# **Broadband Rotational Spectroscopy with Applications to Molecular Structure and Intermolecular Interactions**

Nathan Andrew Seifert Harrisonburg, Virginia

B. A., Amherst College, 2010

A Dissertation presented to the Graduate Faculty of the University of Virginia in Candidacy for the Degree of Doctor of Philosophy

Department of Chemistry

University of Virginia May, 2015

# **Table of Contents**

Abstra	ct	
Acknow	wledgements	4
Preface	2	6
Chapte	er 1: Introduction	
I. N	Microwave Spectroscopy	8
1)	Introduction	8
2)	Theory	9
Ret	ferences	
II.	Electronic Structure Theory	
1)	Introduction	
2)	Multi-Electron Systems	
3)	Hartree-Fock (HF) Theory	
4)	Perturbative Corrections to HF Theory	
5)	Density Functional Theory	40
6)	Basis Sets	44
Ret	ferences	49
III.	A Computational Approach to Intermolecular Interactions	52
1)	Introduction	52
2)	Elements of the SAPT0 Partition	56
3)	SAPT0 Results for Simple Models	61
Ret	ferences	66
IV.	Chirped-Pulse Fourier Transform Microwave Spectroscopy	68
1)	Historical Background	68
2)	The Chirped-Pulse FTMW Technique	71
3)	Applications of CP-FTMW Spectroscopy	81
Ret	ferences	84

-	er 2: AUTOFIT, an Automated Fitting Tool for Broadband Rotational	
-	a, and Applications to 1-Hexanal	
/	Introduction	
	Methodology	
a)	Experimental	
b)	Quantum Chemistry	
c)	AUTOFIT	
	Results	
	Discussion	
a)	1-Hexanal Conformational Energies	
b)	Decreasing AUTOFIT Run Times	
/	Conclusion	
	rences	
-	er 3: Phenol Dimer and Trimer – Structures from Broadband Rotation oscopy	
-	Introduction	
,	Experimental	
<i>,</i>	Results	
4)	Discussion	130
a)	Phenol Dimer	130
b)	Phenol Trimer	141
5)	Conclusion	145
Refe	rences	147
_	er 4: The Structure and Internal Rotation Dynamics of Sevoflurane…] 150	Benzene
1)	Introduction	150
2)	Experimental	152
a)	Spectroscopy	152
b)	Computational	154
3)	Results	155
4)	Discussion	163
a)	Structure	163
b)	Internal Rotation	169

5)	Conclusion	176	
Ref	ferences	179	
Chap	Chapter 5: The Structure of the Sevoflurane Dimer		
1)	Introduction	182	
2)	Experimental & Theoretical	184	
3)	Results & Discussion	187	
4)	Conclusion	196	
Ref	ferences	197	
Chap	oter 6: Structure Determination of the Carbonyl Sulfide (OCS) Tetrame	er 199	
1)	Introduction	199	
2)	Experimental	201	
3)	Spectroscopic Results	202	
4)	Structure Determination	206	
5) (	Conclusions	213	
Ref	ferences	215	
Conc	lusions	216	
1)	Summary	216	
2)	Future Developments	217	
Ref	ferences	219	

# Abstract

Chirped-pulse Fourier transform microwave (CP-FTMW) spectroscopy is a powerful technique for molecular detection, identification and structural characterization for gas-phase molecules and molecular complexes. The broadband nature of the technique allows for simultaneous detection of all species populated in the pulsed jet sample with sufficient intensity. Recent developments of the CP-FTMW technique at low frequency, namely 2-18 GHz, have ushered in a new generation for microwave spectroscopy-enabled structure determination of molecules and molecular complexes, where systems with as many as 20 heavy atoms can be structurally characterized to better than 0.1 Å precision.

A selection of studies illustrating the power of CP-FTMW spectroscopy in the 2-18 GHz region for structure determination are described. These studies include structural characterization of a molecule with high conformational flexibility as well as molecular clusters containing strong interplay between electrostatic and dispersive interactions, as well as weakly bound van der Waals binding characteristics. The sensitivity afforded by the CP-FTMW technique enables detection of isotopologues in natural abundance, which enables direct structure determination of a target molecular system by means of Kraitchman's equations or least-squares structure fitting algorithms. Use of modern quantum chemistry techniques for analyzing intermolecular interactions, such as symmetry adapted perturbation theory (SAPT) and dispersion corrected DFT, is emphasized, as well as its use for accurate and efficient structure prediction in the context of automated broadband spectral assignment.

## Acknowledgements

Although this document is a reflection of a personal journey through scientific research, the struggles and successes of such a journey is undoubtedly a communal experience; none of my development as a scientist could be attributed to solely myself.

First and foremost, I thank my advisor, Brooks Pate. I entered Brooks' lab with basically no knowledge of spectroscopy, and through Brooks' counsel and guidance I would have never discovered a field of chemical science that I have since gained so much respect for. His ambition, clarity of vision, and passion for high quality science has been an endless source of continuing inspiration and education for me. I have built my knowledge of physical chemistry from nearly nothing to something that I can be proud of, and there is no one as important as Brooks in my time at UVa for fostering a research environment where I could develop my skills as a budding scientist and spectroscopist.

I also give my eternal gratitude to my parents, who have spent tireless years supporting me through my intellectual and emotional development. To this end, I could not ask for a better support network. Even though our opinions have occasionally fallen on opposite poles throughout our lives together, my parents have never suppressed my thoughts or desires to rash ends and have always allowed me to build my own path towards adulthood. Although her presence in my life did not fully come to full fruition until mid-way through graduate school, I doubt I could have kept my head and spirits up through the toughest moments of my time at UVa without the endless support of my wife, Kat. Even when I return home in the evening bitter and frustrated, I can always find relief in her presence.

I must also give my deep appreciation to the people who I have worked with closely throughout these years. Amanda Steber, Cristobal Perez and Daniel Zaleski have always been a constant presence in my life and a source of support, laughter and cosponsored frustration. I hope that our paths will continue to cross throughout our careers. This gratitude, of course, extends to many others who I've shared my time at UVa with – Luca Evangelisti, Lorenzo Spada, Matt Muckle, Brent Harris, Justin Neill, amongst others.

So much of this work could not have come to fruition without the ambition and support of my research collaborators. Collaborators such as Alberto Lesarri, Sean & Rebecca Peebles, Gamil Guirgis, Emilio Cocinero, and many others, have lended their respect and support of my work and professional development, and I would not have been so successful without their involvement.

Finally, I thank the Department of Chemistry, the National Science Foundation and Brightspec for funding my years at UVa as well as the vast majority of the research presented in this dissertation.

# Preface

Anybody who has known me for a long enough time to receive correspondence from me probably knows that my truest love, above all else, is exposition. I've never claimed to be a fantastic writer, but I take the most pride in my best pieces of writing and find the source of my greatest disappointments and regrets in the worst. I find it extremely liberating to be at the keyboard with only my thoughts. Perhaps science was not the best choice for me in this light; science insists on, above all else, precision and succinctness. I find both to be a bit difficult to achieve since I have a great passion for expounding on the topics and ideas I care about -- maybe with a bit too much flourish. Unfortunately, science is one of the topics I love the most, so at times my relationship with science can be a bit frustrated.

