid revisited: carbonic acid monomethyl ester as a solid and its conformational isomerism in the gas phase”, *Chemistry—A European Journal* **26**, 285–305 (2020).

- ⁴⁵S. Reddy and S. Balasubramanian, “Carbonic acid: molecule, crystal and aqueous solution”, *Chemical Communications* **50**, 503–514 (2013).
- ⁴⁶S. Reddy, H. Chidambar, and S. Balasubramanian, “Theoretical investigations of candidate crystal structures for β -carbonic acid”, *The Journal of Chemical Physics* **134**, 124511 (2011).
- ⁴⁷J.-D. Chai and M. Head-Gordon, “Systematic optimization of long-range corrected hybrid density functionals”, *The Journal of chemical physics* **128**, 084106 (2008).
- ⁴⁸J.-D. Chai and M. Head-Gordon, “Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections”, *Physical Chemistry Chemical Physics* **10**, 6615–6620 (2008).
- ⁴⁹T. Clark, J. Chandrasekhar, G. W. Spitznagel, and P. V. R. Schleyer, “Efficient diffuse function-augmented basis sets for anion calculations. iii. the 3-21+ g basis set for first-row elements, li–f”, *Journal of Computational Chemistry* **4**, 294–301 (1983).
- ⁵⁰R. Ditchfield, W. J. Hehre, and J. A. Pople, “Self-consistent molecular-orbital methods. ix. an extended gaussian-type basis for molecular-orbital studies of organic molecules”, *The Journal of Chemical Physics* **54**, 724–728 (1971).
- ⁵¹P. C. Hariharan and J. A. Pople, “The influence of polarization functions on molecular orbital hydrogenation energies”, *Theoretica chimica acta* **28**, 213–222 (1973).
- ⁵²W. J. Hehre, R. Ditchfield, and J. A. Pople, “Self—consistent molecular orbital methods. xii. further extensions of gaussian—type basis sets for use in molecular orbital studies of organic molecules”, *The Journal of Chemical Physics* **56**, 2257–2261 (1972).
- ⁵³M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr.,

- J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, *Gaussian16 Revision C.01*, Gaussian Inc. Wallingford CT, 2016.
- ⁵⁴P. Karamanis, “The importance of the dft method on the computation of the second hyperpolarizability of semiconductor clusters of increasing size: a critical analysis on prolate aluminum phosphide clusters”, *International Journal of Quantum Chemistry* **112**, 2115–2125 (2012).
- ⁵⁵C. Adamo and D. Jacquemin, “The calculations of excited-state properties with time-dependent density functional theory”, *Chem. Soc. Rev* **2013**.
- ⁵⁶A. D. Laurent, C. Adamo, and D. Jacquemin, “Dye chemistry with time-dependent density functional theory”, *Physical Chemistry Chemical Physics* **16**, 14334–14356 (2014).
- ⁵⁷A. D. Becke, “Density-functional thermochemistry. iii. the role of exact exchange (1993) j”, *Chem. Phys* **98**, 5648–5652.
- ⁵⁸W. T. Yang, R. G. Parr, and C. T. Lee, “Various functionals for the kinetic energy density of an atom or molecule”, *Phys. Rev. A* **34**, 4586–4590 (1986).
- ⁵⁹C. Lee, W. T. Yang, and R. G. Parr, “Development of the colle-salvetti correlation-energy formula into a functional of the electron density”, *Phys. Rev. B.* **37**, 785–789 (1988).
- ⁶⁰R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople, “Self-consistent molecular orbital methods. xx. a basis set for correlated wave functions”, *The Journal of chemical physics* **72**, 650–654 (1980).
- ⁶¹A. I. Krylov, “Equation-of-motion coupled-cluster methods for open-shell and electronically excited species: the hitchhiker’s guide to fock space”, *Annu. Rev. Phys. Chem.* **59**, 433–462 (2008).
- ⁶²I. Shavitt and R. J. Bartlett, *Many-body methods in chemistry and physics: mbpt and coupled-cluster theory* (Cambridge university press, 2009).

- ⁶³J. F. Stanton and R. J. Bartlett, “The equation of motion coupled-cluster method. a systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties”, *The Journal of chemical physics* **98**, 7029–7039 (1993).
