Theoretical and Experimental Investigations in Chemistry:

Part 1- S_N2 Reactivity at C-6 in Hexopyranosides

Part 2- Polarizability and Raman Intensities in Hydrocarbons

Richard Dawes
Department of Chemistry
University of Manitoba

Table of Contents

List of Figures, Tables, Mathcad Sheets and Reaction Schemes

Figures

Tables

MathCad Sheets

Reaction Schemes

Abstract

In this thesis, two rather different types of problems were investigated. The first was a case of anomalous reactivity in the area of carbohydrate synthesis. Hexopyranosides having the *galacto* configuration (i.e. C-4–OR *axial*) display very low S_N2 reactivities towards anionic nucleophiles, whereas the corresponding *gluco*-configured C-6 sulfonates (C-4–OR *equatorial*) react at rates typical of primary centers. The accepted explanation for this difference involves the repulsive interaction of local dipoles in the transition structure of the *galacto* compound. This interaction is thought to destabilize the transition structure, making this reaction difficult. However, there are numerous inconsistencies in the application of this simple model (cases where the model fails to predict the observed behavior). Thus, a computational project was undertaken to examine six model systems of this type. The energetics and equilibria of the reactants were determined, including solvation. Reaction pathways and kinetics for various displacements were computed. Analyses of the calculated charge densities allowed for evaluation of any electrostatic interactions. This study revealed a number of important factors affecting the rates of reaction, while clearly showing that dipole-dipole interactions are very limited in these systems.

The second project was in the related areas of molecular polarizability and vibrational spectroscopy. Descriptive models have been sought relating the structure and connectivity of molecules to their electronic properties. Researchers in the areas of non-linear optics and of conducting polymers require a better understanding of the effects of structural variations on electronic properties. The

simplest models of molecular polarizability and its derivatives with respect to molecular vibrations are grossly inadequate. The highest-level calculations are generally reliable but are not applicable to even moderately sized systems. Thus trends in these properties were investigated for a large series of molecules. Calculations were performed at a variety of theoretical levels in order to determine the ranges of predicted behavior. Since these calculations predicted unusual properties in bicyclo-[1.1.1]-pentane, an experimental project was completed on this molecule. The results of Raman scattering intensity experiments on bicyclo-[1.1.1]-pentane allowed for evaluation of the performance of various computational methods. More importantly, it allowed for the confirmation of some qualitative structure/property relationships.

General Introduction

It was proposed that the application of computational chemical methods combined with other theoretical considerations would provide insight into problems of interest in chemistry. This very general thesis statement is indicative of how broadly applicable computational chemistry is today. Soon after the development of quantum mechanics it was stated that, "In so far quantum mechanics is correct, chemical questions are problems in applied mathematics" (Eyring *et al.* 1944). Although this is somewhat true, a more accurate statement might be that "…chemical questions are *intractable* problems in applied mathematics".

The dynamics of chemical systems are invariably those of the many body problem and are therefore not approachable by analytical means. In spite of this immediate limitation, a tremendous amount of progress has been made. Currently, many sophisticated methods of approximation are in use. A wide range of these methods will be described in the course of this thesis.

The power of computational chemistry is of course directly linked to the progress in computing power. The vast increases in available computing resources over the past twenty years or so have now made it possible to examine systems of real chemical interest. While some of the theoretical foundations for quantum chemical calculations have been in place for many years, their application without computers is impossible. This is well-known, and most people are under the impression that brute-force computing, with faster and faster computers, is how advances are made. Unfortunately the scaling of the most accurate methods in quantum chemistry is so unfavorable that even if computers

continue to increase in speed at the current rate, systems of more than six atoms will still be beyond our capabilities for many years. Fortunately theoretical chemists have not sat idly by waiting for better computers to come along. In fact, advances in algorithms and computational methodologies have greatly outpaced the advances in data storage and raw computation. In particular, methods for calculating the properties of very large matrices have improved dramatically. Furthermore, while the most accurate (and costly) methods are still important for benchmarking purposes, most of the current effort is directed towards improving the accuracy of more widely applicable approximate methods.

Granting that it is possible to calculate the energy and physical properties of chemical systems to various degrees of accuracy, it is more important to know how best to direct one's computational efforts in order to acquire the maximum insight and "chemical intuition". Simply to recover an experimental result, "yes, I could predict that your reaction wouldn't work", is not enough. Since one could never compute the dynamics of every possible reaction combinatorially, then trends, correlations, and consistent frameworks must be elucidated, and continually improved. It is the role of computational chemistry to probe apparent inconsistencies and through various types of analyses provide new understanding of observed phenomena.

The richness of chemistry is due to its endless complexity. Chemistry has always been understood in terms of trends and simple models. The aforementioned apparent inconsistencies are simply the failures and breakdown of models applied beyond their range. Simple models for chemical behavior are still

useful and necessary, as is the recognition of the limits of their applicability. As chemists in all areas seek to tune the properties of molecules and materials on an atomic basis, increasingly detailed insight is required.

In this thesis two such models were investigated and dissected, and the insight gained was applied to extending the models. The models described, and the chemical physics involved, are so different that this thesis is split into two entirely separate sections. However, in this introduction I hope to convey that this diversity is the strength of theoretical chemistry. While the most appropriate theoretical methods for investigating different phenomena in chemistry differ considerably, there are often some essential similarities. It is probably easier for a theoretical chemist to consider different types of chemistry than it would be for experimentalists. It is through interaction with experimentalists that theoretical chemists become aware of interesting problems for analysis. This has certainly been the case for me. The projects may be quite unrelated, but the connection is in the approach. For each question, the appropriate tools of quantum chemistry were considered. Model systems were conceived and studied, the behavior dissected, and the models extended or discarded. When models are extended, based on the results of theoretical considerations, it is important to plan experiments designed to test the new models for the most extreme cases of their applicability. This experimentation was a large aspect of the work described in the second part of this thesis. As is often the case, these experiments produced data which required theoretical considerations for their interpretation. This is

illustrative of the continuous interplay between theory and experiment, each supporting the other.

Notes on Presentation

Both of the projects discussed in this thesis were rather lengthy investigations which included a number of different approaches. The Raman intensities project was a continuation of previous work in the Gough research group, particularly that of Jason Dwyer. It would have been difficult to imagine the scope which this study eventually encompassed. The project went from theoretical to experimental and back again. It began with survey calculations and went on to synthesis and recording of spectra, then an anharmonic force-field analysis was performed and finally inverse-eigenvalue algorithms. Similarly the reactivity study covered much ground and many forms of analysis.

Thus to maintain a reasonable level of continuity, the presentation through the main parts of this thesis is similar to that of manuscripts published in these areas. A relatively terse style is used, and some familiarity with theoretical chemistry is assumed. This will allow the main focus and direction of the research to remain clear. In order to provide additional information about some of the methods which are used, but only mentioned in passing in the text, a number of appendices have been included. Each of the two parts of the thesis has its own set of appendices which describe particular methods in more detail. Actual MathCad^{TM} worksheets have been reproduced to provide explicit examples of some of the analysis. Sample calculations are included for each type of calculation. Overall the relative size of the appendices is quite large. The appendices themselves are not intended to be just simple data tables or references, but rather, fully developed discussions about each relevant topic.

Computational Methodologies A: Post-SCF Methods

The Hartree-Fock self-consistent field (SCF) method is the foundation of all the methods used in the work described in this thesis. For a closed-shell system, the Hartree-Fock wavefunction is a variationally optimized single Slater determinant. The method is size-extensive and thus allows post-SCF methods based on the SCF the possibility of remaining so (an expansion of size-extensive terms could also be size-extensive). The SCF wavefunction is constructed with the assumption that the electrons are independent particles. Therefore the SCF method lacks all electron correlation except Fermi. All of the higher-level methods discussed here make some attempt to describe the electron correlation and thus improve upon the SCF wavefunction. For a closed-shell system with only one important electron configuration, the SCF method is surprisingly good. When multiple configurations are important, the method can also be extended to multiconfigurational SCF. None of the systems described in this thesis required such an extension. Under the closed-shell single-configuration conditions (which are most usual), the SCF method recovers total electronic energies to within about 1%, and many physical and electronic properties to within about 10%.[1] In some senses, this is a triumph and is a good reflection of the physics of the system. However, it is not a good reflection of the *chemistry* of the system. From a physics standpoint, to develop the electronic structure of a molecule, and capture 99% of the total energy is quite good. However, the chemistry of a system is determined by very small energy differences, which determine conformer populations, reaction barriers etc. Thus the drive in theoretical chemistry is to

develop affordable highly accurate methods. Since the Hartree-Fock method is so fundamental, some details of the Fock operator, the orbitals, and the energy are listed below.

The Hamiltonian operator is written as:[1]

$$\hat{H} = \hat{h} + \hat{g} + \hat{h}_{nuc} = \sum_{pq} h_{pq} E_{pq} + \frac{1}{2} \sum_{pqrs} g_{pqrs} e_{pqrs} + h_{nuc} \tag{A1}$$

where h and g are the one and two electron terms and h_{nuc} is a constant nuclear-nuclear repulsion term.

The Fock operator is written as:

$$\hat{f} = \hat{h} + \hat{V}$$

$$\hat{V} = \sum_{pq} V_{pq} E_{pq} = \sum_{pq} \sum_{i} \left(2g_{pqii} - g_{piiq} \right) E_{pq} \tag{A2-A3}$$

Thus the Fock potential V replaces the two-electron part g. This is an effective one-electron potential in which each electron interacts with the remaining electrons through their average density.

The orbital energies are:

$$\varepsilon_p = h_{pp} + \sum_i \left(2g_{ppii} - g_{piip} \right)$$

$$\varepsilon_p = \int \phi_p(r1) \left(-\frac{1}{2}\nabla^2 - \sum_K \frac{Z_K}{r1_K} \right) \phi_p(r1) dr1$$

$$+ 2\sum_i \int\int \frac{\phi_p(r1)\phi_p(r1)\phi_i(r2)\phi_i(r2)}{r12} dr1 dr2$$

$$- \sum_i \int\int \frac{\phi_p(r1)\phi_i(r1)\phi_i(r2)\phi_p(r2)}{r12} dr1 dr2 \tag{A4-A5}$$

As seen explicitly above, this expression assumes that the orbitals are real. This is the most common situation and simplifies the actual computation. If necessary

one can test for real to complex instabilities in a wavefunction, and then re-optimize the wavefunction if such an instability is found.

The Hartree-Fock energy is the expectation value of the Hamiltonian operator in the Fock state: $E_{HF} = \langle HF | \hat{H} | HF \rangle$ (A6)

This is not the expectation value of the Fock operator. The HF wavefunction is an eigenfunction of the Fock operator, with an eigenvalue equal to the sum of the orbital energies. It is of key importance to recognize that the sum of the orbital energies does not equal the total Hartree-Fock energy, but is only the eigenvalue of the Fock operator. The fluctuation potential can be defined as:

$$\hat{\Phi} = \hat{g} - \hat{V}$$ (A7)

Thus the fluctuation potential is the difference between the two-electron operator and the Fock potential. This allows the Hamiltonian H to be expressed in terms of the Fock operator f, the fluctuation potential, and the nuclear repulsion term.

$$\hat{H} = \hat{f} + \hat{\Phi} + h_{nuc}$$ (A8)

This gives the following expression for the Hartree-Fock energy:

$$E_{HF}^{(0)} = 2 \sum_i \varepsilon_i + \langle HF | \hat{\Phi} | HF \rangle + h_{nuc}$$ (A9)

The factor of 2 in front of the orbital energies reflects that for a closed shell system, the canonical spin orbitals are doubly occupied. The superscript (0) is to indicate that this formulation can now be considered as the Hartree-Fock orbital energies corrected to first order in the basis of the Hartree-Fock state by the fluctuation potential. This does not include electron correlation, but foreshadows the natural extension to higher order corrections, which do.

Møller-Plesset Perturbation Theory

The MP2 method is based on the Hartree-Fock state. The Hartree-Fock energy can be considered to be that of the HF orbital energies corrected to first order in the fluctuation potential. The MP methods can then be viewed as higher-order corrections in the fluctuation potential. It is well known in perturbation theory that a first-order wavefunction correction allows determination of the second-order energy correction. In this case the Hartree-Fock energy already uses the first-order energy correction with the zeroth-order wavefunction (the HF state). Thus by obtaining the MP1 wavefunction, the MP2 energy may be computed. Shown below are the MPn energies through MP2.[1]

$$E_{MP}^{(0)} = \left\langle HF \left| \hat{f} \right| HF \right\rangle = \sum_{I} \varepsilon_{I}$$

$$E_{MP}^{(1)} = \left\langle HF \left| \hat{\Phi} \right| HF \right\rangle \qquad\qquad\qquad (A10\text{-}A12)$$

$$E_{MP}^{(2)} = - \sum_{A>B,I>J} \frac{\left| g_{AIBJ} - g_{AJBI} \right|^2}{\varepsilon_A + \varepsilon_B - \varepsilon_I - \varepsilon_J}$$

As will be illustrated in more detail shortly, the second-order energy correction is the expectation value of the fluctuation potential operator using the Hartree-Fock and MP1 states. Expressing the energy terms this way allows the Hartree-Fock and MP2 energies to be written as:

$$E_{HF} = E_{MP}^{(0)} + E_{MP}^{(1)} + h_{nuc} = \left\langle HF \left| \hat{H} \right| HF \right\rangle \qquad (A13\text{-}A14)$$
$$E_{MP2} = E_{MP}^{(0)} + E_{MP}^{(1)} + E_{MP}^{(2)} + h_{nuc}$$

The second correction to the energy as shown above can be written as:

$$E_{MP}^{(2)} = \langle HF | \hat{\Phi} | MP1 \rangle \tag{A15}$$

It is clear how this arises in terms of a typical perturbation theory expansion. However, since the next post-SCF method to be discussed is the coupled-cluster method, this is a good place to introduce excitation operators. Again, the second quantized form is used to describe wavefunctions. Introducing an excitation operator, the second energy correction can now be written in terms of the Hartree-Fock state.[1]

$$E_{MP}^{(2)} = \langle HF | \hat{\Phi} | MP1 \rangle = \langle HF | \hat{H} \hat{T}_2^{(1)} | HF \rangle \tag{A16}$$

The equality of the second and third terms should not imply that the Hamiltonian and fluctuation operators are equivalent. The details of some operator manipulations have been omitted. This expression shows how the MP1 state is constructed from the HF state through an excitation operator:

$$| MP1 \rangle = \hat{T}_2^{(1)} | HF \rangle \tag{A17}$$

As will be shown explicitly, this operator populates the virtual orbitals and depopulates the occupied orbitals in the same way that creation and annihilation operators work on field modes. The product of operators $\hat{H} \hat{T}_2$ can be replaced by their commutator. This will not be used here since its usefulness is only seen when the MP3 and higher corrections are considered and appear as a series of nested commutators.

The single and double excitation operators are as follows:

$$\left|{A \atop I}\right\rangle = \hat{\tau}_I^A \left|HF\right\rangle = a_A^{\mp} a_I \left|HF\right\rangle$$

$$\left|{AB \atop IJ}\right\rangle = \hat{\tau}_{IJ}^{AB} \left|HF\right\rangle = a_A^{\mp} a_I a_B^{\mp} a_J \left|HF\right\rangle \qquad \text{(A18-A19)}$$

Thus using the creation and annihilation operators the various singly and doubly excited states may be formed. The single excitation operator above can be seen to depopulate orbital I, while populating orbital A. Higher excitations are performed the same way. Excitations higher than single require a restriction on the indices. For double excitations, the conditions A>B and I>J ensure that no duplicate states are created.

Finally it can be shown how the MP1 state is constructed. The operator T has a subscript 2 denoting that these are double excitations, and a superscript 1, indicating that the amplitudes of the created states are derived from first-order perturbation theory. Thus the operator is a summation over all double excitations, weighted by an amplitude factor.

$$\hat{T}_2^{(1)} = \sum_{A>B, I>J} t_{IJ}^{AB(1)} a_A^{\mp} a_I a_B^{\mp} a_J$$

$$t_{IJ}^{AB(1)} = -\frac{\left\langle HF \left|\left[a_J^{\mp} a_B a_I^{\mp} a_A, \hat{H} \right]\right| HF \right\rangle}{\varepsilon_A + \varepsilon_B - \varepsilon_I - \varepsilon_J} \qquad \text{(A20-A21)}$$

When the operator T is used to create the MP1 state, it becomes clear how the MP2 energy is obtained. The higher MP states and energies can also be expressed as a set of excitation operators acting on the HF state. This method

16

differs from the coupled-cluster method in that there is always a *linear* combination of excitation operators. This means that the highest excitation is that of the highest excitation operator. Due to the tremendous possible number of triple and quadruple excitations, and the complexity of evaluating the higher-order amplitudes necessary to incorporate them, the higher methods are only practical for small numbers of electrons. The MP2 method scales as the 5th power of the number of electrons, while MP3 and MP4 scale as the 6th and 7th powers, respectively.[1] The MP methods do maintain size-extensivity but they are no longer variational. Their convergence is not guaranteed and is usually oscillatory. Despite these considerations, corrections as high as MP4 and even MP5 are commonly used and are included in the Gaussian program code. Since each correction beyond HF requires the construction of perturbative wavefunctions, this method rapidly becomes uneconomical. In the research described in this thesis, the MP2 and MP4 methods were evaluated. The MP4 method was found to be too costly, while the MP2 method showed no advantage over B3LYP for the applications considered here.

Coupled-Cluster Theory

The Coupled-cluster method was the highest level of theory used in the research described in this thesis. It is widely known as one of the most accurate theoretical methods available. Although it is non-variational and scales as the 7th power of the number of electrons, the coupled-cluster method (specifically CCSD(T)), is widely used for benchmarking and the development of new

methods (DFT). As with the MP methods, the coupled-cluster method is based on the application of excitation operators to an initial Hartree-Fock state. Thus this method again relies on there being only one important electron configuration, which is well represented by the Fock state. In fact, there is an economical diagnostic calculation that may be run on the HF state, to determine whether or not this is correct.[2] This prevents the running of a lengthy coupled-cluster calculation whose results would be poor. Most common systems encounter no such difficulties, making the method widely useful when a definitive answer is required.

Within the uncorrelated Hartree-Fock state, the electrons occupy the lower energy spin orbitals and don't interact. The correlation of electrons (their interaction) is reflected in excitations to the virtual orbitals. Each possible excitation has its own amplitude, the probability of this excitation occurring. As more possible excitations are considered, the electron correlation is described more completely. In the limit of full-CI, all possible excitations are considered. In coupled-cluster theory, a truncated set of excitation operators is used to treat the excitations. If the coupled-cluster excitation series were not truncated, it would describe the same state as the full-CI. With truncation, and in contrast to the MP series, it will be shown how the CC method incorporates many higher excitations indirectly. This is seen as the reason for its superiority over the MP and CISD methods. Without considering the amplitudes, the excitation expansion may be written for the CC state $|CC\rangle$ as:

$$|CC\rangle = \exp\left(\hat{T}\right)|HF\rangle$$

$$= \sum_{i=0}^{N} \hat{C}_i |HF\rangle \tag{A22}$$

The C's in the expansion are the set of excitation operators T, which cause excitation to the level of the index of C.

$$\hat{C}_0 = 1$$

$$\hat{C}_1 = \hat{T}_1$$

$$\hat{C}_2 = \hat{T}_2 + \frac{1}{2}\hat{T}_1^2 \tag{A23}$$

$$\hat{C}_3 = \hat{T}_3 + \hat{T}_1 \hat{T}_2 + \frac{1}{6}\hat{T}_1^3$$

The key to the method is in the truncation. Looking at C_3, one can see three mechanisms for triple excitation. The direct triple excitation through T_3 is called the connected excitation, while the product of T_1 and T_2, and the cube of T_1 are called disconnected excitations. Thus the triple and higher excitations may be achieved in terms of lower excitation operators. When the series is truncated, it is not truncated at a level of excitation; it is truncated at a level of excitation operator. Thus the series is still summed over all of the C's, but including only some of the operators. If, for example, the CCSD method were used, only the T_3 term would be omitted from the above expression. This allows at least some contribution from all configurations, whereas truncating the excitation level directly in a linear expansion (CISD) does not. Furthermore the method can be shown to be size-extensive no matter which truncation level is chosen.[1]

The most commonly used coupled-cluster method and the one used in this thesis is CCSD(T). This method is truncated after the double excitation operators although higher excitations are still included through the products of the single and double operators. The effect of the connected triple excitations, through the triple excitation operator, is included via perturbation theory rather than direct inclusion of the triples operator. Fortuitously, the implementation through perturbation theory often produces a result closer to that obtained by the full-CI method than by explicit inclusion of the triples operator (CCSDT). This accounts for the tremendous success of the method relative to its cost (which is still extremely high).

B3LYP

Another very popular method also used here is the hybrid-Density Functional Theory (DFT) method known as B3LYP.[3] Developed in the early 1990's by Axel Becke at Queens University in Canada, B3LYP has rapidly become one of the most widely used methods in all electronic structure calculations. Density functional theory, like most theoretical methods, has a hierarchical structure of methods. B3LYP uses functionals at what is known as the generalized gradient approximation (GGA) level as well as ones at the LDA (local density approximation) level. A functional is a function of a function. For example, if the electron density is a function of the nuclear coordinates (as will be discussed regarding the Born-Oppenheimer approximation), and the energy is a function of the density, then the energy can be said to be a *functional* of the

nuclear coordinates. Implementation is straightforward and is included in most commercial codes such as Gaussian. First an HF self-consistent field state is computed. Some exact exchange energy from the HF calculation is incorporated into the DFT calculation. Then during the DFT calculation, the density and its gradient are evaluated at a large number of points on a grid through space. The fineness of the grid may be specified. An iterative solution is found to a pseudo eigenvalue equation, including three functionals of the density and its gradient throughout the grid. The total energy and its gradient with respect to nuclear displacements are then computed. Depending on the type of calculation (e.g. geometry optimization) the calculation proceeds just as it would with an HF calculation. The functionals Becke used to replace the Hartree-Fock exchange energy term are shown below followed by a more detailed explanation:[4-6]

$$E_X(HF) \rightarrow E_{XC} = E_{XC}^{LDA} + 0.2(E_X^{HF} - E_X^{LDA}) + 0.72\Delta E_X^{B88} + 0.81\Delta E_C^{PW91} \quad (A24)$$

This method uses the Hartree-Fock method, which lacks correlation, but then has the exchange-energy term replaced by a term called the exchange-correlation term. This is how some account for correlation is introduced. As can be seen in the first term of the above expression, the exchange-correlation is added from a local-density approximation functional. The second term introduces some of the Hartree-Fock exchange, while removing some of the LDA exchange. The third and fourth terms add a parameterized contribution from two gradient-based functionals. The current implementation in Gaussian software is with slightly different functionals, but the parameterization is the same.

The method works quite well, but is deserving of a few qualifying comments. Although there are physical arguments for using these functionals to describe the correlation, there are no physical grounds for the specific parameterization chosen. In a sense, the method is semi-empirical, since the parameters were chosen in order to obtain the best possible agreement with experiment. Also the capacity for improvement at the GGA level seems to have been exhausted. Becke's experiments with up to 14 parameterized functionals at the GGA level have produced no significant improvement over B3LYP. The next level of DFT does appear promising, and as a natural extension to a method that includes the density and its gradient, meta-GGA's now include functionals based on the Laplacian of the density. At this time the B3LYP method is usually considered to be comparable in performance with the MP2 method, while significantly less costly.

Computational Methodologies B: Basis-Set Superposition Error and Size-Extensivity

Often in theoretical chemistry it is desirable to calculate the interaction energy between molecules or fragments. In this thesis, the energies of the model sugar compounds were computed separately from those of the chloride atoms. This is equivalent to considering them to be at infinite separation. When the chloride ion approaches the sugar molecule during the displacement reaction, there is a real physical interaction between the two systems. Whether attractive or repulsive, the computational method should capture this interaction and its energy. Unfortunately, an additional spurious interaction is computed when finite basis-sets are used (always). This non-physical interaction appears to be attractive since it always lowers the system energy (for variational wavefunctions), and is proportional to the distance between systems. The interaction energy between two systems A and B might be computed as:

$$E_{INT(A \bullet B)} = E_{A \bullet B} - E_A - E_B \qquad \text{(B1)}$$

This is simply the difference between the energy of the complex and that of the isolated systems. The non-physical interaction energy, called basis-set superposition error (BSSE), is due to the lack of completeness of the finite basis-set.[7] The basis set is the set of functions generally centered at the nuclei, used to describe the electron density. Depending on the size of the basis set, assuming that the variational method is used, there will be a difference between the calculated energy and that which would be reached in the lower limit of a complete basis (the Hartree-Fock limit). This is because the functions chosen are

not flexible enough to describe the density optimally. With smaller basis sets this difference can be quite large. When two fragments are brought into proximity the functions centered on the nuclei of fragment A begin to overlap with the functions of fragment B. This is important in order to describe the density of the total system and thus capture the real interaction. However it also means that fragment A, using the functions of fragment B in addition to its own, may enjoy a larger and more flexible basis for description of the density of A. Similarly fragment B benefits from the presence of fragment A. The BSSE energy may be computed and removed by the method of Boys and Bernardi.[7] Their approach is simple; one calculates the energy of fragment A as before and compares this with the energy of fragment A calculated with the additional functions of fragment B, placed where the nuclei of B appear in the complex, but without the nuclear charges. These ghost atoms, with their associated functions, allow calculation of the spurious lowering of A's energy by the functions of B. Similarly one corrects the energy of fragment B using the ghost atoms of A. This correction is known as the counterpoise correction. Due to some early theoretical debates, this is sometimes called the *complete counterpoise correction*, in contrast to other proposed corrections.

The relative importance of this interaction energy depends on the type of systems studied.[1] For covalent bonding interactions as described in this thesis, the effect is relatively small. When hydrogen-bonding interactions are the dominant interaction, such as with the water dimer, this effect grows in importance. BSSE could introduce error of more than 80% into the association energy for this

system. For the argon dimer, BSSE could easily far exceed the true interaction energy. One cannot then simply remove the spurious contribution because the minimum energy separation distance will also be seriously in error. Rather, one must plot the BSSE as a function of distance, and then remove it from the computed interatomic potential, leaving the true potential, thus allowing identification of the true minimum-energy geometry. Noble-gas dimers are good model systems for these calculations since the true interaction energies are extremely small, and yet have been accurately measured by various spectroscopies. Furthermore, when the mechanism of the true attraction is dispersion forces, then this attraction is an electron correlation effect and cannot be recovered by the Hartree-Fock method alone, even at the Hartree-Fock limit. The BSSE-corrected Hartree-Fock limit potential is repulsive at all distances. Therefore any noble-gas dimer structures computed by non-correlated methods are functions of an entirely spurious BSSE interaction.

In practice the magnitude of this effect is dependent on the system studied but in general, using the correlation-consistent basis sets (cc-pVXZ; X=2..7), the correction decreases by a factor of roughly 2-4 for each increase in the cardinal (zeta) number.

The importance of using size-extensive computational methods should be noted from the outset. A size-extensive method is one for which the calculated energy of a system of two non-interacting fragments would be equal to that of the sum of the energies of the individual fragments. This sounds like an obvious property to demand of a computational method but it is not realized by several

commonly used methods. The Hamiltonian of a system of two non-interacting

fragments A and B can be written as:[1]

$$\hat{H}_{AB} = \hat{H}_A + \hat{H}_B \tag{B2}$$

Using second-quantization and the wave-operators for creation of the states A and

B one can write:

$$
\begin{aligned}
|A\rangle &= \hat{\psi}_A |vac\rangle \\
|B\rangle &= \hat{\psi}_B |vac\rangle \\
|AB\rangle &= \hat{\psi}_{AB} |vac\rangle = \hat{\psi}_A \hat{\psi}_B |vac\rangle \\
E_{AB} &= E_A + E_B
\end{aligned}
\tag{B3-B6}
$$

These operators are applied to the vacuum, or empty Fock state. Their

action populates and depopulates the orbitals similarly to the way in which

creation and annihilation operators act on field modes. These requirements make

the wavefunction *multiplicatively separable*, and the energy *additively separable*.

This is not unlike the Born-Oppenheimer approximation discussed in appendix C,

where the separability of nuclear and electronic wavefunctions allowed for

product function solutions and added energy contributions. It is easy to see that

one cannot even begin to calculate interaction energies with a non-size-extensive

method. For example, Hartree-Fock, MPn, and coupled-cluster methods are size-

extensive, while CISD is not.

Computational Methodologies C: Atoms In Molecules

Developed over several decades by Richard Bader, the quantum theory of atoms in molecules[8] is seen by some as simply an alternative way of calculating atomic populations and therefore charges. Although it can indeed be used to determine atomic charges within a molecule, it is much more than that. Atoms In Molecules (AIM) is a completely rigorous formalism with which a molecule may be partitioned into proper atomic subsystems. The method of partitioning a molecule into its constituent atoms is such that within each atomic subsystem a consistent operator algebra is maintained for all of the typical operators of quantum mechanics. Thus, for example, one might evaluate the energies and charges for each of the atoms in a molecule. Many other physical observables are commonly treated. As would be expected, these atomic quantities would sum to equal those of the molecular system. It should be noted that there is a conceptual superiority to this method over many others. Methods of population analysis, such as Mulliken's, partition the electron density in an arbitrary way according to basis-set-dependent atomic and molecular orbitals. The AIM partitioning scheme is often applied to calculated electron densities, but does not rely on them. That is to say that the AIM method partitions real systems in real space, and can be applied to experimental densities obtained from scattering experiments.

The partitioning scheme is simple. Within the Born-Oppenheimer approximation, the electron density is a function of the nuclear coordinates. For a given nuclear configuration, there is a corresponding electron density. Each point in space has an associated density (a scalar value). Since the electron density is a

scalar field, then the gradient field ($\nabla \rho(r)$) is a vector field. Atoms are partitioned by surfaces of zero-flux through the gradient field.[8]

$$\nabla \rho(r) \bullet n(r) = 0 \qquad\qquad (C1)$$

Where n(r) is a vector normal to the gradient field contours.

Equation (C1) is simply the equation for a surface which crosses no gradient paths and thus separates the molecule into atoms. This is best shown by the example below. The molecule bicyclo-[1.1.1]-pentane is shown with critical points (a), and gradient field paths (b). It is clear in (b) where the atomic boundary lines of zero-flux through the gradient field would lie.

a)   b)
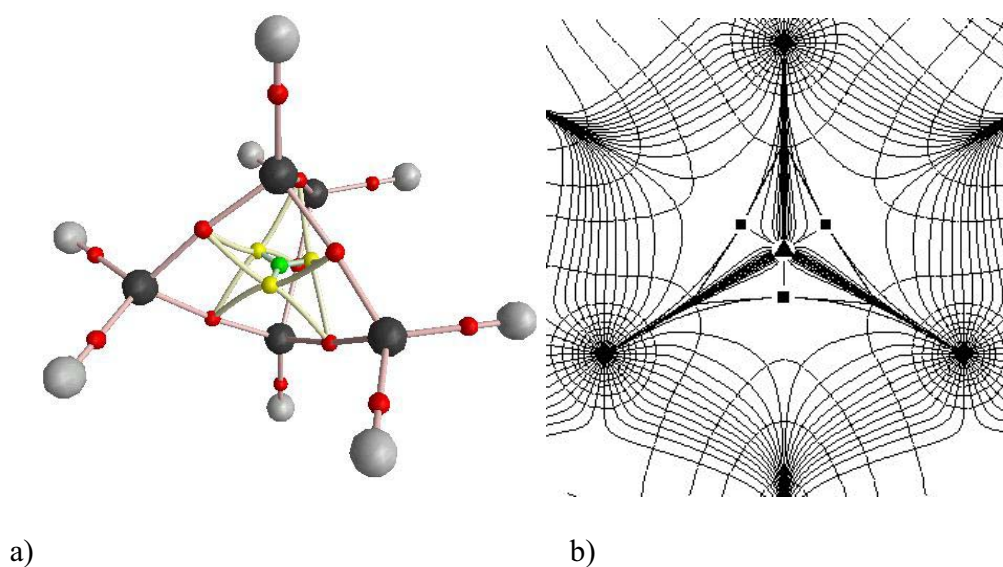
Figure C1: (a) density critical points (red, yellow, and green), (b) contour plot, and gradient field plot of bicyclo-[1.1.1]-pentane. Slice is in the plane of the methylene nuclei.

Once the molecule has been partitioned into its constituent atoms then the properties of the atoms themselves may be considered. Of primary importance to

this investigation is the fact that each atom is the source of a multipole expansion. Any deviation from a totally symmetric density surrounding any nucleus results in a corresponding atomic dipole moment.

The treatment of dipoles within the AIM formalism requires some explanation. Although the total molecular energy is recovered simply by summing the atomic contributions, the recovery of molecular dipoles is not quite as straightforward. Since the charge is partitioned into atomic basins, then each "atom" has a net charge (summing to zero for a neutral molecule), and each atom has an atomic dipole representing asymmetry of charge within the atomic basin. Thus, to recover a molecular dipole from atomic charges and dipoles, one adds the contribution from the relative positions of the charged nuclear attractors to that of the sum of the atomic dipoles. When a dipole is induced by a perturbation such as an applied field, or by changes in conformation, there are two possible contributions. Firstly, there is the change in the atomic dipoles representing distortion of charge within the basins. Secondly, there is a change in the charge-separation term, which includes both transfer of charge across the inter-atomic boundary, changing the amount of charge being separated, as well as geometry relaxations which change the distance by which charge is separated. The charge-separation contribution to the dipole is directed precisely between the two nuclear attractors, while in general the atomic dipole contribution can be in any direction.

What follows below are some worksheet examples of calculations involving the changes in atomic dipoles induced by conformational changes in the *galacto* and *gluco* model compounds. The details of this study are in part 1 of this thesis. These calculations are an example of how AIM might be used.   If molecular fragments were repelling each other through electrostatic interactions, then AIM analysis should reveal the extent of this on an atom-by-atom basis.
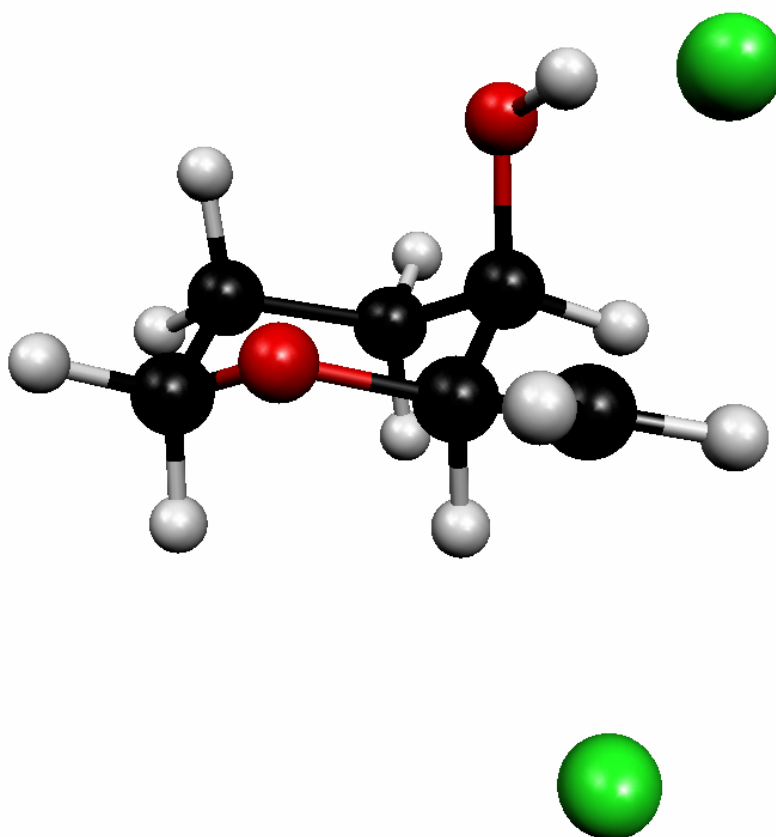
Figure C2: Model *galacto* compound **1TS**. Proposed dipole-dipole interaction is between bond dipole at C4-OH and that between C6 and the chlorine in the **gg** position.   A relatively parallel alignment is found in this structure.   Complete details of this study are found in Part 1 of this thesis.

Vector analysis of AIM derived atomic properties

Galacto-**Fluoro system** (low energy *tg* rotamer)

Cartesian coordinates of nuclear attractors units of bohr
$C4 := (-0.53122106 \quad 1.83791717 \quad -0.39211464)$
$F := (-0.42397502 \quad 2.74003343 \quad 2.1300081)$
$H5 := (2.76263308 \quad -0.65414661 \quad 2.86292133)$
$Cl := (5.47207978 \quad 0.11315015 \quad -0.6416103)$
$C6 := (2.82855736 \quad -1.40112481 \quad 0.94678476)$

Cartesian coordinates of bond critical points between C6-Cl, and C4-F, units of bohr
$CpClC6 := (3.9420865 \quad -0.7536074 \quad 0.2647736)$
$CpFC4 := (-0.4974708 \quad 2.1373013 \quad 0.4709218)$

Atomic dipoles units of electron·bohr
$C6D := (-0.076834 \quad -0.077341 \quad 0.101461)^T$
$ClD := (0.153007 \quad 0.084499 \quad -0.118789)^T$
$FD := (-0.008517 \quad -0.036441 \quad -0.114411)^T$
$C4D := (-0.030848 \quad -0.172064 \quad -0.546468)^T$

Defining normalized vector from C4 to F
$$FC4 := (F - C4) \cdot \frac{1}{\left| \sqrt{(F - C4) \cdot (F - C4)^T} \right|}$$

Defining normalized vector from H5 to F, this is the direction of the anticipated induced dipole, when Cl is rotated up into alignment. The projection of dipoles at C4, and F onto this vector will be followed through the rotamers.
$$FH5 := (F - H5) \cdot \frac{1}{\left| \sqrt{(F - H5) \cdot (F - H5)^T} \right|}$$

Defining normalized vector from C6 to Cl
$$ClC6 := (Cl - C6) \cdot \frac{1}{\left| \sqrt{(Cl - C6) \cdot (Cl - C6)^T} \right|}$$

Establish total magnitude of F atomic dipole
$$\sqrt{FD^T \cdot FD} = (0.120376)$$

Projection of F atomic dipole along C4-F bond
$$FC4 \cdot FD = (-0.120245)$$

Projection of F atomic dipole along vector from H5 to F, defined above.
$$FH5 \cdot FD = \left[-2.6933412879410^{-3} \right]$$

Naming the above quantity for calculation of perturbation in higher energy rotamers.
$$Gal2Ftrans := FH5 \cdot FD$$

Establish total magnitude of C4 atomic dipole
$$\sqrt{C4D^T \cdot C4D} = (0.5737463675)$$

Projection of C4 atomic dipole along C4-F bond
$$FC4 \cdot C4D = (-0.5732685)$$

Projection of C4 atomic dipole along vector from H5 to F, defined above.
$$FH5 \cdot C4D = (-0.01807821211)$$

Naming the above quantity for calculation of perturbation in gg rotamer.
$$Gal2C4trans := FH5 \cdot C4D$$

Establish total magnitude of Cl atomic dipole
$$\sqrt{ClD^T \cdot ClD} = (0.21133397638)$$

Projection of Cl atomic dipole along C6-Cl bond
$$ClC6 \cdot ClD = (0.20988729807)$$

Projection of C4 atomic dipole along vector from H5 to F, defined above.
$$FH5 \cdot ClD = (-0.02412629066)$$

Establish total magnitude of C6 atomic dipole
$$\sqrt{C6D^T \cdot C6D} = (0.14892759435)$$

Projection of C6 atomic dipole along C6-Cl bond
$$ClC6 \cdot C6D = (-0.14011233127)$$

Projection of C6 atomic dipole along vector from H5 to F, defined above.
$$FH5 \cdot C6D = (-0.01952728975)$$

Galacto-**Fluoro** (*gt* rotamer)

Cartesian coordinates of nuclear attractors units of bohr
$C4 := (1.72765232 \quad 2.05213808 \quad 0.81087222)$
$F := (1.88477606 \quad 3.32687401 \quad -1.53946147)$
$H4 := (-2.74308667 \quad 1.79574958 \quad -2.20965735)$
$Cl := (-5.64086887 \quad -0.58072023 \quad 0.20858362)$
$C6 := (-2.86916068 \quad 1.40132349 \quad -0.23633847)$

Cartesian coordinates of bond critical points between C6-Cl, and C4-F, units of bohr
$CpClC6 := (-4.043006843 \quad 0.5462193814 \quad -0.040030549)$
$CpFC4 := (1.77825024 \quad 2.480214883 \quad 0.00582354)$

Charge transfer (change in charge) relative to low energy tg rotamer. units of electrons
$chtC4 := 0.002362$
$chtF := -0.000906$
$chtCl := 0.012304$
$chtC6 := 0.004224$

Atomic dipoles units of electron·bohr
$C6D := (0.087631 \quad 0.115657 \quad -0.040365)^T$
$ClD := (-0.197609 \quad -0.093482 \quad 0.059917)^T$
$FD := (-0.005382 \quad -0.052678 \quad 0.112564)^T$
$C4D := (-0.020818 \quad -0.277986 \quad 0.50478)^T$

Defining normalized vectors as above (it is now H4 which is across from F)

$$FC4 := (F - C4) \cdot \frac{1}{\left| \sqrt{(F - C4) \cdot (F - C4)^T} \right|}$$

$$FH4 := (F - H4) \cdot \frac{1}{\left| \sqrt{(F - H4) \cdot (F - H4)^T} \right|}$$

$$ClC6 := (Cl - C6) \cdot \frac{1}{\left| \sqrt{(Cl - C6) \cdot (Cl - C6)^T} \right|}$$

Establish total magnitude of F atomic dipole
$$\sqrt{FD^T \cdot FD} = (0.124397)$$
Projection of F atomic dipole along C4-F bond
$$FC4 \cdot FD = (-0.124164)$$

Projection of F atomic dipole along vector from H4 to F, defined above.

$$\text{FH4} \cdot \text{FD} = \left[-6.12219038502 \cdot 10^{-3} \right]$$

Naming the above quantity for calculation of perturbation in higher energy rotamers.

$\text{Gal1Ftrans} := \text{FH4·FD}$

Defining and output of projection of charge transfer contribution to induced dipole onto H4-F vector. Charge transfer contribution to dipole can be seen as distance from critical point to nuclear attractor, multiplied by change in charge "$(\text{CpFC4} - \text{F}) \cdot \text{chtF}$" This is then projected via dot product with FH4.

$$\text{cht21F} := \text{FH4} \cdot ((\text{CpFC4} - \text{F}) \cdot \text{chtF})^{\text{T}}$$
$$\text{cht21F} = \left[1.38776005265 \cdot 10^{-4} \right]$$

Calculation and output of projection of induced dipole along H4-F normalized vector, relative to low energy tg rotamer.

$\text{Delta21Ftrans} := (\text{Gal1Ftrans} + \text{cht21F}) - \text{Gal2Ftrans}$

$$\text{Delta21Ftrans} = \left[-3.29007309182 \cdot 10^{-3} \right]$$

All other AIM dipole analysis was performed in a completely analogous fashion.

S$_N$2 reactivity at C-6 in hexopyranose derivatives

Introduction:

The S$_N$2 displacement may appear to be among the simplest organic reactions, but recent experimental and theoretical investigations have revealed surprising aspects of this ubiquitous process.[9, 10] The availability of accurate computational methods has made it relatively straightforward to investigate rigorously anomalies in reactivity that had previously been interpreted only through qualitative descriptions. A striking example of such an anomaly comes from synthetic carbohydrate chemistry.

It has been known for many years that C-6 sulfonate derivatives of hexopyranosides having the *galacto* configuration (i.e. C-4O̶R *axial*) display very low reactivities towards anionic nucleophiles, whereas the corresponding *gluco*-configured C-6 sulfonates (C-4O̶R *equatorial*) react at rates typical of primary centers.[11] In a rare example in which *galacto-* and *gluco-* compounds were subjected to a comparative kinetic study, the second-order rate constant for the reaction of azide with methyl-2,3,4-tri-*O*-acetyl-6-*O*-*p*-tolylsulfonyl-α-D-glucopyranoside was observed to be 32-fold greater than that for the reaction of the analogous galactoside.[12] In the majority of cases, attempted displacement of a C-6 sulfonate from a galactopyranoside fails completely and instead leads to 5,6-elimination or to the formation of a 3,6-anhydrosugar in competition with the expected S$_N$2 product, which is obtained in very low yield.[13] This is but one illustration of a group of cases in which the rates of S$_N$2 displacement apparently depend on the relative geometries of remote polar substituents. While chemists

have found some ways to minimize these difficulties in practice, these kinetic anomalies remain fascinating case studies in physical organic chemistry.

In 1969, A.C. Richardson proposed a persuasive qualitative explanation for reactivity differences of this type.[14] This rationalization assumed that differences in reactivity between the axial (*galacto*) and equatorial (*gluco*) configurations were a consequence of the differences in the energies of the respective transition structures. He argued that the transition structure for displacement would have a geometry in which the scissile C-6–X bond was orthogonal to the C-5–O-5 bond, i n order to minimize dipolar repulsion from interactions with the endocyclic oxygen. In the absence of a bulky substituent on O-4, he suggested that dipole-dipole interactions destabilized the S_N2 transition structure (Figure 1). In a transition structure of this type, the developing negative charge on the leaving group X^- would encounter unfavorable dipolar interactions with an axial electronegative group at C-4 (as in the *galacto* configuration). In the *gluco* configuration, the equatorial C-4–OR group would not create this destabilizing effect. When O-4 carried a bulky blocking group, the low reactivity of *galacto*-configured sulfonates was deemed to have a steric origin. Richardson used a similar dipole rationalization to explain differences in the S_N2 reactivities of *O*-sulfonate derivatives of the secondary hydroxyl groups in hexopyranosides.
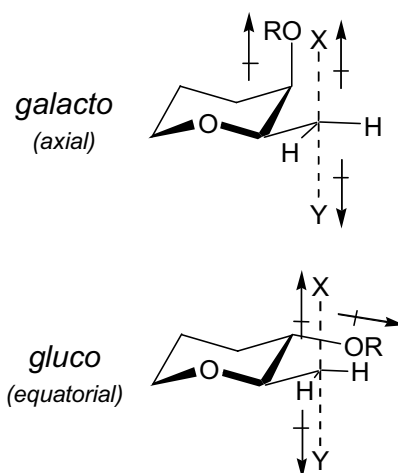
Figure 1: Postulated Dipole Interactions in S$_N$2 Displacements at C-6 in Galactopyranosides.

This steric/dipolar explanation has been nearly universally adopted by the carbohydrate community and is incorporated with very little comment into current textbooks.[15] It continues to be extensively cited and has been applied outside carbohydrate chemistry.[16] Craig and Brauman have recently restated a version of this rationale to explain rates of gas-phase S$_N$2 displacements in a series of ω-substituted primary *n*-alkyl chlorides, although they suggested that transition structure dipoles *enhanced* reaction rates.[17]

Despite its intuitive appeal and the breadth of its applicability, the fundamental correctness of the steric/dipole model is not obvious; moreover, it has never been examined in light of modern electronic theory. Although the initial description of the model only addressed reactions of sugars, it is apparent that the question of its validity has implications for understanding S$_N$2 reactivity in general.