Therefore, if the reader does not mind me to be a tad cliché, I have written this dissertation as a love letter to the research and education I have undertaken in the past five years. Although the research is an important part, I have spent much effort in developing the background information of the presented research since I find great pleasure in the theoretical developments of the generations of scientists both before and contemporary to me. As such, I have tried to explain most of the relevant concepts with an appropriate level of rigor and detail.

This is especially the case in Chapter 1, Section II, which focuses on electronic structure theory. Since I began my journey in spectroscopy, I have developed a deep appreciation for the incredible effort undertaken by theoreticians in devising the computational foundation of chemistry. As such, I have written a significant portion of text to explain the basic foundations of electronic structure theory, perhaps more so than expected in the usual spectroscopy dissertation. I hope the reader finds this instructive and not extraneous – I present it only because I have found my study and discovery of electronic structure theory to be nearly as important as the research presented in this dissertation to my personal development as a scientist. I dream of one day writing a textbook to summarize these incredible developments.

Finally, I would like to shortly reflect on my experience from my first five years in spectroscopy. One of the largest frustrations I had with synthetic organic chemistry – the object of my affection in my undergraduate studies and my first few months at UVa – is the lack of predictive power. In synthetic work, I often felt like I had to blindly search for answers and a way forward in a completely dark room. I found the trial and error nature of the work to be incredibly frustrating. However, the art of *molecular structure*, the construction and analysis of which is central to synthetic organic chemistry, is one of my favorite aspects in all of science.

Microwave spectroscopy, however, resolved all of these issues. Exploitation of the predictive power of computational chemistry and the innately high experimental precision of the technique are intoxicating. The ability to observe molecular structure directly with few or no assumptions is a central driving force for my continuance in this field. There is also the simplicity of the technique, the unambiguous experimental results, and the innate connection with the most fundamental phenomena in chemistry. It's almost guaranteed these are the reasons why it seems that most people who start in microwave spectroscopy never seem to leave.

# **Chapter 1: Introduction**

## I. Microwave Spectroscopy

#### 1) Introduction

Rotational spectroscopy is suited particularly for the study of molecular structure, as the rotational spectrum of a polyatomic molecule is directly encoded by the geometric distribution of its atoms. The only requirement for an allowed rotational spectrum is that the molecule has a permanent dipole moment. Especially for large molecules with more than 3 heavy atoms, rotational motion is generally uncoupled from the other molecular degrees of freedom, or at worst any couplings can be treated perturbatively. As such, the rotational spectrum of a molecule is often very easy to analyze and only requires a relatively simple effective Hamiltonian to achieve a quantitative experimental fit.<sup>1,2</sup>

Typically, pure rotational transitions are observed in the microwave or millimeterwave region (v < 1 THz), and large molecular systems with 10 or more heavy atoms can be structurally characterized using low-frequency microwave radiation below 18 GHz, as will be shown in the research presented in this dissertation. Experimentally, this is highly advantageous as the low frequencies required for detection allow for high frequency precision on observed transitions ( $\Delta v/v \approx 10^{-6}$ ) and there is a great availability of commercially available intense excitation sources, stable frequency modulators, and sensitive detectors thanks to the economy of scale provided by the RADAR and telecommunications industries.

#### 2) Theory

## a) Construction of the Rigid Rotor Hamiltonian

As the observed mechanics deal with overall rotation of a *rigid* molecular body, the Hamiltonian must involve quantization of the total angular momentum of a rotating body. Angular momentum is parameterized by the body's moment of inertia, defined for a general 3D body using the moment of inertia tensor. For a molecule with nuclei masses  $m_i$  and nuclear coordinates ( $x_i$ ,  $y_i$ ,  $z_i$ ), the elements of the inertial tensor can be calculated as:

$$\begin{cases} I_{ij} = \sum_{i} m_i (b_i^2 + c_i^2), & i = j \\ I_{ij} = -\sum_{i} m_i a_i b_i, & i \neq j \end{cases}$$
(1)

The resulting 3x3 matrix is symmetric (since  $I_{ab} = I_{ba}$ ), and for a general asymmetric body the off-diagonal elements are not necessarily zero. Now, consider a rigid molecule in the center of mass frame – implying that translational motion w.r.t. the center of mass is fully uncoupled from the rotational motion about the center of mass. Again, for a general molecular orientation in space, the moment of inertia tensor is not necessarily diagonal, so there will be off-diagonal contributions to the classical angular momentum, e.g.  $P_{ab} = I_{ab}\omega_{ab} \neq 0$ . However, since the inertial tensor is real and symmetric, it is guaranteed to be diagonalizable.

By applying the appropriate transformation matrix R, a new diagonal inertial tensor can be generated,  $I_{PA} = R^{-1}IR$ . This diagonal tensor represents a transformation of the original coordinate system to one that has only three non-zero moments of inertia,

each lying along the axes of the new system. This coordinate system is called the **principal axis** system, where the axes are typically notationally represented as {**a**, **b**, **c**}. This rotation allows representation of the total (rigid) rotational Hamiltonian as a sum of three angular momentum operators along the each of the three principal axes. The classical rotational energy for this diagonalized system is shown in equation (2).

$$E_{rot} = \frac{1}{2} \left( \frac{P_a^2}{I_a} + \frac{P_b^2}{I_b} + \frac{P_c^2}{I_c} \right)$$
(2)

With equation (2) in hand, the system can now be quantized by considering the eigenvalues of the equivalent quantum mechanical angular momentum operators.

From this point on, the chosen molecular system will be an *asymmetric top*, where all three diagonal moments of inertia are non-identical, and that the system is in the *prolate* case where  $I_a < I_b < I_c$ .

The problem will then ultimately consist of a Hamiltonian that contains momentum terms along all three principal axes, and then diagonalizing this Hamiltonian matrix using a basis of orthonormal *symmetric top* wavefunctions, represented by the basis |J, K, M>. In this basis, J is the quantum number associated with the magnitude of the total angular momentum (associated with the operator  $\hat{J}^2$ ), and K arises from the eigenvalues of the *a*-axis projection (molecule-fixed frame) of  $\hat{J}$  (operator  $\hat{J}_a$ ). M is correspondingly the quantum number associated with the space-fixed, z-axis projection of  $\hat{J}$ . Therefore for a rotation in the absence of any external field, the M states are degenerate and the energies are fully independent of the M quantum number, as the fieldfree Hamiltonian is rotationally invariant in the space-fixed frame. The Hamiltonian for the asymmetric top is,

$$H_{rot} = \sum_{m \in \{a,b,c\}} \frac{\hbar^2}{2I_m} \hat{J}_m^2 = \sum_{m \in \{a,b,c\}} B_m \hat{J}_m^2$$
(3)

Where the operator  $\hat{J}_m^2$  is the total angular momentum operator along one of the principal axes, and can be expressed as the standard raising and lowering operators in the standard  $[J^2, J_z] = 0$  representation, as shown below:

$$\widehat{J_a} = \frac{1}{2}(J_+ - J_-)$$
$$\widehat{J_b} = \frac{1}{2i}(J_- - J_+)$$
$$\widehat{J_c} = J_z$$

The matrix elements of these operators in the |J, K, M> basis are tabulated in other works, and are not necessary for this discussion. Since this Hamiltonian is infinite, the typical choice is to factor it into  $(2J+1) \times (2J+1)$  blocks for each value of the J quantum number, where K = -J, -J + 1, ..., 0, ..., J-1, J. For J = 2, the relevant block of the Hamiltonian is a 5x5 matrix.