- ⁶⁴H. J. e. a. Werner, “2”, a package of ab initio programs **2019** (2019).
- ⁶⁵R. A. Kendall, T. H. Dunning Jr, and R. J. Harrison, “Electron affinities of the first-row atoms revisited. systematic basis sets and wave functions”, *The Journal of chemical physics* **96**, 6796–6806 (1992).
- ⁶⁶T. H. Dunning Jr, “Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen”, *The Journal of chemical physics* **90**, 1007–1023 (1989).
- ⁶⁷J. Murillo, J. David, and A. Restrepo, “Insights into the structure and stability of the carbonic acid dimer”, *Physical Chemistry Chemical Physics* **12**, 10963–10970 (2010).
- ⁶⁸A. D. Zapata–Escobar, J. A. Murillo–López, C. Hadad, and A. Restrepo, “Understanding the nature of bonding interactions in the carbonic acid dimers”, *Journal of Molecular Modeling* **25**, 1–10 (2019).
- ⁶⁹W. K. Peters, D. E. Couch, B. Mignolet, X. Shi, Q. L. Nguyen, R. C. Fortenberry, H. B. Schlegel, F. Remacle, H. C. Kapteyn, M. M. Murnane, et al., “Ultrafast 25-fs relaxation in highly excited states of methyl azide mediated by strong nonadiabatic coupling”, *Proceedings of the National Academy of Sciences* **114**, E11072–E11081 (2017).
- ⁷⁰A. M. Wallace and R. C. Fortenberry, “Computational uv spectra for amorphous solids of small molecules”, *Physical Chemistry Chemical Physics*, 24413–24420 (2021).
- ⁷¹A. G. G. M. Tielens and W. Hagen, “Model Calculations of the Molecular Composition of Interstellar Grain Mantles”, *Astronomy and Astrophysics* **114**, 245–260 (1982).
- ⁷²E. Gibb, W. D.C.B., A. Boogert, and A. Tielens, “Interstellar ice: the infrared space observatory legacy”, *Astrophysical Journal, Supplement Series* **151**, 35–73 (2004).
- ⁷³K. I. berg, A. C. A. Boogert, K. M. Pontoppidan, S. van den Broek, E. F. van Dishoeck, S. Bottinelli, G. A. Blake, and N. J. Evans, “The spitzer ice legacy: ice evolution from cores to protostars”, *The Astrophysical Journal* **740**, 109 (2011).

- ⁷⁴A. A. Boogert, P. A. Gerakines, and D. C. Whittet, “Observations of the icy universe”, *Annual Review of Astronomy and Astrophysics* **53**, 541–581 (2015).
- ⁷⁵K. Hiraoka, T. Miyagoshi, T. Takayama, K. Yamamoto, and Y. Kihara, “Gas-grain processes for the formation of CH_4 and H_2O : reactions of H atoms with C, O, and CO in the solid phase at 12 K”, *The Astrophysical Journal* **498**, 710–715 (1998).
- ⁷⁶R. Martín-Doménech, K. I. Öberg, and M. Rajappan”, “Formation of NH_2CHO and CH_3CHO upon UV photoprocessing of interstellar ice analogs”, *The Astrophysical Journal* **894**, 98 (2020).
- ⁷⁷L. J. Karssemeijer, S. Ioppolo, M. C. van Hemert, A. van der Avoird, M. A. Allodi, G. A. Blake, and H. M. Cuppen, “Dynamics of CO in amorphous water-ice environments”, *The Astrophysical Journal* **781**, 16 (2013).
- ⁷⁸M. P. Collings, J. W. Dever, H. J. Fraser, M. R. S. McCoustra, and D. A. Williams, “Carbon monoxide entrapment in interstellar ice analogs”, *The Astrophysical Journal* **583**, 1058–1062 (2003).
- ⁷⁹L. J. Allamandola, S. A. Sandford, and G. J. Valero, “Photochemical and thermal evolution of interstellar/precometary ice analogs”, *Icarus* **76**, 225–252 (1988).