This is an ideal model for investigation. The model is constructed simply in terms of the molecular structure and is therefore easy to apply. It is well

accepted and widely used, yet there are some known inconsistencies. For example, the tricyclic bis-acetonide **a** reacts with azide an order of magnitude more slowly than does the triacetate **b**,[13] while the permethylated tosylate **c** reacts with NaI twice as fast as **a** (see Figure 2).[18]
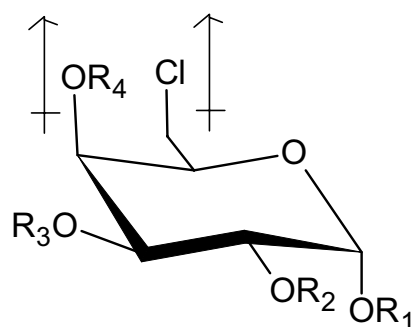
a) $R_1R_2R_3R_4 = C(CH_3)_2$
b) $R_1 = H_3, R_2R_3R_4 = Ac$
c) $R_1R_2R_3R_4 = H_3$

Figure 2: Examples of *galacto* configured compounds with proposed dipole-dipole interactions illustrated

Such observations are fundamentally inconsistent with a generalized transition state dipole effect involving the C4O fragment. Firstly, this local dipole would not differ greatly among various 4-*O*-alkyl galactose derivatives. Secondly, electronegativity considerations suggest that the local C4OAc dipole in more-reactive **b** should be *greater* than the C4OR dipole in the unreactive **a**

Given the inconsistencies of the qualitative model, and the lack of theoretical data, it was determined to investigate some model systems of this type. In the following section I will detail how the theoretical approach was chosen.

38

Theoretical methods

Upon first considering the reactions of the sugar systems, the immediate questions are the number of atoms and electrons and the type of information sought. The simplest molecules containing all of the structure necessary to reflect Richardson's hypothesis are composed of about 20 atoms and over 100 electrons. In addition, the model is described in terms of structural arrangements of functional groups within the molecule as well as their electrostatic interactions. Furthermore, reaction rates require accurate relative energies of all relevant species: reactants, products, and transition structures. These requirements restrict the possible approach from both ends. The systems are too large for the most accurate and sophisticated methods, yet the requirement for accurate structures, electronic structures, and relative energies preclude the use of semi-empirical methods.

The available computing facilities are quite powerful. The High-Performance-Computing facility at the University of Manitoba features a Sunfire 6800 computer. This computer is composed of 20 processors (1050 MHz), 40 GB memory, and nearly 1 terabyte of disk space. Although the computer is shared, I personally have had continuous access to 8 processors, 25 GB of memory, and 400 GB of disk space. Even so, for a benchmark level calculation (e.g. coupled cluster with singles, doubles, and perturbative triples, CCSD(T)), it would take years to calculate the structure and energy of just one reactant species. In contrast, with a typical semi-empirical method such as Austin Model 1 (AM1), the structure and energy of a model sugar compound could be computed in less

than 1 minute on a desktop PC. However, of concern for this project is that the structures themselves (and therefore the alignments of fragments) are known to be unreliable with AM1. Secondly, the electronic information is quite dubious. Finally, the energies obtained are approximate at best and not suitable for kinetics studies. Thus for this project, where electrostatic interactions of molecular fragments are important, as are the kinetics, something much higher is necessary. Although still in use, AM1 is now widely considered unsuitable for research level results.

Based on these considerations, the most accurate affordable *ab initio* method was required. For *ab initio* methods there are generally two considerations: the basis set, and the type of theory. The basis set is the set of functions used to describe the electron orbitals. Here again there is a compromise to be reached. The larger the basis set, the better the description of the electron density, and the more intensive the calculation. Fortunately much is known about the reliability of various basis sets for different applications. Given the requirements of this project and the computers available, it was initially determined that the 6-31+G(d,p) basis set should be used. This is a fairly standard medium-size basis set, commonly used for these purposes, with one exception. The + symbol indicates the inclusion of diffuse functions on all heavy atoms (those of second row, or higher). These diffuse functions are important for the description of longer-range interactions such as the making and breaking of bonds in transition structures.

In terms of the level of theory, the foundation of *ab initio* methods is the self-consistent-field or Hartree-Fock method. Even with a complete basis set, this method lacks inclusion of electron correlation. This often results in poor descriptions of electrostatic properties. All of the higher-level methods used in this thesis are based on the Hartree-Fock method, and begin by computing a self-consistent field. The most common higher-level methods (Post-SCF, see Comp-Meth A) which include electron correlation in some way are those of Density Functional Theory (DFT)[3], Moller-Plesset (MPn, n=2..5),[20] and coupled-cluster (CCSD(T)).[21] As stated previously, the coupled-cluster method is a benchmarking method and is too time consuming for this purpose. The Møller-Plesset methods are quite good, but are still relatively costly compared with the DFT methods. One such DFT method is B3LYP,[3] a hybrid density functional method. This method still uses an SCF but then adds contributions from functionals of the computed electron density. This method is said to include implicitly some electron correlation. It is generally slightly more accurate than MP2 for most properties, while remaining only slightly more costly than the basic SCF.

Thus, for this project, it was decided to perform all calculations at the B3LYP/6-31+G(d,p) level.

Computational Models

The gluco- and galactopyranose systems under consideration have a tremendous number of available conformational minima when all possible rotations of the hydroxyl groups are considered. To facilitate the calculations tetrahydropyran model structures **1** through **6** (Figure 3) were chosen for this study. In these models, hydrogens replace the C-1, C-2, and C-3 hydroxyl groups of the monosaccharides, while the C-4 hydroxyl is preserved or replaced by either fluorine or a methoxy group. The steric environment at C-6 in these simpler structures will not differ significantly from that in the actual monosaccharide derivatives and the electronic effects most relevant to the S_N2 displacement at C-6 are maintained.
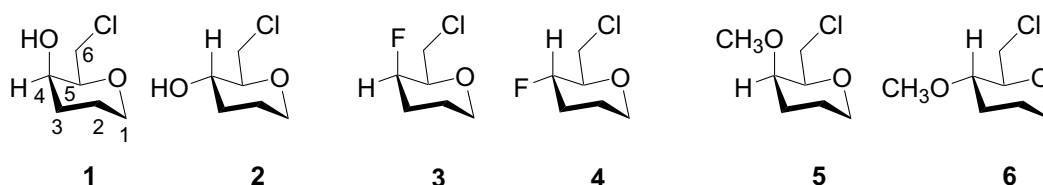
Figure 3: Tetrahydropyran Model Structures. Carbon Atoms are designated using Carbohydrate Numbering.

While most actual examples of these reactions have employed an aryl- or alkylsulfonate leaving group, there are several advantages to using a halide identity reaction. Firstly, there are many fewer degrees of freedom within the leaving group itself. Secondly, the local dipole effect is both simplified and accentuated in these structures. The identity reaction also avoids further complications involving differing electronegativities and steric factors. Chloride

ion's electronegativity (3.0 Pauling, 3.02 Boyd[22]) is reasonably close to that of a methanesulfonate ester group (3.56, calculated using Boyd's method).

In order to study the activation barrier(s) to the displacement reactions, the geometries of the model substrates first had to be determined. The O-5–C-5–C-6–Cl torsion (dihedral ω) defines three rotamers, conventionally identified as *gt*, *tg* and *gg* (Figure 4).
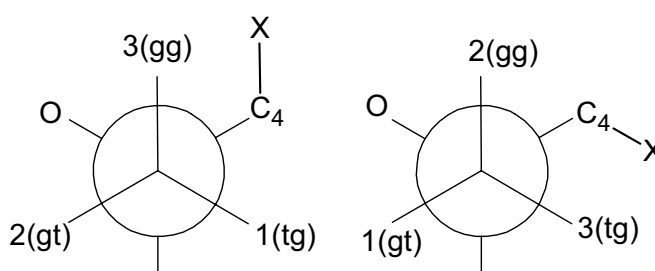
Figure 4: Newman projection showing possible positions for chlorine atom, ranked as (1) preferred, (2) next and (3) least preferred, for *galacto* (left) and *gluco* (right). (X=OH, F, OMe)

Searches for relevant reactant and transition structure geometries were undertaken. Initially, the C-5–C-6 rotamers of the substrates **1** through **6**, including possible intramolecularly hydrogen-bonded species, were identified in a series of geometry optimizations. For **1**, **2**, **5** and **6,** the possibility of permutations of hydroxyl or methoxy rotamers at C-4 was also explored. Transition structures connecting the rotational minima were also obtained. Beginning from these rotational minima, possible reaction coordinates for S_N2 displacements were located. From the S_N2 transition structures, the Intrinsic Reaction Coordinates (IRC's)[23] were obtained, and the free energy was maximized along these paths.

Corrections for basis set superposition errors (BSSE) were calculated using the full counterpoise (see Comp-Meth B) method of Boys and Bernardi.[7]

Since these displacement reactions are typically carried out in DMF solvent at elevated temperature, the effects of solvation and temperature on the calculated energy profiles were estimated. Two solvation models were used (see Appendix B). All the structures and frequencies were recalculated with the Onsager dipole continuum model and a dielectric constant of 36.71 (chosen for DMF at 298 K). The geometries so obtained were then used to calculate solvated energies using the Isodensity Polarized Continuum (IPCM) Model.[24] The vibrations calculated under the Onsager solvation model were retained to provide free energy corrections to the IPCM energies.

Boltzmann statistics were applied to the relative free energies to obtain rotameric populations. The thermal contributions to Gibbs free energies were calculated at each of three temperatures. Energies were evaluated at 298, 373 and 413 K. Room temperature was included to allow comparison with other published results, while 373 and 413 K were chosen to represent realistic reaction conditions. Relative rates of S_N2 displacement (*galacto/gluco)* were calculated for each of the three pairs of systems. Reaction rates were calculated using the free energy maxima along the reaction paths. The reaction path curvature was calculated at the free energy maximum along the IRC (see Appendix D).

Given that the fluoro compounds were anticipated to exhibit the largest dipole-dipole interactions of any of the model compounds, they were selected for further analysis. To probe the potential role of dipole-dipole interactions in the

transition structures, topological analyses of the charge densities were performed according to the theory of Atoms in Molecules (see Comp-Meth C).[8] The full set of atomic properties for each atom was obtained from the previously calculated wavefunctions, using the AIM2000 program.[25]

To estimate the maximum possible classical dipole-dipole repulsion energy in these structures, various hypothetical dipoles of experimental and calculated magnitudes were considered. Dipoles derived from AIM analysis for the calculated transition structures cannot be used due to their origin dependence in charged systems, so model dipoles were used instead. The experimental dipole from hydrogen fluoride and a calculated dipole (B3LYP/6-31+G(d,p)) for hydrogen chloride stretched to match the C–Cl distance found in **1TS** were separated by the distances found in the S_N2 transition structure **1TS**. The dipoles were aligned in parallel and centered at the distance between bond critical points of the C4—F and C6—Cl bonds. Thus, their maximum interaction energy was evaluated.

The possible role of dipole-dipole interactions in determining the reactant rotamer populations was also determined. We calculated the classical electrostatic interactions between idealized local dipoles aligned in parallel and separated by the distance between the C–F and C–Cl bond critical points in **1gg**. The magnitudes of the C4–F and C6–Cl dipoles in **1gg** obtained from AIM analysis (see below) as well as the experimental hydrogen fluoride and hydrogen chloride dipoles were considered.

Results and Discussion

Gas phase calculations located three rotational minima for each compound, corresponding to the expected staggered conformations. The calculated dihedral angles and energies for these rotamers are summarized in Tables 1 and 2. In all of the *galacto* models, the most stable rotamer was ***tg***, while the highest-energy rotamer was ***gg***. In the *gluco* structures, the preferred conformation was always ***gt***, while the least stable rotamer was ***tg***.

A search of structures **1** and **2** for low-energy hydroxyl rotamers at C-4 produced an additional minimum (**1*gg-h***) in which the OH group was rotated to form a hydrogen bond with the Cl. The hydrogen-bonded conformation **1*gg-h*** was considerably more stable than **1*gg***, but was still much higher in free energy than the **1*tg*** or **1*gt*** rotamers. A similar search of **5** and **6** for methoxy rotamers at C-4 produced no additional low-energy structures. For each C-6 rotamer, the methoxy group always preferred a position trans to the C-4—C-5 bond. Proper convergence to minima for the methoxy structures was problematic. The rotational potentials for the methoxy group at C-4 are quite shallow near the global minimum, in some cases necessitating recalculation of force constants at every point.

Rotational transition structures were located between the rotamers in **1** through **4**. These are included in Tables 1 and 2. Due to the convergence problems with **5** and **6**, and their otherwise typical rotamer energies, they were excluded from this analysis. In all cases, the rotational barriers were sufficiently

low that the model systems would be in thermal equilibrium at the temperatures under consideration.

Calculations incorporating solvation by DMF produced significant changes in the relative rotamer free energies for **1** through **6**.

In the axial *galacto* cases, ***gt*** was stabilized considerably at the expense of ***tg***, while ***gg*** remained a minor constituent. For the fluoro and methoxy systems, ***gt*** actually became the preferred rotamer. The hydrogen-bonded geometry **1*gg-h*** was still a minimum, but **1*gg*** was no longer a stationary point. This was a consequence of the computational method, since IPCM solvation energies were calculated for geometries obtained using the Onsager spherical model. With IPCM solvation, the Onsager transition structure connecting ***1gg*** with ***1gg-h*** fell below that of ***1gg***.

In the equatorial *gluco* cases, solvation affected the ***tg*** rotamers differently in each system. The hydroxy model ***2tg*** was strongly stabilized, while the fluoro model ***4tg*** was only slightly stabilized. The already disfavoured methoxy compound ***6tg*** (<1% population in the gas phase) was no longer a minimum on the rotational surface. Of all the ***tg*** rotamers, ***6tg*** is notable for having the smallest dihedral angle ω, nearly eclipsing the hydrogen at C-5. With Onsager solvation, optimizations beginning at ***6tg*** converged to the ***6gt*** conformation. The ***gg*** rotamers were affected to a lesser extent by solvation, and remained well populated (>19%) in all cases.

Table 1: Calculated Dihedral Angles ω (degrees), Free Energies ΔG°_{298} (kcal mol^{-1}), and Percentage Populations for Exocyclic Chloromethyl Rotational Minima and Rotational Transition Structures of *Galacto* Structures 1, 3, and 5.

Structure	B3LYP/6-31+G(d,p)			B3LYP IPCM (DMF)		
	ω	ΔG°_{298}	%Population	ω	ΔG°_{298}	%Population
1*gt*	71.4	0.869	18.00	69.9	0.000	97.05
TS$_{1gt \to 1tg}$	116.8	3.414		120.8	5.223	
1*tg*	169.6	0.000	78.03	167.9	2.192	2.40
TS$_{1tg \to 1gg\text{-}h}$	243.7	5.108		241.6	6.724	
1*gg-h*	293.9	1.765	3.97	294.0	3.064	0.55
TS$_{1gg\text{-}h \to 1gg}$	300.8	6.148		NA	NA	
1*gg*	305.2	5.851	0.004	NA	NA	NA
TS$_{1gg \to 1gt}$	354.4	9.150		353.4	9.352	
3gt	70.9	0.917	17.53	69.1	0.408	33.26
TS$_{3gt \to 3tg}$	116.1	3.349		117.5	3.971	
3tg	169.3	0.000	82.45	170.8	0.000	66.26
TS$_{3tg \to 3gg}$	248.8	8.315		246.5	6.473	
3*gg*	301.2	4.900	0.02	301.1	2.921	0.48
TS$_{3gg \to 3gt}$	355.5	8.901		356.4	7.803	
5gt	72.0	0.949	16.76	69.9	0.000	64.74
5tg	171.0	0.000	83.23	170.2	0.361	35.21
5gg	303.5	5.471	0.008	305.3	4.225	0.05

Table 2: Calculated Dihedral Angles ω (degrees), Free Energies ΔG°_{298} (kcal mol^{-1}), and Percentage Populations for Exocyclic Chloromethyl Rotational Minima and Rotational Transition Structures of *Gluco* Structures 2, 4, and 6.

Structure	B3LYP/6-31+G(d,p)			B3LYP IPCM (DMF)		
	ω	ΔG°_{298}	%Population	ω	ΔG°_{298}	%Population
2gt	72.1	0.000	77.09	71.4	0.000	48.36
TS$_{2gt \rightarrow 2tg}$	125.8	2.948		129.5	2.028	
2tg	147.6	1.568	5.46	147.2	0.239	32.30
TS$_{2tg \rightarrow 2gg}$	223.1	7.683		223.0	7.792	
2gg	295.7	0.880	17.44	295.9	0.543	19.35
TS$_{2gg \rightarrow 2gt}$	359.8	7.437		360.2	6.160	
4gt	71.0	0.000	60.66	72.0	0.000	52.39
TS$_{4gt \rightarrow 4tg}$	128.2	3.420		127.1	2.860	
4tg	157.6	2.085	1.80	158.0	1.281	6.03
TS$_{4tg \rightarrow 4gg}$	225.9	6.351		227.0	6.078	
4gg	295.0	0.284	37.54	294.4	0.137	41.58
TS$_{4gg \rightarrow 4gt}$	360.6	7.180		360.2	6.379	
6gt	70.2	0.000	51.69	70.0	0.000	68.71
6tg	142.7	2.391	0.91	NA	NA	NA
6gg	293.1	0.051	47.39	294.3	0.466	31.29

S_N2 Transition Structures

Transition structures for the S_N2 displacements were sought beginning from each of the rotational minima of the new systems. In each model system, only one S_N2 transition structure was found (see Figures 5 and 6). For the *galacto* compounds, these transition structures connect the **gg** and **tg** geometries. For the *gluco* compounds, the transition structures connect the **gg** and **gt** rotamers. Since these are identity reactions, they could proceed in either direction. However, as **gg** is reactive in both systems, we will refer to the chlorine below the ring system as the nucleophile for discussion purposes. To facilitate structural comparison with the reactant rotamers, we identify the O5–C5–C6–Cl torsion angles as ω_{LE} and ω_{NU}, for the leaving and nucleophilic chlorines, respectively.
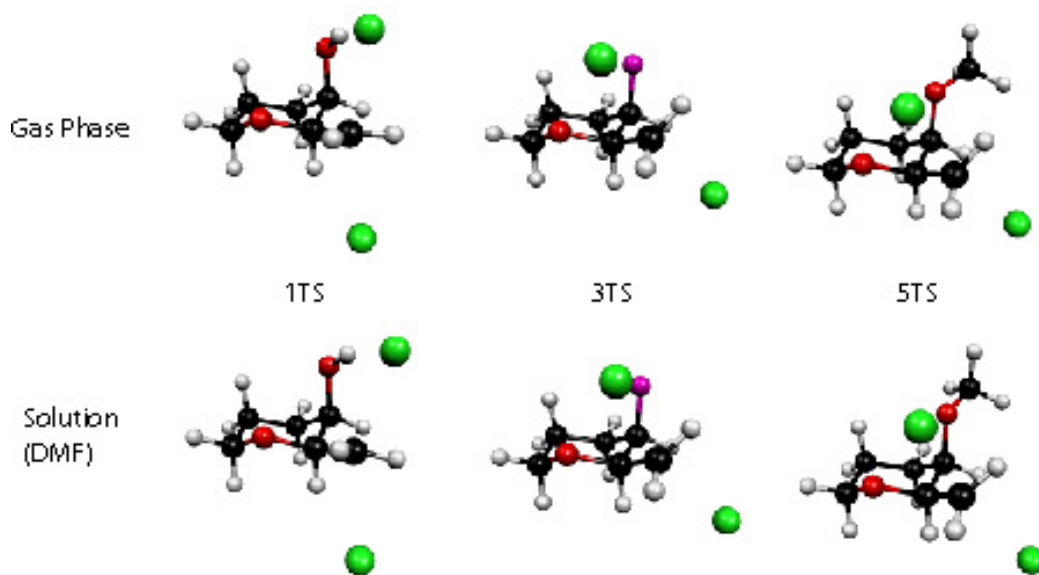
Figure 5: Transition structures for identity S_N2 displacements of *galacto* compounds 1, 3 and 5 in the gas phase and with solvation.

According to the dipole-dipole repulsion model, the Cl-C6-Cl group would be expected to be oriented perpendicular to the ring system, in order to

avoid repulsive interactions with the ring oxygen, although this would create a repulsive interaction with the *galacto* substituent at C4, destabilizing the **TS**. At first glance, it may appear that *galacto* compounds **3TS** and **5TS** are twisted to avoid a repulsive interaction at C4, while compound **1TS** is only prevented from twisting by a hydrogen bond. Interestingly however, with solvation and the consequent decrease in the importance of the hydrogen bond in **1TS**, ω_{LE} actually decreases, with Cl relaxing further from O5 just as in the other cases. The relaxation of the H-bond (gas phase: r_{H-Cl} = 2.066 Å, θ_{O-H-Cl} = 5.5°; solvated: r_{H-Cl} = 2.128 Å, θ_{O-H-Cl} = 18.0°) occurs through a torsional motion of the hydroxyl at C4. The S_N2 transition structures for reactions of the *galacto* structures **3** (fluoro) and **5** (methoxy) do have torsion angles ω_{LE} for the leaving Cl⁻ inclined considerably towards O5, and solvation also produced a slight relaxation away from O5.
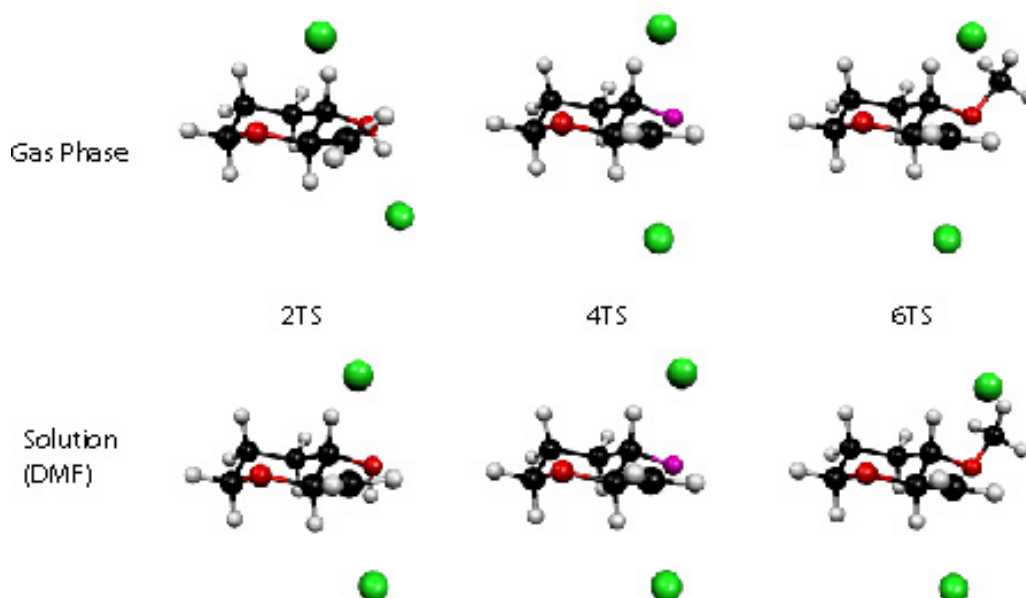
Figure 6: Transition structures for identity S$_N$2 displacements of *gluco* compounds 2, 4 and 6 in the gas phase and with solvation.

The Cl$_{NU}$—C6—Cl$_{LE}$ deflection angles are approximately 150 degrees in all systems; there is no exaggeration of this angle in the *galacto* compounds. Note that these angles are typical of calculated transition structures for S$_N$2 identity reactions involving chloride.[19] The dipole-dipole interaction energy is insufficient to explain the energy difference between the *gluco* and *galacto* systems, even for perfectly aligned dipoles. Because of the ~150 degree deflection angle, this interaction would be further reduced, and insufficient to drive the torsion towards O5. Therefore the torsion angles in **3TS** and **5TS** are not attributed to repulsive dipolar interactions.

This conformational preference can be correlated with the energies and rotational potentials of the reactant rotamers (see Table 1). The orientation of the

Cl_{NU}—$C6$—Cl_{LE} group shows a strong dependence on the reactant C5-C6 rotational potentials in both the *galacto* and *gluco* systems. One chloride (denoted here as Cl_{LE}) must always occupy the *gg* position, and is not a determining factor in either the *galacto* or *gluco* systems. In **1**, the preferred rotamer is gt by a significant margin, and the 1TS torsion angle reflects this. In **3TS** and **5TS**, ω_{NU} is within 15° of that found in the preferred *tg* rotamers. In the *gluco* cases, the *tg* rotamer becomes increasingly disfavoured through the systems **2**, **4** and **6** (see Table 2). Correspondingly, ω_{NU} decreases through the transition structures **2TS**, **4TS** and **6TS** (seen in Figure 6) as the Cl_{NU}—$C6$—Cl_{LE} group twists increasingly into the favored *gt* reactant rotamer orientations.

Atoms In Molecules Analysis

AIM analysis of the charge densities in **3** and **4** affords two different insights into the conformational dependence of the total system energies (Details of this theory are presented in Computational Methods C). Firstly, by partitioning the molecule into its constituent atoms, the energies of each atom can be compared through the various rotamers, revealing the changing relative contributions to the molecular energies. In this way, one can ascertain whether the chlorine and fluorine atoms in the *galacto* **gg** rotamer are perturbing each other relative to the lower energy rotamers. Secondly, one can directly observe distortions of the charge densities. If the dipole model were correct, then rotation at C-6 from a low-energy rotamer into the disfavoured position would induce opposing dipoles into both molecular fragments (Figure 7).

Since only subtle differences in atomic properties were anticipated throughout the various structures, the integration accuracy criteria were stringent. The atomic volume-integrated Laplacian of the charge density defined as:

$$L(\Omega) = -\frac{1}{4} \int_{\Omega} \nabla^2 \rho(\mathbf{r}) d\tau \ , \qquad \text{where } \Omega \text{ is an atomic basin} \qquad (1)$$

vanishes for an exact integration, and can be considered as an error function for numerical integrations.[27] Integrations were performed in natural coordinates with the beta-sphere diameter set at the distance from the nucleus to the nearest critical point. Absolute and relative integration accuracies were set as low as 1×10^{-6}, and the integration path as large as 2.5×10^{6}, as necessary, to achieve a value for $L(\Omega)$ of less than 3×10^{-4} for heavy atoms, and 1×10^{-4} for hydrogens. These criteria[27]

allow recovery of molecular SCF energies to within 0.2 kcal mol^{-1}, and dipoles to within 5×10^{-4} au. Use of these criteria was essential since the use of default criteria results in sufficient accumulated error to misidentify the lowest energy rotamer.
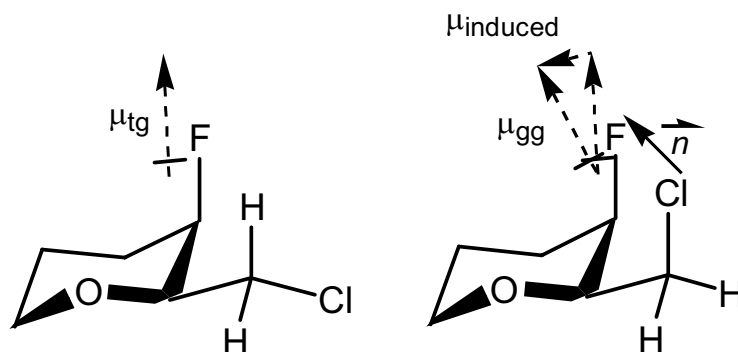
Figure 7: Possible Induced Dipoles arising from Conformational Change (see Table 4 for results).

Table 3: Selected Atomic AIM Properties Relative to Low Energy Rotamers: Relative Energies (kcal mol^{-1}) and Relative Charges (electrons).

		3gt	*3tg*	*3gg***	*4gt*	*4tg***	*4gg*
F	Energy	-2.04	0	-7.24	0	-1.46	2.22
	Charge	-0.001	0	0.001	0	0.003	0.001
C4	Energy	2.45	0	11.83	0	5.94	-3.20
	Charge	0.002	0	0.024	0	0.017	0.000
C5	Energy	3.25	0	8.77	0	-0.71	1.53
	Charge	0.008	0	0.012	0	-0.015	-0.001
C6	Energy	2.93	0	10.65	0	5.11	2.31
	Charge	0.004	0	0.008	0	0.006	0.002
Cl	Energy	3.88	0	6.54	0	-1.75	-2.41
	Charge	0.012	0	0.040	0	0.009	-0.002
O	Energy	-5.04	0	-5.62	0	5.13	-0.85
	Charge	0.001	0	0.000	0	0.003	0.000

** High-energy rotamer

As seen in Table 3, the net charges on atoms only change by at most a few hundredths of an electron.   However, the *energies* of the various atoms undergo relatively large changes. Considering that the ranges of rotamer energies are at most about 5 kcal mol^{-1}, it might be surprising that many individual atomic energies vary by more than twice that much through the rotamers.

The atomic energy changes are not what would be expected from direct electrostatic interactions. For example, in compound **3,** going from the preferred *tg* to the high energy *gg* rotamer, we find the chlorine raised in energy but the fluorine lowered. In compound **4**, both the fluorine and chlorine atoms in the high-energy *tg* rotamer are *lower* in energy than in the lowest energy rotamer. Furthermore, the total system energy changes cannot be recovered from the changes in a few "key" atoms. Energy is redistributed throughout the molecule, including hydrogen atoms and more remote heavy atoms. Some interesting patterns in the energy fluctuations are observed, pointing to complex long-range interactions. For example, the ring oxygen is destabilized by ~5 kcal in both **3** and **4,** when the chlorine atom is *trans* to it. These interactions present an interesting direction for future investigation.

Table 4: Projection of Induced Dipoles[a] (a.u.) onto Normalized Cl→F Vectors $\vec{n}$ (see Figure 7)

Atom	3*gt*	3*tg*	3*gg*	3TS	4*gt*	4*tg*	4*gg*	4TS
F	-0.003	0.000	0.025	0.041	0.000	0.031	0.004	0.024
C-4	-0.019	0.000	0.069	0.088	0.000	0.106	0.030	-0.007

[a] Induced dipoles relative to low energy rotamer.

Finally, we examined the induced dipoles for atoms C4 and F, in the rotamers of **3** and **4** and in the corresponding transition structures. The atomic dipoles of F and C4 are compared through the various structures. As the chloromethyl group is rotated from the lowest energy rotamer into the disfavored rotamer, changes in the local dipoles are observed. Any atomic dipole at F or C4 induced by this rotation is projected onto a vector in the direction of the supposedly perturbing chloride atom. The projections of the induced dipoles onto the normalized Cl→F vectors $\vec{n}$ (shown in Figure 7, and tabulated in Table 4) contain both atomic dipole and charge-transfer contributions (see Comp-Meth C). The latter term is very small, because there is very little transfer between C4 and F, and because this contribution is directed precisely along the C4—F bond, which is nearly orthogonal to the Cl→ F vector. Thus, the dominant contributions to the projections of the induced dipoles were the induced atomic dipoles, which were very small in all cases. Analysis of the induced dipoles and the atomic energies shows that while small dipoles are induced by group alignments, these are not correlated with the energy fluctuations and cannot be primary contributors to the system energies.

Dipole-Dipole Interaction Energy

Richardson's model postulated that the rate differences were largely controlled by the repulsion between parallel dipoles in the transition structures. At 413 K (boiling point of a typical solvent, dimethylformamide), relative rates (*galacto*/*gluco*) of ca. 0.03–0.05 would indicate that the faster reaction must be favored by approximately 2.88–2.46 kcal mol^{-1}. In many cases in the synthetic literature, the relative *galacto* rates must have been even smaller than this, since the *galacto* product could not be isolated in any appreciable yield.[15] In our model systems, the difference in relative electronic energies was over 6.5 kcal mol^{-1}. The distance between the C4-F and the C6-Cl bond critical points in **3TS** is 3.08 Å. At this distance, it would require two dipoles of 3.65 debye in perfect alignment to account for this energy difference. With our calculated dipole for stretched HCl and the experimental dipole of HF, the interaction energy is only 2.28 kcal mol^{-1}. The magnitudes of these model dipoles are greater than those that would occur in the carbohydrate systems. In **3TS** where the local dipoles are not aligned, and in the presence of solvent, any interaction energies due to dipoles would be greatly reduced. It is therefore highly unlikely that this type of interaction could explain a substantial portion of the S$_N$2 reactivity differences.

In both the *gluco* and *galacto* systems, the disfavoured rotamers are those in which the C6–Cl bond is aligned with the C4–F bond. In the galacto case, this is the ***gg*** rotamer while in the gluco case it is the ***tg*** rotamer. It seemed appropriate to determine what the maximum possible repulsion energy might be in these systems. The largest dipole-dipole repulsion found was that using the

experimental HF/HCl dipoles arranged according to the distances found in systems **3** and **4**. These are larger than our carbohydrate local dipoles derived from AIM analysis. Even with these exaggerated dipoles, the maximum repulsion energy was only 1.33 kcal mol^{-1}. This is far less than the range of rotamer energies calculated for our model systems **3** and **4**. Clearly, dipole-dipole repulsion is insufficient to account for the energetic differences among the reactant rotamers.

Factors Controlling Relative Rates

The reaction kinetics were modelled at two realistic temperatures (373 K and 413 K), with consideration of rotamer populations, reaction barriers, free energy barriers and reaction path curvature. The rates of reactions involving rapid pre-equilibria are generally interpreted in terms of Curtin-Hammett/Winstein-Holness kinetics. The rates of conversion between the C5–C6 rotamers of **1** through **6** corresponding to our calculated barriers are much greater than the rates of S_N2 displacement. In such a situation, the Curtin-Hammett Principle has sometimes been interpreted to mean that the outcome of the overall process is wholly independent of the equilibrium populations, depending *only* on the relative rate constants for the subsequent reaction. However, this is erroneous, as Seeman pointed out in his comprehensive 1983 review:[28] "*the ground state conformational preference has a direct (proportional) role*" in the final product ratios. The relative rates of rotamer interconversion and S_N2 reaction derived from our calculations correspond to the "Scheme II" kinetics of the Curtin-Hammett/Winstein-Holness analysis. The net rate constant for a reaction given

these conditions can correctly be expressed as a mole fraction-weighted sum of individual rate constants.[29]

As these are identity reactions, microscopic reversibility dictates that we consider nucleophilic approach to both of the rotamers connected through the IRC. Thus, the total reaction rate is the sum of two equal contributions. When comparing the total reaction rates between systems, the population of the third rotamer, not directly participating in the reaction, becomes important. Since the rotational barriers are much lower than the S_N2 barriers, the populations of reactive rotamers are constantly maintained. However, the total reactive population at any time is only the sum of the two reactive rotamers for each system.

The relative rate of S_N2 displacement (*galacto/gluco*) can be written as the ratio of two sums, the mole-fraction weighted rates for each system. This formulation is equivalent to more conventional expressions in terms of ensemble average free energies.[28] We prefer equation (2) because it provides insight into the contributions of both the reactant states and the transition state to the overall barrier to S_N2 reaction.

$$k_{Rel} \, galacto/gluco = \frac{p_{Galgg}(T) \bullet e^{-\frac{\Delta G^{\ddagger}_{Galgg}}{RT}} + p_{Galtg}(T) \bullet e^{-\frac{\Delta G^{\ddagger}_{Galtg}}{RT}}}{p_{Glugg}(T) \bullet e^{-\frac{\Delta G^{\ddagger}_{Glugg}}{RT}} + p_{Glugt}(T) \bullet e^{-\frac{\Delta G^{\ddagger}_{Glugt}}{RT}}} \qquad \textbf{(2)}$$

Table 5: Mole-Fractions of Reactive Rotamers, Activation Free Energies $\Delta G^{\ddagger}$ (kcal mol^{-1}), Curvature and Reaction Rates in S_N2 Identity Displacement Reactions of 1 through 6 with Cl^{-} Ion.

Structure	T (K)	Mole-fraction of rotamer	$\Delta G^{\ddagger}$ TS [a]	Sum of reactive mole-fractions	Curvature[b] of IRC (a.u.)	Net relative S_N2 rate *galacto/gluco*
1gg		0.0133	27.80	0.0631	1.08	
1tg	373.15	0.0498	28.78			
						0.020
2gg		0.1987	26.92	0.6583	4.57	
2gt		0.4596	27.54			
3gg		0.0119	29.76	0.6387	513.94	
3tg	373.15	0.6268	32.67			
						0.062
4gg		0.3977	30.27	0.9044	1.26	
4gt		0.5067	30.45			
5gg		0.0020	26.94	0.3915	505.75	
5tg	373.15	0.3895	30.27			
						0.061
6gg		0.3319	28.64	1.0000	2.98	
6gt		0.6681	29.16			
1gg		0.0185	28.76	0.0838	1.08	
1tg	413.15	0.0653	29.80			
						0.031
2gg		0.2005	27.88	0.6512	4.57	
2gt		0.4507	28.54			
3gg		0.0169	30.62	.6283	513.94	
3tg	413.15	0.6114	33.57			
						0.084
4gg		0.3889	31.17	.8864	1.26	
4gt		0.4975	31.37			
5gg		0.0034	27.86	0.4073	505.75	
5tg	413.15	0.4039	31.17			
						0.083
6gg		0.3394	29.59	1.0000	2.98	
6gt		0.6606	30.14			

[a] Includes BSSE corrections [b] effect of curvature not included in rate constant

In Table 5 the relative rates of S_N2 displacements for each of the three model systems are compared at two temperatures. The barriers are expressed in kcal mol^{-1} from the reactive rotamers. As seen in Table 3, for hydroxy structures **1** and **2**, the barriers to reaction from the reactive rotamers are slightly higher for the *galacto* system. However, while the barriers do contribute to the low net relative rate in this case, the larger effect arises from the rotamer populations.

The *gluco* system has a roughly ten-fold larger population of reactive rotamers. Considering systems **3** through **6** (fluoro and methoxy), we see that the barriers to reaction from the *gg* rotamers are actually lower for the *galacto* cases. However, this rotamer is not well-populated, and the barriers from the *tg* rotamers are considerably higher. The total reactive populations are somewhat lower for the *galacto* cases, and the wider spread of rotamer energies means that the more populated rotamers have a larger barrier to overcome.

As expected for S_N2 displacement reactions, the maxima in free energy along the IRC were essentially at the transition structures. Thus, variationally optimizing the transition state produced no change in the calculated rates. The furthest that any variational transition state was found from the transition structure was 0.0045 amu$^{1/2}$·bohr for system **5**. However, it was determined that the reaction paths for *galacto* compounds **3** and **5** include massive curvature near the transition state.[30] In systems of lower dimensionality, a series of trajectory calculations using the reaction path Hamiltonian might be performed. In this case, we could apply a simple approximation as described by Miller.[31] A multiplicative constant κ is calculated from the reaction path curvature. This constant tends to 1 for low curvature, and to 1/2 for high curvature. This would account for a nearly two-fold decrease in rate for the *galacto* systems **3** and **5**. It has been suggested that large reaction path curvature may reduce rates by much more than this.[32] In addition, since the reactions are performed in solvent, a more correct treatment would include coupling of vibrational modes to the solvent. Due to the excessive dimensionality of these systems, it would not be practical or

accurate to implement these methods. Which structural features may cause such high curvature to arise in these systems is an interesting question for future research.

Conclusions

At the outset of the investigation, it was anticipated that if the low reactivity of *galacto-* systems could be attributed to a high-energy transition structure, then electrostatic contributors to this energy could be identified through the partitioning of molecular properties within the AIM formalism. The complete set of AIM-derived properties revealed no direct correlation between changes in the energies of arbitrarily defined structural fragments and their relative configurations. A thorough analysis of the transition structures revealed no significant perturbations. AIM analysis of the reactant rotamers likewise did not give any indication that the population differences should be attributed to dipole-dipole interactions.

The AIM analysis showed that small changes in atomic properties occurred throughout the molecular structures in the various rotamers and in the S_N2 transition structures. The properties of the total system could not readily be related to changes in the properties of individual atoms or simple structural fragments. The subtle long-range interactions evidenced by our AIM results point towards the possibility of a rich scheme of influences even from somewhat remote sites.

The maximum possible energy that can be attributed to dipole-dipole interactions in these systems has been estimated. At the distances found in these structures, they cannot be the dominant cause of either the difference in reaction barriers, or the difference in rotamer populations. Such interactions are clearly

only a minor contributor to the reactivity differences between *gluco* and *galacto* systems of this type.

The calculated rotamer populations and the barriers to their interconversion are in general accord with several recent experimental and theoretical studies of C5–C6 rotamer populations in D-glucopyranose and D-galactopyranose.[33] The order in galactose derivatives typically is $P_{tg} \approx P_{gt} > P_{gg}$, the energies are spread over quite a wide range, and the population of the *gg* rotamer depends strongly on substituent and solvent effects. In contrast, glucose derivatives have generally been found to display a relatively narrow range of rotamer energies, with relative populations $P_{gg} > P_{gt} > P_{tg}$ in most cases. Significantly, in *gluco*-configured compounds, the *gg* rotamer remains well populated as substituent groups on other positions are changed, and as solvents are varied.

These calculations reproduced the spread and qualitative orderings of rotamer energies for both the *galacto-* and *gluco-* cases, despite the absence of the hydroxyl groups at C1, C2 and C3. In fact, the populations for **4** are remarkably similar (± 3%) to those recently calculated by Hoffmann and Rychlewski for 4-deoxy-4-fluoro-D-glucopyranose, once their further subdivided results are summed into the three C5–C6 rotamer categories.[34] It is therefore clear that models **1** through **6** were appropriate choices for the question under discussion.

In the gas phase, the ranges of relative energies calculated for the three *galacto-* minima were about 5 kcal mol^{-1}. The calculated populations of the *gg* rotamers in the *galacto-* models were almost negligible in the gas phase. Even

though solvation stabilized the **gg** rotamers slightly, and taking into account the effect of higher temperatures typical of practical reaction conditions (373 and 413K), calculated populations of the **gg** rotamers were never greater than 2%. This contributes to the low total reactive conformer populations. In contrast, the *gluco-* model rotamers were all within about 2 kcal mol^{-1} of one another. The reactive conformer populations ranged from about 65% to 100% in the three model systems at 413K. As seen in Table 3, our results do produce low relative *galacto/gluco* rates in agreement with experiments. However, it is clear that the low rates are not simply due to a perturbed transition structure in the *galacto* case. Obviously, relative barriers will vary to some extent depending on the specific systems under study. These results illustrate the importance of considering the energetics of the reactant, and not focusing primarily on stabilizing/destabilizing features of the transition structure when dealing with conformationally mobile reactants. These results show that dipole-dipole interactions in the transition structures do not determine the relative reactivities in the model S_N2 reactions that were studied. Based on the calculated rotamer energies, the reactive rotamers will be present at a significantly higher equilibrium concentration in the *gluco-* cases. The differences in S_N2 reaction rates must be attributed to a combination of factors including reactant rotamer populations, solvation effects, relative barriers, and reaction path curvature.

It may seem unfortunate that such an intuitive way of regarding structural reactivity relationships must be discarded. However, the transition structure dipole model was essentially a yes/no argument. It implied that chemists not even

attempt reactions with certain substrates, which may actually be feasible. Consideration of the interplay of contributing factors highlighted here may suggest changes to substituents and leaving groups or modifications to reaction solvents, which in combination could optimize the reactive rotamer population and the reaction barrier. Clearly, this could be accomplished in many ways, including *destabilization* of the normally dominant conformers. Experimental determination of rotamer populations using recently improved NMR methods will continue to provide insight into these factors.[33a-c, 35] The possible influence of reaction path curvature is large, and it provides great insight into reactions that are much slower than their barriers would indicate. At present it seems difficult to predict *a priori* how different substituents will influence the curvature.

The dipole model was also proposed to explain stereochemically-dependent reactivity differences in displacement reactions at secondary centres. It is in this connection that it has usually been cited outside carbohydrate chemistry. In these situations, there is much less conformational mobility, and dipolar arguments may well be more valid.[36] Other factors such as differential solvation of transition structures, entropic effects, or electronic factors such as hyperconjugation might also influence reactivity in such cases. However, given the R^{-3} distance dependence of dipole-dipole interactions, they could play a much larger role in vicinal relationships.

S_N2 Reactivity Appendix A: Thermo-Chemistry

This section deals with the calculation of free-energies and rates of reaction within the harmonic approximation, using Conventional Transition State Theory.[37] Free-energies may be calculated following the vibrational analysis of a fixed point. Note that at a fixed point all the gradients vanish and the 3N-6 (3N-5 for linear systems) normal modes are orthogonal coordinates. As will be discussed in detail in Appendix C, vibrational analysis may also be performed at a point on the IRC path.[23] For this, the IRC coordinate must be projected out, leaving 3N-7 orthogonal coordinates, each of which may not be orthogonal to the IRC, and may couple to it.

Assuming that one is treating a non-linear molecule at a fixed point, then within the harmonic approximation, there are 3N-6 orthogonal modes with evenly spaced vibrational levels. This has two important consequences when considering the vibrational thermodynamic partition function. Firstly, due to orthogonality, inter-mode coupling is forbidden. Secondly, the even spacing of the levels allows for a tremendous simplification (see equations below) when summing over the energy levels. Below are the complete set of thermodynamic partition functions commonly used to compute free-energies of molecular systems.

Electronic Partition Function

The electronic partition function can be written as:[38]

$$Q_{elec} = \omega_0 e^{-\varepsilon_0 \beta} + \omega_1 e^{-\varepsilon_1 \beta} + ...$$

ω_n = degeneracy of energy level

ε_n = energy of nth level

$\beta = 1/kT$

(A1)

It can be seen that energy is partitioned into the electronic states according to their spacing and degeneracy. In practice, using commercial code such as Gaussian, the electronic levels are assumed to be widely spaced. Consequently, for a non-degenerate ground-state, by setting the zero of energy at ε_0, the entire electronic partition function can be set to 1.

Rotational Partition Function

Linear Molecule

$$Q_{Rot} = \frac{2I}{\hbar^2 \beta} \; ; \text{ where } I = \text{moment of inertia}$$

Non - Linear Molecule   (A2-A3)

$$Q_{Rot} = \left[\left(\frac{2}{\hbar^2 \beta} \right)^3 \pi I_A I_B I_C \right]^{1/2}$$

Where $I_A I_B I_C = \mu^3 \det[I]$

This is the most common way to express the rotational partition function.[38] This classical expression is based on the moments of inertia rather than an attempt

to sum over discrete quantized levels. This is an especially good approximation for polyatomic molecules with heavy atoms. Even for a linear triatomic, with light atoms, the error introduced by this classical approach has been shown to be less than 1%.

Vibrational Partition Function

Harmonic Approximation

$$\varepsilon_{\text{vib}}(n_1, n_2, \ldots n_{3N-6}) = \sum_{m=1}^{3N-6} \left(n_m + \frac{1}{2} \right) \hbar \omega_m$$

where n is # of quanta in
each of the m modes, for m = 1, 2, ... 3N - 6   (A4-A6)
Moderate temperature : Harmonic levels

$$Q_{\text{vib,m}} = \frac{\exp(-\hbar \omega_m \beta / 2)}{1 - \exp(-\hbar \omega_m \beta)}$$

$$Q_{vib} = \prod_{m=1}^{3N-6} Q_{\text{vib,m}}$$

Using the harmonic approximation, energy is partitioned according to the above formulae.[38] This expression exploits the even level structure for simplification. At moderate temperature only the lowest vibrational levels will be populated, making the true level spacing inconsequential. At high temperature this assumption will cause the predicted behavior to deviate significantly from that observed.

Following a vibrational analysis, since the nuclear masses and coordinates determine the moments of inertia, using the above partition functions one may

calculate the free-energy. This may be implemented into the following equations of conventional transition state theory in order to compute rates of reaction.

Conventional Transition State Theory

$$k^{TST}(T) = \frac{\sigma}{\beta h} \frac{Q^{\ddagger}(T)}{Q^{R}(T)} \exp(-\beta V^{\ddagger})$$

$$\sigma = \# \, of \, equivalent \, paths$$

$$\beta = (k_B T)^{-1}$$

$$Q^R(T) = Q_{Rel}(T) Q^A(T) Q^B(T)$$

$$Q_{Rel} = (2\pi \mu_{AB} / \beta h^2)^{3/2}$$

$$k^{TST}(T) = A \exp(-\beta V^{\ddagger})$$

(A7)

These equations express the rate (k) in terms of the partition functions of the reactants and the transition structure.[37] A relative translational motion (between reactants) partition function is now included (Q_{Rel}). Q (at the transition state or for the reactants) is now a total partition function composed of those electronic, vibrational, and rotational functions shown in the previous pages. This model is limited in that it does not account for a possible free-energy bottle-neck not at the transition structure. This possibility, as well as reaction path curvature, is treated in other appendices.