The set of constants  $\{m \in \{a, b, c\} | B_m = \frac{\hbar^2}{2I_m}\}$  are called the **rotational constants** of the system, and they uniquely define the geometry and, consequently the eigenvalues, of the molecule as a rotating rigid body. In microwave spectroscopy, the rotational constants are typically reported in units of MHz, with the notation A / B / C for  $B_a / B_b / B_c$ , respectively. A common conversion factor found in the literature is  $I_m$  (amu Å<sup>2</sup>) = 505379.15/ $B_m$  (MHz).

At this point, the finite "J-block" Hamiltonian matrices can be diagonalized in order to compute the energy values for each value of K. However, the symmetric top quantum number K is not a good quantum number for asymmetric systems. Instead, each eigenvalue of the J-block Hamiltonian is assigned a value  $\tau = K_a - K_c$ , where  $K_a$  is the value of K in the *prolate* symmetric top basis ( $I_a < I_b = I_c$ ) and  $K_c$  is the value in the *oblate* symmetric top basis ( $I_a > I_b = I_c$ ). Therefore, each eigenvalue is associated with a specific state with quantum numbers J,  $\tau$  or more commonly in the literature (or exclusively in this dissertation) as J,  $K_a$  and  $K_c$ . For a given asymmetric top rotational transition (J',  $K_a$ ',  $K_c$ ')  $\rightarrow$  (J'',  $K_a$ '',  $K_c$ ''), the common notation is **J'** $K_a$ '' $K_c$ ''-**J''** $K_a$ '' $K_c$ ''.

Historically, diagonalization of matrices for large values of J was not computationally feasible, and certain tricks have been developed to approximate the eigenvalues of the asymmetric top Hamiltonian for large J. However, in the modern day, a simple input of a Hamiltonian matrix for a large J is computationally fast using multithreaded linear algebra libraries, so the historical approaches to calculating energy levels are largely irrelevant for discussion in this summary.<sup>3</sup>

#### b) Intensities

In general, the intensities of a rotational spectrum of a rigid (or near-rigid) polyatomic molecule can be separated into two distinct categories: the thermal statistical weights of a given rotational state (e.g. the Boltzmann factor) and the "intrinsic" quantum mechanical intensity, which is contained in a transition's dipole matrix element.

For a given transition (J',  $K_a$ ',  $K_c$ ')  $\rightarrow$  (J'',  $K_a$ '',  $K_c$ '') with frequency v(J'), the thermal weighting is equal to

$$I_{thermal} = \frac{(2J'+1)}{Q_{rot}(T, \{I_a, I_b, I_c\})} e^{\frac{-\nu(J')}{k_B T}}$$

where Q<sub>rot</sub> is the rotational partition function,

$$Q_{rot} = \frac{\pi^{1/2}}{\sigma} \prod_{j \in \{a,b,c\}} \left(\frac{8\pi^2 I_j kT}{h^2}\right)^{\frac{1}{2}}$$

and T is the "rotational temperature" of the experiment. In the case of a typical pulsed jet expansion experiment such as those described in this dissertation,  $T \sim 1-2$  K.

The intrinsic intensity is also calculable, but it requires knowledge of the matrix elements of the dipole operator in the asymmetric top basis. In general, for a pure rotational transition to occur, the molecule (or molecular system) must have a non-zero permanent dipole moment. For an asymmetric top it will be important to consider not the total magnitude of the dipole moment but the magnitude of its projections along each of the principal axes, which are typically written as { $\mu_a$ ,  $\mu_b$ ,  $\mu_c$ } and expressed in Debye (10<sup>-18</sup> statC m).

Since the dipole vector does not necessarily lie along **J**, one must consider how the dipole moment rotates along the molecular space-fixed axes. To represent the dipole moment along a space-fixed axis I, one can rewrite the dipole moment as a dot product of its molecule-fixed vector and a set of *direction cosines*  $\Phi_{Ij}$ , which measure the angular deviation between the dipole projection along the principal axis *j* and a space-fixed axis I, as follows:

$$\boldsymbol{\mu}_{I} = \begin{pmatrix} \mu_{a} \\ \mu_{b} \\ \mu_{c} \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{\Phi}_{Ia} \\ \boldsymbol{\Phi}_{Ib} \\ \boldsymbol{\Phi}_{Ic} \end{pmatrix} = \mu_{a} \Phi_{Ia} + \mu_{b} \Phi_{Ib} + \mu_{c} \Phi_{Ic}$$

Summing over all the space-fixed axes and using this as an operator, the *dipole matrix elements* for a given transition can be defined:

$$\langle J', \tau', M' | \boldsymbol{\mu}_{I} | J'', \tau'', M' \rangle = \sum_{i \in \{a, b, c\}} \mu_{a} \langle J', \tau', M' | \boldsymbol{\Phi}_{Ig} | J'', \tau'', M''' \rangle$$

These direction cosine matrix elements have been tabulated in the literature, and most tools for predicting rotational spectra can calculate them efficiently for an arbitrary molecule.

Selection rules arise by noting that the dipole matrix integral must be even overall. The  $\Delta J$  selection rules are carried over from the symmetric top case, so  $\Delta J = 0$ (Q-branch) or ±1 (R- and P-branch, respectively) hold for asymmetric tops. The  $\Delta \tau$  rules are slightly more complicated due to the asymmetry of the molecule. Specific selection rules for  $\Delta K_a$  and  $\Delta K_c$  arise depending on which dipole component is chosen along the molecule-fixed principal axes. These selection rules are tabulated below.<sup>1</sup>

Transition Type	$\Delta K_a$	$\Delta K_c$
$\mu_a \neq 0 \ (\mu_a \text{-type})$	0, ±2, ±4	1, ±3, ±5
$\mu_b \neq 0 \ (\mu_b$ -type)	1, ±3, ±5	1, ±3, ±5
$\mu_c \neq 0 \; (\mu_c \text{-type})$	1, ±3, ±5	0, ±2, ±4

 Table 1. K selection rules for an asymmetric rigid rotor

Finally, in the absence of an external electric field, all the M components of a given  $(J, \tau)$  rotational state are all degenerate. Therefore, an accurate measure of the intrinsic intensity of a rotational transition of the form  $(J', \tau') \rightarrow (J'', \tau'')$  requires an averaged sum over all the M components of the dipole matrix element, for a final intrinsic intensity of

$$I_{intrinsic} = \frac{\mu_j^2}{(2J+1)} \sum_F \sum_{M',M''} |\langle J', \tau', M' | \Phi_{Fj} | J'', \tau'', M'' \rangle|^2$$

Therefore, the total intensity of a given transition is the product Iintrinsic x Ithermal.