- ⁸⁰L. J. Allamandola, M. P. Bernstein, S. A. Sandford, and R. L. Walker, “Evolution of interstellar ices”, in *Composition and origin of cometary materials*, edited by K. Altwegg, P. Ehrenfreund, J. Geiss, and W. F. Huebner (1999), pp. 219–232.
- ⁸¹P. Gerakines, W. Schutte, and P. Ehrenfreund, “Ultraviolet processing of interstellar ice analogs. I. pure ices.”, *Astronomy and Astrophysics* **312**, 289–305 (1996).
- ⁸²R. Hudson and M. Moore, “Laboratory studies of the formation of methanol and other organic molecules by water+carbon monoxide radiolysis: relevance to comets, icy satellites, and interstellar ices”, *Icarus* **140**, 451–461 (1999).
- ⁸³M. H. Moore and R. K. Khanna, “Infrared and mass spectral studies of proton irradiated $\text{H}_2\text{O} + \text{CO}_2$ ice: evidence for carbonic acid”, *Spectrochimica Acta Part A: Molecular Spectroscopy* **47**, 255–262 (1991).

- ⁸⁴Ioppolo, S., Kanuchová, Z., James, R. L., Dawes, A., Ryabov, A., Dezalay, J., Jones, N. C., Hoffmann, S. V., Mason, N. J., and Strazzulla, G., “Vacuum ultraviolet photoabsorption spectroscopy of space-related ices: formation and destruction of solid carbonic acid upon 1 keV electron irradiation”, *Astronomy & Astrophysics* **646**, A172 (2021).
- ⁸⁵R. I. Kaiser, S. Maity, and B. M. Jones, “Synthesis of prebiotic glycerol in interstellar ices”, *Angewandte Chemie International Edition* **54**, 195–200 (2015).
- ⁸⁶Öberg and K. I., “Photochemistry and astrochemistry: photochemical pathways to interstellar complex organic molecules”, *Chemical Reviews* **116**, 9631–9663 (2016).
- ⁸⁷K. K. Singh, P. Tandon, R. Kumar, A. Misra, Shivani, M. Yadav, A. Ahmad, and M. K. Chaudhary, “Formation of aminomethanol in ammonia-water interstellar ice”, *Monthly Notices of the Royal Astronomical Society*, 10.1093/mnras/stab1778 (2021).
- ⁸⁸G. Fedoseev, K.-J. Chuang, E. F. van Dishoeck, S. Ioppolo, and H. Linnartz, “Simultaneous hydrogenation and UV photolysis experiments of n- and o-isomers of interstellar ice analogues; linking HNC, OCN, NH₂CHO, and NH₂OH”, *Monthly Notices of the Royal Astronomical Society* **460**, 4297–4309 (2016).
- ⁸⁹E. Herbst and E. F. van Dishoeck, “Complex organic interstellar molecules”, *Annual Review of Astronomy and Astrophysics* **47**, 427–480 (2009).
- ⁹⁰M. P. Bernstein, J. P. Dworkin, S. A. Sandford, G. W. Cooper, and L. J. Allamandola, “Racemic amino acids from the ultraviolet photolysis of interstellar ice analogues”, *Nature* **416**, 401–403 (2002).
- ⁹¹D. E. Woon, “Pathways to glycine and other amino acids in ultraviolet-irradiated astrophysical ices determined via quantum chemical modeling”, *The Astrophysical Journal* **571**, L177–L180 (2002).
- ⁹²R. I. Kaiser and K. Roessler, “Theoretical and laboratory studies on the interaction of cosmic-ray particles with interstellar ices. III. Suprathermal chemistry-induced formation of hydrocarbon molecules in solid methane (CH₄), ethylene (C₂H₄), and acetylene (C₂H₂)”, *The Astrophysical Journal* **503**, 959–975 (1998).

- ⁹³V. L. Deringer, N. Bernstein, A. P. Bartk, M. J. Cliffe, R. N. Kerber, L. E. Marbella, C. P. Grey, S. R. Elliott, and G. Csnyi, “Realistic atomistic structure of amorphous silicon from machine-learning-driven molecular dynamics”, *The Journal of Physical Chemistry Letters* **9**, 2879–2885 (2018).