Below are some examples employing the Boltzmann statistical formula to calculate populations of molecular conformers at equilibrium. The free-energies have been calculated with Gaussian software using the above partition functions.

Calculation of Rotamer Populations

Gluco-Fluoro-Gas-Phase 298K (Model Compound **4**)

$R := 8.314510$

$i := 0, 1 .. 2$

$T := 298.15$

Define R, and temperature T Input free energies from Gaussian frequency calculations. Matrix elements 0, 1 and 2 correspond to rotamers gt, tg and gg respectively.

$$G := \begin{bmatrix} -869.826875 \\ -869.82355200 \\ -869.82642200 \end{bmatrix} \cdot 627.5095 \cdot 4184$$

Free-energy is converted

from Hartrees to joules, for

consistency with RT.

$$F(A) := \frac{e^{\frac{-G_A}{R \cdot T}}}{\sum_{i=0}^{2} e^{\frac{-G_i}{R \cdot T}}}$$

Calculate populations from Boltzmann formula

$\%(A) := 100 \cdot F(A)$

$F(0) \rightarrow .6065978854200780298$

Output mole-fraction amounts

$F(1) \rightarrow 1.7965407948703199819 \cdot 10^{-2}$

$F(2) \rightarrow .3754367066312287703$

Confirm normalization

$$\sum_{A=0}^{2} F(A) \rightarrow 1.0000000000000000000$$

Output rotamer percent populations

GT

$\%(0) \rightarrow 60.6597885420078029\text{i}$

$$
\begin{array}{l}
\quad \text{TG} \\
\%(\, 1 \,) \rightarrow 1.79654079487031998
\end{array}
$$

$$
\begin{array}{l}
\quad \text{GG} \\
\%(\, 2 \,) \rightarrow 37.5436706631228770
\end{array}
$$

Thus using Mathcad and the free-energies from Gaussian, it is very straightforward to compute the equilibrium populations. When considering a different temperature, one must recalculate the free-energy as this is temperature dependent. This does not necessitate a new vibrational analysis as the force constants do not depend on temperature. Using the previous vibrational frequency data, the new free-energy is computed trivially by inputting the new temperature into the partition functions. Then the new free-energies are re-inputted into the Boltzmann formula using the appropriate temperature.

When solvation is included, the procedure is much the same except that the free-energies used include the solvation energy. The details of the solvation models are discussed in Appendix B. Below is an example of how solvated equilibrium populations are calculated at three different temperatures. Eipcm refers to the solvated energies of the three rotamers using the IPCM solvation model.[24] Gcorr298 refers to the free-energy correction applied at 298 K. The free-energy corrections were calculated using the Onsager solvation model, since it allows analytical calculation of frequencies.

Gluco-Fluoro-Solvated (Model Compound **4**)

$R := 8.31451 \cdot 0$

$i := 0, 1.. 2$

Define three temperatures

$T1 := 298.15$

$T2 := 373.15$

$T3 := 413.15$

Input Solvated IPCM energies from Gaussian output. Matrix elements 0, 1 and 2 correspond to gt, tg and gg respectively.

$$\text{Eipcm} := \begin{bmatrix} -869.96007840 \\ -869.95768500 \\ -869.96013820 \end{bmatrix} \cdot 627.5095 \cdot 4184$$

Input free energy corrections from Gaussian Onsager frequency output. Matrix elements 0, 1 and 2 correspond to gt, tg and gg, respectively. Corrections calculated at all three temperatures using the "freqchk" utility in Gaussian.

$$\text{Gcorr298} := \begin{bmatrix} 0.123165 \\ 0.122813 \\ 0.123443 \end{bmatrix} \cdot 627.50954184$$

$$\text{Gcorr373} := \begin{bmatrix} .111553 \\ .111131 \\ .111899 \end{bmatrix} \cdot 627.50954184$$

$$\text{Gcorr413} := \begin{bmatrix} .104944 \\ .104483 \\ .105326 \end{bmatrix} \cdot 627.50954184$$

Add free energy corrections to IPCM solvation energies

$\text{G298} := \text{Eipcm} + \text{Gcorr298}$

$\text{G373} := \text{Eipcm} + \text{Gcorr373}$

$\text{G413} := \text{Eipcm} + \text{Gcorr413}$

Calculate populations from Boltzmann formula

$$F298(A) := \frac{e^{\frac{-G298_A}{R \cdot T1}}}{\displaystyle\sum_{i=0}^{2} e^{\frac{-G298_i}{R \cdot T1}}}$$

$$\%298(A) := 100 \cdot F298(A)$$

$$F373(A) := \frac{e^{\frac{-G373_A}{R \cdot T2}}}{\displaystyle\sum_{i=0}^{2} e^{\frac{-G373_i}{R \cdot T2}}}$$

$$\%373(A) := 100 \cdot F373(A)$$

$$F413(A) := \frac{e^{\frac{-G413_A}{R \cdot T3}}}{\displaystyle\sum_{i=0}^{2} e^{\frac{-G413_i}{R \cdot T3}}}$$

$$\%413(A) := 100 \cdot F413(A)$$

Output rotamer percent populations with solvation at 298K

GT
$\%298(0) \rightarrow 52.3903250385035824$
TG
$\%298(1) \rightarrow 6.02946074567507486$
GG
$\%298(2) \rightarrow 41.5802142158229268$

Output rotamer percent populations with solvation at 373K
GT
$\%373(0) \rightarrow 50.6719732157830253$
TG
$\%373(1) \rightarrow 9.55535622203640325$
GG
$\%373(2) \rightarrow 39.7726705621815713$

Output rotamer percent populations with solvation at 413K
GT
%413(0) → 49.7499252699395043

TG
%413(1) → 11.3595982125594291

GG
%413(2) → 38.8904765175010665

Using the populations of reactants in solution and the barriers to S_N2 displacement, one can calculate the relative rates of reaction using the equation for conventional transition state theory shown above. In practice, when computing *relative* rates, many factors are cancelled, leaving equation (2) from Results and Discussion.

$$
k_{Rel}\ galacto/gluco = \frac{p_{GalGG}(T)\bullet e^{-\frac{\Delta G^{\ddagger}_{GalGG}}{RT}} + p_{GalTG}(T)\bullet e^{-\frac{\Delta G^{\ddagger}_{GalTG}}{RT}}}{p_{GluGG}(T)\bullet e^{-\frac{\Delta G^{\ddagger}_{GluGG}}{RT}} + p_{GluGT}(T)\bullet e^{-\frac{\Delta G^{\ddagger}_{GluGT}}{RT}}} \tag{2}
$$

This equation was formulated for the identity reactions keeping in mind the principle of microscopic reversibility. Thus the numerator and the denominator are each the sum of two equal terms. Therefore, simply to compute the relative rates, a further simplification is possible.

$$
k_{Rel}\ galacto/gluco = \frac{2 p_{GalGG}(T)\bullet e^{-\frac{\Delta G^{\ddagger}_{GalGG}}{RT}}}{2 p_{GluGG}(T)\bullet e^{-\frac{\Delta G^{\ddagger}_{GluGG}}{RT}}} = \frac{p_{GalGG}(T)\bullet e^{-\frac{\Delta G^{\ddagger}_{GalGG}}{RT}}}{p_{GluGG}(T)\bullet e^{-\frac{\Delta G^{\ddagger}_{GluGG}}{RT}}} \tag{3}
$$

This equation no longer includes all of the relevant processes but it will compute the correct relative rate. Below is an example of the calculation of the relative rates (*galacto/gluco*) of S_N2 displacement.

Calculation of relative rates of reaction

Fluoro Gluco compound

Solvated energy of reactive gg rotamer
$$\text{GG2ipcm} := -\,869.9601382 \cdot (627.5095 \cdot 4184)$$

Free energy corrections for gg rotamer at three temperatures
$$\text{GG2corr} := \begin{bmatrix} .123443 \\ .111899 \\ .105326 \end{bmatrix} \cdot (627.5095 \cdot 4184)$$

Free energies of reactive gg rotamer
$$\text{GG2} := \text{GG2ipcm} + \text{GG2corr}$$

Solvated energy of nucleophilic Cl ion
$$\text{Clipcm} := -\,460.3845862 \cdot (627.5095 \cdot 4184)$$

Free energy corrections for nucleophilic Cl ion at three temperatures
$$\text{Clcorr} := \begin{bmatrix} -0.015023 \\ -0.019465 \\ -0.021884 \end{bmatrix} \cdot (627.5095 \cdot 4184)$$

Free energies of nucleophilic Cl ion
$$\text{GCl} := (\text{Clipcm} + \text{Clcorr})$$

Solvated energy of SN2 transition structure
$$\text{TSgluipcm} := -\,1330.308438$$

Free energy corrections for SN2 transition structure at three temperatures
$$\text{TSglucorr} := \begin{bmatrix} .117234 \\ .103963 \\ .096411 \end{bmatrix}$$

Correction for Basis Set Superposition error
$$\text{BSSEGLU} := 0.0004198$$

Free energies of SN2 transition structure
$$\text{TSglu} := (\text{TSgluipcm} + \text{TSglucorr} + \text{BSSEGLU}) \cdot (627.5095 \cdot 4184)$$

Free energy barriers to SN2 reaction for Gluco compound at three temperatures

$$DeltaGlu := TSglu - GG2 - GCl$$

Percent population of reactive rotamer for Gluco compound at three temperatures

41.58021421582292688
39.77267056218157131
38.89047651750106657

Free energy barriers to SN2 reaction for Gluco compound at three temperatures

$$DeltaGlu = \begin{bmatrix} 1.19513273629 \cdot 10^5 \\ 1.266415054448 \cdot 10^5 \\ 1.304222250817 \cdot 10^5 \end{bmatrix}$$

Fluoro Galacto Compound

Solvated energy of reactive gg rotamer
$$GG1ipcm = -869.9558297$$

Free energy corrections for gg rotamer at three temperatures
$$GG1corr := \begin{bmatrix} .123356 \\ .111814 \\ .105245 \end{bmatrix}$$

Free energies of reactive gg rotamer
$$GG1 := (GG1ipcm + GG1corr) \cdot (627.5095 \cdot 4184)$$

Note Cl ion data is valid for all reactions and will not be redefined

Solvated energy of SN2 transition structure
$$TSgalipcm := -1330.305144$$

Free energy corrections for SN2 transition structure at three temperatures
$$TSgalcorr := \begin{bmatrix} .117294 \\ .103994 \\ .096430 \end{bmatrix}$$

Correction for Basis Set Superposition error
$$BSSEGAL := .0004630$$

Free energies of SN2 transition structure
$$\text{TSgal} := (\text{TSgalipcm} + \text{TSgalcorr} + \text{BSSEGAL}) \cdot (627.5095 \cdot 4184)$$

Free energy barriers to SN2 reaction for Galacto compound at three temperatures
$$\text{DeltaGal} := \text{TSgal} - \text{GG1} - \text{GCl}$$

Percent population of *gg* rotamer for Galacto compound at three temperatures
.4789865775550650878
1.1917216746359895 8
1.68535453760624396

Free energy barriers to SN2 reaction for Galacto compound at three temperatures

$$\text{DeltaGal} = \begin{bmatrix} 1.17349074187 & \cdot 10^5 \\ 1.243959155109 & \cdot 10^5 \\ 1.281346271517 & \cdot 10^5 \end{bmatrix}$$

Calculation of relative rates Galacto/Gluco using equation (3)

$$\text{GalvsGlu298} := \frac{.47898657755506508788 \cdot e^{\frac{-\text{DeltaGal}_0}{R \cdot T1}}}{41.580214215822926880 \cdot e^{\frac{-\text{DeltaGlu}_0}{R \cdot T1}}}$$

Output of relative rate Galacto/Gluco at 298K
$$\text{GalvsGlu298} = 0.02757946255$$

Calculation of relative rates Galacto/Gluco at 373 K

$$\text{GalvsGlu373} := \frac{1.1917216746359895831 \cdot e^{\frac{-\text{DeltaGal}_1}{R \cdot T2}}}{39.772670562181571352 \cdot e^{\frac{-\text{DeltaGlu}_1}{R \cdot T2}}}$$

Output of relative rate Galacto/Gluco at 373K
$$\text{GalvsGlu373} = 0.06179117776$$

Calculation of relative rates Galacto/Gluco at 413 K

$$\text{GalvsGlu413} := \frac{1.6853545376062439677 \cdot e^{\frac{-\text{DeltaGal}_2}{R \cdot T3}}}{38.890476517501066574 \cdot e^{\frac{-\text{DeltaGlu}_2}{R \cdot T3}}}$$

Output of relative rate Galacto/Gluco at 413K

$\text{GalvsGlu413} = 0.08434555391$

S$_N$2 Reactivity Appendix B: Solvation Models

Although the most convenient and rigorous theoretical treatment of molecular electronic structure and energy is that of the isolated gas-phase system, often it is desirable to estimate the effects of solvation. In practice, most reactions in organic chemistry occur in solution. Thus while certain insights may be obtained from gas-phase calculations, ultimately for physical organic applications the solvent must be considered. The overwhelming complexity of molecules in solution must be appreciated. The carbohydrate systems discussed in this thesis have approximately 60 degrees of freedom, and 100 electrons. The trade off between accuracy and affordability for this system alone has already been discussed. DMF is the solvent that is commonly used for reactions of these systems. Each molecule of solvent adds 36 degrees of freedom and 40 electrons to the equation. Even a slightly dilute solution would have at least 200 solvent molecules per solute molecule. Thus an explicit consideration of the solvation effects would require the treatment of an extra 7200 degrees of freedom and 8000 electrons. This is clearly out of the question as far as an accurate treatment is concerned.

The next level down in the hierarchy of solvation modeling would be an explicit first solvation shell, surrounded by a continuum model representing the average bulk properties of the solvent.[39] The first solvation shell is called the cybotactic region. It is recognized that through specific interactions between the solute and solvent, the properties of the solvent may differ considerably from that of the bulk in this region.[39] For the carbohydrate systems studied, the real

systems have hydroxyl groups at the C1, C2 and C3 positions. These groups are known to hydrogen-bond directly to solvent molecules in an ordered arrangement. This interaction is strong enough to effect changes in the electronic properties of the solvent molecules. For the systems discussed in this thesis, the anomalous behavior that was investigated was unrelated to these hydroxyl groups. Since they are present in both sets of systems that were compared, they could be judiciously removed. In the absence of such groups known to interact directly with the solvent, the inclusion of an explicit first shell becomes less critical. This is very important since inclusion of even one shell of solvent molecules would reduce the affordable level of theory drastically.

This leaves the lowest levels of solvation modeling, the implicit models. The use of implicit continuum solvent models should not be viewed as a terrible compromise. As discussed by Cramer and Truhlar in their review of this topic,[39] there are some considerable advantages to using an implicit model. The explicit 200-molecule system discussed above must have its properties averaged over a vast number of configurations in order to reflect the observed behavior. In contrast, the properties used for the continuum model are already those of the average bulk behavior. Thus a sort of dynamic averaging which is extremely difficult to implement explicitly is already implicit in the model.

The simplest solvation model used was the Onsager spherical cavity model.[40] In this model, the molecule is placed into a spherical cavity whose radius is determined by first calculating an approximate molecular volume. An appropriate radius is defined as that necessary to enclose the electron density to a

certain density cut-off value. Then the molecule is allowed to interact with the solvent through its molecular dipole. Various solvents have been parameterized, and interactions are calculated through their dielectric constant and polarizability. The molecule's dipole induces in the solvent a dipole determined by the dielectric constant and the polarizability of the solvent. The induced dipole and the molecular dipole interact to stabilize the total system. The equation for the solvation energy with the Onsager model is:[40]

$$\Delta E_{solv} = -\frac{\varepsilon - 1}{2\varepsilon + 1} \frac{\mu^2}{R^3} \left[1 - \frac{\varepsilon - 1}{2\varepsilon + 1} \frac{2\alpha}{R^3} \right]^{-1} \qquad (B1)$$

Thus the energy of solvation is dependent on the radius of the cavity R, the dielectric constant of the solvent ε, and the dipole and polarizability of the solute (μ and α). This is an improvement over the now obsolete Born model in which the molecule interacts with the solvent only through its net charge. The Onsager model can be thought of as including the next term in the multipole expansion. From the above expression, it is clear that a molecule with no net dipole, as often occurs due to symmetry, will undergo no interaction whatsoever and the calculated solvation energy will be zero.

The main usefulness of this model lies in its simplicity. The calculated solvation energies are too crude to use for conformer populations, but the equation is so simple that solvated geometry optimizations may be performed as well as solvated frequency calculations for free-energy corrections. These solvated geometries may be used as input for more sophisticated solvation models with which geometry calculations are not feasible. The solvated frequencies may

be retained to provide the vibrational contributions to the free-energy. These were the purposes for which the Onsager model was used in the research described in this thesis.

It is easy to imagine how the geometry of a molecule with a net dipole is varied during a geometry optimization using this model. Trial variations from the gas-phase minimum energy configuration will be retained if they increase the dipole enough to increase the solvation energy by more than they raise the non-solvated structural energy. That is to say, that any variation of the geometry will increase the system's non-solvated energy (since it was a minimum). This variation will be retained if the corresponding change in dipole moment causes a net stabilization through the Onsager solvation energy term. Thus the geometry is optimized with the inclusion of this term. In practice, the changes in geometry observed with this model are usually relatively small. Interestingly, molecules with large dipole derivatives with respect to internal coordinates (those with large IR absorptions) should exhibit large structural changes with solvation (since the dipole changes quickly as the nuclei are moved). More closely associated with structural changes due to solvation are those molecules having freely rotating fragments in a shallow potential (again with some variation of the dipole associated with this movement). The inclusion of the Onsager solvation energy in optimizations of one of the carbohydrate systems, (**6**), was sufficient to destabilize one of the rotamers entirely, leaving only two stable minima.

The Isodensity Polarizable Continuum Model (IPCM)[24] was the most sophisticated solvation model used here. In this model the shape of an electron

density iso-surface is used instead of a spherical cavity. While surrounding the molecule by a polarizable continuum, a new molecular self-consistent-field (SCF) is converged. Then the Isodensity surface corresponding to the new SCF is used to define a new cavity. This process is repeated until the cavity shape is self-consistent. The solvation energies obtained with this model are generally considered to be of sufficient accuracy for quantitative estimates of conformer populations in solution. It should be noted that a conceptually superior method exists (SCIPCM[24]) in which the solvent field is coupled directly into the SCF equations. At this time, for all of its conceptual superiority, the implementation in Gaussian is fraught with numerous problems, making this method unusable. Thus the generally reliable IPCM method was used.

S_N2 Reactivity Appendix C: Variational Transition State Theory

Most electronic structure calculations today are performed with the Born-Oppenheimer approximation.[1] In this approximation, the electronic and nuclear wavefunctions are assumed to be separable, which allows product-function solutions. The time-independent Schrödinger equation for a molecule can be written as:

$$[T_{nuc}(R) + T_{el}(r) + V_{nn}(R) + V_{ne}(R,r) + V_{ee}(r)]\Psi(R,r) = E\Psi(R,r) \quad (C1)$$

Where T and V represent the kinetic and potential energy operators, respectively. Assuming the wavefunctions to be separable allows solutions of the form:

$$\Psi(R,r) = \Phi(R,r)\chi(R) \quad\quad\quad (C2)$$

This means that electronic wavefunctions are computed for each configuration of the nuclei and are thus considered to be a function of the nuclear configuration. This is the same as in AIM theory, wherein the electron density (arising from the wavefunction solution) is considered to be a function of the nuclear coordinates. The nuclear wavefunctions are then calculated as the solutions for particles subject to the potential arising from variation of the electronic energy with coordinate displacements.

The Morse potential is shown below. If the electronic energy of a diatomic molecule were to follow the Morse potential as the internuclear distance was varied, then the *nuclear* wavefunctions could be represented by the energy levels shown. The lowest level possesses the well-known zero-point energy.
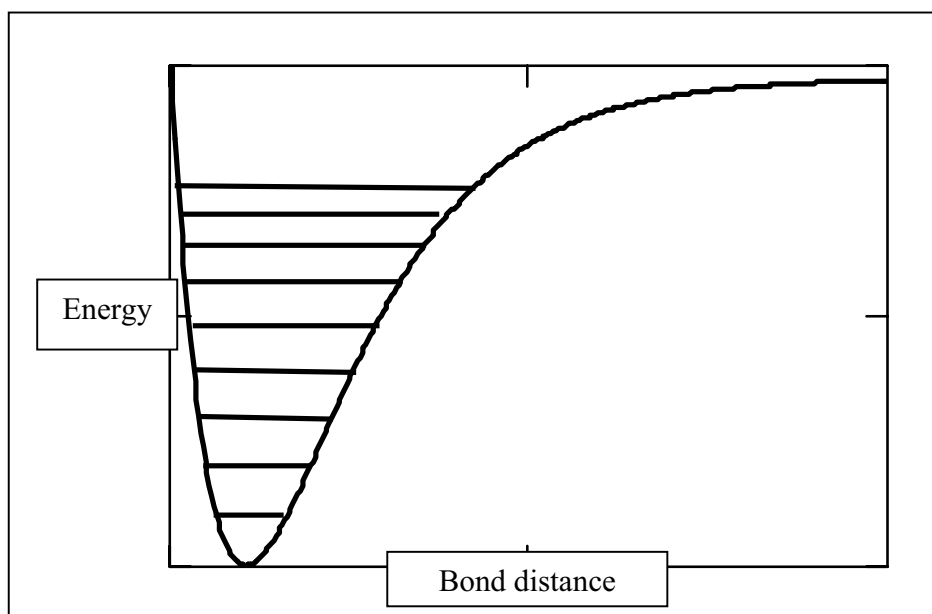
Figure C1: Morse potential with associated nuclear energy levels
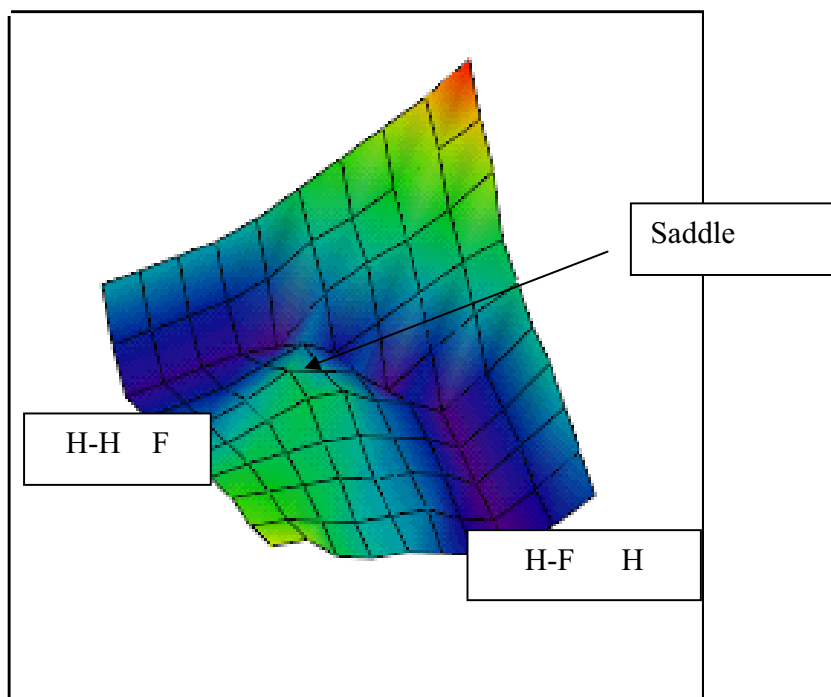
Figure C2: Illustrative potential energy surface for co-linear H₂F (hypothetical values)

In Figure C2 this concept is extended to display a typical reaction surface for a co-linear system such as H_2F. Each bond length is displayed on a different axis, with the saddle-point between the minima shown with an arrow.

Once the potential surface for a reaction has been obtained, then a choice of dynamical theory is made. In the section on thermo-chemistry, the harmonic potential approximation and the conventional transition state theory were used to compute rates of reaction for given potentials. The advantages of the harmonic approximation for evaluation of the partition functions were also discussed. The harmonic approximation can be maintained while still greatly improving on the conventional transition state theory. Recall that with the conventional theory a rate is computed using the free energy barrier from the reactants to the transition structure. It is worth noting the distinction that the transition *structure* is the *structure* corresponding to the saddle-point in the potential surface, while the transition *state* is the highest free energy through which any particular microstate must pass to reach the product states. This will be discussed in more detail as the variational theories are described.

The main failing of the conventional theory is not a consequence of the harmonic approximation; in fact one could go beyond the harmonic description in the free energy calculation and still use the conventional rate equation. The main failing is the assumption that any reactant molecule with sufficient energy to cross the saddle-point will do so, irreversibly. This fails to account for so-called re-crossing trajectories.[37] Reaction trajectories may in fact cross the barrier multiple times, eventually either continuing on to products or being reflected back to

reactants. Figure C3 below illustrates two reaction trajectories which each pass the conventional saddle-point more than once. In one case the eventual result is productive, while in the other it is not.
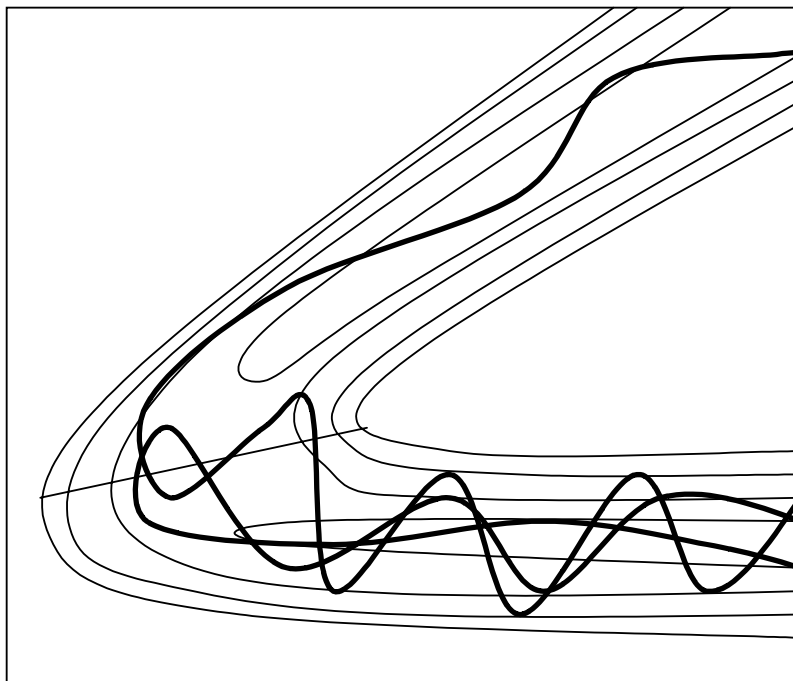
Figure C3: Energy contour plot showing re-crossing trajectories

The reason that this occurs is that the saddle-point represents the highest point in the *potential*, but not necessarily the highest point in the free energy. If one examines the thermodynamic partition functions, then it is clear that they might vary along the reaction path. The moments of inertia are changing and in particular, due to the evolving electronic structure as bonds are formed and broken, the vibrations change considerably. Figure C4 shows an illustration of how vibrational frequencies might vary along the reaction path. Although hypothetical, this figure is similar to those appearing in the literature. Following the reaction coordinate s from left-to-right in the figure, the highest frequency

mode in the starting compound could be related to an N-H stretching mode. Due to a change in the bonding arrangement at this nitrogen center during the reaction, this frequency could change considerably. Thus in general the maximum in free energy along the reaction path may occur to one side or the other of the potential-saddle. In Appendix D the details of the reaction path itself will be discussed. For this discussion the reaction path will be considered to be the minimum energy path between reactants and products, passing through the transition structure.
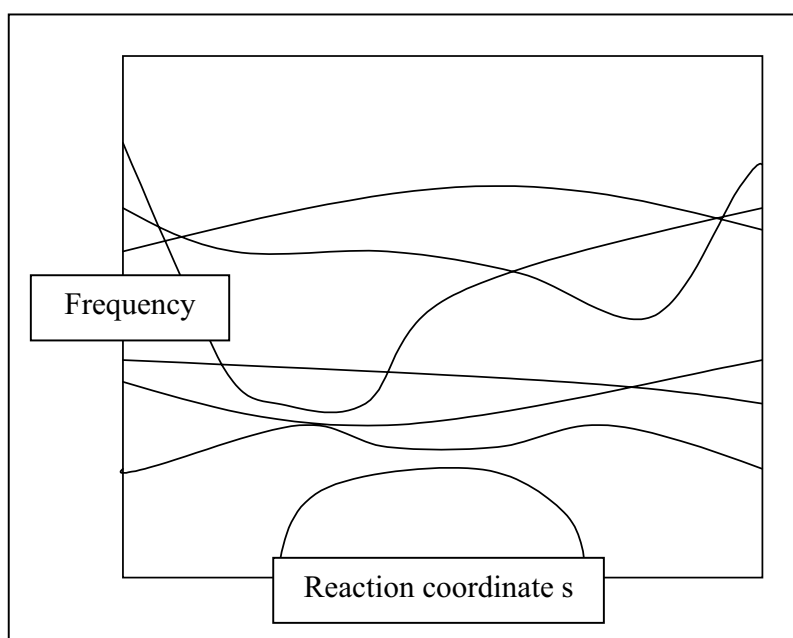
Figure C4: Illustration of how vibrational frequencies might vary along reaction path

Given that the maximum free energy may not be at the transition structure and that this is the cause of re-crossing trajectories and an over-estimation of reaction rates by the conventional theory, improved theories are desirable. A hierarchy of variational transition state theories has been developed by Donald Truhlar at the University of Minnesota.[41-44] Only two of them will be described

here, since only the simplest versions of the theory are easily applied to larger systems.

Canonical Variational Transition State Theory

The canonical variational transition state theory proceeds by calculating ensemble free-energies at the temperature of interest along the reaction path. By maximizing the free energy along the reaction path, the calculated rate is minimized variationally. This is expressed as:

$$k_C^{GT}(T,s) = \frac{\sigma}{\beta h} \frac{Q_C^{GT}(T,s)}{Q_R(T)} \exp\left[-\beta V_{MEP}(s)\right]$$

$$k_C^{CVT}(T) = \min_s k_C^{GT}(T,s)$$

(C3-C4)

where the rate k (canonical rate for a generalized transition state) is calculated as a function of temperature T and reaction coordinate s (along the minimum energy path MEP). The free energy bottleneck for the reaction is found by minimizing the rate along s. In practice, using commercial code such as Gaussian, the free energy is calculated at a series of points along the IRC and then the maximum is interpolated. Of course the step-size may be controlled, allowing this to be a very good approximation. This was the level of variational transition state theory used in the work described in this thesis. It is limited in that, with the adiabatic approximation (that the reaction coordinate involves a smooth, slow change in energy, leaving the microstate quantum numbers of each reactant unchanged as they proceed along the IRC), the transition state for reactants maintaining an excited vibrational state may be higher in energy than that predicted by the ensemble calculation.

Micro-Canonical Variational Transition State Theory

Micro-canonical variational transition state theory was formulated with this adiabatic behavior in mind. For each slice of total energy in the canonical ensemble, a generalized transition state is optimized. A molecule from an ensemble at a certain temperature, within a slice of a certain total energy, will have rotational and vibrational quantum numbers determined by the partition functions. The energy of each state evolves adiabatically along the IRC as:

$$V_a(n,k,s) = V_{MEP}(s) + \varepsilon_{\text{int}}^{GT}(n,k,s) \tag{C5}$$

Where the potential energy V (adiabatic), is a function of vibrational quantum numbers n, rotational quantum numbers k, as well as the potential due to the position s along the MEP.

The net rate is calculated from the rates of each energy slice, each reacting at its own rate. Depending on how thin the slices are, this approach can become extremely complicated. Usually a compromise is used in which microcanonical theory is used in slices up to the maximum energy of the adiabatic ground state. Since, for the ground state the rotational quantum numbers would be zero, this amounts to a vibrational zero-point-corrected potential maximum as seen in the following expression.

$$V_a^G(s) = V_{MEP}(s) + \varepsilon_{\text{int}}^G(s)$$
$$\varepsilon_{\text{int}}^G(s) = \frac{1}{2}\hbar \sum_{m=1}^{3N-7} \omega_m(s) \tag{C6-C7}$$

Reactants from the ensemble with total energy greater than this zero-point corrected potential maximum are treated by canonical variational theory. This

hybrid approach (called improved microcanonical variational theory) is seen as a good compromise between accuracy and complexity.

These methods are still limited in that they do not account for the effects of tunneling or reaction path curvature. Accurate tunneling corrections are difficult to apply for systems of high dimensionality and are usually most important at low temperatures. When the thermal rate is very low, the addition of tunneling is more significant. Since the systems described in this thesis are quite large, and the reactions are modeled at relatively high temperature from a tunneling standpoint, no corrections for tunneling were applied. Non-adiabatic behavior induced by large reaction path curvature is discussed in appendix D.

Survey of region of configuration space

During the calculation of the transition structures for the model sugar compounds it was decided to survey the potential energy associated with approach by the chloride nucleophile to both the reactive gluco compound **6** and the unreactive galacto compound **5**. The results seemed to indicate that the free energy bottle-neck might lie to one side of the transition structure. This did not prove to be the case. This survey involved holding the geometry of the model compound fixed at equilibrium and thus did not represent the full dimensionality of the problem. System **5** did exhibit large reaction path curvature however. The results of the 3D potential map are included here for completeness, although they were omitted from the submitted manuscripts.

Single-point gas phase total system energies for a truncated conical grid of 144 points revealed that most trajectories towards **5*tg*** are strongly disfavoured relative to trajectories approaching **5*gg*** (Figure C5). Despite the preference for **5*tg*** over **5*gg*** (~5 kcal mol^{-1} in the reactant alone), most Cl$^-$ approaches to **5*tg*** rise above the energy of the S$_N$2 transition structure **5TS** itself. Energies higher than **5TS** begin to appear at 6 Å within the cone approaching **5*tg***. At a distance of 4 Å, the majority of points are higher in energy than **5TS.** Only a few points closest to the association complex structure on that side of the reaction profile remain below the energy of **5TS** (by ~10 kcal mol^{-1}).

In contrast, approach to **5*gg*** is unimpeded, with only one point in the entire cone higher in energy than **5TS**, caused by the steric effect of H-5. At a distance of 4 Å, system energies are approximately 16–21 kcal mol^{-1} lower than **5TS**. The association complex with **5*gg*** is considerably lower in energy than that with **5*tg***. The same analysis was also applied to *gluco-* model **6*gg***, revealing no disfavoured trajectories and relative energies similar to the **5*gg*** case.
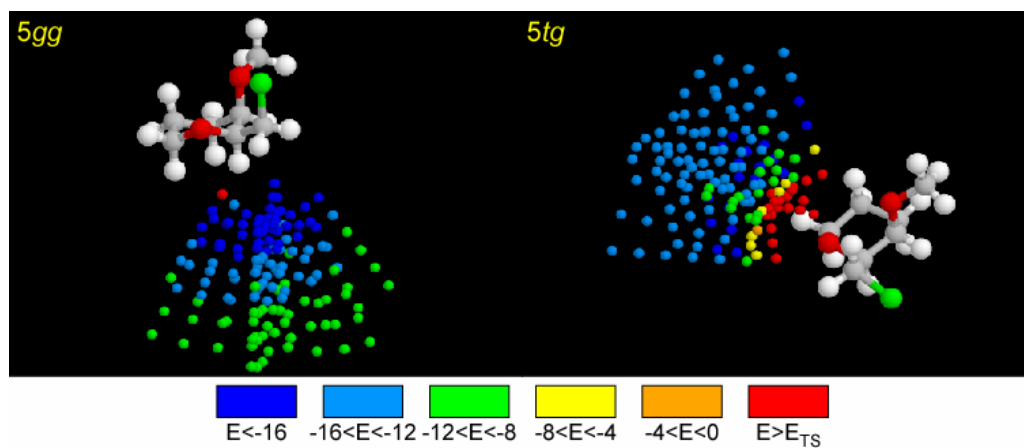
Figure C5: Single-point total system energies relative to 5TS (kcal mol^{-1}) for approach to 5*gg* and 5*tg* by Cl$^-$. Color-coded spheres show energies for chloride at that position.

S$_N$2 Reactivity Appendix D: Intrinsic Reaction Path

The Intrinsic Reaction Coordinate (IRC)[23] is a mathematical construct used in simplified dynamical theories for calculation of rates of reaction. Also known as the Minimum Energy Path (MEP), it describes the minimum-energy enthalpic path connecting two minima. At any point along the reaction coordinate all other degrees of freedom are at their minimum possible energy. Consideration of this path may seem unnecessary for use of the conventional transition state theory, where rate calculation only requires the energies of the reactants and the transition structure. However the IRC is what establishes the connection of the minima by the transition structure. Without this path one can only say that there is a transition structure somewhere between the two minima. It does not preclude the existence of a second, higher transition structure. The IRC path can only be followed downhill and must be constructed from a transition structure. When a structure is at a minimum, 3N-6 degrees of freedom are minimized. There is no way to determine which hill to climb.

Thus beginning at a transition structure, a vibrational analysis is performed. A true transition structure will have 3N-7 positive curvatures (and hence real, positive frequencies), and one negative curvature corresponding to an imaginary frequency. The eigenvectors of the force constant matrix define 3N-6 vibrational coordinates. The 3N-6 total vibrational coordinates are orthogonal at this point within the harmonic approximation since all gradients vanish. The eigenvector (coordinate) corresponding to the negative eigenvalue (imaginary frequency) defines the reaction coordinate at this point. After an infinitesimal

displacement along the reaction coordinate, the energy derivative with respect to this one coordinate is no longer zero since the movement is away from the stationary point. Because all the coordinates are orthogonal at the transition structure, this infinitesimal displacement along the reaction coordinate leaves the other coordinates at their minima. In theory, the IRC is defined by a series of infinitesimal displacements following the steepest descent of the non-zero gradient while continuously minimizing the other degrees of freedom in order to maintain their zero gradients. Thus the IRC runs smoothly downhill along the reaction coordinate from a transition structure to the next minimum. Then it stops. As mentioned previously, the path cannot be followed uphill.
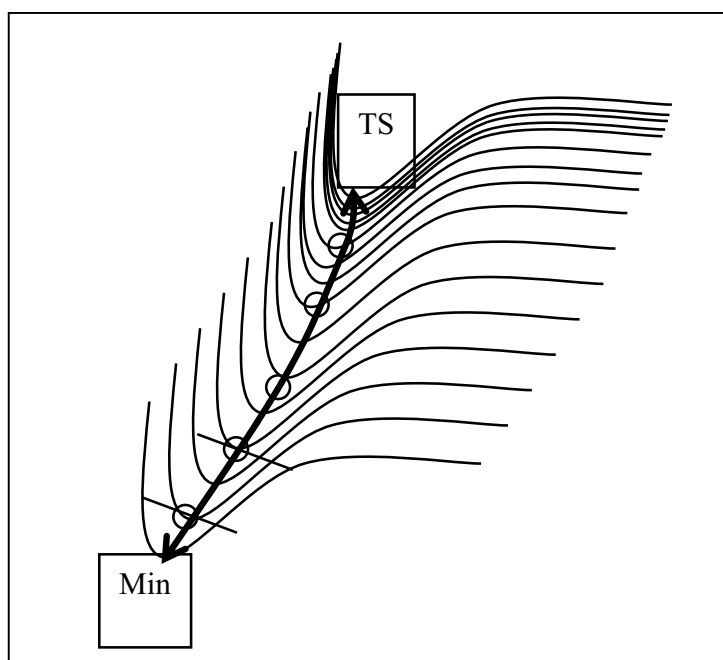
Figure D1: IRC connecting a transition structure with a minimum

The path is also followed down the other side, ideally connecting the transition structure to the products. If the path were to reach a previously unknown minimum, then a search for a missing transition structure must be undertaken and the new path followed until a continuous connection has been established between reactants and products.

In practice, a small *finite* step is taken, after which a new vibrational analysis is performed. This updates the force-constant matrix since, due to the evolving electronic structure along the path (bond forming and breaking), the vibrations change as a function of the path. Then the coordinate of the reaction path is projected out before determining the remaining 3N-7 coordinates and frequencies. This is important since, due to the non-zero gradient along the reaction coordinate, the reaction coordinate is no longer orthogonal to the other 3N-7 mutually orthogonal coordinates. Having established the other coordinates, due to the finite step size and anharmonicity of the true potential, the structure must then be minimized with respect to those 3N-7 coordinates. A series of these finite steps is performed until the structure arrives at a minimum. In theory, by definition the force-field should be updated continuously. In practice, it may not be computed at each step, but should be done often.

Once the IRC has been established, connecting the reactants with products through the transition structure, this information is applied to the dynamic theory. In the conventional transition state theory, one simply inputs the energies of the reactants and transition structure having excluded the possibility of any higher barriers. Most higher-level dynamical theories short of full reactive scattering

simulations rely on the IRC. As discussed in appendix C, variational transition state theories maximize free-energy along the IRC. During computation of the IRC, all of the vibrational analyses are retained, allowing evaluation of the free-energy as a function of s (the position along the reaction path). This provides the necessary information for the variational theories detailed in Appendix C.

As stated above, the IRC is in general not orthogonal to the other 3N-7 coordinates (except at a TS, or a minimum). This means that any mode that is not orthogonal to the IRC is coupled to it. In terms of the variational theories this means that the assumption of adiabacity breaks down. This may be true of some modes in particular. As a result, energy that is directed along the reaction path may be scattered into these modes. This can result either in vibrationally excited products, or possibly in reflection and a loss of reactivity. This may also account for the selective enhancement of reaction rates through specific excitations. This loss of adiabacity through coupling is defined as reaction path curvature:[30]

$$\kappa(s) = \left[\sum_{k=1}^{F-1} B_{k,F}(s)^2 \right]^{\frac{1}{2}}$$

$$B_{k,l}(s) = \sum_{i=1}^{F} L'_{i,k}(s) L_{i,l}(s)$$

(D1-D2)

In these equations F is the total number of vibrational coordinates (3N-6, or 3N-5 linear). The curvature κ is defined as a function of the position along the IRC s. It is obtained through summation over the couplings to the F-1 (3N-7) remaining coordinates. The coupling element B is defined here for a general point in configuration space where any number of gradients may be non-zero. It is simply the inner product of the two coordinates. As discussed above, along the

IRC, the F-1 coordinates are orthogonal to each other, making all of the B's zero except for those in the expression for curvature ($B_{k,F}$). Once the couplings have been evaluated and the curvature obtained, the choice on how to proceed depends on the application. For simple systems with low dimensionality, the strongly coupled coordinates might be included in some trajectory calculations with a reduced dimensional reaction path Hamiltonian (reduced by discarding weakly coupled modes). For the systems studied in the work described in this thesis, the dimensionality was too high to include any explicit treatment of these results. However, the theory does provide a qualitative explanation for some of the differences in reactivity.

Introduction

Since the early development of quantum mechanics, efforts have been made to predict and understand the electronic properties of atoms and molecules. At first, quantum mechanical considerations allowed researchers the ability to predict certain types of phenomena in a general way. After appreciation of the type of energy level structure possessed by atoms and molecules, the ideas of radiative transitions and spectroscopies followed. Many advances in understanding and predicting new non-classical phenomena were made despite the lack of both lasers and computers. In 1923, following the discovery of Compton scattering, Smekal predicted (via second-order perturbation theory) the possibility of inelastic scattering of photons from molecules.[45a] In 1928 C.V. Raman and K.S. Krishnan published their experimental results in Nature.[45b] They had filtered and focused light from the sun through a telescope into a solution of chloroform, and observed the production of a different wavelength of light. The green light entering the sample gave rise to a yellow secondary light due to interaction with the molecular states of the chloroform molecules. Raman was awarded the Nobel Prize and Raman spectroscopy was born.

This item has been removed due to copyright issues. To view it, refer to its source.

Figure 1: C.V. Raman
http://www.deltanu.com/tutorial.htm

At the time, the sometimes-counterintuitive predictions of quantum mechanics were being tested as carefully as possible. This was necessary to establish the resilience of the postulates. Thus the observation of Raman scattering was an important confirmation of part of the framework of our understanding of matter. Today the predictions of quantum mechanics have been confirmed to greater accuracy than those of any other theory.

Raman scattering is understood in terms of details of the molecular states and properties. Now, rather than confirming that quantum mechanics is a good description of matter, scattering experiments can be thought of as measurements of molecular electronic properties. One of the biggest drives in science today is to exploit the possible properties of materials. Researchers in many areas seek to tune the properties of materials literally atom-by-atom. Molecular electronics, quantum computing, and non-linear optical materials are three areas currently of great interest for which a complete understanding of the relationships between molecular structure and electronic properties is essential. Fundamental research is required in many areas in order to realize any of these goals.

The best description of the inelastic Raman scattering process is through Quantum Electro-Dynamics (QED).[46] In this treatment the electromagnetic field is quantized as well as the states of the molecule. Thus a Hamiltonian is constructed for the complete molecule-radiation-field system. This approach is necessary in order to recover some of the finer details of the physical process. In practice, for chemical-physics applications, most researchers use a semi-classical description, where the molecule is described quantum mechanically while the

field is described classically. This approach neglects the effect of the molecular processes on the field. Where there are discrepancies between the two models, QED should be considered as the arbiter. More will be said about the quantum theories of Raman scattering in the appendices. The semi-classical description is generally adequate for most purposes including the processes discussed in this thesis. The reader will notice that the terminology from the two approaches becomes somewhat mixed. It is common to discuss an incident "photon" even within the semi-classical description. This should not cause any confusion, as it simply reflects the fact that people are aware that excited field modes are what actually interact with the molecule. This has allowed the process to be understood in terms of the electronic properties of the molecule, and its interaction with applied fields. The work described in this thesis is in the area of *vibrational* Raman scattering. Thus for the inelastic scattering processes studied, the difference in energy between the incident and scattered photons corresponds to vibrational state transitions. Electronic and pure-rotational Raman scattering are similar processes but are not considered here.

Using a semi-classical description, the differential scattering cross-section was derived by Placzek as:[47]

$$\frac{\partial \sigma}{\partial \Omega} = \frac{\pi^2}{90\varepsilon_0^2}(v_0 - v_i)^4 \left[1 - \exp\left(\frac{-hcv_i}{kT}\right)\right]^{-1} \left[45\left(\frac{\partial \overline{\alpha}}{\partial q_i}\right)^2 + 7\left(\frac{\partial \gamma}{\partial q_i}\right)^2\right] \tag{1}$$

As seen in the above equation, the scattering intensity depends on the fourth power of the frequency of the incident photon. This is followed by a thermal

population term. Finally there are two parts to a term describing the polarizability derivative with respect to the vibrational modes of the molecule. The first part (isotropic) is the derivative of 1/3 of the trace of the polarizability tensor ($\overline{\alpha}$) with respect to the vibrational coordinates ($\overline{\alpha} \equiv$ mean molecular polarizability). The second part analogously reflects the anisotropic contribution of the polarizability tensor. It is this combined term that allows insight into the details of the electronic structure of the molecule, its variation with geometric changes, and its response to external fields. Thus much information can be derived from a Raman spectrum.

A quick review of some of the fundamental assumptions that are implicit in the above equation will illustrate the range of applicability of the semi-classical approach. It is assumed that the incident photon is much lower in energy than any excited electronic states. This is the non-resonance assumption that allows the intensity expression to be constructed in terms of only the vibrational states, neglecting excited electronic states. Clearly, if the frequency of the incident photon were increased into resonance with an electronic state, this expression would not apply.   Next are the assumptions of electrical and mechanical harmonicity. The above expression is most often used with the normal-mode coordinates derived from a harmonic force-field. Also the second and higher derivatives of the polarizability are considered to be zero (electrical harmonicity). However, unlike the electronic resonance condition, these assumptions do not limit the model. One can easily extend the definition of the vibrational coordinate to include that of an anharmonic coordinate derived from a higher-order force-

field analysis. Similarly the polarizability can be extended, to include higher derivatives. Where the interaction with an external field is concerned, one may consider the static polarizability (the dipole induced by a static field), or more realistically the dynamic polarizability (that induced by a dynamic field). The fact that vibrational states for which intensity is recorded are often not accessible by other spectroscopies makes Raman scattering useful for refining potential surfaces. Thus when fully extended, this model may be used to examine details of the anharmonic potential surface and the associated vibrational states, as well as the dynamic polarizability and its variation with geometry.