#### c) Common Perturbations to the Rigid Rotor Approximation

## *i) Centrifugal Distortion*

For most asymmetric tops, especially in the ground vibrational state at low values of J, the rigid rotor approximation is appropriate, and qualitative fits can be achieved by fitting a set of pure rotational transitions to only the three rotational constants. However, to reach experimental uncertainty when fitting most rotational spectra, perturbative corrections to the rigid rotor approximation are often required.

The most common correction is associated with **centrifugal distortion**, arising from vibrational distortions of the rotating molecular structure. For molecules with vibrations that are not large-amplitude in nature, these corrections are small in magnitude (< 100 kHz) with respect to the rotational constants ( > 100 MHz), so they can be quantitatively fit using only the rigid rotor Hamiltonian and a series of perturbative correction terms in a so-called "distortable rotor" Hamiltonian. The most common of these effective Hamiltonians is Watson's reduced Hamiltonian<sup>4,5</sup>, used almost exclusively in modern rotational spectroscopic studies. Use of the Watson Hamiltonian is sufficient enough for quantitative assignment of pure rotational spectra far into the millimeter-wave, assuming an appropriate leading order for the centrifugal distortion corrections is chosen.

For the research presented in this dissertation, only centrifugal distortion corrections in leading order of  $\hat{J}^4$  ("quartic" distortion constants) are necessary to experimentally fit pure rotational spectra.<sup>6</sup> The overall Watson reduced Hamiltonian corresponding to the quartic corrections is shown in equation (4),

$$\widehat{H_{rot}} = \underbrace{A\widehat{f_a^2} + B\widehat{f_b^2} + C\widehat{f_c^2}}_{rigid\ rotor} - \underbrace{\Delta_J \widehat{f^4} - \Delta_{JK} \widehat{f_a^2} - \Delta_K \widehat{f_a^4} - 2\delta_J \widehat{f^2} (\widehat{f_b^2} - \widehat{f_c^2}) - \delta_K \{\widehat{f_a^2}, (\widehat{f_b^2} - \widehat{f_c^2})\}}_{quartic\ centrifugal\ distortion} + O(\widehat{f^6}) \quad (4)$$

Where  $\{\hat{A}, \hat{B}\}$  is the anti-commutator and  $\{\Delta_J, \Delta_{JK}, \Delta_K, \delta_J, \delta_K\}$  are the set of quartic centrifugal distortion constants in the so-called A-reduction, which is appropriate for near-prolate molecules that are "sufficiently" asymmetric. There is an equivalent S-reduction set of distortion constants which are numerically related to the A-reduction, but are instead suited for spectra of very near-symmetric tops and planar molecules where the A-reduction can sometimes fail.<sup>7</sup>

#### *ii)* Internal Rotation

In some cases, the vibrational motion of the molecule cannot be treated as a perturbation on a rigid frame. This is most commonly manifested in molecules that exhibit large-amplitude motion, where the displacement vector associated with a large amplitude mode is near-magnitude to the atomic displacement about the center of mass. A classic example arises from the inversion tunneling of ammonia, first discovered using microwave spectroscopy by Cleeton & Williams in 1934.<sup>8</sup> In this motion, ammonia inverts between two equivalent  $C_{3v}$  structures with a planar  $D_{3h}$  transition state. This generates a symmetric double well potential energy surface, with a barrier to inversion of

approximately 1774 cm<sup>-1</sup>.<sup>9</sup> Tunneling across the barrier causes a doubling of rotational states with opposite parities, and transitions can be observed between the corresponding symmetric and antisymmetric rovibrational states. Consequently, one cannot treat this inversion motion as a perturbation of the rotation; rather, the rotation of ammonia is treated as a motion parameterized on the double-well potential.

Another form of internal rotation more relevant to this discussion arises from the hindered rotation of a symmetric functional group, such as the C<sub>3</sub>-symmetric methyl (-CH<sub>3</sub>) group, or a C<sub>6</sub>-symmetric species such as benzene (see Chapter 4 for details). In these cases, internal rotation effects can be observed via the appearance of a fine structure on the rotational spectrum of the system, arising from an interaction between the internal and overall rotation of the molecule.<sup>10</sup> The manifestation of this interaction is largely defined by the barrier height of the potential energy surface of the internal rotation, and typically falls into one of two cases: 1) low barrier, where the torsional energy levels are of higher energy than the barrier to internal rotation (e.g. free or near-free rotation), or 2) high barrier, where the energy levels of the torsion are well below the potential barrier (hindered rotation).

In general, one can write the potential energy surface of an N-fold symmetric internal rotor as a Fourier series with period  $2\pi/N$ , as shown in Equation 5a below.

$$V(\theta) = a_0 + \sum_{k=1}^{\infty} a_k \cos(kN\theta)$$
 (5a)

If  $a_0$  is taken to be  $-\sum a_k$  then V( $\theta$ ) can be written as,

$$V(\theta) = \sum_{k=1}^{\infty} \frac{V_{kN}}{2} (1 - \cos(kN\theta))$$
(5b)

Where, for the case of a 3-fold symmetric internal rotor to first order (k = 1, N = 3), V<sub>3</sub> is the barrier height of the torsional potential energy surface. Typically, most experimental data for three-fold internal rotors can be represented quantitatively by the simple potential truncation  $V(\theta) = \frac{V_3}{2}(1 - \cos(3\theta))$ , though there have been instances where the V<sub>6</sub> term is required to fit rotational spectroscopic data to experimental accuracy.<sup>11–13</sup>

The overall torsion-rotation Hamiltonian can be written as follows<sup>14,15</sup>,

$$\widehat{H} = \underbrace{\widehat{H_{rot}} + F\widehat{\mathbf{P}^2}}_{rotation} + \underbrace{F\widehat{\mathbf{p}^2} + V(\theta)}_{torsion} - \underbrace{2F\widehat{\mathbf{P}}\widehat{\mathbf{p}}}_{torsion-rotation}$$
(6)

where  ${}^{1}\widehat{P} = \sum_{j \in \{a,b,c\}} \frac{\lambda_{j}I_{\theta}}{I_{j}}$ ,  $\widehat{p}$  is the angular momentum operator of the internal rotor, and F is the "reduced" rotational constant of the rotor. If rotation were barrier-less, the torsional spectrum would obey the eigenvalues for a free rotor, e.g.  $\widehat{H}|m\rangle = Fm^{2}$  ( $m = 0, \pm 1, \pm 2, ...$ ).