- ⁹⁴G. C. Sosso, V. L. Deringer, S. R. Elliott, and G. Csnyi, “Understanding the thermal properties of amorphous solids using machine-learning-based interatomic potentials”, *Molecular Simulation* **44**, 866–880 (2018).
- ⁹⁵J. Mavrai, F. C. Mocanu, V. L. Deringer, G. Csnyi, and S. R. Elliott, “Similarity between amorphous and crystalline phases: the case of TiO_2 ”, *The Journal of Physical Chemistry Letters* **9**, 2985–2990 (2018).
- ⁹⁶C. Arasa, M. C. van Hemert, E. F. van Dishoeck, and G. J. Kroes, “Molecular dynamics simulations of CO_2 formation in interstellar ices”, *The Journal of Physical Chemistry A* **117**, 7064–7074 (2013).
- ⁹⁷V. L. Deringer, M. A. Caro, R. Jana, A. Aarva, S. R. Elliott, T. Laurila, G. Csnyi, and L. Pastewka, “Computational surface chemistry of tetrahedral amorphous carbon by combining machine learning and density functional theory”, *Chemistry of Materials* **30**, 7438–7445 (2018).
- ⁹⁸L. Chen and D. E. Woon, “A theoretical investigation of the plausibility of reactions between ammonia and carbonyl species (formaldehyde, acetaldehyde, and acetone) in interstellar ice analogs at ultracold temperatures”, *The Journal of Physical Chemistry A* **115**, 5166–5183 (2011).
- ⁹⁹E. Escamilla-Roa and C. I. Sainz-Daz, “Amorphous ammonia-water ice deposited onto silicate grain: effect on growth of mantle ice on interstellar and interplanetary dust”, *The Journal of Physical Chemistry C* **118**, 3554–3563 (2014).
- ¹⁰⁰G. Bovolenta, S. Bovino, E. Vhringer-Martinez, D. A. Saez, T. Grassi, and S. Vogt-Geisse, “High level ab initio binding energy distribution of molecules on interstellar ices: hydrogen fluoride”, *Molecular Astrophysics* **21**, 100095 (2020).

- ¹⁰¹Giuliano, B. M., Escribano, R. M., Martín-Doménech, R., Dartois, E., and Muñoz Caro, G. M., “Interstellar ice analogs: band strengths of H_2O , CO_2 , CH_3OH , and NH_3 in the far-infrared region”, *Astronomy & Astrophysics* **565**, A108 (2014).
- ¹⁰²S. Bottinelli, A. C. A. Boogert, J. Bouwman, M. Beckwith, E. F. van Dishoeck, K. I. berg, K. M. Pontoppidan, H. Linnartz, G. A. Blake, N. J. Evans, and F. Lahuis, “The c2d spitzer spectroscopic survey of ices around low-mass young stellar objects. iv. NH_3 and CH_3OH ”, *The Astrophysical Journal* **718**, 1100–1117 (2010).
- ¹⁰³J. Bouwman, W. Ludwig, Z. Awad, K. I. Öberg, G. W. Fuchs, E. F. van Dishoeck, and H. Linnartz, “Band profiles and band strengths in mixed $\text{H}_2\text{O}:\text{CO}$ ices”, *Astronomy & Astrophysics* **476**, 995–1003 (2007).
- ¹⁰⁴J. S. Holt, D. Sadoskas, and C. J. Pursell, “Infrared spectroscopy of the solid phases of ammonia”, *The Journal of Chemical Physics* **120**, 7153–7157 (2004).
- ¹⁰⁵N. J. Mason, A. Dawes, P. D. Holtom, R. J. Mukerji, M. P. Davis, B. Sivaraman, R. I. Kaiser, S. V. Hoffmann, and D. A. Shaw, “Vuv spectroscopy and photo-processing of astrochemical ices: an experimental study”, *Faraday Discussions* **133**, 311–329 (2006).
- ¹⁰⁶E. F. van Dishoeck, E. Herbst, and D. A. Neufeld, “Interstellar water chemistry: from laboratory to observations”, *Chemical Reviews* **113**, 9043–9085 (2013).