It became clear soon after the Raman process was discovered that different types of molecules could produce very different scattering intensities. This means that their electronic properties have different field responses and dependencies on geometric variations. Later it became clear that even in a group of similar molecules such as the saturated hydrocarbons, the influence of subtle structural differences could be significant. It is this observed range of behavior that inspires researchers to imagine what properties might be obtained if the right molecular or material structures were built.

As will be discussed in more detail in the following sections, these properties may also be calculated. This is another area in which the simple capability of calculating the properties of a particular system does not suffice. The essentially infinite number of possible molecular and material structures cannot be considered individually. The relationships between structure and electronic properties must be elucidated and recognized in terms of trends and

underlying patterns. Calculation of polarizabilities, as with the computation of any molecular properties such as structure and energy, involves the inevitable trade off between accuracy and cost. The most accurate methods are too costly to apply towards systems with enough atoms to exhibit the most unusual behavior. Thus it is important to benchmark the lower level, more widely applicable calculations. Where lower level calculations hint at certain trends or relationships, experimental data must be collected to verify and guide the theoretical progress.

The work described in this part of the thesis is a combined theoretical and experimental investigation into the Raman scattering intensities from saturated hydrocarbons. In the next sections I will describe some initial polarizability survey calculations. These low-level *ab initio* calculations were expected to reveal trends in the polarizability and its derivatives in these systems.

After the sections describing the survey calculations will be a number of sections describing the experimental work performed with bicyclo-[1.1.1]-pentane, followed by some higher-level theoretical treatments of this system.

Theoretical Methods

In choosing the best computational method for a large set of survey calculation, previous approaches by other authors were considered. These approaches have ranged from empirical and semi-empirical to *ab initio*. The bond polarizability model,[48] for example, is largely empirical and represents the polarizability of the molecule as that from an arrangement of polarizable bonds. Thus each type of bond in the molecule contributes some additive amount of polarizability to the system. A model with an approximate polarizability ellipsoid for each bond type (e.g. C-H, C-C, Si-H) would predict that the electronic properties of any molecule containing these bonds would change by the same amount with the motion of the same atom type. This allows for no variation at all in recorded intensities due to different structural factors. The inadequacies of this type of crude model have been the motivation for more detailed calculations. Semi-empirical methods such as AM1 are of no use whatsoever for the calculation of higher-order electronic properties such as polarizability. Although some semi-empirical methods have been proposed, and purely empirical models have been used in the past, they suffer the same failing for this application. Because they make simplified assumptions about the bonding of each type of atom, they will never reveal subtle variations in the electronic structure introduced by unusual bonding arrangements.

Thus *ab initio* calculations are the best way to proceed. The problem here is that complex structures likely to exhibit the most extreme variations in their properties have a large number of electrons and nuclei. Most methods scale quite

dramatically in cost relative to the number of electrons, while the number of nuclei increases the number of degrees of freedom for a vibrational analysis. This restricts the use of correlated methods for survey calculations. Yet, a good description of the net response of a molecule's electron density to an external field requires electron correlation.[49-56] Furthermore, accurate polarizability calculations are known to require an extensive representation of the polarization space.[49-53,57] Thus large basis sets augmented with high angular momentum polarization functions are preferred. This also increases the computational cost considerably. The best approach is therefore a compromise in both basis set and method. This raises the question of the reliability of the calculations. When correlation effects are neglected, the absolute magnitude of the results will be dubious. In addition, when smaller basis sets are employed, systematic deficiencies may arise in the results. Thus the only possibly fruitful way to proceed with survey calculations in these circumstances is to treat one particular group of molecules. It is hoped that the deficiencies of the method will be similar in a relative sense throughout the group. This *does* mean that any unusual molecules identified by the survey calculation will warrant a more detailed theoretical study and also experimental confirmation if possible.

The Gough research group was first considering this project in the early to mid-1990's.[58-61] The available research computing power at that time was less than that of most PC's today. A series of basis-set and method evaluation studies performed on test sets of molecules indicated that MP2 was the best affordable method for small systems and that the D95(d,p) basis set was the best compromise

for hydrocarbons. This was before the widespread acceptance of the B3LYP hybrid-DFT method. Due to the large number of molecules in the planned study set, it was decided to forego the MP2 method for the basic Hartree-Fock SCF method. Thus the survey type polarizability calculations were undertaken using the Hartree-Fock method and the D95(d,p) basis set.

Survey Calculations

My entry into the Gough research group and participation in this project began in 1999 and coincided with the second part of the survey calculations for the saturated hydrocarbon project. The first of the two papers had already been published, and calculations for the second set were underway. I completed the calculations for the second paper as well as sought molecular descriptors that might correlate with the calculated derivatives. The presentation here is more complete since the trends have since been evaluated in even more detail. However, this project is best described chronologically so I will first describe some of the experimental observations that motivated this ongoing research.

The experimental observation that polarizability derivatives with respect to the internal coordinates of a molecule (intensity parameters) are not readily transferable motivated further investigation.

Methane[62]	$\partial\bar{\alpha}/\partial r$
C-H	1.26
Ethane[63]	
C-H	1.33
C-C	1.32
Propane[64]	
C-H *Me*	1.37
C-H *ip*	1.44
C-H *op*	1.25
C-C	1.21

Table1: Experimental intensity parameters for Methane Ethane and Propane (Cm/V x 10^{-30})

Insight into the origin of large intensity variations observed even within a group of similar compounds requires that results be placed into a common framework (coordinate system). Force-field methods are employed to transform the data into intensity parameters with respect to internal coordinates. This allows intensity parameters to be described in terms of the same coordinates used to describe the geometries and greatly facilitates the recognition and development of structure/intensity relationships.

As seen in Table 1, there is considerable variation in the intensity parameter for C-H stretching modes alone, especially when one considers that the recorded intensity is proportional to the square of this parameter. The data set is too small to draw any conclusions but it was noticed that the C-H bonds at the end of the chain in propane, oriented in the direction of the length of the molecule (C-H *ip*), have a significantly larger intensity parameter. Those attached to the same carbons but oriented out-of-plane (C-H *op*) have values that are quite reduced. The methylene hydrogens attached to the central carbon are oriented out-of-plane and yet their values are intermediate in this system and higher than the values found in methane or ethane. To explore this further as well as to detect any other trends, it was decided that the C-H stretching intensity parameters for as large a set of molecules as possible should be calculated.[65, 66] The set was to include all of the straight-chain hydrocarbons out to C15 and also C25, as well as many ringed, caged, bicyclic, and branched molecules (see Figures 2 and 3). The goal was to include as many unusual bonding arrangements as possible.
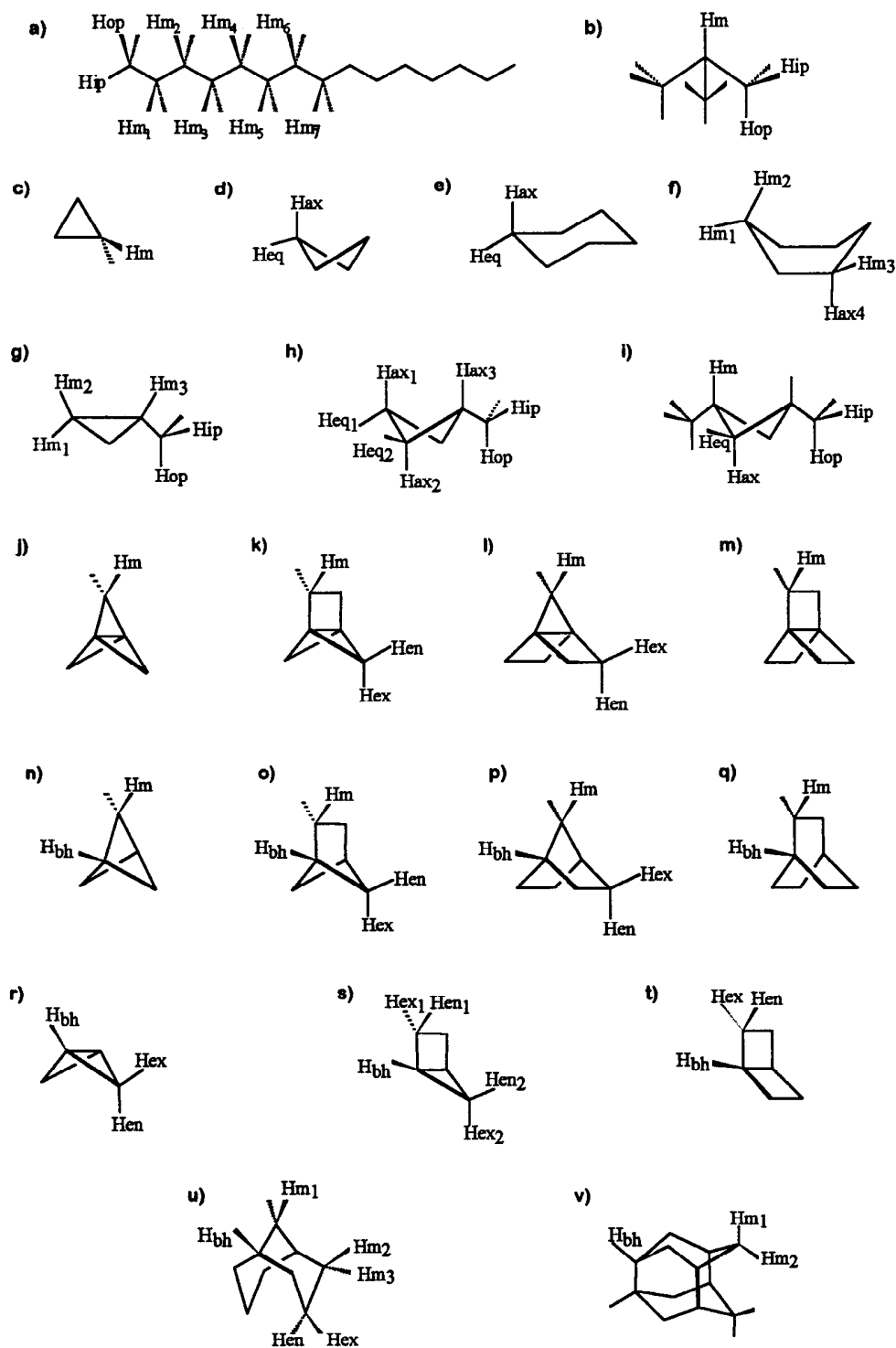
Figure 2: Set of molecules (SET 1) considered for C-H intensity parameter study
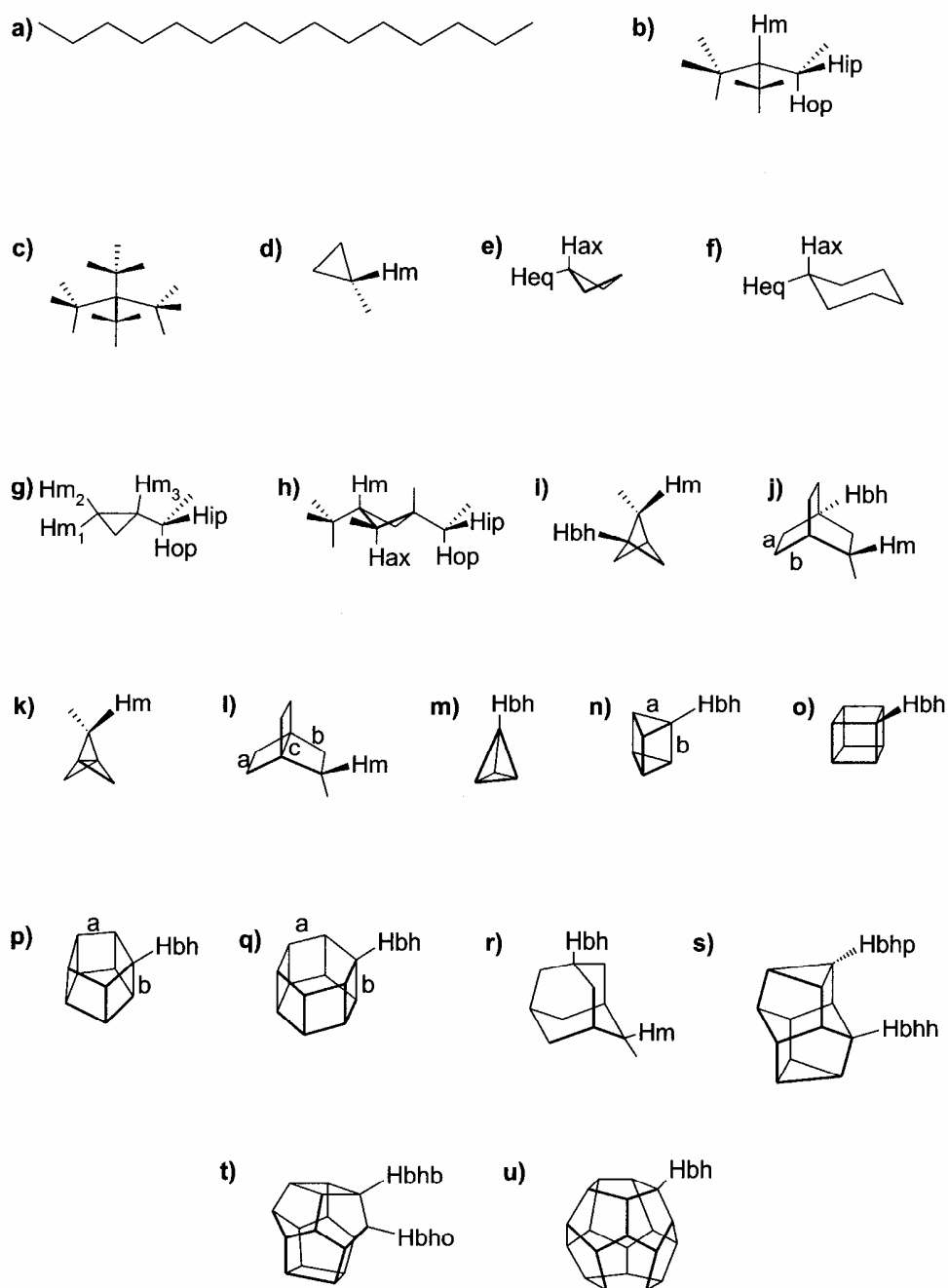
Figure 3: Set 2 molecules studied for C-C and C-H stretches.

Although it is straightforward to treat just about any molecule or type of vibration theoretically, the limitations of the experimental procedures dictate what can be confirmed. It was decided to compute intensity parameters for the type of vibration that is amenable to experiment.[63, 64] Experimentally, trace (isotropic) spectra are chosen because i) there are fewer bands: the derivative of the mean molecular polarizability, $\partial \overline{\alpha} / \partial q_i$ is non-zero only for totally symmetric normal modes; ii) bands are narrow and well resolved: Q branches only; and iii) trace scattering intensity parameters obey sum rules upon isotopic substitution: no rotational correction is required. Absolute intensities are then obtained by scaling relative to N_2 as an internal standard.

Based on these experimental considerations, the survey calculations were planned to model totally symmetric vibrational modes in internal coordinates.

Procedure:

 The geometry of each molecule was optimized at the same level of theory (d95(d,p)) as all of the polarizability calculations. The molecular polarizabilities were then calculated using the analytic Coupled-Perturbed-Hartree-Fock (CPHF) method.[67, 68]

As can be seen in Figure 4, this method is reasonably accurate for molecular polarizabilities of hydrocarbons in a relative sense. The actual numbers could be scaled to produce good agreement with experimental values. This type of performance was essential for the equilibrium molecular polarizabilities before numerical derivatives with respect to stretching motions could be obtained.
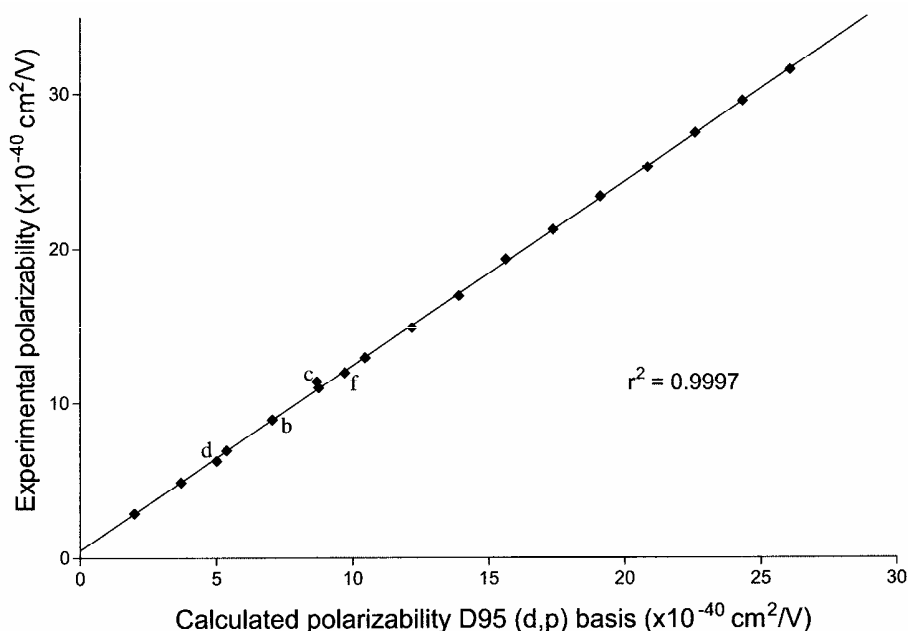
Figure 4: Experimental molecular polarizability vs. calculated with HF/D95(d,p) Data points correspond to n-alkanes (# of Carbons on x-axis). Points b, c, d, and f are molecules from Set 2 in Survey Calculations.[ref]

The derivatives of the polarizability with respect to symmetric stretching motions were calculated numerically by central difference.[65, 66] Symmetrically equivalent bonds were stretched and contracted a distance of $\Delta r = 0.010$ Å.

$$\frac{\partial \overline{\alpha}}{\partial r} = \frac{\Delta \alpha}{\Delta r} = \frac{1}{n}\left(\frac{\overline{\alpha}_+ - \overline{\alpha}_-}{2\Delta r}\right) \tag{2}$$

Thus, from polarizability calculations at these geometries (α_+ and α_-) in which n symmetrically equivalent bonds are stretched or contracted, the first derivative is obtained. The second derivatives of the polarizability tend to be very small. This means that polarizability derivatives calculated using formula (1) are very stable with respect to the numerical step-size. Tests including step-sizes of 0.001 to

0.015 Å have shown very little variation in this range. This does not exclude the possibility that the second derivative could be significant for some molecules. A numerical scheme for calculating the second derivative would be:

$$
\begin{aligned}
\frac{\partial^2 \overline{\alpha}}{\partial r^2} &= \frac{1}{n\Delta r}\left[\left(\frac{\overline{\alpha}_+ - \overline{\alpha}_{eq}}{\Delta r}\right) - \left(\frac{\overline{\alpha}_{eq} - \overline{\alpha}_-}{\Delta r}\right)\right] \\
&= \frac{1}{n}\left(\frac{\overline{\alpha}_+ + \overline{\alpha}_- - 2\overline{\alpha}_{eq}}{(\Delta r)^2}\right)
\end{aligned}
\tag{3}
$$

This is simply the change in the first derivative obtained from the stretched and contracted geometries. While the second derivatives were not explicitly computed using equation (3), the possibility of their being significant was discounted in the following way. The first derivative was first evaluated between the equilibrium and the stretched geometry. Since the stretch was 0.010 Å, the obtained derivative corresponds to that at the geometry stretched by 0.005 Å (the center of the difference). Next the derivative was evaluated between the equilibrium and the contracted geometry. This corresponds to the equilibrium geometry contracted by 0.005 Å. Thus by comparing the first derivative from the stretch and from the contraction, any significant difference is the second derivative (change in first derivative). The most that any calculated first derivative changed through the stretching motion in these molecules was less than 2%. Thus the step-size was deemed relatively unimportant, and the second derivatives were neglected. The results of these C-H stretching calculations will be shown in tables and figures to follow.

Based on the interesting results and trends observed in the C-H stretching calculations from the first set of molecules, a second set of calculations was begun. In this new study, derivatives for C-C stretching vibrations were computed for another set of molecules. This set was again chosen for its variety, while trying to maintain separability of the internal coordinates. Some of the molecules in the first set had hydrogen atoms bonded to the exterior atoms of a carbon framework. These could be selectively stretched without coupling to any other coordinates. The same could not be done for the C-C bonds of the carbon skeleton. In order to extract internal-coordinate polarizability-derivatives for some of these structures, one would have to perform a complete force-field and polarizability-field analysis. Using a complete set of derivatives for all the normal modes in the molecule, the data could then be transformed into internal coordinates using the L-matrix.[69] This would have been too computationally demanding for such a large set. The second set of molecules was determined to have enough variety in its bonding arrangements to elucidate trends, while maintaining convenient separability of the coordinates. While primarily constructed to compute C-C stretches, any C-H bonding arrangements not included in Set 1, were also computed from this set. Thus the complete set of data includes the C-H stretches from both sets, and the C-C stretches from set 2.

Survey Calculations: Results and Discussion

The results of the two sets of survey calculations are reproduced from the published manuscripts.[65, 66] Tables 2 to 4 show the raw data for the C-H stretch.

molecule	Hip	Hop	Hm1	Hm2	Hm3	Hm4	Hm5	Hm6	Hm7
methane	1.045								
ethane	1.096								
propane	1.181	1.053	1.126						
butane	1.214	1.046	1.068						
pentane	1.253	1.036	1.057	1.014					
hexane	1.274	1.037	1.046	1.004					
heptane	1.292	1.035	1.046	0.991	0.992				
octane	1.303	1.036	1.043	0.991	0.981				
nonane	1.313	1.036	1.045	0.989	0.979	0.968			
decane	1.319	1.036	1.044	0.990	0.978	0.968			
undecane	1.324	1.036	1.044	0.990	0.979	0.966	0.968		
dodecane	1.328	1.036	1.044	0.991	0.979	0.967	0.967		
tridecane	1.331	1.036	1.044	0.990	0.979	0.967	0.967	0.965	
tetradecane	1.333	1.036	1.044	0.991	0.979	0.967	0.967	0.966	
pentadecane	1.335	1.036	1.044	0.991	0.979	0.967	0.967	0.965	0.966

[a] Units: 10^{-30} C m/V.

Table 2: Polarizability derivatives (C-H) in n-alkanes, methane to pentadecane (Set 1)

molecule[b]	type	$\partial\bar{\alpha}/\partial r_{CH}$	molecule[b]	type	$\partial\bar{\alpha}/\partial r_{CH}$
isobutane (b)	Hip	1.107	[2.2.2]propellane (m)	Hm	1.130
	Hop	1.000	bicyclo[1.1.1]pentane (n)	Hbh	1.386
	Hm	1.118		Hm	1.124
cyclopropane (c)	Hm	1.133	bicyclo[2.1.1]hexane (o)	Hbh	1.298
cyclobutane (d)	Hax	1.111		Hen	1.145
	Heq	1.213		Hex	1.161
cyclohexane (chair) (e)	Hax	1.021		Hm	1.115
	Heq	1.259	bicyclo[2.2.1]heptane (p)	Hbh	1.264
cyclohexane (boat) (f)	Hm1	1.275		Hen	1.056
	Hm2	0.982		Hex	1.144
	Hm3	1.187		Hm	1.147
	Hm4	1.043	bicyclo[2.2.2]octane (q)	Hbh	1.285
methylcyclopropane (g)	Hip	1.203		Hm	1.103
	Hop	0.954	bicyclo[1.1.10]butane (r)	Hbh	1.216
	Hm1	1.073		Hen	1.128
	Hm2	1.184		Hex	1.156
	Hm3	1.144	bicyclo[2.1.0]pentane (s)	Hbh	1.250
methylcyclobutane (h)	Hip	1.178		Hen1	1.153
	Hop	0.967		Hen2	1.072
	Hax1	1.127		Hex1	1.180
	Hax2	1.049		Hex2	1.215
	Hax3	1.106	bicyclo[2.2.0]hexane (t)	Hbh	1.229
	Heq1	1.250		Hen	1.057
	Heq2	1.208		Hex	1.205
dimethylcyclobutane (i)	Hip	1.200	bicyclo[3.3.1]nonane (u)	Hbh	1.268
	Hop	0.940		Hen	0.908
	Hm	1.113		Hex	1.325
	Hax	0.995		Hm1	1.108
	Heq	1.200		Hm2	1.085
[1.1.1]propellane (j)	Hm	1.039		Hm3	1.162
[2.1.1]propellane (k)	Hm	1.107	iceane (v)	Hbh	1.233
	Hen	1.016		Hm1	1.130
	Hex	1.168		Hm2	1.063
[2.2.1]propellane (l)	Hm	1.105			
	Hen	1.081			
	Hex	1.257			

Table 3: Polarizability derivatives for C-H stretch (Set 1)

Molecule[a]	Bond[b]	$\partial\bar{\alpha}/\partial r_{CH}{}^a$ (10^{-30} cm/V)
Neopentane (**c**)		1.032
Tetrahedrane (**m**)		1.161
Prismane (**n**)		1.328
Cubane (**o**)		1.365
Pentaprismane (**p**)		1.317
Hexaprismane (**q**)		1.306
Adamantane (**r**)	H_{bh}	1.287
Adamantane (**r**)	H_m	1.106
Octahedrane (**s**)	H_{bho}	1.308
Octahedrane (**s**)	H_{bhh}	1.202
Decahedrane (**t**)	H_{bhb}	1.291
Decahedrane (**t**)	H_{bho}	1.218
Dodecahedrane (**u**)		1.220

Table 4: Polarizability derivatives (C-H) for Set 2

The three tables above list the polarizability derivatives calculated for C-H stretching motions in all of the molecules of sets 1 and 2. Through the n-alkanes the in-plane bonds at the ends of the chain are predicted to have increasingly large derivatives. This had already been observed experimentally for propane (see introduction). According to the calculations, this derivative increases non-linearly with chain length but eventually levels off. The methylene hydrogens in this series exhibit decreasing values with chain length. This trend also levels off at long chain length (see Figure 5).
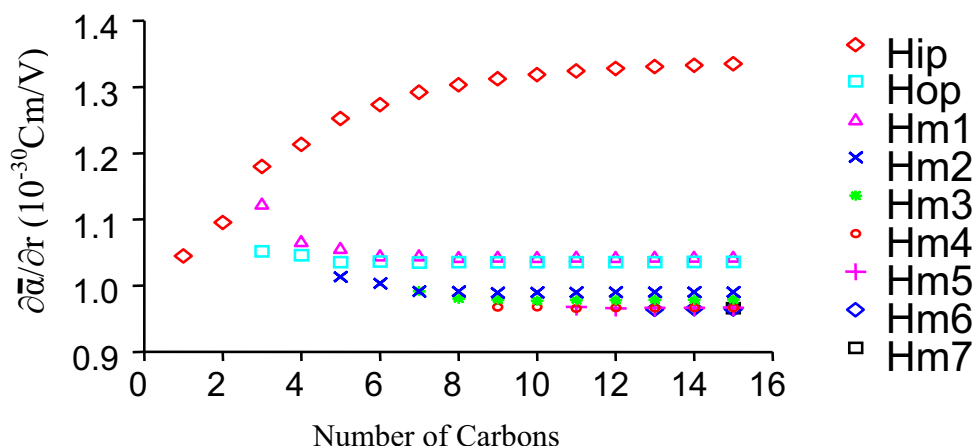
Figure 5: Values of polarizability derivatives for C-H stretches in n-alkanes

Some results from AIM analysis provide a reasonable explanation for the behavior of the derivatives with the C-H stretches in n-alkanes. It was found that stretching of the terminal C-H bonds provoked a large charge-transfer term between the terminal carbons when the field was applied down the length of the molecule. This behavior was dependent on the alignment of these bonds with the length of the molecule as well as their position at the ends of the chain. The polarizability was found to be sensitive to distortions of bonds in these positions, hence a large derivative occurred for stretch of the terminal in plane C-H bond. The net dipole caused by this charge transfer was opposed by atomic dipoles in the interior atoms. Thus with increasing chain length the damping by the interior atoms causes the derivatives to level off (see Figure 5). In contrast, the methylene hydrogens are oriented away from this molecular axis. The extent of the charge density perpendicular to the length of the molecule is similar throughout the chain and is much less than that over the length. It is also much less sensitive to C-H stretching motions. Thus the methylenes have similar, and much smaller, values for $\partial \overline{\alpha} / \partial r_{CH}$.

Analysis of the C-H stretch data for the other hydrocarbons in sets 1 and 2 yields some interesting patterns. In ring systems, such as cyclohexane, the equatorial bonds have larger derivatives than the axial, and this trend increases with ring size. This is easily understood by analogy with the n-alkanes. The equatorial bonds are aligned with the longer molecular axis, allowing greater charge transfer. This difference in axis length increases with ring size.

120

C-H stretching motions that encounter steric hindrance have significantly decreased values for their derivatives. Examples occur in the boat form of cyclohexane as well as the *gauche* conformer of butane. Where the two symmetrically equivalent atoms are brought together, the charge transfer is in the same region of space. This means that the induced dipole (charge x distance) is reduced because the distance is reduced. Otherwise this result can be understood in terms of the reduced change in molecular volume, which occurs when the motion is hindered.
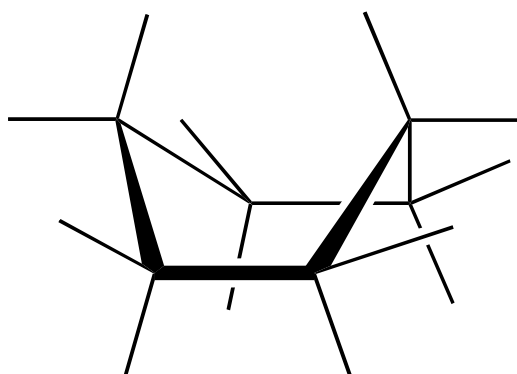
Figure 6: Boat conformation of cyclohexane

The largest derivative of polarizability with C-H stretching found throughout sets 1 and 2 is that for the bridgehead bonds in bicyclo-[1.1.1]-pentane. This can be understood as a superposition of two trends in the intensity parameters. Firstly, bridgehead bonds in general have significantly larger derivatives than most other bonds. This can be related to the results for the n-alkanes in which the terminal bonds have large derivatives. A bridgehead bond is

at the end of three chains, and although it is not perfectly aligned with each chain, there can be significant charge transfer from along each chain. Secondly, hydrogens attached to carbons with strained bonding arrangements tend to have large derivatives. This can be seen from tetrahedrane through to hexaprismane. With stretching of the attached hydrogen, the density of the strained carbon atom is more easily perturbed. In bicyclo-[1.1.1]-pentane the bridgehead hydrogens are attached to strained carbon atoms, yet there are also methylene groups between them to create polarizable chains. Bicyclo-[2.2.2]-octane has less-strained bridgehead carbons. In addition, the extra methylene groups mean that the C-H bonds are in much poorer alignment with the chains. Tetrahedrane and cubane are highly strained but there are no polarizable chains connecting the strained carbons. Also the connecting C-C bonds are not as closely aligned with the C-H bonds. Consequently their C-H derivatives are quite high, based solely on the strain factor, but not as high as those in bicyclo-[1.1.1]-pentane. Bicyclo-[1.1.1]-pentane also possesses a unique structural factor which may contribute significantly. The distance across the cage between the two non-bonded carbons is the shortest non-bonded carbon-carbon distance known for an equilibrium structure. This manifests itself in strong through-space interactions between the bridgehead carbons. Extremely strong couplings have been observed in NMR spectra of molecules based on this structure. Thus as well as the other factors already mentioned, this molecule exhibits direct electronic interactions across the cage.

In summary, the trends for derivatives of the polarizability with respect to C-H stretching vibrations are as follows:

i)   large when at the ends of chains

ii)   large when oriented with longest molecular axis

iii)   very large for bridgehead bonds

iv)   very large when attached to strained carbon atoms

v)   small with steric hindrance

vi)   medium with lack of above circumstances for large derivatives

vii)   total result is a superposition of the influence of the above trends

This completes the discussion of the derivatives of the mean molecular polarizability with respect to C-H stretching motions. The following reproduced tables include the data for the C-C stretches completed for the molecules in Set 2.

Molecule	$\partial\bar{\alpha}/\partial r_{CC}$ (10^{-30} cm/V)												
	CH$_3$-CH$_3$	1,2	2,3	3,4	4,5	5,6	6,7	7,8	8,9	9,10	10,11	11,12	12,13
Ethane	1.388												
	1.32 ± 0.04[a,b]												
Propane		1.329											
		1.21 ± 0.02[a,c]											
Butane		1.461	1.280										
		1.47 ± 0.06[a,d]	*1.27 ± 0.06*[a,d]										
Pentane		1.514	1.423	1.576									
Hexane		1.558	1.470	1.622									
Heptane		1.581	1.512	1.670									
Octane		1.604	1.533	1.670	1.670								
Nonane		1.614	1.553	1.691	1.713								
Decane		1.626	1.563	1.713	1.738	1.758							
Undecane		1.632	1.575	1.723	1.760	1.779							
Tridecane		1.643	1.587	1.741	1.781	1.816	1.823						
Pentadecane		1.649	1.595	1.752	1.794	1.832	1.850	1.857					
Penticosane		1.660	1.607	1.769	1.814	1.857	1.879	1.899	1.909	1.919	1.924	1.929	1.927

[a]Experimental values are shown in italics below the calculated value.

Table 5: Polarizability derivatives for C-C stretches in n-alkanes. Numbers in italics are the experimental values.

The derivatives obtained for the C-C stretches in the n-alkanes produced another interesting trend. The derivative for stretching of a C-C bond at any given position on the chain increases with the overall length of the molecule. For example the value for stretching the bond between carbon numbers 2 and 3 is larger for nonane than for heptane. This increase is non-linear at first but then levels off. Figure 7 shows this trend clearly, along with the steep non-linearity of the effect.
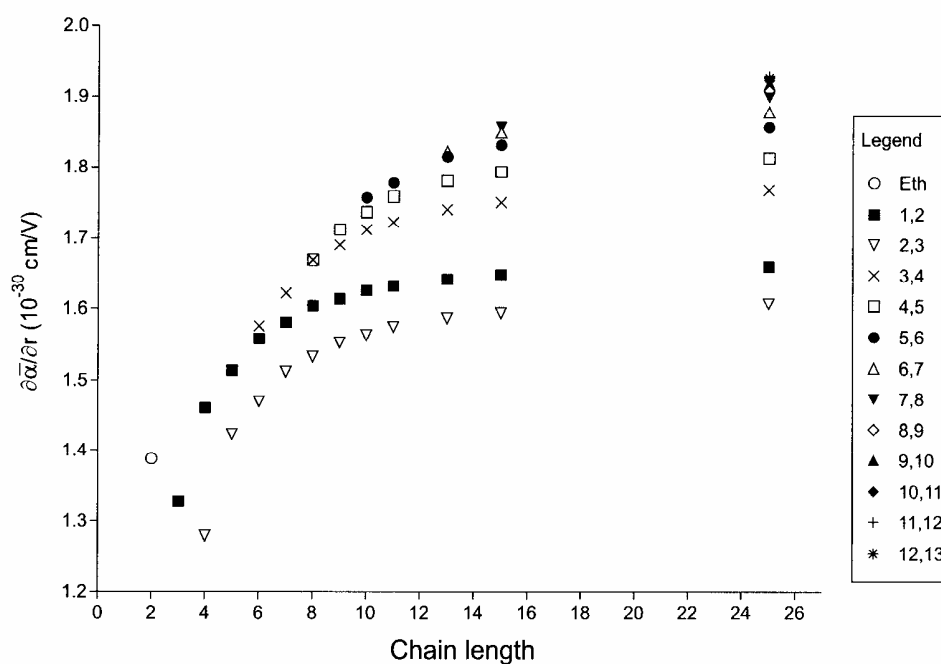
Figure 7: Calculated Polarizability derivatives for C-C stretches in n-alkanes. Legend identifies the two carbons participating in the bond. Numbering begins at the end of the chain.

Hidden in these data is another trend, that for the values along the chain for any particular hydrocarbon, the derivatives for the terminal C-C bond are larger than those for the next most interior. After this depressed value, the values

towards the center of the molecule increase non-linearly until the maximum is reached at the center. This trend was initially observed with the D95(d,p) basis set but was confirmed in 2003[70] by calculations with the B3LYP/aug-cc-pVDZ method. The results are shown in Figure 8 at both levels of theory for nonane and pentadecane and at the D95(d,p) level only for penticosane. The agreement between the two methods is surprisingly good.
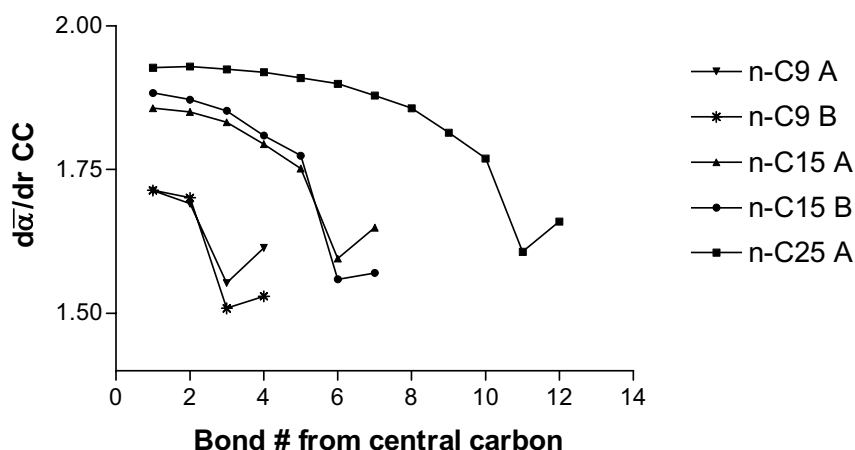
Figure 8: Polarizability derivatives for C-C stretches along carbon chains for C9, C15, and C25. (A) corresponds to the HF/D95(d,p) method, while (B) corresponds to the B3LYP/aug-cc-pVDZ method.

The results for C-C stretches in the other systems in set 2 are shown in Table 6. Calculated values for the strain energy of the carbons involved in the bond are included with these intensity parameters. The carbon strain energy was previously found to play a large role in the values for C-H intensity parameters. While the n-alkanes exhibit interesting trends of their own, some effort was made to bring order to the other results. Testing was undertaken for correlations

between many molecular properties and the C-C intensity parameters. Bond lengths, atomic populations, charge-densities at bond-critical-points, bond-ellipticities, atomic-volumes, and strain-energies were all considered. No significant correlations were observed except for that with the strain energies.

Molecule[a]	Stretch	$\partial\bar{\alpha}/\partial r_{CC}$[b]	Strain energy[c]
Isobutane (**b**)		1.280	0.00
Neopentane (**c**)		1.271	0.00
Cyclopropane (**d**)		1.196	18.76
Cyclobutane (**e**)		1.261	12.82
Cyclohexane (**f**)		1.217 (*1.03 ±.02*)[d]	0.12
Methylcyclopropane (**g**)	methyl ring	1.200	9.12
Methylcyclopropane (**g**)	ring	1.221	18.22
Dimethylcyclobutane (**h**)	methyl ring	1.370	5.84
Dimethylcyclobutane (**h**)	ring	1.293	11.70
Bicyclo[1.1.1]pentane (**i**)		1.218	27.30
Bicyclo[2.2.2]octane (**j**)	bonds labeled a	1.276	2.54
Bicyclo[2.2.2]octane (**j**)	bonds labeled b	1.145	2.54
[1.1.1]Propellane (**k**)	mixed symmetric stretch	1.157	40.90
[2.2.2]Propellane (**l**)	bonds labeled b	1.288	23.46
[2.2.2]Propellane (**l**)	mixed[e] stretch bonds a + c	1.092	23.46
Tetrahedrane (**m**)		0.981	69.92
Prismane (**n**)	bonds labeled a	0.978	47.60
Prismane (**n**)	bonds labeled b	1.085	47.61
Cubane (**o**)		1.053	38.72
Pentaprismane (**p**)	bonds labeled a	1.051	26.60
Pentaprismane (**p**)	bonds labeled b	1.007	26.60
Hexaprismane (**q**)	bonds labeled a	1.172	27.00
Hexaprismane (**q**)	bonds labeled b	0.904	27.00
Adamantane (**r**)		1.141	0.62
Octahedrane (**s**)	all bonds	1.037	12.96
Decahedrane (**t**)	all bonds	1.058	8.50
Dodecahedrane (**u**)		1.079	4.04

Table 6: Polarizability derivatives and strain energies for C-C stretches (Set 2)

The strain energy for hydrocarbons was calculated in the following way:[66]

i)   The molecular energy was calculated at the optimized geometry

ii)   The strain-free energies of the molecular fragments were computed as:

a. CH_3 = 1/2 E(ethane) = -39.6246205 au

b. $CH_2 = 1/2 \{[E(\text{propane})-2 \times E(CH_3)]+[E(\text{cyclohexane})/6]\} = -39.0431489$ au

c. $CH = E(\text{isobutene}) - [3 \times E(CH_3)] = -38.4623977$ au

d. Quaternary $C = E(\text{neopentane}) - [4 \times E(CH_3)] = -37.8810579$

iii) The difference between the molecular energy and the strain-free energy was deemed to be the strain-energy

iv) The total molecular strain-energy was simply averaged into the number of carbon atoms in the molecule

Figure 9 shows the rough correlation between the average strain energy per carbon and the calculated C-C intensity parameters. This correlation might have been better if the total strain energy of each molecule could have been distributed among the different carbon atoms more systematically. This could be accomplished through AIM analysis though at great cost. An obvious problem with this descriptor restrained efforts in this direction. The C-C intensity parameters for the prismane series of molecules are different for stretching of the two different types of C-C bonds. Viewing these molecules as two stacked rings, the two bond types are: the vertical bonds between rings, and the bonds within the rings themselves. Any description of the carbon atom itself will be the same since these bonds originate at the same atom. Thus no such descriptor can correlate well to these results. This correlation was retained because it is easily and generally applicable. It also has the advantage of being roughly estimated without any calculations at all.
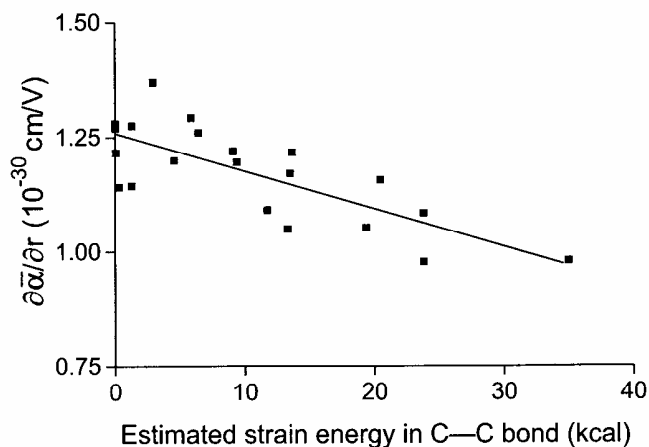
Figure 9: Polarizability derivative for C-C stretch vs. calculated strain energy (see Table 6)

The problem with trying to associate some properties of the equilibrium structure with the calculated intensity parameters is a fundamental one. While it would be ideal to identify easily recognizable structural features that correlate to a molecule's dynamic electronic properties, there is an inherent difficulty. The intensity parameter is a dynamic property, a measure of the system's response to changes in the structure. Thus we are actually measuring the change in an electronic property (polarizability) induced by small changes in the structure. This will not necessarily correlate to any property of the equilibrium structure.

In spite of this difficulty, many recognizable behavioral trends were identified. Since all of these factors discussed above were found in the results of the HF/D95(d,p) method, it remained to confirm that these trends were real, and not just artifacts of a deficient method. As mentioned in the theoretical methods section, the HF/D95(d,p) method lacks electron correlation and suffers from a limited basis set. Thus the best candidate for experimental study was chosen from

sets 1 and 2: bicyclo-[1.1.1]-pentane (Figure 10). This molecule was by far the most interesting molecule in the two sets, from both theoretical and experimental viewpoints. From a theoretical standpoint it was interesting since it has two types of C-H bonds, one predicted to be the largest intensity parameter known for a C-H bond, and the other very ordinary, typical of methylenes in other alkanes. Thus this molecule could benchmark the upper range of C-H intensity parameters, while showing a large contrast within one molecule. Experimentally this molecule meets many necessary criteria. For example, it is gas-phase at room temperature. This is important since one requires the response of the isolated system in order to compare experimental results with calculations. The spectra of solids do not reflect the type of vibrations treated here. It is highly symmetric, which limits the number of totally symmetric bands recorded. High-symmetry also facilitates the subsequent force-field analysis used to transform the recorded intensities into the desired internal coordinate intensity parameters.
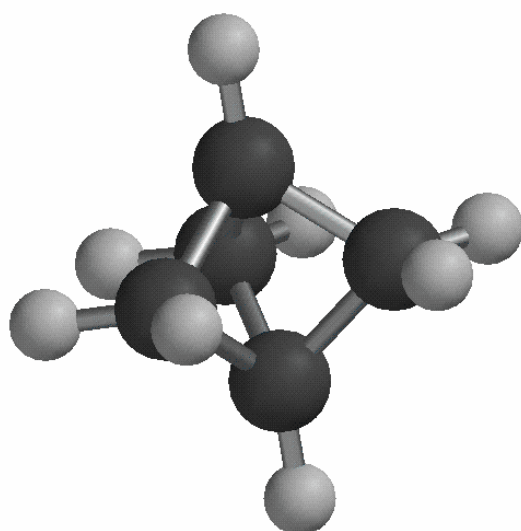
Figure 10: Bicyclo-[1.1.1]-pentane

From the two sets surveyed, this molecule was thus the most desirable candidate from which to obtain experimental absolute intensity Raman scattering spectra. Although it is quite stable, no companies sell it in any quantity. While custom synthesis companies advertise that they will make quantities of virtually any compound upon request, at least six declined the contract to prepare a sample of bicyclo-[1.1.1]-pentane. In the next section I will detail my own efforts to prepare a sample of this molecule.

Synthesis of Bicyclo-[1.1.1]-pentane

As mentioned in the previous section, the synthesis of bicyclo-[1.1.1]-pentane (BCP) was undertaken out of necessity. It could not be purchased, even through companies specializing in custom syntheses. However, in terms of experimental candidates, this was by far the most important molecule. Furthermore, based on its symmetry and physical properties (e.g. gas-phase at room temperature), once obtained it was anticipated to be quite convenient for handling and recording of spectra. Although unavailable, this molecule was not unknown. Kenneth Wiberg first reported its synthesis in the 1960's.[71] His approach was a complicated photochemical procedure which resulted in a number of products including BCP, though in very poor yield. This made isolation of a highly pure sample very difficult.

My own attempts to synthesize BCP included a few different methods before a suitable route was established. It should be emphasized that the reaction steps used in my procedures were not new, except for some unsuccessful attempts to produce BCP directly from a dilithio compound. Since the goal was solely to obtain a sample, it was simply a matter of seeking the most efficient literature preparation. Soon after some early failed attempts were completed, a review article was published detailing much of what is known about these systems including their synthesis.[72] A new approach made it quite straightforward to synthesize the related compound [1.1.1]-propellane.[73] This compound provides a number of routes to derivatives of BCP. Ironically, although a vast number of functionalized bicyclo-[1.1.1]-pentanes were then accessible, the parent

hydrocarbon was still the most difficult to obtain. This system also exhibits some counterintuitive behavior from a synthesis standpoint. If one simply adds iodine (I_2) to a solution of [1.1.1]-propellane in ether, then 1,3-di-iodo-BCP precipitates out as crystals.[74] However, a subsequent attempt to reduce this compound to BCP using tri-n-butyl-tin hydride would only reform the propellane. Ultimately the successful procedure required some eleven steps to produce a pure sample of BCP.