In practice, treatment of this Hamiltonian to provide spectroscopic results depends on the nature of the torsional potential energy surface (PES). When the barrier of the PES is small with respect to the torsional energy level splittings, the coupling term -2F**Pp** cannot be treated as a perturbation, so an effective spectroscopic fit requires a "global" analysis where the pure rotational transitions are treated using a coupled basis containing multiple torsional states. However, if the barrier is much larger than the torsional level splittings, then the coupling term can be treated perturbatively and satisfactory

<sup>&</sup>lt;sup>1</sup> A clarification:  $\lambda_j$  is the direction cosine between the internal rotor axis and the j<sup>th</sup> principal axis, I<sub>0</sub> is the moment of inertia of the rotor, and I<sub>j</sub> is the moment of inertia for the molecule along that axis.

spectroscopic fits can be achieved in a "local" fashion by using a basis consisting of only one torsional energy level.<sup>16</sup>

In the high barrier case, which is relevant for the results presented in this dissertation, the effective Hamiltonian can be written as a sum of H<sub>rot</sub> and a perturbative sum of coefficients that are dependent on the torsional energy level (with quantum number v) and the symmetry number  $\sigma$  (for a three-fold symmetric rotor,  $\sigma = 0$  specifies the A symmetry state and  $\sigma = \pm 1$  correspond to the doubly-degenerate E state). This approach was originally developed by Herschbach<sup>17</sup> using a Van Vleck<sup>18</sup> perturbative treatment of equation (6). The results from this treatment give, to n<sup>th</sup> order, the Hamiltonian of the form,

$$\widehat{H}_{v\sigma} = \widehat{H}_{rot} + F \sum_{n} W_{v\sigma}^{(n)}(s) \widehat{P}^{n}$$
(7)

where  $W_{\nu\sigma}(s)$  are perturbative coefficients dependent only on the value of the dimensionless reduced barrier height,  $s = \frac{4V_3}{9F}$ . As expected, the magnitude of the W coefficients decreases as s increases, so the magnitude of the  $(v, \sigma) = (v, 0) \rightarrow (v, 1)$  splitting decreases to zero as the barrier approaches infinity.

To further simplify things, evaluation of equation (7) to first order (relevant leading order for most high-barrier  $C_{3v}$  rotors) shows that  $W^{(1)}$  is non-zero only for states with *E* symmetry.<sup>19</sup> As such, the spectrum of the rotor state with *A* symmetry can be evaluated as if no internal rotation affects the spectrum. This is significantly helpful for structural studies on systems with internal rotation, as the rotational constants of the *A* symmetry state can provide meaningful structural data without consideration of the torsional-rotational interactions.

#### d) Structure Determination

The key to unlocking direct structure determination via rotational spectroscopy lies in the Born-Oppenheimer (B-O) approximation. Assuming B-O, changes in an atom's nuclear mass (e.g. isotopic substitution) will effect no change in the electronic parameters of the bound state it is contained in. Since the rotational constants of a species are determined solely by the inertial tensor – a product of masses and distances – the measurement of the rotational spectrum of a molecule's isotopologues can provide direct constraints on the distance portion of the rotational constants.

In particular, Kraitchman's equations<sup>20</sup> provide a direct means of calculating the position in Cartesian space of an isotopically substituted atom in a molecule. Consider a molecule with experimental moments of inertia in the principal axis system,  $I_{xx}$ ,  $I_{yy}$ ,  $I_{zz}$ , and the experimentally determined moments of inertia for the same molecule but with a single atom substituted with an isotope (mass shift  $\Delta m$ ),  $I_{xx}$ ',  $I_{yy}$ ',  $I_{zz}$ '. Assuming the B-O approximation, the radial part of the inertial moments will stay constant. The geometric shift of the center of mass upon substitution can then be calculated, and  $I_{xx}$ ' is written as

$$I'_{xx} = I_{xx} + \Delta m(y^2 + z^2) - \frac{\Delta m^2(y^2 - z^2)}{M + \Delta m} = I_{xx} + \frac{M\Delta m}{M + \Delta m}(y^2 + z^2)$$

where M is the total mass of the molecule and  $\{x, y, z\}$  are the coordinates of the substituted atom. By applying this same correction to the other two moments of inertia, there are now three linear equations for three unknowns (the coordinates of the substituted atom). A 3x3 secular determinant can then be solved to acquire the *absolute* 

values (since the moments of inertia are functions of  $x_i^2$ ) of the Cartesian coordinates for the substituted atoms. For a non-planar molecule, the coordinates have the form

$$|i| = \left[\frac{\Delta I_i}{\mu} \left(1 + \frac{\Delta I_j}{I_i - I_j}\right) \left(1 + \frac{\Delta I_k}{I_i - I_k}\right)\right]^{1/2}$$

Where  $\{i, j, k\} \in \{x, y, z\}$  in cyclic order,  $\Delta I_i = I_{ii}$ , and  $\mu = \frac{M\Delta m}{(M + \Delta m)}$ . If the experimental spectra of a species and at least one isotopic substitution of each atom of the species can be assigned, then one can derive the complete direct ground state structure of a molecule. Simplified forms of Kraitchman's equations can also be derived for more symmetric and planar species.<sup>2</sup>

Since rotational constants can typically be determined to fractional errors as good as or better than  $\Delta B_i/B_i = 10^{-6}$ , Kraitchman coordinates (also called substitution, or r<sub>s</sub>, coordinates) for polyatomic molecules are typically precise to better than 0.01 Å. However, the primary source of error is generally less due to poor determination of constants and rather the treatment of vibrational averaging in Kraitchman's method. Kraitchman's equations assume the zero-point effects on the ground state moment of inertia are constant and isotropic<sup>21</sup>, so calculating differences of moments will cancel out any of the zero-point corrections. For atoms affected by a large amplitude motion in a molecule, the r<sub>s</sub> coordinates can sometimes be significantly different than either the equilibrium r<sub>e</sub> or ground-state r<sub>0</sub> structures. This is also the case for hydrogen substitution coordinates, since the large fractional difference in mass between <sup>1</sup>H and <sup>2</sup>D can cause a significant change in zero-point motion. In general, the situation worsens for atoms close to a principal axis, where at least one of the  $\Delta I_i$  terms are near zero; this results in imaginary coordinates, which are set to zero. The distance cut-off for this effect varies widely on the molecule, but it generally occurs when an atom is within 0.1 Å of a principal axis. However, for well-behaved atoms (far from the principal axes), the r<sub>s</sub> parameters are typically  $\approx \frac{1}{2}(r_e + r_0)$ . The error arising from zero-point averaging is commonly estimated using Costain's rule, or  $\Delta z \approx 0.0015/|z|$  Å.<sup>22–24</sup>

There are additional natural limitations to this method, due to isotopic considerations. Some atoms, such as fluorine, have no stable isotopes. Others, such as hydrogen, have isotopes in very low abundance; deuterium has a fractional abundance of  $1.1 \times 10^{-4}$ . However, since carbon-13 is at 1% abundance relative to carbon-12, and nitrogen-15 and oxygen-18 have fractional abundances greater than  $1 \times 10^{-3}$ , the heavy atom backbones of polar organic molecules can typically be determined using isotopic spectra observed in natural abundance with relative ease.

For data sets that include rotational constants of molecular species with multiple isotopic substitutions, Kraitchman's equations are not typically used. In these cases, the abundance of distance constraints arising from a data set of spectra of single and multiple isotopologues can be used to determine a least-squares fitted effective structure. In some cases, these data sets can be used to determine the positions of atoms without stable isotopes to precisions as good, if not better, than Kraitchman's method would otherwise provide. The most common form of least-squares fitting is for determination of the effective ground-state r<sub>0</sub> structure<sup>25,26</sup>, which only requires input of the ground-state rotational constants or moments of inertia. In many cases this can provide accurate structures for systems with imaginary coordinate and isotope issues when Kraitchman's method is applied. However, since zero-point motion is entirely neglected in this

determination, the precision of the determined structure is generally bounded from below by the magnitude of the zero-point effects.