- ¹⁰⁷I. R. Cooke, E. C. Fayolle, and K. I. berg, “ CO_2 infrared phonon modes in interstellar ice mixtures”, *The Astrophysical Journal* **832**, 5 (2016).
- ¹⁰⁸P. A. Gerakines and R. L. Hudson, “First infrared band strengths for amorphous CO_2 , an overlooked component of interstellar ices”, *The Astrophysical Journal* **808**, L40 (2015).
- ¹⁰⁹M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr.,

- J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, *Gaussian16 revision c.01*, Gaussian Inc. Wallingford CT, 2016.
- ¹¹⁰W. J. Hehre, R. Ditchfield, and J. A. Pople, “Selfconsistent molecular orbital methods. xii. further extensions of gaussian type basis sets for use in molecular orbital studies of organic molecules”, *The Journal of Chemical Physics* **56**, 2257–2261 (1972).
- ¹¹¹J.-D. Chai and M. Head-Gordon, “Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections”, *Phys. Chem. Chem. Phys.* **10**, 6615–6620 (2008).
- ¹¹²M. Matsumoto and T. Nishimura, “Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator”, *ACM Transactions on Modeling and Computer Simulation* **8**, 330 (1998).
- ¹¹³A. D. Becke, “Density functional thermochemistry. iii. the role of exact exchange”, *The Journal of Chemical Physics* **98**, 5648–5652 (1993).
- ¹¹⁴C. Adamo and V. Barone, “Toward reliable density functional methods without adjustable parameters: the pbe0 model”, *The Journal of Chemical Physics* **110**, 6158–6170 (1999).
- ¹¹⁵T. Yanai, D. P. Tew, and N. C. Handy, “A new hybrid exchange correlation functional using the coulomb-attenuating method (cam-b3lyp)”, *Chemical Physics Letters* **393**, 51–57 (2004).
- ¹¹⁶S. Grimme, S. Ehrlich, and L. Goerigk, “Effect of the damping function in dispersion corrected density functional theory”, *Journal of Computational Chemistry* **32**, 1456–1465 (2011).
- ¹¹⁷A. D. McLean and G. S. Chandler, “Contracted gaussian basis sets for molecular calculations. i. second row atoms, z=1118”, *The Journal of Chemical Physics* **72**, 5639–5648 (1980).

- ¹¹⁸R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople, “Selfconsistent molecular orbital methods. xx. a basis set for correlated wave functions”, *The Journal of Chemical Physics* **72**, 650–654 (1980).
- ¹¹⁹D. Pines, J. Ditkovich, T. Mukra, Y. Miller, P. M. Kiefer, S. Daschakraborty, J. T. Hynes, and E. Pines, “How acidic is carbonic acid?”, *The Journal of Physical Chemistry B* **120**, PMID: 26862781, 2440–2451 (2016).
- ¹²⁰N. Stolte and D. Pan, “Large presence of carbonic acid in co₂-rich aqueous fluids under earth’s mantle conditions”, *The Journal of Physical Chemistry Letters* **10**, PMID: 31411889, 5135–5141 (2019).
- ¹²¹C. Li and H. Ji, “Chemical weathering and the role of sulfuric and nitric acids in carbonate weathering: isotopes (13c, 15n, 34s, and 18o) and chemical constraints”, *Journal of Geophysical Research: Biogeosciences* **121**, 1288–1305 (2016).
- ¹²²S. Kanakiya, L. Adam, L. Esteban, M. C. Rowe, and P. Shane, “Dissolution and secondary mineral precipitation in basalts due to reactions with carbonic acid”, *Journal of Geophysical Research: Solid Earth* **122**, 4312–4327 (2017).
- ¹²³M. Abbaszadeh, M. Nasiri, and M. Riazi, “Experimental investigation of the impact of rock dissolution on carbonate rock properties in the presence of carbonated water”, *English, Environmental Earth Sciences* **75**, 10.1007/s12665-016-5624-3 (2016).
- ¹²⁴W. Zheng and R. I. Kaiser, “On the formation of carbonic acid (h₂co₃) in solar system ices”, *Chemical Physics Letters* **450**, 55–60 (2007).