All efforts to synthesize BCP started with the synthesis of [1.1.1]-propellane (see scheme 1). All of the syntheses described here were performed under inert atmosphere conditions (under Argon). All glassware was oven baked overnight and solvents were freshly distilled in an effort to exclude water from interfering with the desired chemical processes. A solution of propellane was prepared as follows:[73]
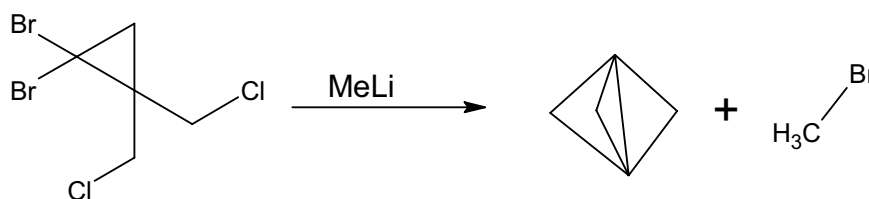
Scheme 1: Preparation of [1.1.1]-propellane

i)   prepared a solution of 1,1-di-bromo-2,2-di-chloromethyl-cyclopropane in dry ether

ii)   cooled solution to -42^0 C

iii)   added two equivalents of methyl lithium solution in ether, dropwise

iv)   allowed solution to warm to room temperature for 1 hour

v)   distilled into flask at -78^0 C, warmed to -15^0 C

vi)   added excess Mg powder, raised to room temperature

vii)   distilled into flask at -78^0 C

The first few times this solution was prepared, a small amount of benzene was added as an internal standard for the NMR spectra. This allowed the yield for a small-scale batch to be estimated at about 60%. Based on a method described by Uwe Bunz in his PhD thesis[75], it was decided to attempt a nearly direct approach to BCP through formation of 1,3-dilithio-BCP. The strategy was to react the propellane solution with lithium 4,4'-di-t-butylbiphenyl (LiDBB) in a solution of refluxing Me_2O at -23^0 C. The procedure reported by Bunz worked reasonably well, although the pink crystals corresponding to the desired dilithio compound proved to be extremely sensitive to any amount of air which might find its way into the system. Once the crystals were formed, the ether could be pumped off until the crystals were dry. Although any subsequent reactions performed by Bunz had involved solvating the crystals in benzene, it was hoped that if the dilithio compound could be dissolved in a high boiling solvent, and protonated with a high boiling alcohol, then BCP (b.p. 35^0 C) could easily be distilled off. With this in mind, the crystals were dissolved in decahydronaphthalene, which boils at about 190^0 C. Two equivalents of n-pentanol were added with stirring. Attempts to distill any BCP into a cold (-196^0 C) trap failed. The crystals did not seem to be very soluble in

134

decahydronaphthalene and no visible change occurred with addition of 1-pentanol.

Since the dilithio compound had formed very much as described by Bunz, it was decided to adopt his methods to at least see if BCP could be produced. The dilithio compound was dried under vacuum, and then solvated with clean degassed benzene. Two equivalents of water were added to protonate the crystals. The NMR spectra showed that a solution of BCP in benzene had indeed formed, although a considerable amount of ether was still present. Thus the desired BCP had formed but there was no way to separate it from the ether and benzene. This was encouraging enough to make another attempt with a high boiling solvent. The dilithio crystals were prepared again and solvated with xylene. Amyl alcohol was added for the protonation. Again no product could be distilled off.

Since BCP had been successfully synthesized but not isolated, it was decided to construct a preparatory scale gas-chromatography (GC) apparatus as Wiberg had used in his preparation. A 20' copper-tubing column with a 1 cm internal diameter was wound around a six inch mandrel. This was then packed with column material coated with Carbowax 20M. This was adapted to an analytical GC instrument on loan from Dr. H.D. Gesser. A mild flash-volatilization of 100^0 C was used with a column temperature of 40^0 C. Helium was used as the carrier gas, thus making it possible to use liquid nitrogen for the collection trap. Test solutions of pentane: ether: benzene were easily separated. The pentane and ether were well separated and eluted at about 3 and 4 minutes, respectively. These conditions caused a considerable delay (25 minutes) before

the benzene came through. It was thought that heating the column after passage of the ether would speed up the benzene elution, although this would necessitate additional cooling time before the next injection. Thus since good separation was achieved with the test mixture, it was decided to prepare a large sample of BCP and try to collect it. Several attempts to collect BCP failed despite clear evidence that it had been produced. Even at low temperature the BCP was too easily swept from the collection vessel by the carrier gas. Despite the use of some highly efficient collection devices specially constructed by Ian Ward (glass-blower, Dept. of Chemistry), no useable amount of BCP could be isolated. This was frustrating since the test mixtures with n-pentane in place of BCP could be separated trivially. There was also some evidence that the BCP was decomposing during the flash volatilization.

At this time it was realized that Dr. E.W. Della had recently prepared a pure sample of BCP for an NMR study.[76] His previous work with these systems had appeared many times in the literature.[77-79] Some e-mail correspondence with Dr. Della suggested that the best route to a pure sample was somewhat roundabout. He suggested a series of steps which should ultimately result in a pure sample of BCP. The strategy (shown in scheme 2) was to attach various bulky substituent groups to the bridgehead positions. This allowed recovery of a non-volatile stable solid at the end of each intermediate step. These intermediate compounds were known to be relatively insensitive to oxygen and water. Finally the last bulky group could be removed to produce BCP in the non-volatile tri-n-

butyl-tin hydride. Ideally at this stage the pure BCP could be swept into a cold collection flask.
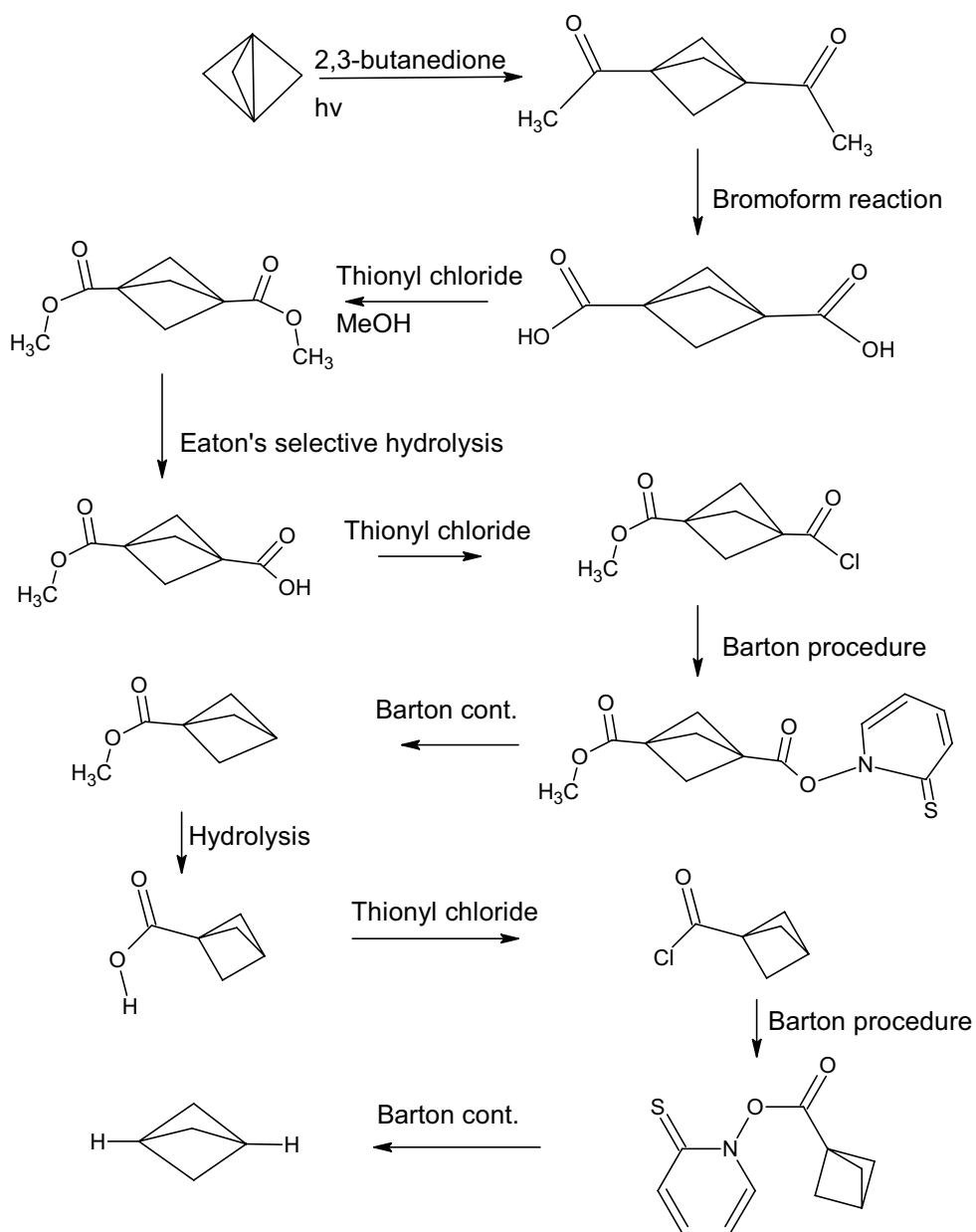
Scheme 2: Complete synthesis of BCP from propellane

In practice, most of the steps worked quite well as described in the literature. Yields comparable to those reported by the various authors[73, 77-82] were realized. The best part of this strategy was that material could be stored at any stage while another batch was brought along from the beginning. Thus while losses did eventually diminish the amounts carried through, more could soon be added. The only step to proceed more slowly and in lower yield than expected was the final one. Enough of the mono-acid was used to produce 3 ml of BCP assuming that the yield was similar to that in the literature. This was expected to occur within 30 minutes. In actuality, less than 1 ml of BCP was isolated, having slowly accumulated over 8 hours. This is not too surprising since the reaction was photochemical and required a bright lamp. In my procedure, a bright lamp (500 W) was placed next to the reaction flask. The literature procedure used a 300 W lamp, but may have used a special highly reflective reaction chamber. An additional problem arose from sweeping the reaction for so long. At the end of 8 hours a small but visible portion of the tin hydride had come over with the BCP. This was easily removed through repeated cryo-distillations. Thus at last a usable sample of bicyclo-[1.1.1]-pentane was obtained. Some additional details are found in the supplementary information section.

Absolute Intensity Raman Spectra of Bicyclo-[1.1.1]-pentane

Once a sample of pure bicyclo-[1.1.1]-pentane had been prepared, it remained to record the Raman spectra. The goal was to derive the experimental intensity parameters for comparison with those computed at various levels of theory. In particular the internal-coordinate intensity-parameter for stretching of the bridgehead C-H bonds was sought for benchmarking. As discussed previously, the survey calculations predicted this to be the largest parameter for any C-H bond.[65] Furthermore, due to the unusual charge density and consequent poor agreement between theoretical methods, this has become an important way of assessing some high-level theoretical methods.

Experimental Setup

Given the assumptions under which the semi-classical intensity expression was derived (see Introduction), it was important to adhere to the necessary requirements as much as possible. The differential scattering cross-section intensity expression is again shown here for convenience.

$$\frac{\partial \sigma}{\partial \Omega} = \frac{\pi^2}{90\varepsilon_0^2}(v_0 - v_i)^4 \left[1 - \exp\left(\frac{-hcv_i}{kT}\right)\right]^{-1} \left[45\left(\frac{\partial \bar{\alpha}}{\partial q_i}\right)^2 + 7\left(\frac{\partial \gamma}{\partial q_i}\right)^2\right] \qquad (1)$$

Foremost is the assumption of non-resonance. The frequency of the laser used in the scattering experiment must not approach resonance with excited electronic states of the molecule for this expression to hold. On the other hand, the observed intensity varies with the fourth-power of the laser frequency. Given

that Raman scattering is a very weak process, a higher frequency laser provides much more recorded intensity. The inelastically scattered Raman photons are approximately 10^8 fold less intense than the elastically scattered Rayleigh photons. This is illustrated in a perturbation theory treatment, wherein the Rayleigh scattering is seen to be a first-order process, while Raman scattering is a second-order process.[46] For saturated hydrocarbons, the preferred laser source is the green line of an Argon ion laser at 514.5 nm (19435 cm^{-1}). Choosing this frequency favors caution with respect to the electronic resonance condition. Some researchers have used another line of the Argon ion laser at 488 nm. This line is also far enough from resonance in most saturated hydrocarbons and will produce more scattering intensity on a watt per watt basis (about 25%). However, the actual output of the Argon ion laser is much less at this frequency (about 1ß), thus limiting any advantage.

When linearly polarized light is used as a scattering source, the above intensity expression can be split according to the polarization of the scattered light as follows:[63, 64]

$$\frac{\partial \sigma}{\partial \Omega} = \frac{\pi^2}{90\varepsilon_0^2}(v_0 - v_i)^4 \left[1 - \exp\left(\frac{-hcv_i}{kT}\right)\right]^{-1} \left[45\left(\frac{\partial \overline{\alpha}}{\partial q_i}\right)^2 + 7\left(\frac{\partial \gamma}{\partial q_i}\right)^2\right]$$

$$\frac{\partial \sigma_/}{\partial \Omega} = \frac{\pi^2}{90\varepsilon_0^2}(v_0 - v_i)^4 \left[1 - \exp\left(\frac{-hcv_i}{kT}\right)\right]^{-1} \left[45\left(\frac{\partial \overline{\alpha}}{\partial q_i}\right)^2 + 4\left(\frac{\partial \gamma}{\partial q_i}\right)^2\right] \qquad (4)$$

$$\frac{\partial \sigma_\perp}{\partial \Omega} = \frac{\pi^2}{90\varepsilon_0^2}(v_0 - v_i)^4 \left[1 - \exp\left(\frac{-hcv_i}{kT}\right)\right]^{-1} \left[3\left(\frac{\partial \gamma}{\partial q_i}\right)^2\right] \qquad (5)$$

Light scattered with polarization parallel to that of the incident photons contains contributions from both the isotropic and anisotropic parts of the polarizability tensor. The photons with perpendicular polarization contain only a contribution from the anisotropic term. This should be seen as very convenient when one considers that the desired intensity parameter is derived strictly from the isotropic term. By recording the scattered light with each type of polarization (using a filter), then subtracting 4β of the perpendicular contribution from the parallel, the remaining intensity, solely that of the isotropic trace scattering term, is obtained.

So far it has been decided to use the 514.5 nm line from an Argon ion laser and to record the scattering from both orthogonal polarizations. The next consideration is that of recording sufficient intensity and particularly signal-to-noise ratios. As mentioned previously, the Raman scattering process is very weak in general, and is particularly so from a dilute gas. This necessitates the use of a multi-pass setup,[83, 84] in which the laser is reflected back and forth many times through the sample. Use of a multi-pass setup does indeed greatly enhance the scattering process but it also introduces two other concerns. Firstly, the mirrors employed to create the multipass do not reflect the light perfectly and the precise number of passes is difficult to determine. This means that the intensity of the laser is slowly decaying with each pass and the total intensity at the focal point is not precisely known. Secondly, the light is reflected from a number of points along a curved surface. This means that the polarization direction of the light varies slightly through the different passes. Fortunately geometric considerations

have shown that the error introduced into the scattering analysis from the deviation in polarization direction is negligible.[85] The problem of the unknown intensity of the source is overcome through the use of internal standards. Typically researchers record gas-phase spectra with an internal standard such as nitrogen gas. This allows the recorded intensities to be calibrated with the well-known intensity of the nitrogen signal. This also lowers the demands on a detection device. A suitable detector need only provide accurate *relative* intensities, since adjustments can be made based on the nitrogen signal. Based on these considerations, the Gough lab's Raman instrument for recording gas-phase spectra consists of the following:

i)   Coherent Innova 200 λ5-Argon ion laser tuned to a single line at 514.5 nm

ii)   Multi-pass optical arrangement

iii)   Small Brewster-angled sample cell

iv)   Spex 14018 -Double-monochromator wavelength selector

v)   RCA-31034 Water-cooled, nitrogen-swept, photomultiplier detector

The double-monochromator photomultiplier detection system does not detect lower energy photons as efficiently as it does high-energy photons. Thus the intensities of recorded spectra must be corrected differently over the typical wavelength range. Proper calibration of the instrument was a project in itself for previous students (Hemant Srivastava and Jason Dwyer). Standard lamps with known emission characteristics (e.g. blackbody) were used for this. Happily these previous efforts have resulted in correction curves in the form of simple

polynomials which can be applied to spectra. These correction curves were generated separately for incident light of parallel and perpendicular polarization. Below are the actual correction curves reproduced from a MathCad worksheet.

Intensity Correction

Parallel

$x := 4000, 3999 ... 500$

$A := 0.9956735$

$B := 1.395 \cdot 10^{-4}$

$C := 2.042 \cdot 10^{-8}$

$D := 1.013 \cdot 10^{-11}$

$E := -2.737 \cdot 10^{-16}$

$parallel(x) := A + B \cdot x + C \cdot x^2 + D \cdot x^3 + E \cdot x^4$

Perpendicular

$A := 1.006045$

$B := 1.448 \cdot 10^{-4}$

$C := 5.942 \cdot 10^{-8}$

$D := -3.9 \cdot 10^{-12}$

$E := 1.563 \cdot 10^{-15}$

$perpendicular(x) := A + B \cdot x + C \cdot x^2 + D \cdot x^3 + E \cdot x^4$
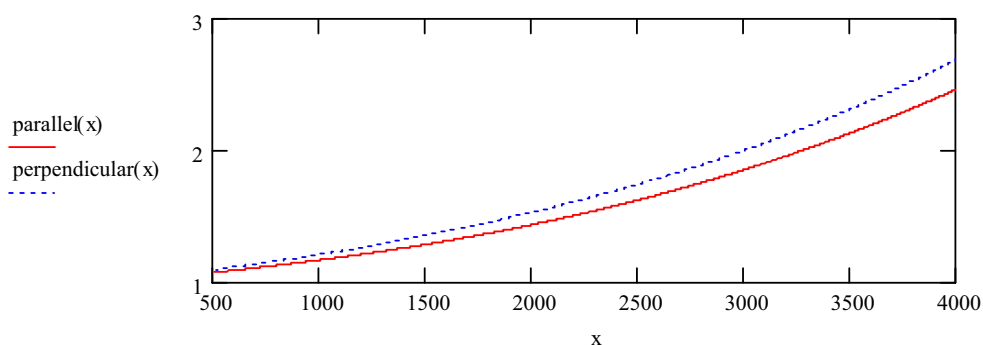
MathCad sheet 1: Intensity correction factors

Note that bands recorded at 3000 cm^{-1} with this instrument receive a considerably larger correction than those recorded at 1000 cm^{-1}. The

perpendicularly polarized spectrum receives a slightly larger correction throughout the wavelength range. Intensity parameters recorded with this instrument have been in excellent agreement with those recorded by other labs around the world. The spectrum is obtained by slowly scanning the wavelength range and recording the counts of the photomultiplier at each wavelength. This process is rather lengthy and must be repeated several times before sufficient intensity has been recorded. To allow the instrument to maintain consistency over this extended time, a dust shroud covers the multi-pass setup. Since a hole allows the laser to pass through the shroud into the multipass setup, a positive pressure of filtered air is maintained.

Figure 11: Raman multi-pass setup with dust shroud. The air filter above the shroud and some of the optics are visible. A lens focuses the laser source from below into the multi-pass mirrors.
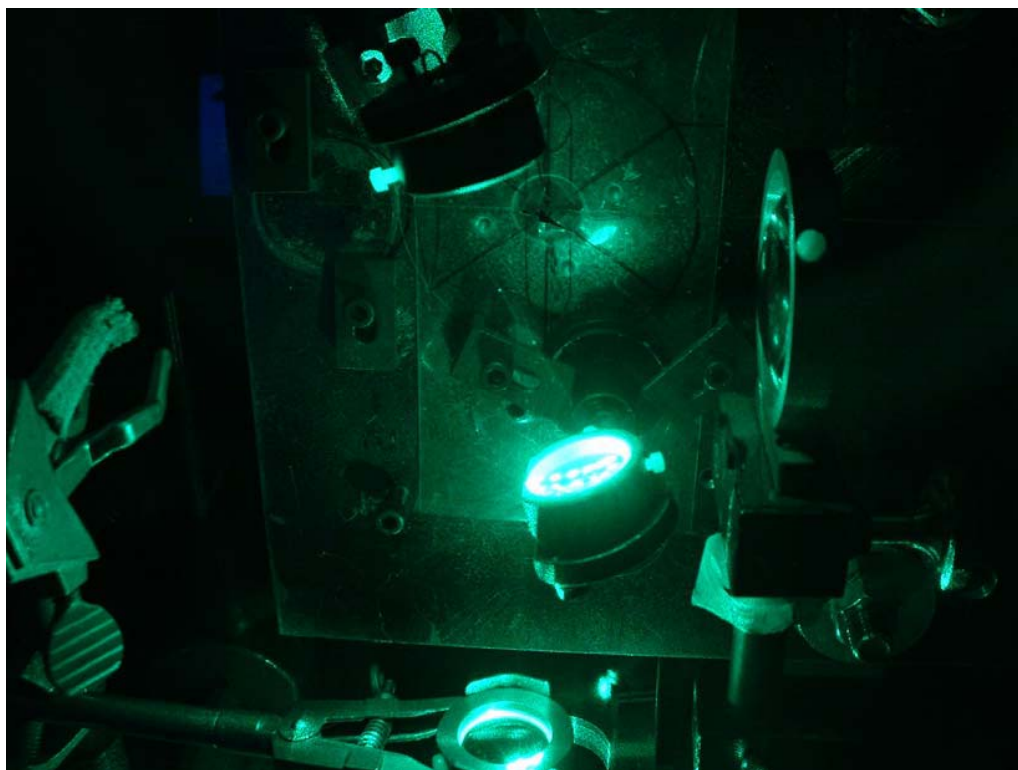
Figure 12: Multi-pass setup. Spots are visible on the lower mirror from the multiple passes of the laser. A large lens focuses light from the focal point of the multipass region into the double-monochromator (off to right).

When the instrument has first been assembled, the nitrogen in air provides a convenient test. Since our atmosphere is approximately 80% nitrogen, a strong signal from nitrogen can be recorded from the laser multipass operating without a sample cell. The system can be tested using both orientations of the polarization filter. This ensures that the setup is calibrated correctly and will provide data similar to those of other labs throughout the world. Furthermore, the optics can be optimized in terms of the recorded intensity. There is almost an art to aligning the tiny mirrors to create the best possible multipass. Minute adjustments are made until the number of counts per second from the photomultiplier is

145

maximized. This occurs when a large number of passes are achieved, aligned in one plane, and each passing through a focal point at the center. Next the focusing mirror which sends the scattered light into the double-monochromator is adjusted to capture the maximum amount of signal. Finally, the double-monochromator is calibrated using a Ne lamp. For quick confirmation, a line from the mercury atoms present in the fluorescent lights in the room can be observed, and also the signal from nitrogen will appear at the accepted 2330 cm^{-1}.

At this time the sample may be prepared in a sample cell. Since the molecules of interest are gasses at room temperature, this is done on a vacuum line. The Gough lab has a special set of high-vacuum lines and mixing flasks attached to a pressure gauge and vacuum pumps. This allows convenient handling of gas-phase samples. Multiple cycles of freeze-pump-thaw are employed to degas the sample. Then the sample gas and any nitrogen standard gas are allowed into a mixing flask. After the gasses have been adequately mixed, a precise pressure of the gas mixture is allowed into a small sample cell. The cell itself is designed to hold a vacuum for extended periods and to maximize transmission of laser light in the required frequency range. This means using high-quality Brewster-angled windows of a precise thickness. Next the Brewster-angled sample cell is inserted within the multipass optics. Ideally, the cell should be oriented such that it does not disturb the passes of the laser, which now must transmit through the cell with each pass. In practice, the presence of a cell does perturb the laser path somewhat, and can shift the exact position of the focal point slightly. This necessitates a re-optimization of the signal with the cell in place. If

a partial pressure of nitrogen is present in the cell, the signal from nitrogen at 2330 cm^{-1} may still be used;otherwise a str ong band from the sample gas must be located.

Using a pressure of 143 Torr of pure sample, a series of scans were recorded for both parallel and perpendicular polarization directions. Based on the aperture widths of the double-monochromator (180μ180μ), which determ ines the resolution, steps of 1 cm^{-1} were taken. This is approximately equal to the maximum resolution allowed by the double-monochromator. Enough scans were recorded to total 48 seconds of integrated intensity at each wavenumber point, for each polarization. This allowed a reasonable ratio of signal to noise to be achieved. Subtraction of $\frac{4}{3}$ tim es the perpendicularly polarized spectrum yields the isotropic spectrum necessary to obtain the desired intensity parameters.

In order to compute the absolute intensity of the recorded Raman spectrum, two mixtures were prepared using nitrogen as an internal standard. Mixtures of BCP:N$_2$ of (120.8 : 123.7), and (121.4 : 339.9) Torr were used. For each sample, the intensities of the signal for nitrogen and that of the strongest band in the bicyclo-[1.1.1]-pentane spectrum were recorded. The second mixture included sufficient nitrogen for the two peaks to be of comparable intensity. Thus once the absolute intensity of the strongest band was determined, the intensities of all other bands could be calculated relative to the strongest band from the pure-compound spectra.

Figure 13: Raman scattering spectrum: Bicyclo-[1.1.1]-pentane. Sample pressure 143 Torr, laser power 7.0 watts. Complete spectrum with parallel polarization.

All of the spectra shown in figures 13 through 22 have been intensity corrected using the correction curves shown above. Figure 13 shows the parallel polarization spectrum of bicyclo-[1.1.1]-pentane through the entire wavelength range. There are clearly a large number of features, many of them broad and overlapped. Fortunately, computation of the internal coordinate intensity parameters only requires integration of the trace (or isotropic) spectrum.[63, 64] As discussed above, this spectrum is obtained by subtracting 4/3 of the perpendicularly polarized spectrum (anisotropic) from that of the parallel. In the isotropic spectrum, only bands corresponding to totally symmetric (A_{1g}) states will have non-zero intensity. As will be shown in detail in the following section on computation of intensity parameters, bicyclo-[1.1.1]-pentane has 5 fundamental modes with this symmetry. Thus one might expect 5 narrow, well-resolved peaks in the trace spectrum, unless resonantly enhanced higher-order transitions are found to complicate matters. Figures 14 and 15 show the C-H stretching region of the spectrum for the two polarizations in more detail. Figure 18 shows the isotropic spectrum obtained by subtracting out the anisotropic contribution. While there are only two fundamental C-H stretching motions anticipated to occur in this region, figure 18 clearly shows 5 peaks. This is due to resonantly enhanced higher-order transitions (combinations and overtones) with the correct symmetry. Considerable computational effort was required to distribute the recorded intensity properly. The sections on higher-order force-field methods, resonance interactions, and inverse-eigenvalue algorithms provide the details of how this was done.

Figure 14: Raman scattering spectrum: Bicyclo-[1.1.1]-pentane. Sample pressure 143 Torr, laser power 7.0 watts. C-H Stretching region with parallel polarization.

Figure 15: Raman scattering spectrum: Bicyclo-[1.1.1]-pentane. Sample pressure 143 Torr, laser power 7.0 watts. C-H Stretching region with perpendicular polarization.

The next region of the spectrum to consider is that of the C-C stretch. Figures 16 and 17 illustrate the parallel and perpendicularly polarized spectra in the wavenumber range of 1200 to 800 cm^{-1}. The isotropic spectrum is shown in figure 19. Two fundamentals of A_1' symmetry are expected to appear in this region, and except for a very weak additional feature, this is what is recorded.

Finally there is the C-H bending region of the spectrum where the last of the five totally symmetric modes is found. Figure 20 shows both the parallel and isotropic spectra. The transition occurs at 1510 cm^{-1} but is extremely weak. This is because bending of the methylene hydrogens does little to increase the molecular volume and does not involve stretching any bonds. Thus only a very slight change in the polarizability is anticipated. Furthermore, while symmetry considerations allow non-zero intensity in the isotropic spectrum, this band has a large depolarization ratio (0.73) and therefore most of the intensity is from the anisotropic term. After subtracting out the anisotropic contribution, the remaining isotropic part is not resolved in terms of signal to noise for a meaningful integration. This does not present problems for the analysis since, if this peak has less than 2000 integrated counts and up to 100% error, the maximum error for this peak is 2000 counts. This much error is much less than 1% of the counts for the most intense C-H stretching bands.
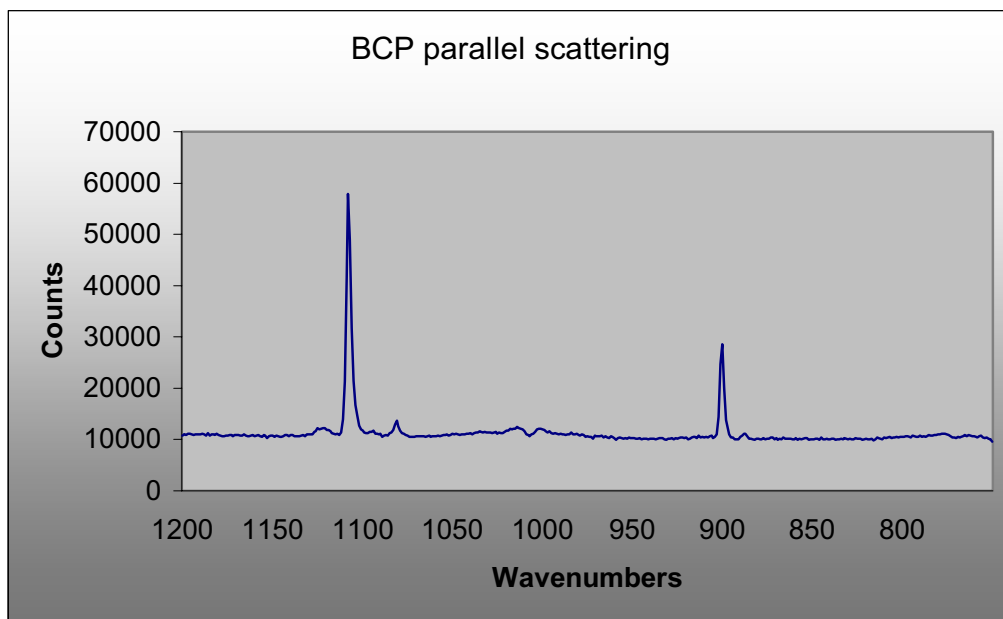
Figure 16: Raman scattering spectrum: Bicyclo-[1.1.1]-pentane. Sample pressure 143 Torr, laser power 7.0 watts. C-C Stretch/Bend region for parallel polarization.
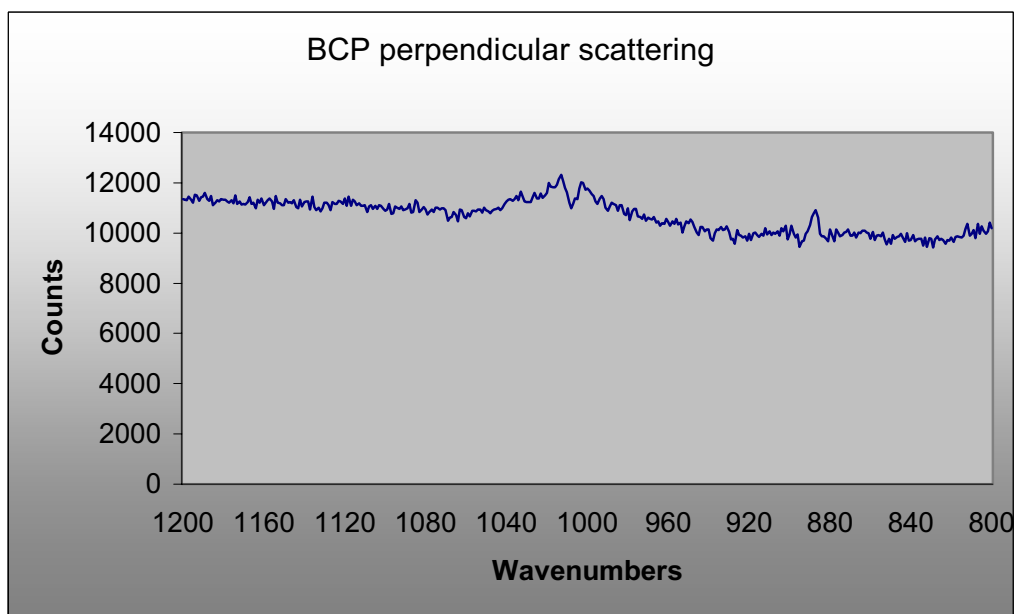
Figure 17: Raman scattering spectrum: Bicyclo-[1.1.1]-pentane. Sample pressure 143 Torr, laser power 7.0 watts. C-C Stretch/Bend region with perpendicular polarization.
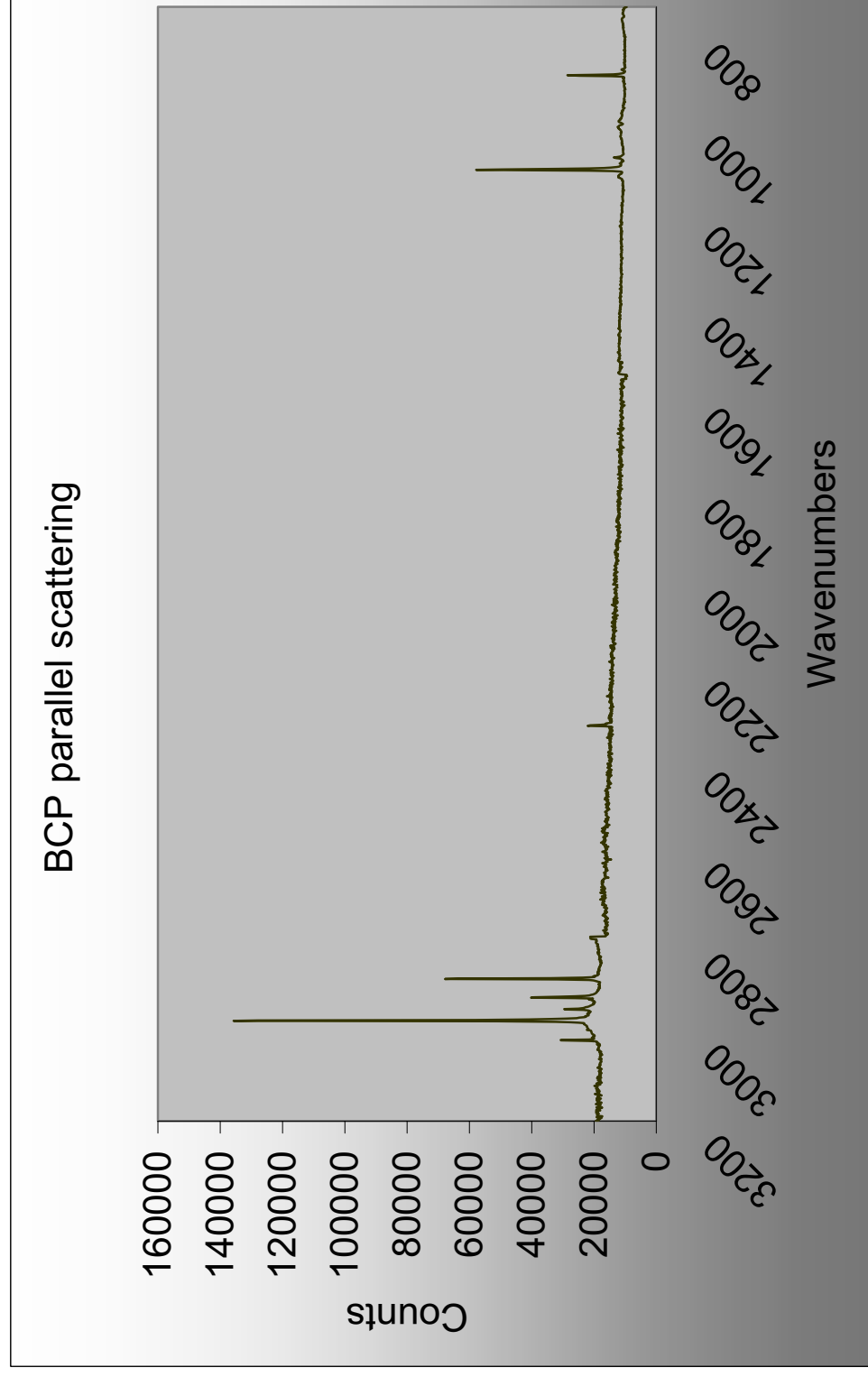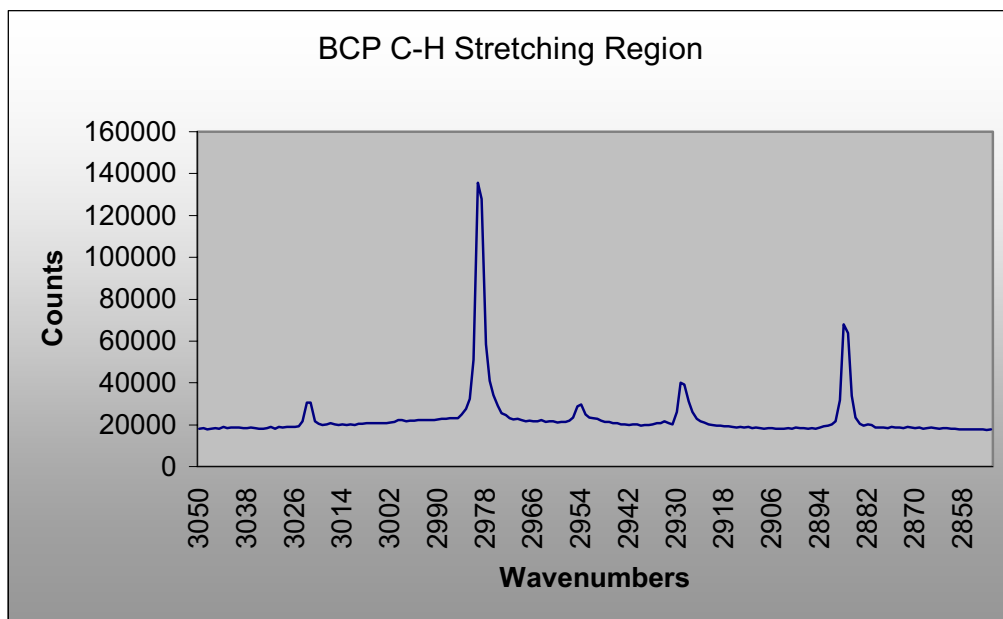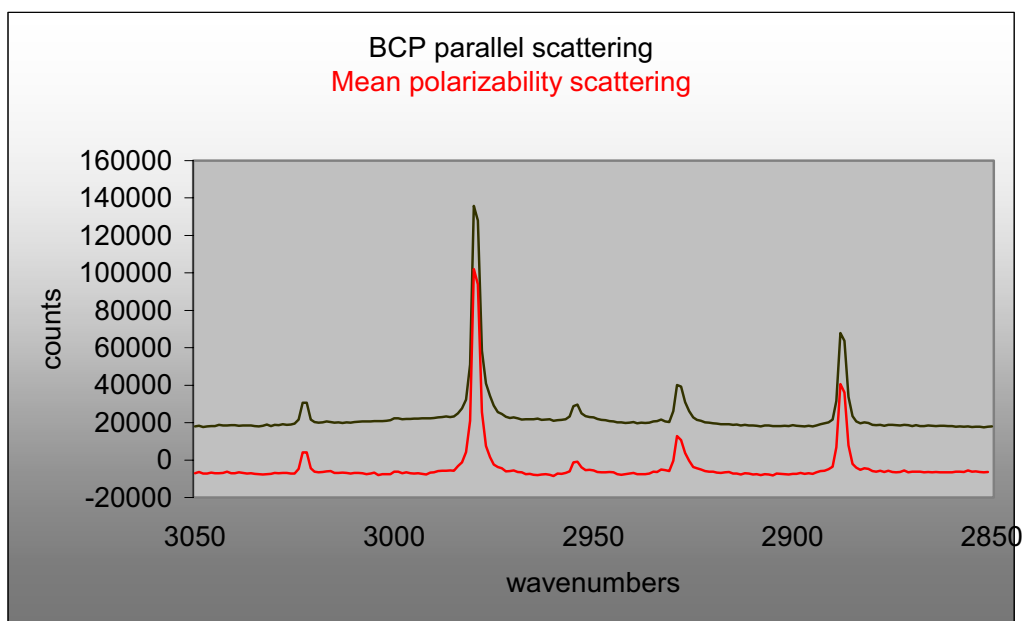
Figure 18: Raman scattering spectrum: Bicyclo-[1.1.1]-pentane. Sample pressure 143 Torr, laser power 7.0 watts. C-H Stretching region with parallel polarization and with 4/3 perpendicular subtracted (mean-polarizability spectrum, shown in red)
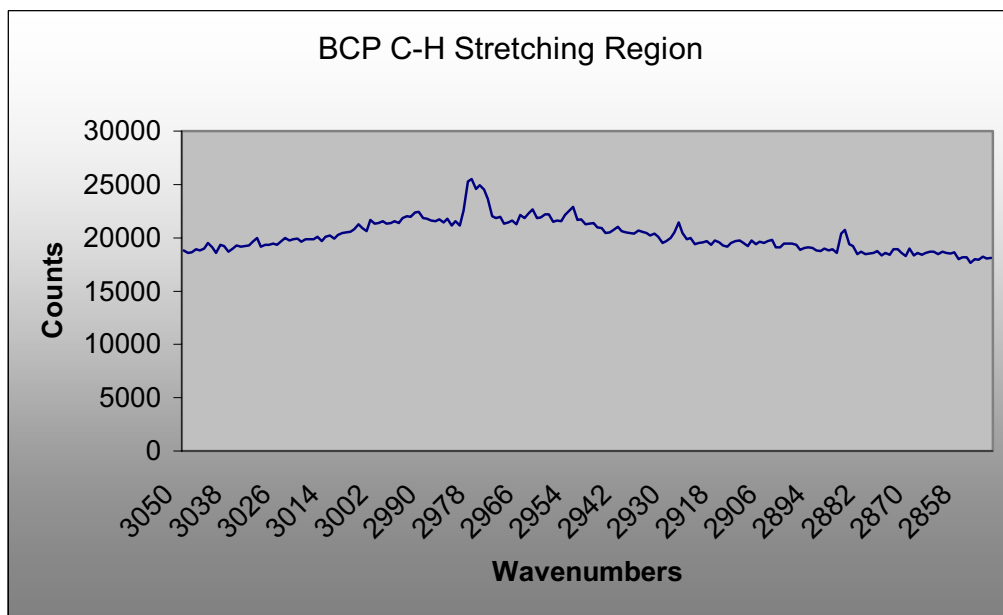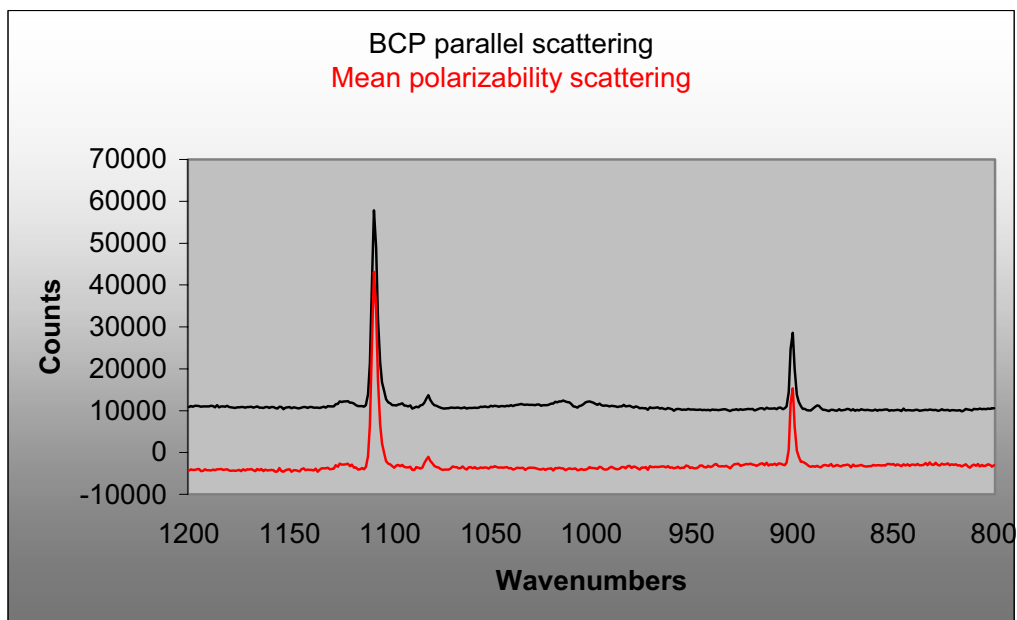

Figure 19: Raman scattering spectrum: Bicyclo-[1.1.1]-pentane. Sample pressure 143 Torr, laser power 7.0 watts. C-C Stretch/Bend region with parallel polarization and with 4/3 perpendicular subtracted (mean-polarizability spectrum, shown in red)
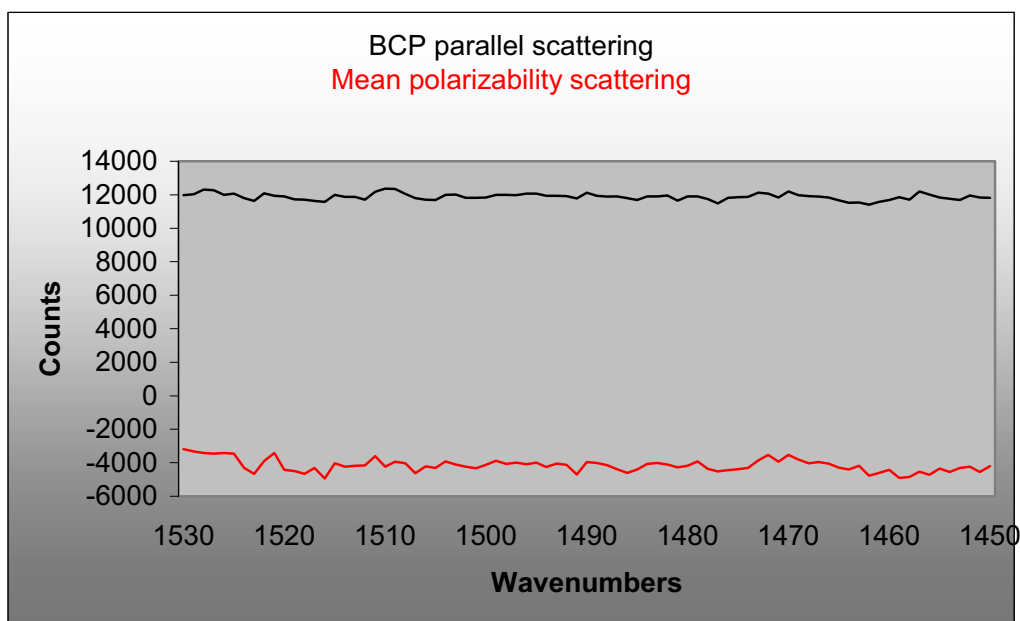
Figure 20: Raman scattering spectrum: Bicyclo-[1.1.1]-pentane. Sample pressure 143 Torr, laser power 7.0 watts. Parallel and mean polarizability spectra. Extremely weak band at 1510 cm^{-1} in parallel spectrum is almost unresolved in mean polarizability spectrum (see discussion).

Figures 21 and 22 show the spectra required for the calculation of absolute intensities. Figure 21 shows the most intense peak of the bicyclo-[1.1.1]-pentane spectrum, while figure 22 shows the signal from nitrogen. It becomes evident just how intense the peak for the bridgehead stretch is when one considers that there is a 2.800 fold excess of N_2. In the following section the data corresponding to these spectra will be used to derive the experimental intensity parameters.