Another commonly used approximation is to use least-squares fitting to determine the equilibrium structure of the molecule by providing the ground state constants and either experimental or theoretical values for the vibration-rotation coupling constants  $\alpha_i$ . In practice, experimental determination of these constants are difficult, especially for large molecules, and theoretical calculations require computationally expensive anharmonic vibrational calculations.<sup>27</sup>

However, a paper by Watson, Roytburg and Ulrich<sup>28</sup> outlines another method, called  $r_m^{(n)}$ , that provides near equilibrium structures by fitting the first order  $(r_m^{(1)})$  and/or second order  $(r_m^{(2)})$  vibrational corrections to the ground state moments of inertia. In the  $r_m^{(1)}$  model, the ground-state experimental moments of inertia I<sup>0</sup> along axis *i* are fit to  $I_i^0 = I_i^m(r_m) + c_i\sqrt{I_i^m(r_m)}$ , where  $c_i$  is a fit parameter (one for each principal axis) to account for zero-point motion proportional to leading order of  $O(I_0^{1/2})$ . The  $r_m^{(1)}$  method gives similar results Kraitchman's method but with additional flexibility in treating the zero-point motion of the molecule, and as such has none of the zero-coordinate issues found with Kraitchman's method. Like other least-squares methods, it can also be used to determine the position of atoms without stable isotopes. Finally, the  $r_m^{(2)}$  model adds an additional set of three constants that arise from corrections using Wilson's **G**<sup>-1</sup> matrix<sup>29</sup>, but its formulation is beyond the scope of this discussion.

## References

(1) Townes, C. H.; Schawlow, A. L. *Microwave Spectroscopy*; Courier Corporation, 1955.

(2) Gordy, W.; Cook, R. L. *Microwave molecular spectra*; Wiley, 1984.

(3) Guest, M. Performance of Various Computers in Computational Chemistry http://www.cfs.dl.ac.uk/benchmarks/compchem.html (accessed Mar 4, 2015).

(4) Watson, J. K. G. Mol. Phys. 1968, 15, 479.

(5) Meyer, H. Annu. Rev. Phys. Chem. 2002, 53, 141.

(6) Watson, J. K. G. J. Chem. Phys. 1967, 46, 1935.

(7) Watson, J. K. G. In *Handbook of High-resolution Spectroscopy*; John Wiley & Sons, Ltd, 2011.

(8) Cleeton, C. E.; Williams, N. H. *Phys. Rev.* **1934**, *45*, 234.

(9) Marquardt, R.; Sagui, K.; Zheng, J.; Thiel, W.; Luckhaus, D.; Yurchenko, S.; Mariotti, F.; Quack, M. J. Phys. Chem. A **2013**, 117, 7502.

(10) Lin, C. C.; Swalen, J. D. Rev. Mod. Phys. 1959, 31, 841.

(11) Lees, R. M. J. Chem. Phys. 1973, 59, 2690.

(12) Moloney, M. J.; Krisher, L. C. J. Chem. Phys. 1966, 45, 3277.

(13) Kidwell, N. M.; Reilly, N. J.; Nebgen, B.; Mehta-Hurt, D. N.; Hoehn, R. D.; Kokkin, D. L.; McCarthy, M. C.; Slipchenko, L. V.; Zwier, T. S. *J. Phys. Chem. A* **2013**, *117*, 13465.

- (14) Burkhard, D. G.; Dennison, D. M. Phys. Rev. 1951, 84, 408.
- (15) Nielsen, H. H. *Phys. Rev.* **1932**, *40*, 445.
- (16) Kleiner, I. J. Mol. Spectrosc. 2010, 260, 1.
- (17) Herschbach, D. R. J. Chem. Phys. 2004, 31, 91.
- (18) Van Vleck, J. H. *Phys. Rev.* **1929**, *33*, 467.
- (19) Kroto, H. W. *Molecular Rotation Spectra*; Dover Publications, 2003.
- (20) Kraitchman, J. Am. J. Phys. 1953, 21, 17.
- (21) Watson, J. K. G. J. Mol. Spectrosc. **1973**, 48, 479.

(22) Harmony, M. D.; Laurie, V. W.; Kuczkowski, R. L.; Schwendeman, R. H.; Ramsay, D. A.; Lovas, F. J.; Lafferty, W. J.; Maki, A. G. *J. Phys. Chem. Ref. Data* **1979**, *8*, 619.

- (23) Tam, H. S.; Harmony, M. D. J. Phys. Chem. 1991, 95, 9267.
- (24) Costain, C. C. J. Chem. Phys. 1958, 29, 864.
- (25) Kisiel, Z. J. Mol. Spectrosc. 2003, 218, 58.
- (26) Nösberger, P.; Bauder, A.; Günthard, H. H. Chem. Phys. 1973, 1, 418.

(27) Demaison, J.; Boggs, J. E.; Csaszar, A. G. *Equilibrium Molecular Structures: From Spectroscopy to Quantum Chemistry*; CRC Press, 2010.

- (28) Watson, J. K. G.; Roytburg, A.; Ulrich, W. J. Mol. Spectrosc. 1999, 196, 102.
- (29) Wilson Jr., E. B. J. Chem. Phys. 1941, 9, 76.

## **II.** Electronic Structure Theory

#### 1) Introduction

"The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved."

- *Paul Dirac*, 1929<sup>1</sup>

Although the behavior of any molecule can be represented exactly and rigorously, the solution set is generally impossible to represent analytically outside of the simplest oneelectron systems. The only means of progress is through approximate techniques which, at the time of their formulation, required a massive paradigm shift in computational methodology in order to become feasible.

Much of the success story can be attributed to the development of computers. Since the end of the "first quantum revolution" in the 1920s, chemists and physicists have worked in unison to build a hierarchical theoretical model of chemistry, with efficient methodology ranging from qualitative trend prediction to models accurate enough to correctly predict and model results for even the most stringent, precise and highresolution experimental techniques. The results of this endeavor have been monumental. For example, in 1955, C. W. Scherr, a Ph.D. student of Clemens Roothan, finished his thesis work on the first all-electron LCAO calculation of molecular nitrogen using a *minimal* basis set. Between Scherr and his two assistants, this single-point calculation took over two years to complete.<sup>2</sup> Now, recreating Scherr's minimal basis calculation of N<sub>2</sub> takes approximately 1 second on a modern desktop computer. In fact, the development of quantum chemistry has been so successful that many of its greatest results have been distilled to a "black box", in that many successful chemists have used its predictive power for research and discovery without direct reference to a mathematically rigorous foundation.<sup>3,4</sup> This is the power of quantum chemistry, in that it provides a fully hierarchical system for prediction for any chemist.