- ¹²⁵B. M. Jones, R. I. Kaiser, and G. Strazzulla, “Carbonic Acid as a Reserve of Carbon Dioxide on Icy Moons: The Formation of Carbon Dioxide (co₂) in a Polar Environment”, *The Astrophysical Journal* **788**, 170 (2014).
- ¹²⁶S. Radhakrishnan, M. S. Gudipati, W. Sander, and A. Lignell, “Photochemical processes in co₂/h₂o ice mixtures with trapped pyrene, a model polycyclic aromatic hydrocarbon”, *The Astrophysical Journal* **864**, 151 (2018).

- ¹²⁷P. A. Gerakines, D. C. B. Whittet, P. Ehrenfreund, A. C. A. Boogert, A. G. G. M. Tielens, W. A. Schutte, J. E. Chiar, E. F. van Dishoeck, T. Prusti, F. P. Helmich, and T. de Graauw, “Observations of solid carbon dioxide in molecular clouds with The infrared space observatory”, *The Astrophysical Journal* **522**, 357–377 (1999).
- ¹²⁸Hoang, M., Altwegg, K., Balsiger, H., Beth, A., Bieler, A., Calmonte, U., Combi, M. R., De Keyser, J., Fiethe, B., Fougere, N., Fuselier, S. A., Galli, A., Garnier, P., Gasc, S., Gombosi, T., Hansen, K. C., Jäckel, A., Korth, A., Lasue, J., Le Roy, L., Mall, U., Rème, H., Rubin, M., Sémon, T., Toubanc, D., Tzou, C.-Y., Waite, J. H., and Wurz, P., “The heterogeneous coma of comet 67p/churyumov-gerasimenko as seen by rosina: h₂o, co₂, and co from september 2014 to february 2016”, *A&A* **600**, A77 (2017).
- ¹²⁹B. J. Buratti, D. P. Cruikshank, R. H. Brown, R. N. Clark, J. M. Bauer, R. Jaumann, T. B. McCord, D. P. Simonelli, C. A. Hibbitts, G. B. Hansen, T. C. Owen, K. H. Baines, G. Bellucci, J.-P. Bibring, F. Capaccioni, P. Cerroni, A. Coradini, P. Drossart, V. Formisano, Y. Langevin, D. L. Matson, V. Mennella, R. M. Nelson, P. D. Nicholson, B. Sicardy, C. Sotin, T. L. Roush, K. Soderlund, and A. Muradyan, “Cassini visual and infrared mapping spectrometer observations of iapetus: detection of co₂”, *The Astrophysical Journal* **622**, L149–L152 (2005).
- ¹³⁰G. Filacchione, A. Raponi, F. Capaccioni, M. Ciarniello, F. Tosi, M. T. Capria, M. C. De Sanctis, A. Migliorini, G. Piccioni, P. Cerroni, M. A. Barucci, S. Fornasier, B. Schmitt, E. Quirico, S. Erard, D. Bockelee-Morvan, C. Leyrat, G. Arnold, V. Mennella, E. Ammannito, G. Bellucci, J. Benkhoff, J. P. Bibring, A. Blanco, M. I. Blecka, R. Carlson, U. Carsenty, L. Colangeli, M. Combes, M. Combi, J. Crovisier, P. Drossart, T. Encrenaz, C. Federico, U. Fink, S. Fonti, M. Fulchignoni, W.-H. Ip, P. Irwin, R. Jaumann, E. Kuehrt, Y. Langevin, G. Magni, T. McCord, L. Moroz, S. Mottola, E. Palomba, U. Schade, K. Stephan, F. Taylor, D. Tiphene, G. P. Tozzi, P. Beck, N. Biver, L. Bonal, J.-P. Combe, D. Despan, E. Flamini, M. Formisano, A. Frigeri, D. Grassi, M. S. Gudipati, D. Kappel, A. Longobardo, F. Mancarella, K. Markus, F. Merlin, R. Orosei, G. Rinaldi, M. Cartacci, A. Cicchetti, Y. Hello, F. Henry, S. Jacquino, J. M. Reess, R. Noschese, R. Politi, and G. Peter, “Seasonal exposure of

- carbon dioxide ice on the nucleus of comet 67p/churyumov-gerasimenko”, *Science* **354**, 1563–1566 (2016).