Since the intensity analysis requires the vibrational frequencies and normal mode vectors, the complete set of experimental and calculated frequencies is listed in Table 7. For reasons that are explained in Appendix C, the force-field employed to derive the normal coordinates is the non-scaled quadratic field computed using B3LYP/aug-cc-pVTZ.
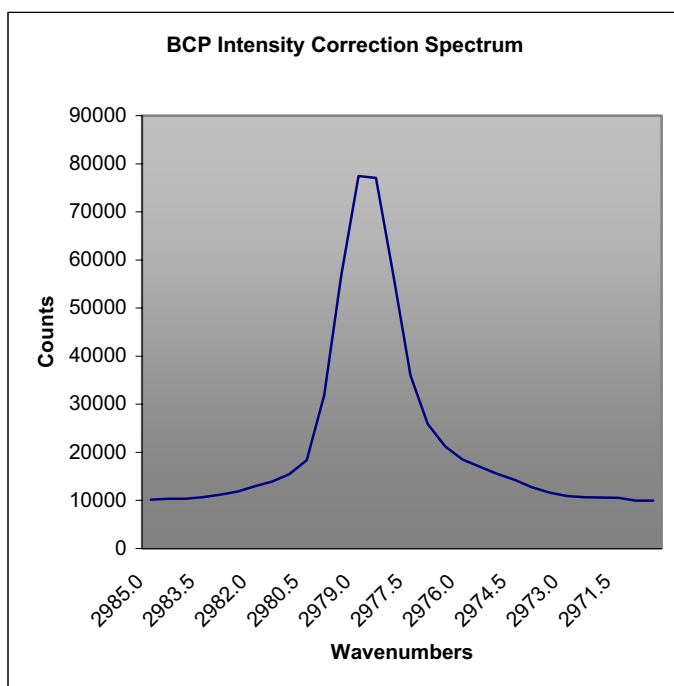
Figure 21: Absolute intensity correction spectrum: BCP:N$_2$ 121.4 : 339.9 Torr. Laser power 7.0 watts, parallel polarization. Band corresponds to bridgehead stretch in BCP
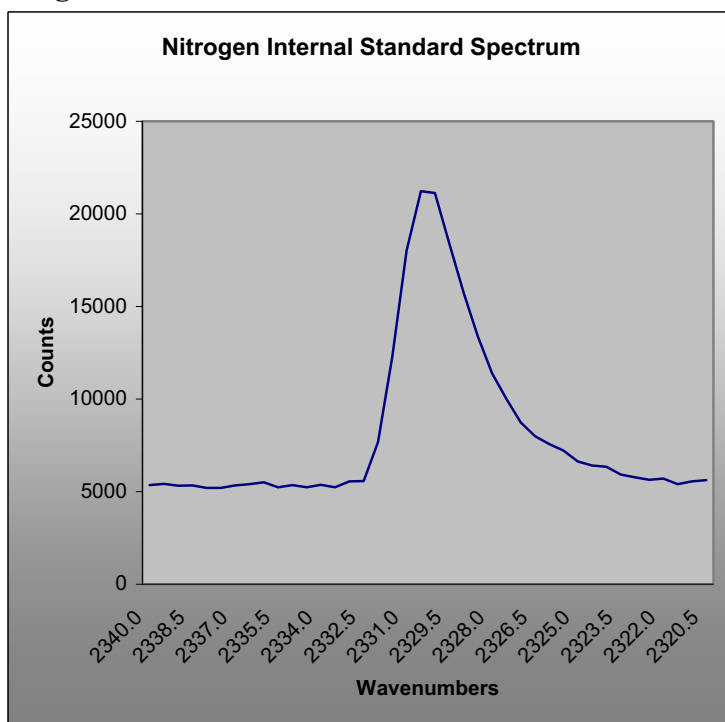
Figure 22: Absolute intensity correction spectrum: BCP:N$_2$ 121.4 : 339.9 Torr. Laser power 7.0 watts, parallel polarization. Band corresponds to signal from nitrogen standard.

A_1'	B3LYP/aug-cc-pVTZ	EXPT
1	3089.6	2982[a]
2	3041.3	2909[a]
3	1553.0	1509
4	1118.3	1108
5	906.0	900
A_2'		
6	3093.9	i.a.
7	968.2	i.a.
E'		
8	3098.7	2976
9	3035.9	2887
10	1503.2	1456
11	1256.7	1232
12	1115.7	1098
13	897.1	886
14	540.9	540
A_1''		
15	1003.7	i.a.
A_2''		
16	3084.1	2976
17	1250.7	1220
18	843.7	832
E''		
19	1214.8	n.o.
20	1142.7	n.o.
21	1028.4	1006
22	773.3	770

Table 7: Calculated and observed frequencies (cm^{-1}) for BCP. Mode numbering and IR transitions correspond to those reported by Wiberg et al.[86]
[a] deperturbed frequency (see appendices F and G)
i.a.: Raman and IR inactive
n.o.: not observed

The vibrational frequencies are listed according to their symmetry blocks.

Modes 1-5 are the totally symmetric Raman active modes whose intensities were

recorded in the above isotropic spectra. The E' and E'' doubly degenerate modes

are considered to be symmetric with respect to the x-z plane when the molecule is

oriented with its principal symmetry axis along the z-axis, and one of the

methylene carbons extended out along the x-axis. For the force-field analyses

(see Appendices) modes 23-29 are the asymmetric degenerate partners to the E'

modes 8-14. Similarly modes 30-33 are the asymmetric degenerate partners to

the E" modes 19-22.

Determination of Intensity Parameters

Once the spectra have been recorded, intensity corrected, and integrated, the absolute intensity of the strongest band is determined by reference to the peak corresponding to the nitrogen internal standard. This is done using the isotropic mean polarizability spectra since these are the type of intensity parameters sought. When the absolute intensity of the strongest band has been established, the other peaks become known relative to it.

The trace scattering spectrum is composed only of contributions arising from totally symmetric modes (A_1'). In a best-case scenario, the peaks will be well separated and free from anharmonic resonant interactions and the integrated intensities can be assigned to the corresponding dimensionless normal modes. A sign ambiguity exists, since these intensities are related to the squares of the intensity parameters. Previously, the ambiguity has been resolved through numerical optimization of intensity data over a set of molecular isotopomers, through which the character of the normal modes varies.[63, 64] It has been found that polarizability calculations with at least a moderate level of theory will always predict signs for the intensity parameters in agreement with these optimizations. If other isotopic data are lacking, one may be guided by theory. Because the isotopic sum rules are poorly obeyed for molecules with large anharmonicities, it may be preferable not to use isotopomers but rather an anharmonic force-field. Using the best available force-field, the L-matrix is used to transform these normal mode intensities into parameters with respect to internal coordinates. Thus in this best case (no anharmonic resonances), using the L-matrix and the

signs determined by theory, it is only necessary to solve a small, linear system of equations to obtain the desired internal coordinate intensity parameters.[63, 64]

The next level of complexity is encountered when the spectrum is disturbed by anharmonic resonances. When there is only one, totally symmetric fundamental mode in each region of the spectrum, the treatment of splitting due to resonances is still relatively straightforward. This is true for ethane in which the C-H stretching mode (A_{1g}) is split into four bands.[63] The first level of approximation is to assign any trace scattering in the region to the fundamental. This approximation neglects second derivatives of the polarizability that would provide inherent intensity to the second-order transitions involved in the resonance. Theoretical papers describing calculated Raman intensities often compare the results with experimental numbers obtained in this way without comment.[87] Montero has published formulae derived from perturbation theory that may be used to correct intensities for one fundamental mode involved in weak or strong resonances. These formulae require a cubic force field and a quadratic polarizability field.[62, 88, 89] Often, as with ethane, the inherent intensity of higher-order transitions is found to be small, indicating that much of the recorded intensity should be attributed to the fundamental. In either method, the L-matrix transformation to internal coordinates is performed as before.

More complex resonance interactions require the complete diagonalization of a matrix describing the coupling. This will be described in detail in Appendix D. In any case the goal is to distribute the recorded intensity into the fundamental modes. Assuming that this has been done (however simple or complicated it may

159

have been) the next step is to transform the normal mode intensity parameters into internal coordinates. MathCad Sheet 2, below, shows the calculation of normal mode intensity parameters using the nitrogen internal standard.

$$i := 0, 1 .. 5 \qquad T := 298 \cdot K$$

$$c := 299792458 \frac{m}{sec}$$

$$h := 6.6260755 \, 10^{-34} \cdot joule \cdot sec$$

$$\varepsilon_0 := 8.854187817 \, 10^{-12} \cdot \frac{C}{m \cdot V}$$

$$k_b := 1.380658 \, 10^{-23} \cdot \frac{joule}{K}$$

$$v0 := 19435 \cdot cm^{-1}$$

$$v := \begin{pmatrix} 2979.6 \\ 2887.6 \\ 1511.0 \\ 1107.8 \\ 900.3 \\ 2331.0 \end{pmatrix} \cdot cm^{-1}$$

$$crosssection_i := \frac{\pi^2}{2 \cdot \varepsilon_0^2} \cdot \left(v0 - v_i\right)^4 \cdot \left[1 - e^{\frac{\left(-h \cdot c \cdot v_i\right)}{k_b \cdot T}} \right]^{-1} \cdot 1 \cdot \frac{C^2}{V^2} \cdot m^4$$

$$40.8 \cdot 10^{-36} \cdot m^2 = \frac{\pi^2}{2 \cdot \varepsilon_0^2} \cdot \left(v0 - v_5\right)^4 \cdot \left[1 - e^{\frac{\left(-h \cdot c \cdot v_5\right)}{k_b \cdot T}} \right]^{-1} \cdot X^2 \cdot \left(\frac{C}{V} \cdot m^2\right)^2$$

$$\frac{1.992776 \times 10^5}{312000.15} \cdot 8.872 = 5.666634671$$

$$5.666634671 \cdot 40.8 \cdot 10^{-36} \cdot m^2 = \frac{\pi^2}{2 \cdot \varepsilon_0^2} \cdot \left(v0 - v_3\right)^4 \cdot \left[1 - e^{\frac{\left(-h \cdot c \cdot v_3\right)}{k_b \cdot T}} \right]^{-1} \cdot X^2 \cdot \left(\frac{C}{V} \cdot m^2\right)^2$$

$$\begin{pmatrix} 1.800028158 \times 10^{-41} \\ -1.800028158 \times 10^{-41} \end{pmatrix} \quad \frac{C \cdot m^2}{V}$$

Solutions for X (note sign ambiguity)

MathCad Sheet 2: Normal mode intensity parameters using nitrogen standard

160

Following through the calculation in Sheet 2, first the physical constants are defined. Next the temperature, laser frequency, and the various transition frequencies are specified. The expression for the cross-section is then listed. This is followed by the same expression explicitly containing the accepted absolute isotropic cross-section for nitrogen ($40.8 \cdot 10^{-36}$ m^2).[63, 64] A variable X^2 has been inserted. The cross-section used in this equation is that for nitrogen multiplied by the relative intensity of the band in question to that of nitrogen. This is calculated from the nitrogen reference spectrum where it was found that the largest band for BCP at 312000 counts was 8.872 times as strong as that of nitrogen (accounting for the differing partial pressures). The solution for X corresponds to the intensity parameter. In this example the intensity parameter for mode 4 (1107.8 cm^{-1}) is computed. The index of the vectors in Sheet 2 begins at 0. Thus v_3 does correspond to 1107.8 cm^{-1}. This is where the sign ambiguity enters. In accordance with the Raman intensity calculation the positive value was retained.

A discrepancy with the magnitude of the solution is immediately apparent since the experimental values discussed previously with regards to the survey calculations are of the order of 10^{-30} Cm/V. The result of $1.800 \cdot 10^{-41}$ Cm^2/V can easily be converted into a result for a normal mode with dimensions of length. This transformation is detailed in Appendix C and shown below in Sheet 3. In this transformation the usual conversion factor for dimensionless normal modes to mass-weighted normal modes is used as in Appendix C. The result is then compensated by the reduced mass to produce normal mode coordinates with units of length only.

$i := 0, 1 .. 4$

$AMU := 1.6605402 \cdot 10^{-27} \cdot kg$

$$redmass := \begin{pmatrix} 1.09287546 \\ 1.05320416 \\ 1.14651462 \\ 4.49001028 \\ 5.43676289 \end{pmatrix} \cdot AMU \qquad force := \begin{pmatrix} 6.1466 \\ 5.7397 \\ 1.6293 \\ 3.3084 \\ 2.6292 \end{pmatrix}$$

$$scale_i := \sqrt{\frac{h}{4 \cdot \pi^2 \cdot c \cdot \omega_i}} \cdot \left(\sqrt{redmass_i}\right)^{-1} \qquad \frac{1}{2 \cdot \pi \cdot c} \cdot \sqrt{\frac{force_i \cdot 10^{-3} \cdot dyne \cdot \left(10^{-10} \cdot m\right)^{-1}}{redmass_i}}$$

$$scale = \begin{pmatrix} 0.099924963 \\ 0.1025946 \\ 0.13760479 \\ 0.081942366 \\ 0.082734409 \end{pmatrix} 10^{-10} \cdot m$$

3089.6328
3041.3265
1553.0501
1118.3045
905.9753

cm^{-1}

Re-calc freqs to check that units are consistant

$$\frac{1.800028158 \cdot 10^{-41} \cdot \frac{C \cdot m^2}{V}}{.081942366 \cdot 10^{-10} \cdot m} = 2.196700249 \times 10^{-30} \frac{C \cdot m}{V}$$

Convert so that parameter is with respect to coordinate in units of length.

MathCad Sheet 3: Normal mode intensity parameters conversion to normal mode coordinates with units of length.

Thus the intensity parameter for normal mode 4 is obtained in the familiar SI units of x10^{-30} Cm/V. The relationship between the forces in millidynes, the reduced masses, and the vibrational frequencies is also shown. The normal mode intensity parameters for the other modes are obtained in precisely the same way.

The next step is to define five symmetry coordinates and their associated internal coordinates. This will include construction of an L-matrix. This will be used for the subsequent transformation into internal coordinate intensity

162

parameters, for the final comparison with those computed at various levels of theory.
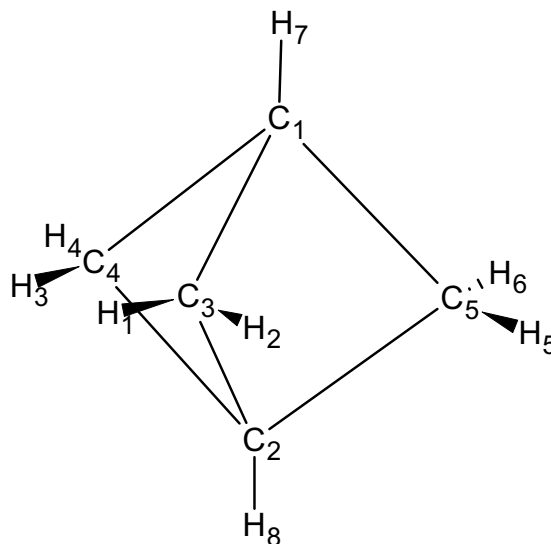
Figure 23: BCP with atom numbering for description of symmetry coordinates

Symmetry Coordinates

The structure of BCP shown in Figure 23 allows five symmetry coordinates to be defined.

$S_1 = 2^{-1/2} \cdot [rC_1H_7 + rC_2H_8]$

$S_2 = 6^{-1/2} \cdot [rC_3H_1 + rC_3H_2 + rC_4H_3 + rC_4H_4 + rC_5H_5 + rC_5H_6]$

$S_3 = 6^{-1/2} \cdot [rC1C3 + rC_2C_3 + rC_1C_4 + rC_2C_4 + rC_1C_5 + rC_2C_5]$

$S_4 = rC_1C_2$

$S_5 = 3^{-1/2} \{ [(rC_3H_1 \cdot rC_3H_2)^{1/2}] \cdot \Delta\alpha H_1C_3H_2 + [(rC_4H_3 \cdot rC_4H_4)^{1/2}] \cdot \Delta\alpha H_3C_4H_4 +$

$[(rC_5H_5 \cdot rC_5H_6)^{1/2}] \cdot \Delta\alpha H_5C_5H_6 \}$

Just as the five totally symmetric normal modes are numbered as 1 through 5 in descending order of frequency, the symmetry coordinates are also numbered 1 through 5. Although normal modes 1 and 2 will be shown to include large components of symmetry coordinates 1 and 2, no particular relationship should be assumed. The other normal modes require significant contributions from all of the symmetry coordinates and thus will not be associated with specific symmetry coordinates. Symmetry coordinates 1 and 2 are simply the bridgehead and methylene C-H stretching motions, respectively. Coordinate 3 is the C-C stretching motion. It is important to notice that the bridgehead-carbon-to-bridgehead-carbon distance has its own coordinate (number 4). Thus coordinate 3 is effected by "breathing" out the methylene groups. This motion extends the C-C bonds while leaving coordinate 4 unchanged. Similarly when coordinate 4 is effected, the methylene groups must "breathe in" to maintain the C-C distances for coordinate 3. Coordinate 5 is much simpler than it appears. It is really just the change in methylene H-C-H angle, weighted by the methylene C-H bond length to provide dimensions which include length. The conventional approach is to use the root of the product of the two lengths. In this case the two lengths are the same anyway. This coordinate is relatively unimportant since it is not one of the desired intensity parameters. Clearly these coordinate choices are not unique but they do span the D_{3h} configuration-space as required.

Once the coordinates have been defined, the Cartesian displacement normal mode vectors for the five totally symmetric coordinates are analyzed to

determine their composition in terms of the symmetry coordinates. A unit normal-mode displacement corresponds to 1 angstrom. The contribution from symmetry coordinate 5 must be established from an infinitesimal motion (actually 10^{-8} angstroms) since this coordinate couples to coordinate 2, due to the length weighting.

$$Q1 := \begin{pmatrix} -1.03128745 \\ -0.32879343 \\ 0.11929042 \\ 0.1174605 \\ 2.35937802 \end{pmatrix} \quad Q2 := \begin{pmatrix} 0.32763122 \\ -0.99688417 \\ -0.03124262 \\ 0.04578931 \\ 3.74528831 \end{pmatrix} \quad Q3 := \begin{pmatrix} -0.01549893 \\ 0.24280639 \\ -0.06114833 \\ -0.16198168 \\ -90.05528586 \end{pmatrix}$$

$$Q4 := \begin{pmatrix} 0.0300736 \\ 0.06011383 \\ 0.73813337 \\ 0.81495723 \\ -26.97543206 \end{pmatrix} \quad Q5 := \begin{pmatrix} -0.01856254 \\ 0.01505446 \\ 0.39094933 \\ -0.24740131 \\ -7.15601052 \end{pmatrix}$$

$$L := \begin{pmatrix} -1.03128745 & -0.32879343 & 0.11929042 & 0.1174605 & 2.35937802 \\ 0.32763122 & -0.99688417 & -0.03124262 & 0.04578931 & 3.74528831 \\ -0.01549893 & 0.24280639 & -0.06114833 & -0.16198168 & -90.05528586 \\ 0.0300736 & 0.06011383 & 0.73813337 & 0.81495723 & -26.97543206 \\ -0.01856254 & 0.01505446 & 0.39094933 & -0.24740131 & -7.15601052 \end{pmatrix}$$

MathCad Sheet 4: Normal modes 1-5 in basis of Symmetry coordinates

As seen in Sheet 4, Q1 through Q5 indicate the composition of the normal modes in terms of the symmetry coordinates. Q1 can be seen to contain a large contribution from symmetry coordinate 1 (bridgehead C-H stretch). The L-matrix for transformation of normal mode intensity parameters into symmetry coordinate

parameters is also shown. This is the transpose of the set of vectors Q1 through Q5.

Sheet 5 (shown below) shows the result of the transformation into symmetry coordinate intensity parameters.

$$
\text{dalphadq} := \begin{pmatrix} -2.956235505 \cdot 10^{-41} \\ -2.205770566 \cdot 10^{-41} \\ -1.809975391 \cdot 10^{-42} \\ 1.800028158 \cdot 10^{-41} \\ -9.412865425 \cdot 10^{-42} \end{pmatrix} \cdot \frac{C \cdot m^2}{V}
\qquad
\text{scale} := \begin{pmatrix} 0.099923841 \\ 0.102594803 \\ 0.137605667 \\ 0.08194246 \\ 0.082734127 \end{pmatrix} \cdot 10^{-10} \cdot m
$$

$$
\text{dalphadqlength}_i := \frac{\text{dalphadq}_i}{\text{scale}_i}
$$

$$
\text{dalphadqlength} = \begin{pmatrix} -2.95848866 \\ -2.14998275 \\ -0.13153349 \\ 2.19669773 \\ -1.13772463 \end{pmatrix} \frac{C \cdot m}{V} \cdot 10^{-30}
$$

$$
L := \begin{pmatrix} -1.03128745 & -0.32879343 & 0.11929042 & 0.1174605 & 2.35937802 \\ 0.32763122 & -0.99688417 & -0.03124262 & 0.04578931 & 3.74528831 \\ -0.01549893 & 0.24280639 & -0.06114833 & -0.16198168 & -90.05528586 \\ 0.0300736 & 0.06011383 & 0.73813337 & 0.81495723 & -26.97543206 \\ -0.01856254 & 0.01505446 & 0.39094933 & -0.24740131 & -7.15601052 \end{pmatrix}
$$

$$
\text{lsolve}(L, \text{dalphadqlength}) = \begin{pmatrix} 2.179134 \\ 3.062118 \\ -0.797032 \\ 3.245013 \\ 4.046021 \times 10^{-3} \end{pmatrix} \frac{C \cdot m}{V} \cdot 10^{-30}
$$

MathCad Sheet 5: Calculation of symmetry coordinate intensity parameters using L-matrix.

166

Once the symmetry coordinate parameters have been obtained as in Sheet 5, the final conversion into internal coordinate parameters is performed.

$$\frac{2.179133582}{\sqrt{2}} = 1.54088$$

$$\frac{3.062117763}{\sqrt{6}} = 1.2501$$

$$-0.797031928 = -0.79703$$

Internal coordinate intensity parameters
units of Cm/V x 10^{-30}

$$\frac{3.245012904}{\sqrt{6}} = 1.32477$$

$$\frac{4.046020608 \cdot 10^{-3}}{\sqrt{3}} = 2.33597 \times 10^{-3}$$

MathCad Sheet 6: Internal coordinate intensity parameters for BCP

The intensity parameters computed in Sheet six correspond to the desired internal coordinates that make up symmetry coordinates 1 through 5. Values of 1.541 and 1.250 (x10^{-30} Cm/V) were obtained for the bridgehead and methylene C-H stretches, respectively.

Results and Conclusions

In this research, the trends in polarizability derivatives and the consequent Raman scattering intensities have been studied for saturated hydrocarbons. A number of structural factors were identified through a large set of survey calculations and the predicted trends were discussed in detail in the sections entitled Survey calculations, and Survey calculations: Results and discussion.

Bicyclo-[1.1.1]-pentane was found to possess a number of structural features and physical properties which together make it the most interesting molecule of this group. The two bridgehead C-H bonds, aligned perfectly with the principal molecular axis, were predicted to have the largest polarizability derivative of any hydrocarbon yet studied.[65] The other C-H bond type was predicted to be quite ordinary in its derivative, thus providing a strong contrast. The non-bonded bridgehead C-C distance is the shortest known for an equilibrium structure.[72] The consequent electronic interactions across the cage are very strong and have been noted in NMR spectra.[90-92] This unusual structural arrangement makes theoretical descriptions of the charge density very difficult. Electron correlation effects take on an increased importance in this system. The electronic wavefunction exhibits instabilities with respect to certain coordinate displacements. As a result this molecule offers excellent potential for insight into the performance of various theoretical methods.

Due to a lack of accurate experimental Raman intensity data, the development of new theoretical methods has lacked confirmation. While theoretical methodologies continue to advance with regard to calculation of

Raman intensities,[87] reliable experimental results are often lacking, or are accompanied by large uncertainties. In other cases, only modest corrections have been applied to problems with anharmonic resonances. Results are reported in various formats and force-field transformations are required in order to make useful comparisons. For example, theoretical calculations that included the influence of dynamic polarizabilities on Raman intensities[87] could not be verified due to uncertainties in the published experimental data. From the data listed in that paper, it is not clear that the dynamic polarizability correction is improving agreement with experiment; rather, the ag reement appears to be worsened for some compounds. The Gough lab's most accurate results for ethane have been transformed into the format and units reported by Van Caillie and Amos.[87] As can be seen in Table 8, any of the better functionals, along with the dynamic correction, produce an excellent agreement with experiment. These results provide an important affirmation of work in this direction.

C_2H_6	Static Raman scattering activity[a]				Dynamic Raman scattering activity[a]				Expt[b]	Expt[c]
	SCF	B3LYP	PBEO	B97-1	SCF	B3LYP	PBEO	B97-1	514.5 nm	514.5 nm
ν_1	354.11	372.17	368.02	370.21	392.06	420.27	413.81	417.06	360.1	410.37
ν_2	0.34	0.02	0.02	0.02	0.48	0.04	0.02	0.04	2.2	0.0
ν_3	15.33	11.78	10.76	11.78	10.95	13.09	11.83	13.09	16.9	14.55

Table 8: Raman scattering activity of ethane ($\text{Å}^4 \text{ amu}^{-1}$)
[a] From [87]
[b] From [93]
[c] From [63], and [Murphy, private communication]

However, the experimental data for bicyclo-[1.1.1]-pentane indicate that the dynamic polarizability corrections (Table 9) are in the wrong direction for this molecule.

C_5H_8	SCF[a]	B3LYP[b]	B3LYP (TD)[c]	CCSD(T)[d]	EXPT[e]
Bridgehead C-H	1.386	1.698	1.826	1.593	1.541
Methylene C-H	1.124	1.347	1.436	1.282	1.250
C-C	1.218	1.104	1.170	---	1.325

Table 9: Raman trace scattering parameters for bicyclo-[1.1.1]-pentane ($\partial\overline{\alpha}/\partial r \times 10^{-30}$ Cm/V), from theory and experiment
[a] HF/D95(d,p), reference 65.
[b] B3LYP/aug-cc-pVTZ, (CADPAC6) [94]
[c] TDDFT (CADPAC6) [94]
[d] CCSD(T)/aug-cc-pVTZ, calculated for this work
[e] Experimental values, this work

This does not vilify the TDDFT method, since the static calculation at the same level of theory (B3LYP/aug-cc-pVTZ) shows systematic error. With this particular molecule and the B3LYP/aug-cc-pVTZ method, the C-H derivatives are overestimated and the C-C derivatives are underestimated while the opposite problem occurs with the HF/D95(d,p) method and the simple hydrocarbons. Perhaps if a dynamic correction were applied to a high-quality correlated calculation such as the coupled-cluster method, it would perform well in this case also. However, the static calculations at the CCSD(T) level are already just slightly above the experimental values and the dynamic correction usually increases the calculated derivatives. In order to probe this question further some

calculations were performed on methane. An experimental value exists for the C-H stretch intensity parameter[62] and due to the low number of electrons and high symmetry, high-level calculations may be done quickly. Furthermore the bonding arrangement is precisely the preferred tetrahedral orientation, so the density should be well described at the triple-zeta basis set level. All calculations were performed using the B3LYP/aug-cc-pVTZ geometry. The results are shown in Table 10.

Methane	B3LYP	B3LYP (TD)	CCSD	CCSD(T)	EXPT[a]
C-H	1.264	1.338	1.261	1.230	1.26

Table 10: Intensity parameter for methane, Expt and aug-cc-pVTZ basis set
[a] Reference [62]

The experimental value is bracketed by the static B3LYP and CCSD(T) calculated values. The CCSD value allows the importance of including a perturbative treatment of triple excitations to be estimated. For methane, the triple's contribution is small and negative. The worsened agreement with the experimental value using CCSD(T) is not significant since both are within experimental error. It is immediately apparent that the static B3LYP calculation performs quite well for saturated hydrocarbons with ordinary charge densities. For methane the static calculation produces a result within experimental error. The dynamic correction increases the value too much, leaving it above the margin for experimental error. For ethane, the static value is low (Table 8), while the dynamic correction overshoots the experimental value. However, the dynamic value does remain closer to experiment. The performance of various DFT methods with or without the dynamic correction has recently been reviewed by

Neugebauer et. al. in a comparison with some coupled cluster methods.[56] Their results included mostly very small systems such as ammonia and carbon dioxide. The saturated hydrocarbons, methane and ethane, were also included. None of the systems considered in their work included unusual charge densities or strong correlation effects. Their conclusion was that the hybrid-DFT methods (especially B3LYP) worked quite well in predicting Raman intensities (most often better than with the dynamic correction). They also found that, for the CCSD method, the results were more strongly dependent on the quality of the basis set. A poor quality basis set caused an overestimation of the polarizability derivatives that was further increased by a dynamic correction.

My own results for saturated hydrocarbons indicate that the static B3LYP method can provide a reasonable estimate of intensity parameters. In some cases the calculated values are very close to experimental (methane). However, the method is still not entirely reliable and can produce values in considerable error. Although this might be expected for an unusual structure such as BCP, it could not be anticipated for simple systems such as ethane (see Table 8 above). When used with this method, the TDDFT dynamic correction also becomes unpredictable, since in some systems the results are much worse (methane), while in others they are much improved (ethane). With no way of knowing *a priori* how the calculation will perform, the results of any hybrid-DFT calculation with or without the dynamic correction must be viewed with caution.

Static calculations using the CCSD(T) method with a large basis set produce excellent values for all systems considered (when compared with data

obtained using a 514.5 nm laser line). The calculated values were always within experimental error. These calculations are considered to be too expensive to perform routinely. It should be noted that in view of the limited experimental data and the poor reliability of other methods, there are no routine Raman intensity calculations. Until all of the trends have been confirmed and understood, it is still critical to use benchmark level calculations. Trends observed in systematically erroneous results will only impair understanding.

The use of the coupled-cluster method was particularly important for BCP since the experiment itself was complicated by resonances. When the resonance interaction matrix was constructed (see Appendix E), certain weak coupling mechanisms were neglected. The strongest Fermi resonance matrix elements were connected to mode 2. Also, the Darling-Dennison constant was strongly dependent on the deperturbed frequency for mode 2. The intensity assigned to mode 1 was less sensitive to the DD constant and the neglected resonances. The fact that the CCSD(T) calculated values are very close to the experimental ones is solid evidence that the resonance treatment and other experimental procedures were appropriate. The calculated intensity parameter for the bridgehead stretch (1.593) was only 3.25% higher than experi mental (1.541). For the methylene stretching intensity parameter (derived largely from mode 2), the calculated value (1.282) was within 2.50% of experim ent (1.250). These values are well within the tightest margin of experimental error accepted for these types of scattering experiments (5‰

It is encouraging that although it took benchmark level calculations to describe the Raman intensities of BCP accurately, the general trends produced with low-level calculations were confirmed. The bridgehead stretch intensity parameter is now the largest known for a saturated hydrocarbon. The alignment of the bridgehead C-H bonds with the molecular axis, and position at the ends of three chains were found to be associated with large derivatives in the survey calculations. These experimental results validate predictions made using the trends observed in the survey calculations.

In the following section on future research, some ideas will be presented on how best to proceed from here. Many researchers are interested in accurately modeling the properties of extended systems, especially those with π-conjugation. These are certainly beyond the reach of the coupled-cluster method. The work presented in this thesis helps to shed light on the relative strengths of the more affordable methods.

Future Research

Now that some of the predicted trends in the influences of structure and conformation on polarizability derivatives have been confirmed, the question is how best to direct future research. Many researchers are interested in the electronic properties of extended systems, especially those with π-conjugation. Most systems of interest in the areas of non-linear optics and materials science are far too large to treat with the only method found to be consistently reliable, coupled-cluster. Many research groups tend to employ DFT calculations to provide insight into their particular systems. The dynamic corrections are often neglected due to computational expense. Furthermore, problems have been identified with the TDDFT method for systems with π-conjugation.[56] Thus the majority of polarizability calculations carried out today on polymers, graphite sheets and other extended systems are at the BLYP/6-31G or similar level.

The best way to use what has been learned in this study would be to probe the specific deficiencies of methods at this level. Where highly accurate coupled-cluster wavefunctions can be obtained, they should be dissected and compared with those computed with lower levels of theory. In this way, the origins of specific problems can be identified. As discussed in the first part of this thesis in the section on S_N2 reactivity, the theory of Atoms In Molecules[19] is a useful tool for dissecting charge densities. The Gough research group has published studies of AIM analysis of Raman intensities before. The flow of charge from one atomic basin to another induced by applied fields can be obtained from calculated wavefunctions at various geometries. Thus the polarizability derivatives can be

recovered in terms of the individual atomic contributions. Some insights have been obtained in the past for straight-chain hydrocarbons up to n-pentane, as well as cyclohexane. These analyses were performed on low-level non-correlated wavefunctions (D95)*.

The main difficulty, which arises when this method is applied to BCP, can actually be used to advantage. The problem is that AIM analysis of any set of calculated wavefunctions will dissect but ultimately recover only the behavior predicted at that level. If B3LYP/aug-cc-pVTZ is grossly overestimating the value for the C-H stretching derivative, then AIM analysis will provide the atomic contributions to this overestimation. BCP has an unusual charge density with a correspondingly great variation in the calculated Raman intensities over different levels of theory. Thus if wavefunctions from a wide variety of theoretical methods were analyzed, the problems associated with each method could be found. Using the highly accurate CCSD(T) wavefunctions as a guide, one might determine, for example, that the BLYP method overestimates charge transfer from carbon atoms. This sort of direct feedback on method performance could be very useful for a researcher trying to understand why a material with promising theoretical properties behaved slightly differently.

I propose that a project in this area could consist of:

i)   Computation of wavefunctions for BCP at equilibrium and at various stretched geometries using the CCSD(T) method.

ii)   Analysis of these wavefunctions for an accurate picture of the physical behavior (at least within the static approximation).

iii) Computation of non-correlated SCF wavefunctions with different sets of polarization functions for the same structures

iv) Computation of wavefunctions using one or two DFT methods

v) Computation of wavefunctions using lower coupled-cluster methods (CCS, CCSD)

A complete AIM analysis of these wavefunctions would provide a wealth of information. At the SCF level, the results would reveal just how the addition of polarization functions helps to complete the polarization space. For the DFT methods, the analysis could probe the detailed performance of some different functionals. The idea of parameterizing functionals to reproduce the coupled-cluster densities has already been put forward. This could be extended to include the best reproduction of response properties. Finally, the analysis of lower levels of the coupled-cluster method could dissect the contribution from excitations of each order. The effect on the polarizability of gradually including more electron correlation could be evaluated. This study could also be extended to other small systems that provoke a wide range in the calculated properties from different methods. It could be used to study systems where the importance of electron correlation effects varies. Appendix B discusses some specific ways to obtain high-level wavefunctions more quickly.

Raman Intensities Appendix A: Modern Computation of Polarizability and

Raman Intensities

The practical aspects of computation of polarizabilities and Raman intensities were recently reviewed by the Gough research group for a special publication on these topics.[70] C. Lupinetti surveyed and summarized the current literature, after which K.M. Gough and I wrote a manuscript detailing the current state of theoretical and experimental work in these areas. Much of the focus of the article was on the methods used in our laboratory to record and treat experimental data. This included higher-order force-field analyses, and inverse-eigenvalue algorithms that I found necessary to account for anharmonic resonances in my study of BCP. These topics will be discussed in detail in the other appendices. Following is a summary of the *ab initio* computational methods that are currently in use in the area of Raman intensity calculations as discussed in the manuscript.

The rapid developments in computational capability in the last decade necessitate a reconsideration of the methodology employed in the past and the improvements that may be incorporated, given the size and type of molecules that are commonly investigated. Numerous basis sets have been developed for the calculation of electronic properties and in particular, the molecular polarizability.[49-55] The accuracy of calculations can be further refined through incorporation of electron correlation, configuration interaction, nuclear polarizability, and dynamic rather than static systems.

Electronic properties are normally obtained from derivatives (analytical or finite difference) calculated at the optimized geometry; however, it has been argued that experimental geometries yield better results.[95, 96] A more rigorous solution to this problem is to employ a higher-order anharmonic force field to obtain geometries within the more realistic anharmonic potential.[97-99] The position expectation values for the nuclei no longer coincide with the potential minimum and the molecular electronic properties are slightly altered.[100-102] Of far greater experimental importance are the extreme effects of various resonances that frequently complicate recorded spectra. As will be discussed in appendices E-G, calculated higher-order force fields, as well as dipole and polarizability fields, yield corrections for resonant "intensity borrowing" from fundamental transitions by higher-order transitions.[98] This correction will be shown to be important for the comparison of calculated intensities with those recorded experimentally. There are many options for improving theoretical computations beyond the level employed in the past. Choices include the size of the basis set, addition of diffuse and/or polarization functions, inclusion of electron correlation, configuration interaction, vibrational responses to the applied field, and the choice between dynamic and static calculations.

The importance of adding diffuse and/or polarization functions to basis sets is well known and has been explored in many studies.[49-56] In general, large bases with split valence and exponent optimization, diffuse functions, and polarization functions are known to improve the accuracy of the results. Diffuse functions appear to be more important than polarization functions; however, if the

latter are optimized for the calculation of electronic properties, the results can be as good as those for much larger bases with diffuse functions. For example, Van Caillie and Amos[87] found that Sadlej's polarization functions[103] yielded results that were just as good for numerous small molecules as those from an augmented basis set (aug-cc-pVTZ) whether calculated as SCF, or with Density Functional Theory (DFT) [Generalized Gradient Approximation (GGA), such as B3LYP; or Local Density Approximation (LDA)]. Inclusion of electron correlation explicitly, through Moller-Plesset perturbation theory or implicitly, with DFT, is desirable, but can lead to unexpected anomalies. With DFT, local field problems have been identified and improved functionals are now being developed.[104, 105] Configuration interaction may be important in some cases but may be computationally too expensive for molecules with more than a few atoms.[106, 107]

Nuclear coordinates may enter at several levels. The polarizability values may be averaged over the zero-point nuclear displacements. Additionally, nuclear displacements in the presence of an electric field, called vibrational relaxation[108] or vibrational polarizability[109], may make a significant contribution to the polarizability in some cases, but the reported values were far from convergence with respect to basis set (since only 6-31G* was used). The contribution to the total polarizability is not likely to exceed 2% for our molecules and the effect on the derivative is probably minimal.

In the survey studies,[65,66] the D95(d,p) basis was used, as its consistency for molecular polarizabilities had been established. However, the results typically overestimate the derivative for a CC stretch and underestimate that for the CH

180

stretch. The B3LYP functionals, coupled with a large basis set (aug-cc-pVTZ), produce $\partial\overline{\alpha}/\partial r$ values that are in excellent agreement with experiment for the CH and CC stretching modes of propane (Table A1).

C_3H_8	SCF[a]	B3LYP[b]	EXPT[c]
In-Plane C-H	1.181	1.401	1.438
Out-of-Plane C-H	1.053	1.240	1.252
Methylene C-H	1.126	1.329	1.366
C-C	1.329	1.186	1.212
C_6H_{12}			
Equatorial C-H	1.259	1.454	1.33(F)[d]
Axial C-H	1.021	1.222	1.17(F)[d]
C-C	1.217	1.110	1.03(F)[d]

Table A1: Raman trace scattering parameters for propane, and cyclohexane ($\partial\overline{\alpha}/\partial r \times 10^{-30}$ Cm/V), from theory and experiment.
[a] HF/D95(d,p), reference 65.
[b] B3LYP/aug-cc-pVTZ, (CADPAC6) [94], calculated for this work
[c] Reference 64.
[d] Reference 110. (F) Strongly affected by Fermi resonances (see discussion)

The B3LYP results for cyclohexane are also much improved over the HF/D95(d,p), but agreement with experiment is less good. The bands in cyclohexane are strongly perturbed by Fermi resonance and the isotopic sum rules are not obeyed well. To resolve this problem, it will be necessary to employ the higher order force field methods described above. Overall, the derivative for the CH stretch is increased by about 18-20%, while that for the CC is lowered by

about 10%, compared with the results from the HF method with the smaller basis set.

Possible improvements to the theoretical methods cover a range of possibilities. For the computation of polarizability and polarizability derivatives, some considerations are more important than others. Clearly, large basis sets with diffuse and polarization functions should always be used, since the improvements are systematic and do not simply result in a scalable increase in property values. We have found that for most small to medium molecules, the aug-cc-pVTZ basis set provides excellent results. Furthermore, basis set tests, up to aug-cc-pV5Z on small, saturated hydrocarbons, have shown that the vibrational eigenvectors and Raman intensities are essentially converged with respect to basis set at the triple zeta level.

The dynamic correction is an exciting development that appears consistently to increase the static values by a few percent. This often produces accuracy within experimental error, but in some instances, worsens agreement with experiment. This depends on the system and the method (see Results and Conclusions). Also, this additional correction comes at great computational expense at this time. For all but small test cases, given the predictable increase, it seems reasonable to accept that the static values may be slightly low. It would certainly not be practical to employ this method for some of the more extended systems (such as conducting polymers) currently of interest. Also it should only be applied when the method used is already producing a good description of the static field properties.

Vibrational relaxation merits some investigation. While we have not found it necessary to include this effect in our calculations to date, there is certainly evidence that it could be important in some cases. Since geometry optimizations and nuclear forces can now be routinely computed in the presence of an applied field (e.g.: Gaussian 2003[70]), it would be useful to perform survey calculations to find a molecule exhibiting anomalously large relaxation effects, as a candidate for experimental confirmation.

In terms of the level of theory appropriate for Raman intensity calculations, most systems of interest to the Gough group are best suited to DFT calculations at the GGA level. While CCSD(T) and similar methods are important for benchmarking and comparison purposes for functional development, the types of systems in which the most unusual electronic behaviors are provoked (anomalous intensities, non-linearities, etc.) are simply too large to treat this way.

From the experimental perspective, it has become important to use theoretical methods more extensively for the treatment of recorded spectra. The use of isotopomers has been valuable[63,64] but anharmonicities limit the transferability of intensity parameters and accuracy is reduced when isotope sum rules are poorly obeyed. Sign ambiguities and anharmonic resonances can often be resolved through theory. The next level of accuracy demands that higher-order force field methods be employed. Results for all of the systems considered in this thesis, including saturated and unsaturated long chain hydrocarbons, as well as others such as those for extended sheets (e.g. the PAH's), exhibit some

remarkably similar trends. Bond length alternation along the system, accompanied by non-linear increases in polarizability derivatives towards the central bonds, finally reaching a linear asymptote, are common themes. Another consistent observation is an "end effect", where stretches of bonds at the end of a chain or at the edge of a sheet produce unusual polarizability derivative values. A wide range of theoretical methods has predicted these effects. The general trends are observed at the HF level with small basis sets (HF/6-31G*), and are maintained at various higher levels (e.g.: B3LYP/aug-cc-pVTZ). A combined theoretical and experimental probe into the similarities between apparently categorically different systems is required. Understanding the mechanisms behind the non-linear regime and the end effect is the current challenge.

Raman Intensities Appendix B: Specific Methods for Numerical

Polarizability Calculations

The previous appendix dealt with the types of methods commonly used for calculation of the polarizabilities of molecules. Most of these methods employ an analytical calculation of vectors using the Coupled-Perturbed Hartree-Fock (CPHF) method. This approach is very convenient for most methods and is the default method in the Gaussian software. However, this method cannot be applied to all levels of theory, most notably the coupled-cluster method. The polarizability for the coupled-cluster methods is actually computed numerically as an energy or dipole derivative with respect to an applied field.

$$\Delta E = f\,\mu - \frac{f^2 \alpha}{2}$$
$$\alpha = \frac{2\Delta E}{f^2} \qquad\qquad (B1)$$
$$\alpha = \frac{\mu}{f}$$

Expression (B1) as given by Bauschlicher is correct to second order.[55] If a small field f is applied (0.001 a.u.), the neglect of higher-order polarizabilities does not significantly affect the results. With the use of small fields the requirement for numerical accuracy is strict. In his work, Bauschlicher used these expressions to compute molecular polarizabilities. His SCF convergence criteria were 10^{-8} (this is now the default in Gaussian). For calculation of polarizabilities usable for computation of derivatives of the polarizability, a reasonable convergence is 10^{-11}. This necessitates only a few more SCF cycles which, for a coupled-cluster calculation, is a very small percentage of the total calculation.

When only the polarizability tensor values and not the perturbed wavefunctions are required, analysis of a system such as BCP can be done in three calculations. This simplification is possible firstly because all of the off-diagonal elements are zero. This means that only the xx, yy, and zz components are needed. Secondly, proper calculation of the yy component, which requires averaging of the results obtained for applied fields in each direction, is unnecessary, since the xx component is determined from one calculation (the other is equivalent by symmetry) and due to the overall symmetry of the molecule, the xx and yy components are equal. Thus, after only the xx has been computed, both are known. For the zz component, the positive and negative directions are related by symmetry, so this requires only one more calculation. Sheet B1 shows how the polarizabilities were obtained from energies for BCP at equilibrium (EQ) and at bridgehead (bh) and methylene (me) C-H contracted geometries.

$$EQ := \begin{pmatrix} -194.90872206 \\ -194.90875072 \\ -194.90875072 \\ -194.90874878 \end{pmatrix} \qquad bh := \begin{pmatrix} -194.90856101 \\ -194.90858965 \\ -194.90858965 \\ -194.90858748 \end{pmatrix} \qquad me := \begin{pmatrix} -194.90815689 \\ -194.90818523 \\ -194.90818523 \\ -194.90818355 \end{pmatrix}$$

$$\frac{\left(EQ_0 - EQ_1\right) \cdot 2}{.000001} = 57.32 \qquad \frac{\left(bh_0 - bh_1\right) \cdot 2}{.000001} = 57.28 \qquad \frac{\left(me_0 - me_1\right) \cdot 2}{.000001} = 56.68$$

$$\frac{\left(EQ_0 - EQ_2\right) \cdot 2}{.000001} = 57.32 \qquad \frac{\left(bh_0 - bh_2\right) \cdot 2}{.000001} = 57.28 \qquad \frac{\left(me_0 - me_2\right) \cdot 2}{.000001} = 56.68$$

$$\frac{\left(EQ_0 - EQ_3\right) \cdot 2}{.000001} = 53.44 \qquad \frac{\left(bh_0 - bh_3\right) \cdot 2}{.000001} = 52.94 \qquad \frac{\left(me_0 - me_3\right) \cdot 2}{.000001} = 53.32$$

MathCad Sheet B1: Numerical polarizability calculations for BCP using CCSD(T) energies.

As shown in Sheet B1, the energies for a particular geometry are lowered in the presence of an applied field. The four matrix elements correspond to no applied field, and fields applied in the x, y, and z-directions. The field induces a dipole moment in the molecule that is opposite to the field direction, causing a net stabilizing interaction with the field. The extent to which the energy is lowered, also manifested in the magnitude of the induced dipole, is a measure of the polarizability. The importance of well-converged energies is obvious when it is realized that the eighth decimal place in the energies determines the second decimal place in the polarizabilities. The application of larger fields would reduce this sensitivity but would raise the significance of neglected higher-order polarizabilities.

Calculation of the polarizability at each of the geometries required a minimum of three calculations (no field, x-field, and z-field). The first calculation with no applied field can be completed rather quickly. As stated previously, the CCSD(T) method scales as N^7 with the number of electrons. This is reduced by symmetry. With such poor scaling, any symmetry that can be used results in huge time saving. An energy calculation for BCP using the full D_{3h} symmetry tends to run about 5 times more quickly than one with no symmetry (1 day as opposed to 5 days, using the HPC facility at the U of M).