Therefore, since this dissertation is primarily focused on molecular structure, it would be a great disservice to the material to ignore the theoretical foundation of electronic structure, which is eternally tied to molecular spectroscopy. This discussion of quantum chemistry is neither fully rigorous nor complete, but many great theoretical chemists in the past have published countless books and review articles summarizing the results discussed here with the rigor and expertise it deserves. Published work as essential complements to this section include textbooks by Szabo and Ostlund<sup>5</sup> and Cook<sup>6</sup>, as well as review articles from Becke<sup>7</sup>, Simons<sup>8</sup>, Slater<sup>9</sup> and Roothan.<sup>10</sup>

In particular, this section will focus primarily on the specific computational techniques exploited in the presented research. A few of the discussed techniques are ubiquitous in nature, such as the Hartree-Fock (HF) method<sup>10</sup> and Møller-Plesset (MPn) perturbation theory<sup>11</sup>, which correspond indirectly to LCAO-MO theory and perturbative corrections to LCAO-MO. Also included is short discussion on density functional theory (DFT) and dispersion corrections to DFT.

#### 2) Multi-Electron Systems

The overarching assumption made in both electronic structure theory and molecular spectroscopy is the Bohr-Oppenheimer (B-O) approximation.<sup>12</sup> The assumption of the B-O approximation implies that nuclear motion is much slower than electronic motion, so the dynamics can be treated as two uncoupled parts. Freezing the nuclear motion reduces the calculation of the nuclear energy as a constant Coloumb repulsion term. The structural problem, therefore, reduces to solving an electronic Hamiltonian, separable from the nuclear part, accounting only for the dynamics of the electrons in a molecule. The general N-electron, M-nuclear (treated as point charges) Hamiltonian is shown below in Equation (1).

$$\widehat{H_{elec}} = \underbrace{-\sum_{i=1}^{N} \frac{\hbar^2}{2m_e} \nabla_i^2}_{electron \, KE} - \underbrace{\sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A e^2}{4\pi\varepsilon_0 |r_i - r_A|}}_{elec-nuc \, Coloumb} + \underbrace{\sum_{i=1}^{N} \sum_{j>i}^{N} \frac{e^2}{4\pi\varepsilon_0 |r_i - r_j|}}_{elec-elec \, Coloumb}$$
(1)

When the Schrödinger equation for this Hamiltonian is solved, an electronic wavefunction that is 3N dimensional is returned, along with the total electronic energy. Since the nuclear structure is frozen in the B-O approxmation, the nuclear Hamiltonian largely reduces to a constant energy corresponding to the total nuclear-nuclear repulsion energy of the molecule. This is the basis of calculating a *potential energy surface*, where the total electronic energy of the system is calculated as a function of its nuclear coordinates.

However, the solutions to equation (1) are not finitely analytic.<sup>13,14</sup> This is due to the electron-electron Coloumb term in equation (1), where the energy of an electron is dependent on the coordinates of every other electron in the system. However, the 1-

electron case is exactly solvable, so the approach undertaken is to treat the total electronic Hamiltonian as a sum of 1-electron terms, where each electron sees an *averaged* electric field generated by the remaining electrons in the system, called the *mean field approxmation*. This allows one to separate the total electronic wavefunction into a product of one-electron spatial orbitals. Following this path, one will find that the meanfield approximation does remarkably well at representing the basic electronic structure of a molecule, and that further improvements can ultimately be made by applying additional corrections through methods such as perturbation theory.

In order to represent the wavefunction appropriately as a product of one-electron wavefunctions, one must account for both the Pauli exclusion principle and the fact that electrons are antisymmetric upon exchange. Accounting for the exclusion principle, a spatial orbital  $\psi_i(\mathbf{r})$  is multiplied with a spin function  $\alpha$  or  $\beta$  corresponding to spin +1/2 or spin -1/2, generating a spin orbital  $\chi_i(\mathbf{r}) = \psi_i(\mathbf{r})\alpha$  or  $\psi_i(\mathbf{r})\beta$ .  $\alpha$  and  $\beta$  are orthonormal spin wavefunctions. For exchange antisymmetry, the total electronic wavefunction  $\psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)$  is written not as a product of spin orbitals  $|\chi_1\rangle|\chi_2\rangle \cdots |\chi_N\rangle$ , which does not account for the antisymmetry, but rather a *Slater determinant* that conveys the proper antisymmetry  $\psi(\mathbf{r}_1, \mathbf{r}_2, ..., -\mathbf{r}_i, ..., \mathbf{r}_N) = -\psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_i, ..., \mathbf{r}_N)$ , as shown in equation (2).

$$\psi(r_1, r_2, \dots, r_N) = \frac{1}{\sqrt{N!}} det \begin{bmatrix} \chi_1(r_1) & \cdots & \chi_N(r_1) \\ \vdots & \ddots & \vdots \\ \chi_1(r_N) & \cdots & \chi_N(r_N) \end{bmatrix}$$
(2)

With this formalism in tow, the solutions to the one-electron Hamiltonian, called the **Hartree-Fock equation**, can be written as the following eigenvalue problem:

$$\widehat{f(j)} = -\frac{\hbar^2}{2m_e} \nabla_j^2 - \sum_{j=1}^M \frac{Z_A e^2}{4\pi\varepsilon_0 |r_j - r_A|} + v(j) \quad (3)$$

where 
$$\widehat{f(j)}\chi(r_j) = [h(j) + v(j)]\chi(r_j) = \varepsilon_j\chi(r_j)$$

In equation (3), f(j) is called the *Fock* operator, which contains the relevant 1-electron terms of the Hamiltonian [h(j)] as well as an unspecified electric potential v(j) that accounts for the electric potential generated by the other electrons in the system in some average way. The mathematical form of this potential will be discussed in the next section.

Since each spatial orbital  $\psi_i$  has two corresponding spin orbitals, there will be in general be more spin orbitals available than there are electrons. These extra orbitals are unoccupied and are typically termed **virtual** orbitals. One can imagine a molecular system where one of these virtual orbitals are very close in energy to an occupied orbital, so that a new Slater determinant corresponding to an excitation between these two closely-spaced orbitals is near-energy to the ground state configuration. An obvious example of this would be a state with triplet spin multiplicity, where there are three degenerate ground state configurations of spin orbitals. In this case, the formalism described here fails, and one has to use a *multiconfigurational* theory in order to accurately represent the system.

But what happens if an excited configuration has a non-zero but small enough amplitude to be treated as a perturbation? Here, use of the ground-state configuration is appropriate but the correction due to this excited configuration might be necessary. In reality, the situation is not so dire. By applying the Hartree-Fock equation, one can derive *Brillouin's Theorem*, which shows that Slater determinants representing a single excitation from the ground state do not interact with the ground state determinant. Therefore, any corrections to the Hartree-Fock formalism will first arise only with doubly-excited configurations, which typically result in small corrections that are treated beginning with second-order perturbation theory.

#### 3) Hartree-Fock (HF) Theory

Since Brillouin's Theorem implies the choice of representing the system with a single Slater determinant is stable to at least first order, one can now evaluate the Hartree-Fock equation with some confidence in the zeroth-order approximation. Recall that the Fock operator f(j) consists of a previously undefined potential v(j) that specifies the average interaction of electron j with every other electron in the system. In order to evaluate this potential, one can apply first order perturbation theory to the electronic Hamiltonian in equation 1, using v(j) as the perturbing potential.