- ¹³¹T. Kuroda, A. S. Medvedev, Y. Kasaba, and P. Hartogh, “Carbon dioxide ice clouds, snowfalls, and baroclinic waves in the northern winter polar atmosphere of mars”, *Geophysical Research Letters* **40**, 1484–1488 (2013).
- ¹³²J. Bernard, M. Seidl, I. Kohl, K. R. Liedl, E. Mayer, Ò. Gálvez, H. Grothe, and T. Loerting, “Spectroscopic observation of matrix-isolated carbonic acid trapped from the gas phase”, *Angewandte Chemie International Edition* **50**, 1939–1943 (2011).
- ¹³³Y. Oba, N. Watanabe, A. Kouchi, T. Hama, and V. Pirronello, “Formation of Carbonic Acid (h_2co_3) by Surface Reactions of Non-Energetic oh Radicals With co Molecules at Low Temperatures”, *The Astrophysical Journal* **722**, 1598–1606 (2010).
- ¹³⁴H. P. Reisenauer, J. P. Wagner, and P. R. Schreiner, “Gas-phase preparation of carbonic acid and its monomethyl ester”, *Angewandte Chemie International Edition* **53**, 11766–11771 (2014).
- ¹³⁵W. Hage, A. Hallbrucker, and E. Mayer, “Carbonic acid: synthesis by protonation of bicarbonate and ftir spectroscopic characterization via a new cryogenic technique”, *Journal of the American Chemical Society* **115**, 8427–8431 (1993).
- ¹³⁶A. Dibenedetto, M. Aresta, P. Giannoccaro, C. Pastore, I. Pápai, and G. Schubert, *On the existence of the elusive monomethyl ester of carbonic acid [ch3oc (o) oh] at 300 k: 1h-and 13c nmr measurements and dft calculations*, 2006.
- ¹³⁷J. P. Wagner, H. P. Reisenauer, V. Hirvonen, C.-H. Wu, J. L. Tyberg, W. D. Allen, and P. R. Schreiner, “Tunnelling in carbonic acid”, *Chemical Communications* **52**, 7858–7861 (2016).
- ¹³⁸M. M. Linden, J. P. Wagner, B. Bernhardt, M. A. Bartlett, W. D. Allen, and P. R. Schreiner, “Intricate conformational tunneling in carbonic acid monomethyl ester”, *The Journal of Physical Chemistry Letters* **9**, PMID: 29544243, 1663–1667 (2018).
- ¹³⁹A. Barnes, W. Orville-Thomas, R. Gaufres, and A. Müller, *Matrix isolation spectroscopy*, Vol. 76 (Springer Science & Business Media, 2012).

- ¹⁴⁰E.-M. Köck, J. Bernard, M. Podewitz, D. F. Dinu, R. G. Huber, K. R. Liedl, H. Grothe, E. Bertel, R. Schlögl, and T. Loerting, “Alpha-carbonic acid revisited: carbonic acid monomethyl ester as a solid and its conformational isomerism in the gas phase”, *Chemistry – A European Journal* **26**, 285–305 (2020).
- ¹⁴¹S. K. Reddy and S. Balasubramanian, “Carbonic acid: molecule, crystal and aqueous solution”, *Chem. Commun.* **50**, 503–514 (2014).
- ¹⁴²J. Murillo, J. David, and A. Restrepo, “Insights into the structure and stability of the carbonic acid dimer”, *Physical Chemistry Chemical Physics* **12**, 10963–10970 (2010).
- ¹⁴³S. K. Reddy, C. H. Kulkarni, and S. Balasubramanian, “Theoretical investigations of candidate crystal structures for β -carbonic acid”, *The Journal of Chemical Physics* **134**, 124511 (2011).
- ¹⁴⁴Y. Oba, N. Watanabe, A. Kouchi, T. Hama, and V. Pirronello, “Experimental study of CO_2 formation by surface reactions of non-energetic OH radicals with CO molecules”, *The Astrophysical Journal* **712**, L174–L178 (2010).
- ¹⁴⁵A. D. Beck, *J. Chem. Phys.* **98**, 5648–6 (1993).