With an applied field, the symmetry of the wavefunction is generally reduced. This could have meant that most of the calculations would take close to 5 days each. An important consideration is that of residual symmetry. This is a lowering but not a complete loss of symmetry. If the applied field causes a

lowering of symmetry to that of an Abelian sub-group of the original symmetry, then the Gaussian program will use that symmetry. For example, when a y-field is applied to BCP, reducing the D_{3h} symmetry to C_{2V} which is the highest Abelian sub-group of D_{3h}, the calculation will proceed using C_{2V} symmetry. This is still much faster than if no symmetry were used. However, when a z-field is applied to BCP, the symmetry is lowered to C_{3V}. This is a higher symmetry than C_{2V}, yet it is not a sub-group of D_{3h}. Thus the calculation will proceed without using any symmetry and be very time consuming. Also, the wavefunction is not constrained to the true symmetry of the system.

Fortunately both problems could be easily overcome by the same solution, described here. Consider the z-field perturbation to BCP. The true symmetry of the system becomes C_{3V}. Keeping in mind that finite stretches of 0.01 Angstrom are being applied in computing the polarizability derivatives, a very small geometric perturbation can be applied. When one of the bridgehead C-H bonds is extended in a positive stretch by 10^{-6} Angstroms, the symmetry is reduced to C_{3V}. When the z-field is applied, the wavefunction maintains its C_{3V} symmetry and the calculation proceeds using that symmetry.

The geometric perturbation is four-orders of magnitude smaller than the effects being explored. Furthermore the use of a *positive* stretch can be shown to reduce this even more. As seen in Sheet B1, a lower energy is manifested as an increased polarizability. By stretching the C-H bond slightly from equilibrium, the energy is raised, reducing the calculated polarizability. However, the positive stretch of the bridgehead bond electronically *increases* the polarizability. Thus

stretching the bond has two very small effects, one mechanically raising the energy (seen as a decrease in polarizability), and the other electronically increasing the polarizability. Thus the two effects cancel even within the order in which they occur, which is four orders lower than the effects being investigated. This approach can often be adapted to any system of interest providing that there is some symmetry.

Another approach that may be used involves replacing one atom by a different isotope in the calculation. This will lower the symmetry without adjusting the geometry at all. Of course, this is not appropriate if the masses are important such as with a frequency calculation.

Raman Intensities Appendix C: Matrix Methods for Quadratic Force-Fields

Bicyclo-[1.1.1]-pentane has 13 atoms and therefore 39 degrees of freedom, 33 of those vibrational. This means that the quadratic force-constant matrices for this molecule are 39-dimensional. After some analysis this can be reduced to 33. The eigenvectors corresponding to the 6 zero-eigenvalues of the mass-weighted force-field are related to rotations and translations. These can therefore be discarded during a vibrational analysis. The methods and concepts are completely general although some complexities are introduced by degenerate modes. Because of the generality of the concepts of dimensionless normal modes and force-constants, some of these concepts will be illustrated using the water molecule as an example, for simplicity.

The dimensionless normal modes of a molecule are derived from the eigenvectors of the mass-weighted force-field.[69] In general, the force-field may not be uniquely determined from the transition energies (spectrum) and thus must be computed. For the purposes of Raman intensity analysis the most accurate possible eigenvectors are required. Thus the highest affordable level of electronic structure theory should be chosen. For the BCP project, all of the force-fields were calculated at the B3LYP/aug-cc-pVTZ level. Higher basis sets were also tested (up to quintuple zeta) but it was found that the eigenvectors had converged with respect to basis set at the triple zeta level. Some researchers have scaled the molecular force-field to produce agreement with all of the experimental transitions. This is generally not recommended if there are resonances present and a higher-order force-field is to be employed.[97] It may be that the best

affordable level of theory is predicting the unperturbed frequencies quite well, and that scaling is worsening the eigenvectors in an attempt to account for the resonances. A good test of the theoretical method is its prediction for transition frequencies that are clearly not perturbed by resonances. If the agreement is good for those modes and slightly off in a sensible fashion for the perturbed modes, this is an acceptable level of theory.

A brief description of the units and conventions for force-field analysis will be followed by a sample analysis using water. This example is convenient since the relevant matrices can be displayed explicitly. Previously, where the derivation of the intensity parameters for BCP was illustrated, the composition of the 5 totally symmetric modes of BCP was described. The full 39-dimensional treatment of BCP is precisely analogous to the 9-dimensional water analysis shown here.

A frequency analysis for deriving the harmonic force-field is generally performed at a minimum-energy geometry. This means that the linear forces are essentially zero. Thus the force-constant matrix in the normal mode basis will be diagonal. However, a frequency analysis performed with most commercial software such as Gaussian, yields a matrix of dimension 3N corresponding to the Cartesian forces. The quadratic force-constants obtained from a frequency analysis are in units of Hartrees per $(Bohr)^2$. This is equivalent to J/m^2 in SI units. Once they have been mass-weighted this becomes J/m^2Kg. The normal modes obtained as eigenvectors of the mass-weighted force-field have dimensions of $kg^{1/2} \cdot m$. The eigenvalues can be converted into wavenumbers and correspond to

the transition energies. Six of the eigenvalues should be very close to zero, corresponding to the translations and rotations. The six associated eigenvectors describe these motions. The translations are easily identified since the center of mass of the molecule is displaced with respect to theses coordinates.

The eigenvectors of the non-mass-weighted force-field have dimension of meters and can be used to create displacements along a normal mode coordinate in Cartesian coordinates. Once the normal coordinates are made dimensionless, the force-constant simply has units of energy. These can be expressed in wavenumbers and for the quadratic constants they are equal to the transition energies. This formulation is very convenient when the force-field is extended to higher orders. Due to the dimensionless coordinates, the cubic and quartic constants are all expressible in wavenumbers. This will be shown in Appendix D.

Below is the quadratic force-field analysis for water. In Sheet C1, the dimension of the problem is defined, followed by the necessary physical constants. Next the nuclear positions X are defined followed by the nuclear masses M. The Cartesian force constants are entered as a column vector (not shown). These are then stacked to fill the lower diagonal of a 9x9 matrix. This is then symmetrized and displayed to 3 decimal places as EQ. This reflects that these constants are for the equilibrium geometry X. Planck's constant is defined as h1.

In Sheet C2 the force constant matrix is mass-weighted and the eigenvalues are obtained and sorted in ascending order. Due to the low tolerances for the SCF and the CPHF vectors used to compute this sample force-field, the

first 6 eigenvalues are not precisely zero. The fineness of the integration grid for hybrid-DFT methods also contributes to this source of error. This is independent of the level of theory and simply reflects the convergence criteria. The CPHF vectors and the SCF for bicyclo-[1.1.1]-pentane were both computed with convergences of 10^{-11}. Also a "veryfine" grid was used for the DFT integrations. This results in much improved eigenvalues. A poor-quality force-constant matrix manifests itself in poor quality eigenvectors. The center of mass of the molecule will move slightly with a poor quality vibrational eigenvector.   Thus the vibrational coordinates are said to be contaminated with rotations and translations.

$l := 1, 2 .. 9$

$m := 1, 2 .. 9$

$AMU := 1.6605402 \cdot 10^{-27} \cdot kg$

$H := 4.3597482 \cdot 10^{-18} \cdot joule$

$c := 29979245800 \cdot \dfrac{cm}{sec}$

$h1 := 6.6260755 \cdot 10^{-34} \cdot joule \cdot sec$

$a_0 := 0.529177249 \cdot 10^{-8} \cdot cm$

$$X := \begin{pmatrix} -0.015358 \\ 0 \\ -0.040856 \\ -0.020352 \\ 0 \\ 0.955758 \\ 0.945768 \\ 0 \\ 1.200455 \end{pmatrix} \qquad M := \begin{pmatrix} 1.00783 \\ 1.00783 \\ 1.00783 \\ 16 \\ 16 \\ 16 \\ 1.00783 \\ 1.00783 \\ 1.00783 \end{pmatrix}$$

X, M correspond to position and masses respectively

$EQ_{1,m} := pEQ_{\frac{(1+1)\cdot1}{2} - 1 + m}$

Stacking algorithm: places forces from column vector (not shown) into matrix

$EQ_{m,1} := EQ_{1,m}$

$$EQ = \begin{pmatrix} 0.053 & 0 & 0.022 & -0.044 & 0 & 0.036 & -0.01 & 0 & -0.058 \\ 0 & 0 & 0 & 0 & -0 & 0 & 0 & 0 & 0 \\ 0.022 & 0 & 0.436 & -0.038 & 0 & -0.416 & 0.017 & 0 & -0.02 \\ -0.044 & 0 & -0.038 & 0.435 & 0 & 0.093 & -0.392 & 0 & -0.054 \\ 0 & -0 & 0 & 0 & 0 & 0 & 0 & -0 & 0 \\ 0.036 & 0 & -0.416 & 0.093 & 0 & 0.483 & -0.128 & 0 & -0.067 \\ -0.01 & 0 & 0.017 & -0.392 & 0 & -0.128 & 0.401 & 0 & 0.112 \\ 0 & 0 & 0 & 0 & -0 & 0 & 0 & 0 & 0 \\ -0.058 & 0 & -0.02 & -0.054 & 0 & -0.067 & 0.112 & 0 & 0.088 \end{pmatrix}$$

MathCad Sheet C1: Harmonic force-field for water

$$M := \sqrt{M}$$

$$k := M \cdot M^T$$

Mass-weighting force-field

$$EQ_{m,1} := \frac{EQ_{m,1}}{k_{m,1}}$$

$$o := sort \ (eigenvals \ (EQ))$$

$$p_1 := eigenvec \ \left(EQ, o_1\right)$$

Sorting and normalizing
eigenvalues and vectors

$$p1_1 := \frac{p_1}{|p_1|}$$

$$o = \begin{pmatrix} -0.0000334322 \\ -0.0000000009 \\ -0.0000000006 \\ 0.0000000003 \\ 0.0000228511 \\ 0.0000521634 \\ 0.1084277574 \\ 0.4414133095 \\ 0.47839281 \end{pmatrix}$$

List of sorted eigenvals. The
lowest 6 are approximately
zero

$$\nu_1 := \sqrt{\frac{o_1 \cdot \dfrac{H}{a_0^2 \cdot AMU}}{4 \cdot \pi^2 \cdot c^2}}$$

Compute frequencies. Set first 6 to
zero.

$$\nu = \begin{pmatrix} 29.7226i \\ 0.154i \\ 0.1211i \\ 0.0952 \\ 24.573 \\ 37.1268 \\ 1692.6786 \\ 3415.2853 \\ 3555.4665 \end{pmatrix} cm^{-1} \qquad \omega := \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1692.6786 \\ 3415.2853 \\ 3555.4665 \end{pmatrix} \cdot cm^{-1}$$

MathCad Sheet C2: Calculation of eigenvectors and vibrational frequencies

In Sheet C2 the eigenvectors of the mass-weighted force-field were obtained and then normalized and renamed. Thus the p1's are the mass-weighted eigenvectors. The largest eigenvalue supposed to be zero was 37 cm^{-1}. The six eigenvalues, which should have been zero, are set to zero when ω is defined as the set of vibrational frequencies.

In Sheet C3 the mass-weighting is removed to create a set of Cartesian displacement coordinates named as the p's. They are then used to compute the center of mass displacement with each coordinate. The last three, which are close to zero, represent the vibrational coordinates. When stricter convergence criteria were used for BCP, the movements of the center of mass for the vibrational coordinates were zero to at least 9 decimal places. The translational coordinates are easily distinguished by their large values. The reduced mass of each coordinate is also computed and will become significant in the cubic and quartic force-field analyses. Next the Cartesian displacement vectors p_7 through p_9 are displayed. An example is then shown of the calculation of the geometry for a normal-mode displaced structure. In this example, since the vector is normalized and carries units of Angstroms, the computed geometry RR corresponds to a net displacement of 0.01 Å.

$$Q_{1, m} := 0$$

$$Q_{1,1} := M_1$$

$$p_1 := Q^{-1} \cdot p_1$$

$$p_1 := \frac{p_1}{|p_1|}$$

> Centermass is the displacement of the center of mass with respect to the normal modes. 3 non-zero values correspond to translations.

$$\text{centermass}_1 := M^2 \cdot (p_1)$$

$$\text{redmass}_1 := M^2 \cdot (p_1)^2$$

$$\text{centermass} = \begin{pmatrix} -0.000012008 \\ 3.380010291 \\ 10.4014152897 \\ 14.3159213572 \\ 0 \\ 0.0000170096 \\ 0.0000000042 \\ 0.0000000006 \\ -0.0000000012 \end{pmatrix}$$

$$\text{redmass} = \begin{pmatrix} 1.0464575368 \\ 6.0053107969 \\ 6.0052998353 \\ 6.0050411991 \\ 1.00783 \\ 1.1258582864 \\ 1.0895730981 \\ 1.0387999957 \\ 1.0812768874 \end{pmatrix}$$

$$p_9 = \begin{pmatrix} -0.003479931928 \\ 0 \\ 0.705363941914 \\ -0.042850911624 \\ 0 \\ -0.055342712435 \\ 0.683767862334 \\ 0 \\ 0.173239988204 \end{pmatrix} \quad p_8 = \begin{pmatrix} -0.086531720991 \\ 0 \\ -0.701055973001 \\ -0.035937204128 \\ 0 \\ 0.027825564804 \\ 0.657059755123 \\ 0 \\ 0.259305837734 \end{pmatrix} \quad p_7 = \begin{pmatrix} 0.7034836692 \\ 0 \\ -0.0488315433 \\ -0.0583847336 \\ 0 \\ 0.0452062927 \\ 0.2234143962 \\ 0 \\ -0.6688496312 \end{pmatrix}$$

$$X2 := X - .01 \cdot p_9$$

$$RR_{\text{ceil}\left(\frac{m}{3}\right), m+3-3 \cdot \text{ceil}\left(\frac{m}{3}\right)} := X2_m$$

> Subtracts 0.01 times normal mode 9 from coordinate vector.

$$RR = \begin{pmatrix} -0.015323200681 & 0 & -0.047909639419 \\ -0.019923490884 & 0 & 0.956311427124 \\ 0.938930321377 & 0 & 1.198722600118 \end{pmatrix}$$

> Coordinate vector is stacked into xyz coordinates for use as input file for Gaussian.

MathCad Sheet C3: Calculation of reduced masses and a normal-mode displaced geometry for water

Sheet C3 shows how normal coordinates, Q, are related to the Cartesian displacements Δx. These coordinates can be formally related as:

$$\Delta x_k = M_k^{-1/2} l_k^n Q_n$$

$$Q_n = l_n^k M_k^{1/2} \Delta x_k$$

(C1-C2)

These were used to obtain the Cartesian displacements in sheet C3. The relationships between dimensionless normal coordinates q, and normal coordinates Q are:

$$q_r = \left(\frac{4\pi^2 c \omega_r}{h} \right)^{1/2} Q_r$$

$$Q_r = \left(\frac{4\pi^2 c \omega_r}{h} \right)^{-1/2} q_r$$

(C3-C4)

In Sheet C4, the factor to convert to dimensionless normal mode coordinates is applied. The force constants in wavenumbers corresponding to the dimensionless normal modes are obtained. The left and right operation by the eigenvectors of the mass-weighted force-field provides the eigenvalue in Hartrees, but with the units of length (squared) and mass still included. Division by hc (Plancks constant and the speed of light c) converts the energy from Hartrees into wavenumbers (cm^{-1}). Next the factors under the square-roots are applied. These are the conventional conversion factors between normal-mode coordinates, and dimensionless normal mode coordinates.

$$F99 := \left(p1_9\right)^T \cdot EQ \cdot p1_9 \cdot \frac{H}{a_0{}^2 \cdot AMU \cdot (h1 \cdot c)} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_9}} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_9}}$$

$$F99 = (\,3555.4665\,)\ cm^{-1}$$

$$F88 := \left(p1_8\right)^T \cdot EQ \cdot p1_8 \cdot \frac{H}{a_0{}^2 \cdot AMU \cdot h1 \cdot c} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_8}} \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_8}}$$

$$F88 = (\,3415.2852\,)\ cm^{-1}$$

$$F77 := \left(p1_7\right)^T \cdot EQ \cdot p1_7 \cdot \frac{H}{a_0{}^2 \cdot AMU \cdot h1 \cdot c} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_7}} \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_7}}$$

$$F77 = (\,1692.6786\,)\ cm^{-1}$$

$$F78 := \left(p1_7\right)^T \cdot EQ \cdot p1_8 \cdot \frac{H}{a_0{}^2 \cdot AMU \cdot h1 \cdot c} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_7}} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_8}}$$

$$F78 = (\,0\,)\ cm^{-1}$$

$$F79 := \left(p1_7\right)^T \cdot EQ \cdot p1_9 \cdot \frac{H}{a_0{}^2 \cdot AMU \cdot h1 \cdot c} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_7}} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_9}}$$

$$F79 = (\,0\,)\ cm^{-1}$$

MathCad Sheet C4: Harmonic force-constants in dimensionless normal coordinates for water

Each of the three non-zero force constants is computed in these dimensionless coordinates, with energy units of wavenumbers. The values should be recognized as the original transition energies. Some off-diagonal elements are

also computed to illustrate that these are indeed zero (this was an equilibrium geometry).

Raman Intensities Appendix D: Anharmonic Force-Fields

The dimensionless normal coordinates q, defined in Appendix C, form a convenient basis in which to express the molecular potential. As discussed in Appendix C, the linear terms vanish for an equilibrium geometry. Previously the series was truncated at the second-order or harmonic level. Extension of this series allows for a more accurate description of the potential, especially when overtones, which have greater amplitudes, are considered. The method can be extended to any order, although truncation at the fourth to sixth order is most common. This approach should be used with caution for systems exhibiting large amplitude torsional motions, although it should be noted that excellent results have been obtained for methanol.[111] There are no such difficulties with BCP which is quite rigid. The molecular potential can be expressed in the normal mode basis as:

$$V = \frac{1}{2}\sum_r \omega_r q_r^2 + \frac{1}{6}\sum_{r,s,t} \phi^{rst} q_r q_s q_t + \frac{1}{24}\sum_{r,s,t,u} \phi^{rstu} q_r q_s q_t q_u \qquad (D1)$$

The factors of 1/2, 1/6 and 1/24 are used to account for the unrestricted summations. This series, truncated at the fourth-order, was used for the calculations described in this thesis. When restricted summations are used, the constants themselves must be weighted depending on the number of possible permutations of the indices. This weighting is less convenient to keep track of for subsequent analyses of resonances for which matrix elements including these constants are evaluated.

Some symmetry restrictions apply to the cubic and quartic constants φ. The direct product of the symmetries of the labels r,s,t,(u) must transform as the totally symmetric representation A_1', otherwise the constant is strictly zero. Also any constant with indices that are a permutation of the indices of another constant is equal to that other constant. This follows from consideration of the potential surface as a smooth and continuous function, for which the order of multiple derivatives is unimportant.

BCP is a symmetric top molecule with D_{3h} symmetry. As such, the complications of degenerate modes and the associated possibility of vibrational angular momenta arise. When quanta of energy are in the degenerate modes, the classical vibrational trajectory is in general an ellipse. This is similar to how circularly and elliptically polarized light are defined using two orthogonal polarization vectors. Thus the vibrational states may possess angular momentum.

Complicated summations are generally performed using software. SPECTRO 3.0 is a software package that was first developed by Dr. Handy's research group in 1990.[112] Dr. Handy was kind enough to provide a free copy of his SPECTRO program for the analyses of BCP. The expression for the total vibrational energy shown below is written according to the conventions used by the Handy group. The total vibrational energy of a symmetric top molecule computed with a quartic force-field can be expressed as:

$$E(v, l) = \sum_s \omega_s (v_s + 1/2) + \sum_t \omega_t (v_t + 1) +$$

$$\sum_{s \leq s'} x_{ss'} (v_s + 1/2)(v_{s'} + 1/2) +$$

$$\sum_{s,t} x_{st} (v_s + 1/2)(v_t + 1) + \qquad \text{(D2)}$$

$$\sum_{t \leq t'} x_{tt'} (v_t + 1)(v_{t'} + 1) +$$

$$\sum_{t \leq t'} g_{tt'} l_t l_{t'}$$

In the above expression s labels the non-degenerate modes, while t labels the degenerate ones. The various anharmonic constants x and g are computed from the quartic force-field.

It is important to note that the software and the above expression are used to compute the anharmonic energies for the various vibrational states. Due to the inclusion of both mechanical and electrical anharmonicities the intensities of combination and overtone bands may also be calculated. However, in this thesis the purpose of computing an anharmonic force-field was not to calculate the energies of particular vibrational states but rather to resolve the resonances which were found to perturb the C-H stretching region in the Raman spectra. While the SPECTRO program will compute resonance interactions between two states, more complicated interactions must be diagonalized by hand. The matrix elements required for evaluation of resonant interactions are derived from the cubic and quartic force constants. Due to the extremely high number of possible resonance interactions for BCP, I decided to compute only the cubic and quartic constants necessary to resolve the C-H resonant polyad, rather than attempt to treat the complete set of molecular vibrational states. This approach focuses on

203

the goal of distributing the integrated intensity from the Raman spectra into the fundamental modes. Once the eigenvectors for the C-H polyad were obtained, the desired intensity parameters could be computed.

The computation of the cubic and quartic force constants as derivatives of the energy with respect to the normal mode coordinates is generally done numerically. Analytic second derivatives are commonly available for many levels of theory in the standard electronic structure software packages such as Gaussian. Methods and algorithms have also been established for analytic third and higher derivatives. These methods are computationally expensive and are not commonly available. The most common approach is to compute analytic second derivatives at various normal mode displaced geometries (Δ_i). Numerical estimators provide the cubic and quartic constants from the changes in the analytic second-order constants. The numerical scheme is as follows:[100, 102]

$$F_{i,jk} = \frac{F_{jk}(+\Delta_i) - F_{jk}(-\Delta_i)}{2\Delta_i}$$

$$F_{ii,jk} = \frac{F_{jk}(+\Delta_i) + F_{jk}(-\Delta_i) - 2F_{jk}(0)}{\Delta_i^2}$$

(D3-D4)

Thus, the changes in the second-order constants F_{jk}, induced by a displacement along normal mode i, provide the cubic constants F_{ijk}. The numerical scheme for the quartic constants produces only the semi-diagonal terms. These are the only ones required for analysis of resonance interactions to this order.[112] Due to the small displacement in the denominator (typically 0.01 Angstroms), the quartic

constants are generally less numerically stable than the cubic. This necessitates strict convergence criteria for the SCF and CPHF vectors (10^{-11}). Furthermore, the various permutations of the indices can all be computed and averaged. The variation among the constants prior to averaging provides a measure of the numerical accuracy of the constants. Finally, based on the direct-product of the symmetry representations of the mode labels, the constants that should be zero, based on symmetry, can be set to zero. This provides another measure of the numerical accuracy. The variations in cubic constants with permuted indices for BCP were of the order of 0.02 cm^{-1}. In view of the numerical scheme, it is not surprising that the variations in the quartic constants were about 2 cm^{-1}.

A sample calculation is shown below, again using water as the example. Since Gaussian03[113] can perform this analysis for some simple systems, the results will be compared.

Constants ω_r	cm^{-1}	Constants φ^{rstu}				cm^{-1}
1	3415.3	1	1	1	1	762.4
2	1692.7	2	1	1	1	−141.2
3	3555.4	2	2	1	1	−240.3
Constants φ^{rst}		2	2	2	1	144.2
1 1 1	−1680.7	2	2	2	2	−166.5
2 1 1	120.2	3	3	1	1	745.9
2 2 1	197.5	3	3	2	1	−186.4
2 2 2	−229.4	3	3	2	2	−267.5
3 3 1	−1642.3	3	3	3	3	748.9
3 3 2	370.1					

Table D1: Output from Gaussian03 of quartic force-field in dimensionless normal coordinates for water at B3LYP/3-21G.

To demonstrate the method, this calculation was performed with a very small basis set for efficiency. While not physically accurate, this allows the procedure to be illustrated. Below are shown the MathCad sheets used to compute these same constants from quadratic force-constants computed at normal mode displaced (0.01 Angstrom) geometries.

$$F11 := (Q1)^T \cdot EQ \cdot Q1 \cdot \frac{H}{a_0^2 \cdot AMU \cdot h1 \cdot c} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_1}} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_1}}$$

$$F11 = (3415.3)\ cm^{-1}$$

$$F22 := (Q2)^T \cdot EQ \cdot Q2 \cdot \frac{H}{a_0^2 \cdot AMU \cdot h1 \cdot c} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_2}} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_2}}$$

$$F22 = (1692.7)\ cm^{-1}$$

$$F33 := (Q3)^T \cdot EQ \cdot Q3 \cdot \frac{H}{a_0^2 \cdot AMU \cdot h1 \cdot c} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_3}} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_3}}$$

$$F33 = (3555.4)\ cm^{-1}$$

MathCad Sheet D1: Calculation of quadratic force constants F_{ii} with conversion to dimensionless coordinates.

$$F111 := (Q1)^T \cdot \frac{(F1a - F1b)}{2 \cdot .01} \cdot Q1 \cdot \frac{H}{a_0^2 \cdot 10^{-8} \cdot cm \cdot AMU \cdot h1 \cdot c} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_1}} \cdot \left(\sqrt{redmass_1}\right)^{-1} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_1}} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_1}}$$

$$F111 = (-1678)\ cm^{-1}$$

$$F211 := (Q1)^T \cdot \frac{(F2a - F2b)}{2 \cdot .01} \cdot Q1 \cdot \frac{H}{a_0^2 \cdot 10^{-8} \cdot cm \cdot AMU \cdot h1 \cdot c} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_2}} \cdot \left(\sqrt{redmass_2}\right)^{-1} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_1}} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_1}}$$

$$F211 = (119.6)\ cm^{-1}$$

$$F332 := (Q3)^T \cdot \frac{(F3a - F3b)}{2 \cdot .01} \cdot Q2 \cdot \frac{H}{a_0^2 \cdot 10^{-8} \cdot cm \cdot AMU \cdot h1 \cdot c} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_3}} \cdot \left(\sqrt{redmass_3}\right)^{-1} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_3}} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_2}}$$

$$F332 = (369.7)\ cm^{-1}$$

MathCad Sheet D2: Calculation of some cubic force constants F_{ijk} with conversion to dimensionless coordinates.

The quadratic constants shown in Sheet D1 agree fully with those from the Gaussian output. This is because these forces were computed analytically for the equilibrium geometry. The normal mode vectors operate on the Cartesian force matrix to produce the normal mode force constant. This is then converted into units of wavenumbers and scaled to produce the value for the dimensionless coordinates. The cubic constants are computed in much the same way, except that the difference between Cartesian force-fields computed at normal mode displaced geometries is used. These values are then divided by the displacement amount as per the numerical scheme. The mode was displaced in units of length; therefore the final scaling factor for this mode, used for conversion into dimensionless coordinates, is compensated by the reduced mass. Due to the default convergence criteria used for these calculations, the agreement between the manual displacements and those performed with Gaussian is to within about 2 cm^{-1}. This is due to the use of different step sizes rather than variation between constants with permuted indices. To demonstrate the accuracy obtained with the default convergence criteria, two elements with permuted indices, that are required by symmetry to be zero, are calculated.

$$F223 := (Q2)^T \cdot \frac{(F2a - F2b)}{2 \cdot .01} \cdot Q3 \cdot \frac{H}{a_0^2 \cdot 10^{-8} \cdot cm \cdot AMU \cdot h1 \cdot c} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_2}} \cdot \left(\sqrt{redmass_2}\right)^{-1} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_2}} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_3}}$$

$$F223 = (0.0002) \, cm^{-1}$$

$$F322 := (Q2)^T \cdot \frac{(F3a - F3b)}{2 \cdot .01} \cdot Q2 \cdot \frac{H}{a_0^2 \cdot 10^{-8} \cdot cm \cdot AMU \cdot h1 \cdot c} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_3}} \cdot \left(\sqrt{redmass_3}\right)^{-1} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_2}} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_2}}$$

$$F322 = (-0.0001) \, cm^{-1}$$

MathCad Sheet D3: Calculation of some cubic force constants with permuted indices (zero by symmetry).

As seen in Sheet D3, the agreement between permuted indices and the observation of the symmetry rules is quite good. Therefore most of the difference between these values and those computed automatically through Gaussian can be assigned to the different step sizes. The quartic constants are computed in a similar fashion and are shown in Sheet D4.

$$F1111 := Q1^T \cdot \frac{(F1a + F1b - 2 \cdot EQ)}{.01^2} \cdot Q1 \cdot \frac{H}{a_0^2 \cdot \left(10^{-8} \cdot cm\right)^2 \cdot AMU \cdot h1 \cdot c} \cdot \left[\sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_1}} \cdot \left(\sqrt{redmass_1}\right)^{-1} \right]^2 \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_1}} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_1}}$$

$$F1111 = (763.2)\, cm^{-1}$$

$$F3321 := Q2^T \cdot \frac{(F3a + F3b - 2 \cdot EQ)}{.01^2} \cdot Q1 \cdot \frac{H}{a_0^2 \cdot \left(10^{-8} \cdot cm\right)^2 \cdot AMU \cdot h1 \cdot c} \cdot \left[\sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_3}} \cdot \left(\sqrt{redmass_3}\right)^{-1} \right]^2 \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_2}} \cdot \sqrt{\frac{h1}{4 \cdot \pi^2 \cdot c \cdot \omega_1}}$$

$$F3321 = (-190.7)\, cm^{-1}$$

MathCad Sheet D4: Calculation of some quartic force constants F_{ijkl} with conversion to dimensionless coordinates.

As seen in Sheet D4, the scaling factor for the displacement mode enters twice (squared term), and now the energy is divided by the square of the displacement distance as in the numerical scheme. These numbers and those computed with Gaussian, which used different step sizes, agree to within about 4 cm^{-1}. Elements that are supposed to vanish by symmetry are now as large as 0.1 cm^{-1}.

The calculation of the cubic and quartic constants for BCP was performed in exactly the same way. The only complication was the initial definition of the normal coordinates, which required some adjustment when describing the degenerate vibrations. When first produced in MathCad, the orthogonal eigenvectors were not constructed according to the conventions of spectroscopy. Linear combinations of eigenvectors were used to form normal mode coordinates for each of the 11 pairs that were symmetric and anti-symmetric with respect to

the x-z plane. These are the modes defined at the end of the section on Raman spectra. Once these conventional normal mode eigenvectors were constructed, the method could proceed as in the water example. The displaced geometries were calculated according to these vectors, which were then used as operators. In this way, any necessary cubic and quartic constants could be obtained. Construction of these symmetrized modes can be automated in SPECTRO. In the next appendix (Appendix E) some of the possible resonance interactions will be considered and matrix elements will be computed. The off-diagonal elements of the interaction matrix will be computed in terms of the anharmonic force constants.

Raman Intensities Appendix E: Anharmonic Resonances

In order to set up a matrix describing the resonance interactions in BCP for the C-H stretching polyad, the states which are involved must be determined. This region (2800-3000 cm^{-1}) contains many possible states.
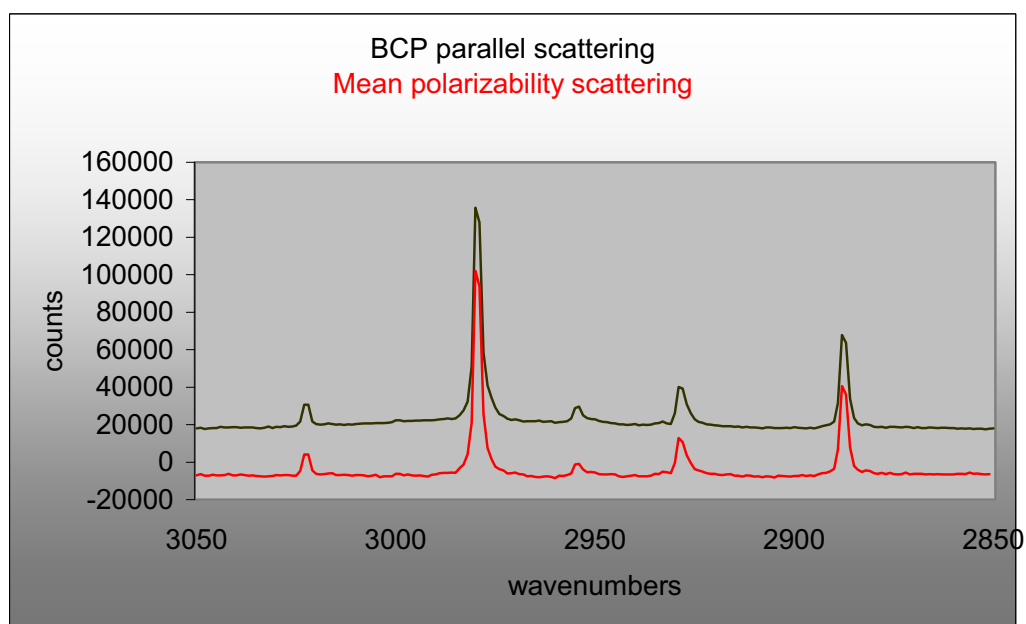
Figure E1: Isotropic Raman spectrum for the C-H stretching region. Figure is reproduced from section of Raman spectra (Figure 18).

Fortunately, the fact that the spectrum under consideration is that of the mean polarizability scattering allows for some simplification. Only states with a totally symmetric symmetry representation are allowed intensity in this spectrum. Only two fundamental transitions with this symmetry appear in the region. These are the expected bridgehead and methylene C-H stretching modes. However, five peaks are recorded in the isotropic spectrum. Wiberg et. al. published a harmonic force-field analysis of BCP in 1992.[86] The frequencies and their assignments are listed in Table E1.

Assigned State[a]	Symmetry[a]	Observed Freq[a]	This Study
$\upsilon 1$	$A_1{}'$	2978	2979.6
$\upsilon 2$	$A_1{}'$	2886	2887.6
$2\upsilon 3$	$A_1{}'$	3020	3022.5
$\upsilon(3+10)$	E'	2952	2954.5
?	?	2926	2928.6

Table E1: Observed states in C-H stretching region.
[a] Wiberg et. al. (1992) reference 86

The assignments by the Wiberg group were made in 1992 without the benefit of an isotropic Raman spectrum or an anharmonic force-field. Thus no attempt was made to account for the resonance interactions. Although the combination band $\upsilon(3+10)$ would be expected to lie in this region, it is of the wrong symmetry to appear in the isotropic spectrum. The overtone at 3020 cm^{-1} was identified as $2\upsilon 1$ in their manuscript, although this is probably a typographical error. It seems reasonable that the authors had identified it as $2\upsilon 3$, since the fundamental $\upsilon 3$ appears at 1509 cm^{-1}, while $\upsilon 1$ occurs at 2978 cm^{-1}. Otherwise the observed frequencies are in good agreement, given that both groups report approximately 1 wavenumber resolution.

Based on the isotropic spectra that were obtained in the work described in this thesis, and my quartic force-field, it was determined that these modes should be reassigned. Considering both one and two quantum transitions, and coupling mechanisms including Fermi (type I and II), Darling-Dennison as well as

rotational and vibrational l-doubling, a list of transitions that might appear in this region was created.[114-120] The states and their symmetries are listed in Table E2.

State	Symmetry	Couplings	Coup. Mech
$\nu 1$	$A_1{}'$	$2\nu 3, 2\nu 10^0$	F
$\nu 2$	$A_1{}'$	$2\nu 3, 2\nu 10^0$	F
$2\nu 3$	$A_1{}'$	$\nu 1, \nu 2, 2\nu 10^0$	F, D-D
$2\nu 10^0$	$A_1{}'$	$\nu 1, \nu 2, 2\nu 3, 2\nu 10^{\pm 2}$	F, D-D
$2\nu 10^{\pm 2}$	E'	$2\nu 10^0, \nu(3+10)^{\pm 1}$	Rot. l-dou, D-D
$\nu(3+10)^{\pm 1}$	E'	$2\nu 10^{\pm 2}, \nu 8^{\pm 1}, \nu 9^{\pm 1}$	D-D, F
$\nu 8^{\pm 1}$	E'	$\nu(3+10)^{\pm 1}, \nu 9^{\pm 1}$	F, vib-l-dou
$\nu 9^{\pm 1}$	E'	$\nu(3+10)^{\pm 1}, \nu 8^{\pm 1}$	F, vib-l-dou

Table E2: States, symmetries and couplings for BCP in C-H stretching region. F, D-D denote Fermi and Darling-Dennison resonances

As seen in Table E2, states $\nu 1$, $\nu 2$, $2\nu 3$, and $2\nu 10^0$ all have the totally symmetric $A_1{}'$ symmetry representation. The fundamental mode $\nu 10$ (E') appears at 1459 cm^{-1}. The overtone $2\nu 10$ is split into three states based on the allowed vibrational angular momenta (0, ±2). The state $2\nu 10^0$ is of $A_1{}'$ symmetry, while states $2\nu 10^{\pm 2}$ are of E' symmetry. The Fermi and Darling-Dennison resonances connect $\nu 1$ and $\nu 2$ with the other $A_1{}'$ states, but provide no mechanism for coupling to the E' states. There is a weak coupling via rotational l-doubling between the $2\nu 10^0$ state and the $2\nu 10^{\pm 2}$ states. This is the only coupling between the states of the two different symmetries.

For the states of E' symmetry, there are possible Darling-Dennison resonances between the $2\nu 10^{\pm 2}$ states and the $\nu(3+10)^{\pm 1}$ state. The $\nu(3+10)^{\pm 1}$ state can in turn couple to the nearby E' states $\nu 8$ and $\nu 9$ via Fermi and vibrational l-doubling mechanisms. There are no direct couplings between $2\nu 10^{\pm 2}$ and $\nu 8$ or

$\upsilon 9$ as the Fermi resonance matrix elements will be seen to vanish. Thus four of the five peaks in the isotropic spectrum are directly explained by the states of A_1' symmetry. These are the peaks at 3022, 2980, 2929, and 2888 wavenumbers. The fifth peak at 2955 can be assigned as the $2\upsilon 10^{\pm 2}$ states coupled to the $2\upsilon 10^0$ state and enhanced through coupling to the $\upsilon(3+10)^{\pm 1}$ state and through that, indirectly to $\upsilon 8$ and $\upsilon 9$. Thus through a cascade of resonances, the eigenvector for this state (2954 cm^{-1}) could be seen to contain some of the basis-states of A_1' symmetry.

It was decided to treat only the resonances between the four states of A_1' symmetry, and neglect the fifth weak transition. This overlooks only a small amount of integrated intensity while reducing the size of the resonance interaction matrix from eight-dimensions to four. Also the weak coupling mechanisms such as rotational and vibrational l-doubling are more difficult to compute accurately. Furthermore, the subsequent implementation of an inverse-eigenvalue algorithm (Appendix F) is greatly simplified, since the number of possible solutions is reduced from 40320 to 24.

Fermi and Darling-Dennison Resonance Matrix Elements

As seen in Table E2, states $\upsilon 1$ and $\upsilon 2$ can couple to states $2\upsilon 3$ and $2\upsilon 10^0$ through Fermi resonance mechanisms. States $2\upsilon 3$ and $2\upsilon 10^0$ are also coupled to each other via a Darling-Dennison mechanism. Thus the 4-dimensional resonance interaction matrix required a total of five off-diagonal elements; 4 Fermi type I, and one Darling-Dennison resonance constant. The possible Fermi

interactions for a D_{6h} point group have been published[118b] and also apply to BCP (D_{3h}). The applicable interactions are as follows:

Interaction	Vibrational A.M.	Matrix Element
$2v_n \sim v_{n'}$		$\dfrac{\phi_{nnn'}}{4}\left[\dfrac{v_n(v_n-1)(v_{n'}+1)}{2}\right]^{1/2}$
$2v_t \sim v_n$	$\Delta l_t = 0$	$\dfrac{\phi_{ntt}}{4}\left[\dfrac{(v_n+1)(v_t^2-l_t^2)}{2}\right]^{1/2}$ (E1-E2)

In the above expressions, the indices n and t label the non-degenerate and degenerate modes, respectively. The first expression applies to the interactions of $v1$ and $v2$ with $2v3$. Thus the cubic constants φ_{331} and φ_{332} are used. The second expression describes the interactions of $v1$ and $v2$ with $2v10^0$. In this example there is no change in vibrational angular momentum ($\Delta l_t=0$). These interactions are governed by cubic constants $\varphi_{1,10,10}$ and $\varphi_{2,10,10}$. All of these cubic constants are allowed to be non-zero since the direct product of the symmetry representations for the indices corresponds to that of the totally symmetric A_1'. In contrast, the cubic constant for interaction of $v1$ with $v(3+10)$ (denoted $\varphi_{1,3,10}$), must be zero since modes 1 and 3 are totally symmetric, while mode 10 is E'.

The Darling-Dennison (DD) constant describing the coupling between $2v3$ and $2v10^0$ is denoted $K_{3\,3,10\,10}$. This is not to be confused with $\varphi_{3\,3\,10\,10}$ although $K_{3\,3,10\,10}$ does contain a contribution from $\varphi_{3\,3\,10\,10}$. The DD interaction involves four quanta of energy as opposed to three for the Fermi interactions. The total

change in vibrational angular momentum must be 0 or ±3. For 2υ3 and 2υ10⁰ this is zero and is therefore allowed. The matrix element given by Mills and Robiette is:[115]

$$\left\langle v_n + 2, v_t - 2, l_t \middle| H/hc \middle| v_n, v_t, l_t \right\rangle =$$
$$\frac{1}{4} K_{nntt} \left[\left(v_n + 1 \right) \left(v_n + 2 \right) \left(v_t^2 - v_t^2 \right) \right]^{1/2}$$

(E3)

Thus the Darling-Dennison constant can be computed as: $\frac{2\sqrt{2}}{4} K_{3\ 3,10\ 10}$

The general equation for DD constants of the type K_{aabb} is given by Lehmann as:[118]

$$K_{aabb} = \frac{1}{4} \phi_{aabb} + \sum_\alpha - B_\alpha (\zeta_{ab}^\alpha)^2 \frac{\left(\omega_a + \omega_b \right)^2}{\omega_a \omega_b} +$$
$$\sum_k \frac{1}{8} \phi_{kaa} \phi_{kbb} \omega_k \left[\frac{1}{4\omega_a^2 - \omega_k^2} + \frac{1}{4\omega_b^2 - \omega_k^2} \right] +$$
$$\sum_k - \frac{1}{2} \phi_{kab}^2 \frac{\omega_k}{\omega_k^2 - \left(\omega_a - \omega_b \right)^2}$$

(E4)

This constant can be computed automatically using SPECTRO. The software provides the contribution from each term. The second term is a rotational contribution through the zeta matrices whose sum was zero for this interaction in BCP.

The non-zero terms were also computed using MathCad to provide more insight into some of the individual contributions. It appears that most of the entire

DD constant stems from a resonance involving mode 2. In term three, both φ_{233} and $\varphi_{2\ 10\ 10}$ are large and $4(\omega_{10})^2$ is very close to $(\omega_2)^2$. Since the total DD constant becomes strongly dependant on the values of the frequencies, the experimental values were used instead of the calculated ones. In Appendix F a discussion is included about self-consistent DD constants since this constant should in fact be computed with the deperturbed frequencies and not the experimental ones.

After evaluation of the various interaction constants the resonance matrix may be set up.

	$2\upsilon3$	$\upsilon1$	$\upsilon2$	$2\upsilon10^0$	Expt
$2\upsilon3$	$2\upsilon3$*	8.673	-23.348	17.815	3022.5
$\upsilon1$	8.673	$\upsilon1$*	0.000	8.967	2979.6
$\upsilon2$	-23.348	0.000	$\upsilon2$*	-24.464	2887.6
$2\upsilon10^0$	17.815	8.967	-24.464	$2\upsilon10^{0*}$	2928.6

Table E3: Resonance interaction matrix (Fermi and Darling-Dennison)
* best calculated/estimated fundamental transition energy (see Appendix F)

The eigenvalues of this interaction matrix should yield the experimental frequencies (rightmost column). The eigenvectors are the composition of each recorded band in the basis of the fundamental states. Rather than using computed frequencies as diagonal elements and then comparing the results with the experimental values, it was decided to solve the inverse-eigenvalue problem for improved diagonal elements in order to obtain the best possible eigenvectors.

After all, the purpose of this resonance analysis was to distribute the recorded intensity as accurately as possible. This procedure is discussed in Appendix F.

It is apparent from the interaction matrix that the strongest couplings are to mode 2. Furthermore, as mentioned previously, the Darling-Dennison constant was also strongly dependant on the frequency of mode 2. The intensity assigned to mode 1 is fairly insensitive to the value of the DD constant. However, the four states which were neglected in the resonance treatment are coupled much more strongly to mode 2 than mode 1. Therefore the intensity parameter obtained for mode 2 is a good test of the appropriateness of neglecting the additional weaker couplings. Whether or not to solve for a self-consistent DD constant is also determined by the accuracy of the intensity parameter for mode 2. A discussion of these factors in view of the results from some benchmarking level calculations is made in the section on conclusions.

Raman Intensities Appendix F: Inverse-Eigenvalue Algorithms

Inverse-eigenvalue algorithms[121-123] are numerical methods for solving the inverse-eigenvalue problem. This problem arises when the eigenvalues of a matrix are known, but some of the matrix elements are either inaccurately determined or unknown. If inaccurate or approximate values are employed, such a matrix will produce eigenvalues in poor agreement with the known values. This situation commonly occurs in physical chemistry when the known eigenvalues are experimentally observed transitions. Often the off-diagonal elements of the matrix (which represent the couplings) are also recorded in the spectrum, or they might be accurately calculated. This leaves the deperturbed transitions represented by the diagonal elements, not directly available from the spectrum. They are often more costly to calculate or may be unknown. The deperturbed transition energies may be desired but usually (as in this thesis) the goal is to obtain the eigenvectors. The eigenvectors are necessary for interpretation of the spectrum in terms of the basis states. Clearly this analysis would only be necessary in cases of strong mixing where significant intensity has been recorded for bands without one dominant basis vector.

Finding the diagonal elements which, when placed into the matrix, will produce the experimental eigenvalues is challenging. For a square matrix of size N, there are N equations in N unknowns. However, these equations may be highly non-linear, depending on the number of different couplings present. For all but the smallest or most sparse matrices, a numerical method is necessary. For vibrational spectroscopy, an inverse-eigenvalue algorithm may be implemented in

a scenario with the following conditions. If a spectrum is complicated by multiple anharmonic resonances, then a higher-order force-field analysis will determine the couplings between the various modes. The resulting resonant polyad must be resolved through diagonalization of a matrix describing all of the resonant interactions through fourth order. As discussed in Appendix E, the off-diagonal elements (the Fermi and Darling-Dennison interaction constants) are derived from the cubic and quartic force fields,[115, 120] and are known to be accurate even when calculated at a relatively moderate level of theory. The diagonal elements are the deperturbed anharmonic frequencies of the involved states. These are less accurately known, since even the initial harmonic frequencies usually require scaling.[100] Furthermore, in complex cases, many resonant denominators must be removed. Ultimately, this causes a breakdown in the perturbative method, and limits the size of system to which it may be applied. The eigenvalues of this resonance interaction matrix should yield the experimentally observed transition energies, while the eigenvectors provide the composition of each transition in terms of the fundamentals and combinations or overtones thereof. This completes the goal of distributing the observed intensity amongst the fundamental modes. In practice, if the diagonal elements are of poor quality, diagonalization will generally not yield particularly good agreement with the observed transition frequencies. Given accurate interaction constants and experimental eigenvalues, the inverse-eigenvalue problem may be solved to obtain a set of diagonal elements that will produce the observed transitions.[124] This approach allows the less accurate diagonal elements to be improved empirically in order to obtain

improved eigenvectors. Finally, using the derived eigenvectors, the intensity may be distributed amongst the fundamentals for transformation into the internal coordinate intensity parameters as before.

The complete solution to this problem, in MathCad is presented at the end of this appendix. The MathCad sheets contain explanatory documentation, however the rationale and approach are first discussed here.

Before implementing any inverse-eigenvalue method, one must decide which basis states to include and where to cut off the coupling strength. This determines the order of the problem. In a complicated molecule such as bicyclo-[1.1.1]-pentane, the number of possible interactions is staggering. The types and strengths of coupling mechanisms must be considered. An accurate description of the states demands that all of the most important resonant interactions be included. At least Fermi and Darling-Dennison resonances, as well as rotational and vibrational l-doubling, should be considered. Based on molecular symmetry considerations and the quartic force-field analysis, only Fermi and Darling-Dennison resonances among a total of four basis states were used for bicyclo-[1.1.1]-pentane (see Appendix E).

The method used in this thesis is based on Newton's method and assumes that an exact solution exists. This means that diagonal elements exist that will produce the desired eigenvalues. Only if the off-diagonal elements are also in error will this not be true. Using the notation of T. Luke,[123] the problem may be solved in the following way:

i)   Set up an n-dimensional matrix **A**, using the accurate off-diagonal elements, and zeros for diagonal elements.

ii)   Set up a matrix **D** as a diagonal matrix of the approximate diagonal values.

iii)   Matrix (**A**+**D**) has eigenvalues $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, .. \lambda_n)^{\text{T}}$

iv)   Desired eigenvalues are $\boldsymbol{\omega} = (\omega_1, \omega_2, .. \omega_n)^{\text{T}}$

v)   Seek new matrix **D** such that $\| \boldsymbol{\lambda} - \boldsymbol{\omega} \|^2$ is a minimum (preferably zero)

There are a number of physical constraints which simplify the problem. Because the quantities involved describe physical observables, the matrices and eigenvalues are real. The first step after the matrices are set up is to apply a spectral shift. The average of the desired eigenvalues is subtracted from them such that:

$$\sum_n \omega_n = 0 \tag{F1}$$

This constrains **D** such that:

$$Tr(D) = 0 \tag{F2}$$

This constraint causes:

$$\sum_n \lambda_n = 0 \tag{F3}$$

Beginning with the approximate values for **D**, these values are updated at each iteration.

$$D^{i+1} = D^i + \Delta D^i \tag{F4}$$

The key to this algorithm is the calculation of **ΔD**. Firstly, at each iteration i, the eigenvectors of the total matrix (A+D^i) are computed.

$$\left(A + D^i\right)X^i_k = \lambda^i_k X^i_k \tag{F5}$$

Next a matrix **J** is defined with components that are the squares of the components (x) of **X** followed by a transpose. The squared components of the eigenvectors are now the associated probabilities.

$$J^i = \left[\left(x^i_{jk}\right)^2\right]^T \tag{F6}$$

Now **ΔD** can be calculated as:

$$J^i \Delta D^i = \omega - \lambda^i$$
$$\Delta D^i = \left(J^i\right)^{-1}\left(\omega - \lambda^i\right) \tag{F7}$$

This method is quadratically convergent when an exact solution exists and the starting point is close enough. In practice this method can become unstable, particularly if an exact solution does not exist. In fact **J** can become singular. Thus the preferred method is:

$$\Delta D^i = (J^i)^T (\omega - \lambda^i) \tag{F8}$$

It can be shown that this method will converge. However, the rate is not guaranteed and may be slow. This is only a concern for large matrices since it was found that, for matrices of order 4-6, a desktop PC could easily run 10,000 iterations in less than 10 seconds. As convergence is achieved, the determinant of **J** tends to zero. If a large number of starting values are tested in an automated

fashion, then the algorithm can be made to exit if $\det(\mathbf{J}) < \varepsilon$. A suitably small value for ε is chosen depending on the desired level of convergence.

Once the algorithm has been run, it is trivial to reverse the spectral shift, once again obtaining eigenvalues in the familiar range. It should be noted that in general the possibility exists for a multiplicity of solutions. If complex solutions are allowed then there could be up to n! solutions. For the bicyclo-[1.1.1]-pentane system, for which 4 states were included, this corresponds to 24 solutions. In practice with the restriction that the eigenvalues are real, there tend to be only one or two. Furthermore, since the solutions represent the physical states of the system, it is generally straightforward to select the solution which corresponds to the physically realistic eigenvectors. If the approximate diagonal elements used for $\mathbf{D^0}$ are not close enough to allow convergence, or result in convergence to an unreasonable solution, then it is desirable to attempt to find all of the real solutions. The use of each possible permutation of the experimentally observed eigenvalues as starting points will generally yield the various real solutions. Note that for a matrix of dimension 4 this corresponds to 24 permutations, but with order 6, there are already 720. It is not difficult to automate the algorithm itself, but it can prove difficult to automate the sorting process. If a large number of solutions are to be examined automatically, then some weight must be assigned to the convergence and some to the physical sensibility of the eigenvectors. If it is known in advance that one or two of the transitions should have a large component of certain basis vectors, then vectors can be constructed which, when multiplied by the computed matrix of eigenvectors, will produce a large value

when these components are large. Thus a test function is created, composed of the eigenvector test, plus a term that is large if the convergence is good. After all of the permutations have been run through the algorithm, the solution corresponding to the maximum value of this test function is viewed. This should be the solution which best fits the criteria of good convergence, and physically realistic eigenvectors.

For the 4-dimensional problem it is easier to "walk in" the solution. In this method, permutations of the experimental eigenvalues or estimated values are entered as the diagonal elements, and the eigenvalues and vectors are obtained immediately without running the algorithm. The permutation which produces realistic eigenvectors, but with eigenvalues different from the experimental values, is retained. These values may not be close enough for the algorithm to converge to a solution. However, one can make up intermediate values close to those obtained, and in the direction of the desired values. If the algorithm converges to those values, the diagonal elements obtained are used as the starting point for the next step. Slowly, diagonal elements are obtained which produce eigenvalues closer and closer to the desired set. At each step, one must make sure that the eigenvectors have not diverged towards a different solution. If they do, then one must go back and use smaller steps. In this fashion, it is relatively straightforward to maintain the correct set of eigenvectors (they evolve as well but should stay realistic), while adjusting the diagonals toward the solution.

Self-Consistent Darling-Dennison Constant

In the work described in this thesis, an additional complication arose. The Darling-Dennison constant, one of the off-diagonal elements, could be computed from the *ab initio* quartic force-field. The computation of this constant is discussed in appendix E. However, it should be computed from the deperturbed frequencies. Since this constant describes a resonance interaction, it generally tends to depend strongly on the deperturbed frequencies of one or two modes. The catch is that the deperturbed frequencies that are necessary for calculation of the Darling-Dennison constant are themselves obtained by diagonalizing the interaction matrix, which *includes* the Darling-Dennison constant. This apparent difficulty is actually rather common in theoretical chemistry. The Hartree-Fock equations are a pseudo-eigenvalue problem in which the operator evolves along with the eigenfunctions. Thus the solution obtained after numerous iterations is that of the self-consistent-field. Although no precedent was found for self-consistent Darling-Dennison constants, this was clearly the correct approach. Subsequent discussions at the CSC-2003 conference revealed that other researchers also make sure to produce self-consistent constants.[125] When the inverse-eigenvalue algorithm is performed using a Darling-Dennison constant computed from the quartic force-field, a set of deperturbed frequencies are obtained. These are in turn used to compute an improved Darling-Dennison constant. Using the new Darling-Dennison constant, a new set of deperturbed frequencies is computed with the algorithm. The new deperturbed frequencies are used to compute a further improved Darling-Dennison constant. Ultimately the

Darling-Dennison constant and the deperturbed frequencies become self-consistent. This procedure required about 15 iterations to achieve self-consistency to three decimal places for the 4-dimensional system describing bicyclo-[1.1.1]-pentane. The convergence was oscillatory in this particular case. Fortunately the Darling-Dennison interaction was not so strong as to influence the convergence of the inverse-eigenvalue algorithm. If the Darling-Dennison changed by too much then this could be a problem. Although this sounds rather complex and almost subjective, it really isn't. The inverse-eigenvalue algorithm only adjusted the diagonal elements by small amounts to produce agreement. The values of the diagonal elements produced, corresponding to overtone transitions, were very reasonable based on the frequencies of the fundamentals. Furthermore, the adjustment of the Darling-Dennison constant was almost an aesthetic touch since the distributed intensity values were very insensitive to the value of this constant. In fact, the Darling-Dennison constant could be varied by 25%, and still produce less than 0.50% variation in the intensity assigned to the two fundamental modes in the polyad.

The inverse-eigenvalue algorithm is quite simple and could be programmed in a number of ways. Shown below is an example of how this was done in MathCad. This example corresponds to the method in which the diagonals were "walked in". Thus the estimated diagonal elements shown below are already refined, since this example was printed after convergence was achieved.

Inverse Eigenvalue Algorithm

$i := 0, 1 .. 3$

Define matrix A of off-diagonals

$$A := \begin{pmatrix} 0 & 8.673 & -23.348 & 17.815 \\ 8.673 & 0 & 0 & 8.967 \\ -23.348 & 0 & 0 & -24.464 \\ 17.815 & 8.967 & -24.464 & 0 \end{pmatrix}$$

Define experimental eigenvalues ω

$$\omega := \begin{pmatrix} 3022.5 \\ 2979.6 \\ 2887.6 \\ 2928.6 \end{pmatrix}$$

Apply spectral shift

$$\omega := \omega - \frac{\displaystyle\sum_i \omega_i}{4}$$

$$\omega = \begin{pmatrix} 67.925 \\ 25.025 \\ -66.975 \\ -25.975 \end{pmatrix}$$

Permutation of eigenvalues, which produced realistic eigenvectors

$$\omega := \begin{pmatrix} 25.025 \\ 67.925 \\ -25.975 \\ -66.975 \end{pmatrix}$$

Refined diagonal elements

$$D0 := \begin{pmatrix} 54.97874 \\ 26.897796 \\ -46.817608 \\ -35.058928 \end{pmatrix}$$

D := diag(D0)

Construction of Matrix J

J1 := $\overrightarrow{\text{eigenvecs}(A + D)}$

J1 := $(J1)^2$

J1 := $J1^T$

λ := eigenvals(A + D)

Inverse eigenvalue algorithm including exit command if det(J) becomes small, signaling convergence. n is the maximum number of iterations.

inverse(n) :=
$\begin{vmatrix} D0 \leftarrow D0 \\ D \leftarrow D \\ J1 \leftarrow J1 \\ \lambda \leftarrow \lambda \\ \text{for } t \in 1..n \\ \quad \begin{vmatrix} D \leftarrow \text{diag}(D0) \\ \lambda \leftarrow \text{eigenvals}(A + D) \\ D0 \leftarrow D0 + J1^T \cdot (\omega - \lambda) \\ J1 \leftarrow \text{eigenvecs}(A + D) \\ J1 \leftarrow \overrightarrow{(J1)^2} \\ J1 \leftarrow J1^T \\ \text{break if } ||J1|| \leq 10^{-6} \end{vmatrix} \\ \left[D0 \quad \lambda + 2954.575 \sum_i \left[(\lambda)_i - (\omega)_i \right]^2 \quad J1^T \right] \end{vmatrix}$

Output includes deperturbed frequencies, re-shifted eigenvalues, sum of square differences from experimental eigenvalues (convergence), and matrix of squared eigenvectors representing probabilities.

$$\text{inverse}(12000) = \left[\begin{pmatrix} 54.97874 \\ 26.897796 \\ -46.817608 \\ -35.058928 \end{pmatrix} \begin{pmatrix} 2979.6 \\ 3022.5 \\ 2928.6 \\ 2887.6 \end{pmatrix} \ 0 \ \begin{pmatrix} 0.063292 & 0.829974 & 0.101644 & 0.00509 \\ 0.930398 & 0.059486 & 0.006134 & 0.003982 \\ 0.004556 & 0.055427 & 0.29905 & 0.640966 \\ 0.001754 & 0.055113 & 0.593172 & 0.349962 \end{pmatrix} \right]$$

The order of basis states expressed here is:

(2 ν3)   (ν1)   (ν2)   (2 ν10)

References

(1) T. Helgaker, P. Jorgensen, J. Olsen. Molecular Electronic Structure Theory. Wiley (2000).

(2) T.J. Lee, P.R. Taylor. *Int. J. Quant. Chem. Symp*. **23**, 199 (1989).

(3) A.D. Becke. *J. Chem. Phys*. **98**, 5648 (1993).

(4) A.D. Becke. *Phys. Rev. A* **38**, 3098 (1988).

(5) J.P. Perdew, J.A. Chevary, S.H. Vosko, K.A. Jackson, M.R. Pederson, D.J. Singh, C. Fiolhais. *Phys. Rev. B* **46**, (1992).

(6) W. Kohn, L.J. Sham. *Phys. Rev*. **140**, A1133 (1965).

(7) (a) F.B. van Duijneveldt, J.G.C.M. van Duijneveldt-van de Rijdt, J.H. van Lenthe. *Chem. Rev*. **94**, 1873 (1994). (b) S.F. Boys, F. Bernardi. *Mol. Phys*. **19**, 553 (1970).

(8) R.F.W. Bader. Atoms in Molecules: a Quantum Theory; The International Series of Monographs on Chemistry 22; Clarendon Press: Oxford (1990).

(9) W. L Hase, L. Sun, K. Song. *Science* **296**, 875 (2002)

(10) C. K. Regan, S.L. Craig, J.I. Brauman. *Science* **295,** 2245 (2002).

(11) D. H. Ball, F.W. Parrish. *Adv. Carbohydr. Chem. Biochem*. **24**, 139 (1969).

(12) M.-C. Wu, L. Anderson, C.W. Slife, L.J.J. Jensen. *Org. Chem*. **39**, 3014 (1974).

(13) (a) L.A. Mulard, P. Kovác, C.P.J. Glaudemans. *Carbohydr. Res*. **259**, 117 (1994). (b) M.A. Ali, L. Hough, A.C. Richardson. *Carbohydr. Res*. **216**, 271 (1991).

(14) A.C. Richardson. *Carbohydr. Res*. **10**, 395 (1969).

(15) P.M. Collins, R.J. Ferrier. Monosaccharides. Their Chemistry and Their Roles in Natural Products. Wiley: Chichester, (1995).

(16) D. Seebach, H.-F. Chow, R.F.W. Jackson, M.A. Sutter, S. Thaisrivongs, J. Zimmermann. *Liebigs Ann. Chem*. 1281 (1986).

(17) S.L. Craig, J.I Brauman. *J. Am. Chem. Soc*. **121**, 6690 (1999).

(18) S. Nadkarni, N.R.J Williams. *J. Chem. Soc*. 3496 (1965).

(19) P. Hohenberg, W. Kohn. *Phys. Rev. B* **136**, 864 (1964).

(20) C. Moller, M. S. Plesset. Phys. Rev . **46**, 618 (1934).

(21) (a) J.A. Pople, R. Krishnan, H.B. Schlegel, J.S. Binkley. Int. J. Quant. Chem. **XIV**, 545 (1978). (b) R.J. Bartlett, G.D. Purvis. *Int. J. Quant. Chem*. **14**, 516 (1978).

(22) R.J. Boyd, S.L. Boyd. *Am. Chem. Soc*. **114**, 1652 (1992).

(23) (a) C. Gonzalez, H.B. Schlegel. *J. Chem. Phys. 90, 2154 (1989). (b)* C. Gonzalez, H.B. Schlegel. *J. Phys. Chem*. **94**, 5523 (1990).

(24) J.B. Foresman, T.A. Keith, J.B. Wiberg, J. Snoonian, M.J.J. Frisch. *Phys. Chem*. **100**, 16098 (1996).

(25) F. Biegler-König, J. Schönbohm, D.J. Bayles. *Comput. Chem*. **22**, 545 (2001).

(26) I. Lee, C.K. Kim, D.S. Chung, B.-S. Lee. *J. Org. Chem*. **59**, 4490 (1994).

(27) F.M. Aicken, P.L.A. Popelier. *Can. J. Chem*. **78**, 415 (2000).

(28) J.I. Seeman. *Chem. Rev*. **83**, 83 (1983).

(29) See equation (20) in reference (28).

(30) W.H. Miller, N.C. Handy, J.E. Adams. *J. Chem. Phys*. **72**, 99 (1980).

(31) W.H. Miller. *J. Chem. Phys*. **76**, 4904 (1982).

(32) H.W. Wang, W.L. Hase. *Chem. Phys*. **212**, 247 (1996).

(33) (a) M.-C. Brochier-Salon, C. Morin. *Magn. Reson. Chem*. **38**, 1041 (2000). (b) G.D. Rockwell, T.B. Grindley. *J. Am. Chem. Soc*. **120**, 10953 (1998). (c) J.I. Padrón, E.Q. Morales, J.T. Vázquez. *J. Org. Chem*. **63**, 8247 (1998). (d) E. Morales, J.I. Padrón, M. Trujillo, J.T. Vázquez. *J. Org. Chem*. **60**, 2537 (1995). (e) S.K. Gregurick, S.A. Kafafi. *J. Carbohydr. Chem*. **18**, 867 (1999). (f) A. Lubineau, M.C. Scherrmann, J. Mentech. European Patent Application EP 0 587 471 A1, Chem. Abs. 121, 158095. (g) F. Zuccarello, G. Buemi. *Carbohydr. Res*. **273**, 129 (1995). (h) J.W. Brown, B.D. Wladkowski. *J. Am. Chem. Soc*. **118**, 1190 (1996). (i) B.D. Wladkowski, S.A. Chenoweth, K.E. Jones, J.W. Brown. *J. Phys. Chem*. **102**, 5086 (1998).

(34) M. Hoffmann, J. Rychlewski. *J. Am. Chem. Soc*. **123**, 2308 (2001).

(35) I. Tvaroska, F.R. Taravel, J.P. Utille, J.P. Carver. *Carbohydrate. Res*. **337**, 353 (2002).

(36) S. Gronert, L.M. Pratt, S. Mogali. *J. Am. Chem. Soc*. **123**, 3081 (2001).

(37) P. Pechukas. *Ann. Rev. Phys. Chem*. **32,** 159 (1981).

(38) Molecular Thermodynamics, McQuarrie and Simon (1999).

(39)   C.J. Cramer, G.G. Truhlar. *Chem. Rev*. **99**, 2161 (1999).

(40)   L. Onsager. *J. Am. Chem. Soc*. **58**, 1486 (1936).

(41)   B.C. Garrett, D.G. Truhlar, R.S. Grev, A.D. Magnusson. *J. Phys. Chem*. **84**, (1980).

(42)   D.G. Truhlar. *J. Chem. Phys*. **53**, 2041 (1970).

(43)   D.G. Truhlar, A. Kuppermann. *J. Am. Chem. Soc*. **93**, 1840 (1971).

(44)   R.T. Skodje, D.G. Truhlar, B.C. Garrett. *J. Chem. Phys*. **77**, 5955 (1982).

(45)   Smekal, *Naturwis*s **11**, 873 (1923) (b) C.V. Raman, K.S. Krishnan. *Nature* **121**, 501 (1928).

(46)   D.P. Craig, T. Thirunamachandran. Molecular Quantum Electrodynamics. Dover (1984).

(47)   G. Placzek. U.S. Atomic Energy Commission, UCRL-Trans-524(L), (1962); [translated from Handbuch der Radiologie, 2nd ed., edited by E. Marks, Vol. **6**, 204 (1934).

(48)   K.J. Miller. *J. Am. Chem. Soc*. **112**, 8533 (1990).

(49)   G. Maroulis. *J. Chem Phys*. **97**, 4188 (1992).

(50)   G. Maroulis. *J. Phys. Chem*. **100**, 13466 (1996).

(51)   G. Maroulis, D. Xenides, U. Hohm, A. Loose. *J. Chem. Phys*. **115**, 7957 (2001).

(52)   G. Maroulis, C. Makris, U. Hohm, W. Wachsmuth. *J. Phys. Chem*. A, **103**, 4359 (1999).

(53)   G.J.B. Hurst, M. Dupuis, E. Clementi. J. Chem. Phys. **89**, 385 (1988).

(54)   M.A. Spackman. *J. Phys. Chem*. **93**, 7594 (1989).

(55)   C.W. Bauschlicher Jr., P.R. Taylor. *Theor Chim Acta* **71**, 263 (1987)

(56)   J. Neugebauer, M. Reiher, B.A. Hess. *J. Chem. Phys*. **117**, 8623 (2002).

(57)   M. Dory, L. Beudels, J.G. Fripiat, J. Delhalle, J. M. André, M. Dupuis. *Internat. J. Quantum Chem*. **42**, 1577 (1992).

(58)   K.M. Gough. *J. Chem. Phys*. **91**, 2424 (1989)

(59)   K.M. Gough, H.K. Srivastava, K. Belohorcová. *J. Chem. Phys*. **98**, 9669 (1993).

(60)   K.M. Gough, H.K. Srivastava, K. Belohorcová. *J. Phys. Chem*. **98** 771 (1994).

(61)   K.M. Gough, H.K. Srivastava. *J. Phys. Chem*. **100** 5210 (1996).

(62)   D. Bermejo, S. Montero. *J. Chem. Phys*. **81**, 3835 (1984).

(63)   K.M. Gough, W.F. Murphy. *J. Chem. Phys*. **85** 4290 (1986).

(64)   K.M. Gough, W.F. Murphy. *J. Chem. Phys*. **87** 3341 (1987).

(65)   K.M. Gough, J. R. Dwyer. *J. Phys. Chem*. A **102** 2723 (1998).

(66)   K.M. Gough, J.R. Dwyer, R. Dawes. *Can. J. Chem*. **78** 1035 (2000).

(67)   J. Gerratt, I.M. Mills. *J. Chem. Phys*. **49**, 1719 (1968).

(68)   P. Pulay. *J. Chem. Phys*. **78**, 5043 (1983).

(69)   E.B. Wilson, J.C. Decius and P.C. Cross. Molecular Vibrations, (1955).

(70)   K.M. Gough, C. Lupinetti, R. Dawes, in press

(71)   K.B. Wiberg, D.S. Conner, G.M. Lampman. *Tet. Lett*. 531 (1964).

(72)   M.D. Levin, P. Kaszynski, J. Michl. *Chem. Rev*. **100**, 169 (2000).

(73)   P. Kaszynski, J. Michl. *J. Org. Chem*. **53**, 4593 (1988).

(74)   K.B. Wiberg, S.T. Waddell. *J. Am. Chem. Soc*. **112**, 2194 (1990).

(75)   Uwe Bunz. Beitrage zur Chemie des Bicyclo(1.1.1)pentans. *Universität München* (1990).

(76)   P. Lazzeretti, M. Malagoli, R. Zanasi, E.W. Della, I.J. Lochert, C.G. Giribet, M.C. Ruiz de Azua, R.H. Contreras. *J. Chem. Soc. Faraday Trans*. **91,** 4031 (1995).

(77)   E.W. Della, J. Tsanaktsidis. *Aust. J. Chem*. **39**, 2061 (1986).

(78)   E.W. Della, D.K. Taylor. *Aust. J. Chem*. **44**, 881 (1991).

(79)   E.W. Della, D.K. Taylor. *J. Org. Chem*. **59**, 2986 (1994).

(80)   D.H.R Barton, N. Ozbalik, B. Vacher. *Tetrahedron* **44**, 7385 (1988).

(81)   E.W. Della, D.K. Taylor. *Aust. J. Chem*. **43**, 945 (1990).

(82)   P.E. Eaton, Y. Xiong, J.P. Zhou. *J. Org. Chem*. **57**, 4277 (1992).

(83)   D. Kaur et. al. *Applied Optics* **29**, 119 (1990).

(84)   D. Herriott, H. Kogelnik, R. Kompfner. *Applied Optics* **3**, 523 (1964).

(85)   S.K. Deb, M.L. Bansal, A.P. Roy. *Appl. Spectroscopy* **38**, 500 (1984).

(86)   K.B. Wiberg, R.E. Rosenberg, S.T. Waddell. *J. Phys. Chem*. **96**, 8293 (1992).

(87)   C. Van Caillie, R. Amos. *Phys. Chem. Chem. Phys*. **2**, 2123 (2000).

(88)   S. Montero. *J. Chem. Phys*. **79**, 4091 (1983).

(89)   S. Montero. *J. Chem. Phys*. **77**, 23 (1982).

(90)   W. Adcock et al. *J. Org. Chem*. **64**, 2618 (1999).

(91)   C.G. Giribet et al. *J. Comp. Chem*. **19**, 181 (1998).

(92)   E.W. Della, I. J. Lochert, J.E Peralta, R.H. Contreras. *Mag. Res. Chem*. **38**, 395 (2000).

(93)   J. Martin, S. Montero. *J. Chem. Phys*. **80**, 4610 (1984).

(94)   CADPAC: The Cambridge Analytic Derivatives Package Issue 6, Cambridge, 1995. A suite of quantum chemistry programs developed by R. D. Amos with contributions from I. L. Alberts, J. S. Andrews, S. M. Colwell, N. C. Handy, D. Jayatilaka, P. J. Knowles, R. Kobayashi, K. E. Laidig, G. Laming, A. M. Lee, P. E. Maslen, C. W. Murray, J. E. Rice, E. D. Simandiras, A. J. Stone ,M.-D. Su and D. J. Tozer

(94)   B. Kirtman, M. Hasan. *J. Chem. Phys*. **96**, 470 (1992).

(95)   S. Millefiori. A. Alparone. *J. Molec. Structure (Theochem)* **422**, 179 (1998).

(96)   W.H. Green Jr, A. Willetts, D. Jayatilaka, N.C. Handy. *Chem. Phys. Lett*. **169**, 127 (1990).

(97)   Willetts, N.C. Handy, W.H. Green Jr., D. Jayatilaka. *J. Phys Chem*. **94**, 5608 (1990).

(98)   W. Schneider, W. Theil. *Chem. Phys. Lett*. **157**, 367 (1989).

(99)   P.E. Maslen, N.C. Handy, R.D. Amos, D. Jayatilaka. *J. Chem. Phys*. **97**, 4233 (1992).

(100)   J.M.L.Martin, T.J. Lee, P.R. Taylor. *J. Chem. Phys*. **108**, 676 (1998).

(101)   D.A. Clabo Jr., W.D. Allen, R.B. Remington, Y. Yamaguchi, H.F. Schaefer III. *Chem. Phys*. **123**, 187 (1988).

(102)   A.J. Sadlej. *Theoret. Chim. Acta*, **79**, 123 (1991).

(103) M. van Faassen, P.L. de Boeij, R. van Leeuwen, J.A. Berger, J.G. Snijders. *J. Chem. Phys*. **118**, 1044 (2003).

(104) Z-L. Cai, K. Sendt, J.R. Reimers. *J. Chem. Phys*. **117**, 5543 (2002).

(105) M. Pecul, S. Coriani. *Chem. Phys. Lett*. **355**, 327 (2002).

(106) M. Pecul, A. Rizzo. *J. Chem. Phys*. **116**, 1259 (2002).

(107) M. Gussoni, M. Rui, G. Zerbi, *J. Molec. Structure* **447**, 163 (1998).

(108) D.M. Bishop. *Reviews of Modern Physics*, **62**, 343 (1990).

(109) K.M. Gough, W.F. Murphy. *J. Chem. Phys*. **87** 1509 (1987).

(110) Miani, V. Hänninen, M. Horn, L. Halonen. Mol. Phys. **98**, 1737 (2000).

(111) SPECTRO: A program for the derivation of spectroscopic constants from provided quartic force fields and cubic dipole fields, J.F. Gaw, A. Willetts, W.H. Green, N.C. Handy, Advances in Molecular Vibrations and Collision Dynamics Vol 1B (1991)

(112) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, J. A. Pople. Gaussian 03, Revision A.1, Gaussian, Inc., Pittsburgh PA, 2003.

(113) H.H. Nielsen. *Rev. Mod. Phys*. **23,** 90 (1951).

(114) I.M. Mills, A.G. Robiette. *Mol. Phys*. **56,** 743 (1985).

(115) G. Amat, H.H. Nielsen. *J. Mol. Spect*. **2**, 152 (1958).

(116) K.K. Lehmann. *J. Chem. Phys*. **84,** 6524 (1986).

(117) K.K. Lehmann. *Mol. Phys*. **66,** 1129 (1989). (b) A. Miani, E Cane, P. Palmieri, A. Trombetti, N.C. Handy, *J. Chem. Phys* **112**, 248 (2000).

(118) R. G. Della Valle. *Mol. Phys*. **63,** 611 (1988).

(119) A.F. Borro, I.M. Mills, A. Mose. *Chem. Phys*. **190**, 363 (1995).

(120) T.M. Luke. *J. Phys. B*. **18**, 589 (1985).

(121) T.M. Luke. *J. Phys. B*. **19**, 843 (1986).

(122) A.S. Deakin, T.M. Luke. *J. Phys. A*. **25**, 635 (1992).

(123) P. Palmieri, R. Tarroni, M.M. Huhn, N.C. Handy, A. Willetts. *Chem. Phys*. **190**, 327 (1995).

(124) Personal discussion with Dr. Allan East regarding Dr. Wesley Allen.

Supplementary Information

The complete set of eigenvectors of the mass-weighted force-field are listed here for BCP. The normal modes are those listed in Table 7 at the end of the Raman spectra section. Here the translations and rotations are included (modes 34-36, and 37-39 respectively), although they have not been constructed in any particular relation to the molecule coordinates since they were not used for any analysis. The vibrational eigenvectors are constructed according to the conventions of spectroscopy as discussed for Table 7.

The geometry is defined in Cartesian coordinates as X below. Atoms 1-5 are carbons while atoms 6-13 are hydrogens.

X=

1	0.000000	0.000000	0.938224
2	0.000000	0.000000	-0.938224
3	0.000000	1.239197	0.000000
4	1.073176	-0.619598	0.000000
5	-1.073176	-0.619598	0.000000
6	0.901553	1.852009	0.000000
7	-0.901553	1.852009	0.000000
8	1.153111	-1.706772	0.000000
9	2.054663	-0.145237	0.000000
10	-2.054663	-0.145237	0.000000
11	-1.153111	-1.706772	0.000000
12	0.000000	0.000000	-2.026812
13	0.000000	0.000000	2.026812

Frequencies	mode 1	mode 2	mode 3	mode 4	mode 5
0.0000	0.000000	0.000000	0.000000	0.000000	0.000000
0.8960	0.000000	0.000000	-0.000001	0.000002	-0.000002
1.0897	-0.197643	-0.052729	-0.098913	0.603353	0.290410
12.7582	0.000000	0.000000	0.000000	0.000000	0.000000
12.7862	0.000000	0.000000	-0.000001	0.000002	-0.000002
14.8306	0.197643	0.052729	0.098913	-0.603353	-0.290410
540.8951	0.000000	0.000000	0.000000	-0.000001	0.000000
540.8981	-0.047706	0.117558	-0.193640	0.198817	-0.490115
773.3130	0.000000	0.000000	0.000000	0.000000	0.000000
773.3131	-0.041325	0.101830	-0.167691	0.172164	-0.424454
843.7057	0.023836	-0.058790	0.096818	-0.099414	0.245057
897.0754	0.000000	0.000000	0.000000	0.000000	0.000000
897.0785	0.041325	-0.101830	0.167691	-0.172164	0.424455
905.9747	0.023836	-0.058790	0.096818	-0.099412	0.245058
968.2267	0.000000	0.000000	0.000000	0.000000	0.000000
1003.6714	0.100247	-0.319086	-0.232035	-0.029840	-0.006749
1028.4003	0.068118	-0.205143	0.300535	0.115191	-0.127855
1028.4006	0.000000	0.000000	0.000000	0.000000	0.000000
1115.6545	-0.100247	0.319086	0.232035	0.029841	0.006749
1115.6555	0.068118	-0.205143	0.300535	0.115193	-0.127854
1118.3158	0.000000	0.000000	0.000000	0.000000	0.000000
1142.7491	0.008865	-0.018119	0.376270	0.114690	-0.107354
1142.7494	-0.120834	0.378983	0.050679	-0.031753	0.069772
1214.7787	0.000000	0.000000	0.000000	0.000000	0.000000
1214.7791	0.109153	-0.337274	0.144246	0.084829	-0.114100
1250.7157	0.052776	-0.173802	-0.351200	-0.083431	0.058081
1256.6953	0.000000	0.000000	0.000000	0.000000	0.000000
1256.6955	-0.109153	0.337275	-0.144246	-0.084830	0.114100
1503.2196	0.052776	-0.173802	-0.351200	-0.083431	0.058082
1503.2212	0.000000	0.000000	0.000000	0.000000	0.000000
1553.0368	-0.008865	0.018119	-0.376271	-0.114689	0.107354
3035.9292	-0.120834	0.378985	0.050679	-0.031753	0.069772
3035.9300	0.000000	0.000000	0.000000	0.000000	0.000000
3041.3445	0.000000	0.000000	0.000000	0.000000	0.000000
3084.0714	0.000000	0.000000	0.000000	0.000005	-0.000002
3089.6428	-0.643005	-0.211344	0.038941	-0.184929	-0.078510
3093.8650	0.000000	0.000000	0.000000	0.000000	0.000000
3098.6811	0.000000	0.000000	0.000000	0.000005	-0.000002
3098.6819	0.643005	0.211344	-0.038941	0.184929	0.078510

mode 6	mode 7	mode 8	mode 9	mode 10	mode 11	mode 12	mode 13
0.000001	0.000003	0.004211	-0.001761	-0.006319	-0.420959	-0.032341	-0.292756
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000001	0.000000
0.000001	0.000003	0.004211	-0.001761	-0.006319	-0.420959	-0.032341	-0.292756
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	-0.000001	0.000000
0.185645	-0.266817	0.259796	-0.008304	-0.008059	0.148274	0.459590	-0.020621
0.000000	0.000000	0.000000	0.000000	-0.000001	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
-0.092796	0.133409	0.067538	0.130467	-0.184213	0.199230	-0.054345	0.372872
-0.160731	0.231079	0.111040	-0.080118	0.101704	-0.029419	0.296719	-0.227178
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
-0.092797	0.133409	0.067538	0.130467	-0.184214	0.199230	-0.054346	0.372871
0.160733	-0.231079	-0.111039	0.080118	-0.101704	0.029418	-0.296720	0.227177
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
-0.319997	0.146981	-0.450879	0.012220	0.002729	-0.048432	-0.132425	-0.005180
-0.217014	-0.330856	-0.309744	0.009057	-0.013372	0.125686	0.387142	-0.004729
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
-0.319997	0.146981	-0.450879	0.012219	0.002731	-0.048432	-0.132424	-0.005180
0.217014	0.330856	0.309744	-0.009056	0.013375	-0.125686	-0.387141	0.004730
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
-0.027933	-0.360023	0.017634	-0.020164	0.480808	-0.021092	-0.235728	0.102662
0.385533	0.038141	-0.258899	0.495108	0.049704	-0.019330	0.069387	-0.067011
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.347839	0.213039	-0.260786	-0.424913	0.187143	0.090826	0.091105	0.102683
0.168571	-0.292716	-0.121486	-0.224040	-0.432194	-0.076855	-0.138796	-0.057529
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.347844	0.213039	-0.260786	-0.424911	0.187144	0.090826	0.091104	0.102683
-0.168574	0.292716	0.121486	0.224039	0.432196	0.076855	0.138795	0.057529
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
-0.027934	-0.360023	0.017634	-0.020164	0.480810	-0.021092	-0.235729	0.102662
-0.385539	-0.038141	0.258898	-0.495106	-0.049704	0.019330	-0.069387	0.067011
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000002	-0.001776	0.003069	0.000680	0.487981	-0.216761	-0.441029
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000002	-0.001776	0.003069	0.000680	0.487981	-0.216761	-0.441029
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

mode 14	mode 15	mode 16	mode 17	mode 18	mode 19	mode 20	mode 21
-0.244278	-0.000001	0.000000	0.000001	0.000001	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000002	0.000000	0.321535	0.236472	-0.080037
0.000000	0.000000	-0.201752	-0.205905	0.490501	0.000001	0.000000	0.000000
-0.244278	0.000001	0.000000	-0.000001	-0.000001	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	-0.000002	0.000000	-0.321535	-0.236472	0.080037
0.000000	0.000000	-0.201752	-0.205905	0.490501	0.000001	0.000000	0.000000
0.434850	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.004432	0.330217	-0.219002	-0.265073	0.313306	-0.095896
-0.056260	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.283530	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	-0.000001	0.004432	0.330218	-0.219002	0.132534	-0.156653	0.047948
-0.056260	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
-0.283530	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000001	0.004432	0.330218	-0.219002	0.132535	-0.156654	0.047948
0.308439	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
-0.267240	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.408251	-0.001485	-0.310699	-0.235207	0.309021	-0.321231	0.222480
0.308439	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.267240	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	-0.408251	-0.001485	-0.310700	-0.235206	0.309021	-0.321231	0.222479
0.141694	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.095115	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.408249	-0.001485	-0.310702	-0.235206	-0.320423	-0.163793	-0.395948
-0.088745	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.230462	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	-0.408245	-0.001485	-0.310700	-0.235207	0.011407	0.485023	0.173468
-0.088745	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
-0.230462	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.408245	-0.001485	-0.310701	-0.235207	0.011406	0.485024	0.173468
0.141694	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
-0.095115	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	-0.408249	-0.001485	-0.310702	-0.235207	-0.320424	-0.163792	-0.395948
-0.074579	-0.000006	0.000000	0.000001	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000002	0.000000	0.381606	0.072076	-0.500334
0.000000	0.000000	0.677687	-0.066581	0.146628	0.000000	0.000000	0.000000
-0.074579	0.000006	0.000000	-0.000001	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	-0.000002	0.000000	-0.381606	-0.072076	0.500334
0.000000	0.000000	0.677687	-0.066581	0.146628	0.000000	0.000000	0.000000

mode 22	mode 23	mode 24	mode 25	mode 26	mode 27	mode 28	mode 29
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.427974	0.004211	0.001761	0.006320	0.420959	0.032346	0.292756	0.244282
0.000000	-0.000017	0.000004	-0.000003	-0.000001	-0.000018	0.000002	-0.000001
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
-0.427974	0.004211	0.001761	0.006320	0.420959	0.032346	0.292756	0.244282
0.000000	0.000017	-0.000004	0.000003	0.000001	0.000018	-0.000002	0.000001
0.000000	-0.000001	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.003427	-0.176734	0.242926	-0.216216	0.225648	-0.504035	0.219960
0.474625	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.111015	0.080110	-0.101706	0.029419	-0.296727	0.227181	-0.283533
0.000000	0.195725	-0.037946	0.066780	-0.165259	-0.288276	-0.110540	-0.271138
-0.237312	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	-0.111014	-0.080110	0.101707	-0.029419	0.296727	-0.227181	0.283533
0.000000	0.195723	-0.037946	0.066780	-0.165259	-0.288276	-0.110540	-0.271138
-0.237313	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	-0.011659	0.476444	0.325719	-0.003546	0.009744	-0.004748	-0.001151
0.000000	-0.011763	0.300808	-0.444386	-0.062633	0.052270	-0.138623	0.067512
0.188940	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.011663	-0.476444	-0.325719	0.003546	-0.009744	0.004748	0.001151
0.000000	-0.011765	0.300808	-0.444386	-0.062633	0.052270	-0.138623	0.067513
0.188940	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.039255	-0.009639	0.262652	0.141470	0.327492	0.057535	-0.363510
0.000000	-0.480334	0.268368	0.033689	-0.035293	-0.051031	-0.030778	-0.099228
0.070986	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	-0.419559	-0.261402	0.119847	-0.045285	-0.258087	0.067002	0.037929
0.000000	-0.201939	-0.136323	-0.259985	0.076624	0.275804	-0.030762	-0.329670
-0.259925	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.419555	0.261401	-0.119848	0.045285	0.258087	-0.067002	-0.037929
0.000000	-0.201937	-0.136323	-0.259987	0.076624	0.275805	-0.030762	-0.329670
-0.259925	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	-0.039255	0.009638	-0.262654	-0.141470	-0.327492	-0.057535	0.363510
0.000000	-0.480329	0.268367	0.033689	-0.035293	-0.051031	-0.030778	-0.099228
0.070986	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
-0.198967	-0.001776	-0.003069	-0.000679	-0.487982	0.216763	0.441028	0.074580
0.000000	-0.000056	0.000018	0.000001	0.000000	0.000005	-0.000001	0.000000
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.198967	-0.001776	-0.003069	-0.000679	-0.487982	0.216763	0.441028	0.074580
0.000000	0.000056	-0.000018	-0.000001	0.000000	-0.000005	0.000001	0.000000

mode 30	mode 31	mode 32	mode 33
0.321537	0.236471	-0.080037	0.427973
0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	-0.000001
-0.321537	-0.236471	0.080037	-0.427973
0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	-0.000001
0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000
0.000000	-0.000001	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000
-0.229558	0.271333	-0.083049	0.411038
0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000
0.229556	-0.271333	0.083049	-0.411037
0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000
0.191584	0.374593	0.328749	-0.191052
0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000
-0.191584	-0.374592	-0.328749	0.191052
0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000
0.171827	-0.465492	0.028291	0.259153
0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000
0.363410	-0.090900	0.357054	0.068101
0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000
-0.363408	0.090899	-0.357054	-0.068100
0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000
-0.171825	0.465492	-0.028291	-0.259152
0.381606	0.072074	-0.500333	-0.198969
0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000
-0.381606	-0.072074	0.500333	0.198969
0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000

Trans1	Trans2	Trans3	Rot1	Rot2	Rot3
0.018974	0.000163	0.419465	-0.059958	-0.384099	0.000730
0.419460	0.000354	-0.018974	0.384102	-0.059958	-0.000130
-0.000008	0.419890	-0.000002	-0.000343	0.000228	0.000000
0.018974	-0.000157	0.419461	0.059959	0.384103	0.000730
0.419460	-0.000338	-0.018974	-0.384102	0.059958	-0.000130
-0.000008	0.419890	-0.000002	-0.000343	0.000228	0.000000
0.018864	0.000003	0.420297	0.000000	0.000002	-0.476428
0.419460	0.000008	-0.018974	0.000000	0.000000	-0.000130
-0.000008	0.419433	-0.000002	-0.507790	0.079440	0.000000
0.019029	0.000003	0.419042	0.000000	0.000002	0.239313
0.419559	0.000008	-0.019701	0.000000	0.000000	0.413111
-0.000008	0.419936	-0.000005	0.321981	0.400085	0.000000
0.019029	0.000003	0.419042	0.000000	0.000002	0.239313
0.419369	0.000008	-0.018247	0.000000	0.000000	-0.413371
-0.000008	0.420302	0.000000	0.184781	-0.478840	0.000000
0.005451	0.000001	0.121922	0.000000	0.000001	-0.206426
0.121584	0.000002	-0.005674	0.000000	0.000000	0.100566
-0.000002	0.121443	-0.000001	-0.203120	0.141373	0.000000
0.005451	0.000001	0.121922	0.000000	0.000001	-0.206426
0.121537	0.000002	-0.005324	0.000000	0.000000	-0.100641
-0.000002	0.121532	0.000000	-0.236528	-0.072643	0.000000
0.005544	0.000001	0.121227	0.000000	0.000001	0.190664
0.121591	0.000002	-0.005725	0.000000	0.000000	0.128617
-0.000002	0.121811	-0.000001	0.223787	0.105234	0.000000
0.005502	0.000001	0.121532	0.000000	0.000001	0.016403
0.121616	0.000002	-0.005901	0.000000	0.000000	0.229227
-0.000002	0.121600	-0.000002	0.055147	0.241174	0.000000
0.005503	0.000001	0.121532	0.000000	0.000001	0.016403
0.121511	0.000002	-0.005096	0.000000	0.000000	-0.229302
-0.000002	0.121803	0.000001	-0.020965	-0.246409	0.000000
0.005541	0.000001	0.121227	0.000000	0.000001	0.190664
0.121533	0.000002	-0.005272	0.000000	0.000000	-0.128692
-0.000002	0.121925	0.000000	0.181083	-0.168332	0.000000
0.005499	-0.000099	0.121561	0.037545	0.240522	0.000212
0.121560	-0.000215	-0.005499	-0.240521	0.037545	-0.000038
-0.000002	0.121685	-0.000001	-0.000099	0.000066	0.000000
0.005499	0.000101	0.121563	-0.037545	-0.240521	0.000212
0.121560	0.000219	-0.005499	0.240521	-0.037545	-0.000038
-0.000002	0.121685	-0.000001	-0.000099	0.000066	0.000000

<center>**Synthesis**</center>

What follows is a description of the synthesis of BCP starting from the 1,3-dicarboxylic acid of BCP. This is easily prepared following the reviewed procedures in Organic Syntheses Vol 77 p250.

1,3-dimethyl ester: Step 1

The di-acid (24.6g, 157mmol) was refluxed (10h) in excess (45ml) thionyl chloride. This was then evaporated off using a water aspirator. Crude product obtained was purified using a Kugelrohr apparatus (120º C/12mmHg). Yield was about 90%.

Step 2: Product from step 1 was slowly added to excess anhydrous MeOH then refluxed for 30 mins. Evaporation of MeOH afforded a quantitative yield of the desired 1,3-dimethyl ester.

Half-ester (1-carboxylic acid,-3-methyl ester) of BCP

NaOH in MeOH (40ml, 2.5M, 100mmol) was added to the dimethyl ester in THF (600ml, .16M, 100mmol) dropwise at RT. Mixture was stirred overnight, then evaporated to dryness (vacuum, no heat). Water (240ml) was added, then extracted with $CHCl_3$. HCl was added to the aqueous layer to pH 3. Extractions were taken using $CHCl_3$ (4x80ml). Extracts were dried ($MgSO_4$), then dried under vacuum. A yield of ~90% of the desired 1-carboxylic acid,-3-methyl ester of BCP was obtained.

1-carboxylic acid of BCP

The half-ester obtained previously was refluxed in excess thionyl chloride for 10h then evaporated to dryness. The yield was quantitative. The acid-chloride product is dissolved in freshly distilled degassed benzene. One equivalent of N-hydroxypyridinethione sodium salt was suspended in benzene with 2 equivalents of t-butyl thiol and a dash of DMAP (catalyst). This was refluxed next to a 500W tungsten lamp for 1h. After workup the crude product was extracted with hot pentane to yield (~60%) the desired mono-acid.

Barton ester of BCP

The previously obtained monoacid was dissolved in dry CH_2Cl_2 (1.0 M). Slightly more than one equivalent of hydroxypyridinethione was added under Ar. The solution was cooled to 5º C and protected from light. 1 equivalent of DCC was added with stirring for 1.5h. Solution was filtered then dried to yield (~90%) the desired Barton ester of BCP.

BCP

All available quantities of the Barton ester of BCP were added to neat tri-n-butyl tin hydride. The reaction was gently stirred, and irradiated with a 500W tungsten lamp. The flask was swept with N_2, allowing BCP to be collected in a cold (-78º) trap. As stated previously, several cryo-distillations were required to remove the tri-n-butyl tin hydride that was also swept over.

Acknowledgements

My most grateful acknowledgement to all the people who made this thesis

possible:

Glossary of Terms

$\overline{\alpha}$	isotropic polarizability
ν_0	laser frequency (cm^{-1})
$\partial\sigma / \partial\Omega$	differential scattering cross-section
a_0	Bohr
AIM	Atoms In Molecules
AM1	Austin Model 1
AMU	atomic mass units
aug-cc-pVTZ	augmented correlation consistent triple zeta basis set
B3LYP	Becke's 3 parameter functional (Lee, Yang and Parr)
BSSE	basis set superposition error
c	speed of light
C	Coulomb
CCSD(T)	coupled-cluster with singles, doubles, and perturbative triples
dalphadq	$\partial\overline{\alpha} / \partial q$
dalphadqlength	$\partial\overline{\alpha} / \partial q$ units of length
DFT	density functional theory
f	field
GGA	generalized gradient approximation
h	Planck's constant
H	Hartrees
IPCM	isodensity polarizable continuum model

IRC	Intrinsic reaction coordinate
J	joules
k	rate
k_b	Boltzmann constant
LDA	local density approximation
m	meter
MEP	minimum energy path
MP	Moller-Plesset
NMR	nuclear magnetic resonance
q	dimensionless normal coordinate
redmass	reduced mass
SCF	self-consistent field
SCIPCM	self-consistent isodensity polarizable continuum model
T	temperature
TDDFT	time-dependent density functional theory
TS	transition structure
V	Volt
γ	anisotropic polarizability
ΔE	change in energy
ε_0	permitivity of free space
κ	reaction path curvature
μ	dipole
ν	frequency (cm^{-1})