The resulting first-order Hartree-Fock energy is,

$$E_{HF} = \sum_{i} \langle i | \hat{h} | i \rangle - \sum_{i > j} [ii|jj] - [ij|ji]$$
(4)

where

$$[ij|kl] = \iint dr_1 dr_2 \ \chi_i^*(r_1)\chi_j(r_1) \frac{1}{r_{12}}\chi_k^*(r_2)\chi_l(r_2)$$
(5)

Due to orthonormality, only the two-electron integrals with index symmetry [ii|jj] and [ij|ji] survive in the expansion. The first term, [ii|jj], corresponds to the Coloumb

repulsion of electron in spin orbital  $\chi_i$  to the electron in  $\chi_j$ , often called the Coloumb integral  $\hat{f}(i)$ . This is an averaged effect, as expected; if one index is fixed and the other is summed over all possible values,  $\hat{f}(i)$  can be written as an operator:

$$\hat{f}(i)\chi_i(r_1) = \left[\sum_{j \neq i} \int |\chi_j(r_2)|^2 \frac{1}{r_{ij}} dr_2\right]\chi_i(r_1)$$

Therefore, the Coloumb integral for a given electron averages overall all of the repulsions to itself from every other electron in the system.

The other term, [ij|ji], corresponds to the matrix elements of the exchange integral  $\hat{K}(i)$  of which there is no classical analogue. This is the energy arising from swapping two electrons with positions  $r_1$  and  $r_2$  between two spin orbitals  $\chi_1$  and  $\chi_2$ . Similar to the Coloumb integral, this effect is averaged for a given electron over all the other electrons, and partially depend on the value of the overlap integral  $\langle i|j \rangle$  (no overlap implies no exchange effect between two occupied orbitals).

In theory, equation (4) provides all the tools needed to calculate the Hartree-Fock energy of an arbitrary molecule using an antisymmetrized linear combination of its atomic orbitals. However, for a molecule of any reasonable size the integro-differential equation (4) is computationally infeasible, as it requires numerical integration along a finely spaced 3D grid.<sup>15</sup>

However, Roothan<sup>10</sup> and Hall<sup>16</sup> presented a different method by applying the variational method to the spin orbitals. By expanding the spin orbitals into a well-defined, but non-orthonormal finite basis set of functions, equation (4) can be reduced to a matrix eigenvalue problem, which is far less computationally expensive than evaluating

equation (4) directly. Consider expanding the spatial component  $\psi_i$  of each spin orbital (e.g.  $\chi_i = \psi_i \alpha$  or  $\psi_i \beta$ ) in the Slater determinant in a finite basis  $\{\phi_i\}$ :

$$\psi_i = \sum_{m=1}^M C_{mi} \phi_m \quad (6)$$

For an appropriate choice of basis set, completeness guarantees that as  $M \rightarrow \infty$  the evaluated energy converges on the Hartree-Fock limit, e.g. the energy determined by equation (4). The choice of basis  $\{\phi_j\}$  appropriate for accurate representation of molecular orbitals and accurate values of the ground state energy is an essential problem in quantum chemistry.

Substituting the expansion in equation (6) into the Hartree-Fock expression into equation (3) and multiplying both sides of (3) with  $\phi^*$ , a generalized eigenvalue equation is generated:

$$\sum_{n} C_{ni} \int dr_{1} \phi_{m}^{*} \hat{f} \phi_{n} = \varepsilon_{i} \sum_{n} C_{ni} \int dr_{1} \phi_{m}^{*} \phi_{n} \qquad or ..$$
$$\sum_{n} F_{mn} C_{ni} = \epsilon_{i} \sum_{n} S_{mn} C_{ni} \qquad (7)$$

Where equation (7), called the **Roothan** equation, is more generally written as  $FC = SC\varepsilon$ , where **F** is the Fock matrix, **S** is the overlap matrix, **C** is a M x M matrix of the basis set coefficients, and  $\varepsilon$  is a diagonal matrix of orbital energies.

All that is now missing from the Hartree-Fock toolbox is an explicit form of the Fock matrix, as one must consider the matrix elements of the Coloumb and exchange integrals. Consider the Fock operator in equation (3) as a separate sum of its one-electron and two-electron terms. Since the introduced basis set  $\{\emptyset\}$  is not orthonormal, one must consider the off-diagonal matrix elements for both sets of terms. The matrix element  $\langle i|\hat{h}|j\rangle$  is typically written as  $H_{ij}^{core}$ , and the two electron terms can be expanded using a density matrix formalism,

$$F_{ij} = H_{ij}^{core} + \sum_{i}^{\frac{N}{2}} \sum_{m,n} C_{ni}^{*} C_{mi}(2[ij|nm] - [im|nj])$$
  
=  $H_{ij}^{core} + \sum_{m,n} P_{mn} \left( [ij|nm] - \frac{1}{2} [im|nj] \right) = H_{ij}^{core} + G_{ij}$  (8)

where **P** is the density matrix made up of elements of the coefficient matrix **C** and the matrix **G** is the two-electron part of the Fock matrix. The factor of ½ found in the expression arises from index symmetry. The density matrix **P** actually constitutes the electronic probability density in the chosen basis  $\{\emptyset\}$  --- an orbital product  $\phi_i^*\phi_j$  has a corresponding element  $P_{ij}$  that is zero valued if there is no charge density overlap between the two orbitals.

With equation (8) in hand, one can now calculate the Hartree-Fock energy of an arbitrary molecular system. However, since the HF process is variational in the density matrix **P**, one must optimize the energy with respect to the values contained in the coefficient matrix, which ultimately specify the specific weighting of the atomic orbitals in the molecular orbital basis. This process is done iteratively, using the **self-consistent field** (SCF) method. Using an initial guess **P**<sub>0</sub> for the density matrix, one can create a newly improved density matrix **P'** using the coefficients calculated by solving the Roothan equations using **P**<sub>0</sub>. This process is summarized briefly as Algorithm 1 below.

**Input:** Molecular geometry specification (nuclear coordinates  $\{\mathbf{R}\}$ , number of electrons, etc.) and basis set  $\{\emptyset\}$ .

- 1) Calculate S, H<sup>core</sup>, and two electron integrals using input specification.
- Diagonalize S in order to obtain a transformation matrix R that orthogonalizes {Ø}.
- Generate density matrix guess P (e.g. Po for first iteration). Typically Po is a null matrix, which implies only H<sup>core</sup> contributes exclusively to the first iteration of the HF energy.
- 4) Calculate **G** using **P**, and form the Fock matrix  $\mathbf{F} = \mathbf{H}^{core} + \mathbf{G}$ .
- 5) Transform **F** using orthogonalization matrix **R**:  $\mathbf{F}' = \mathbf{R}^{\dagger} \mathbf{F} \mathbf{R}$ , and diagonalize **F**' to obtain **C**' (coefficients in orthogonalized basis) and orbital energies  $\varepsilon$  (*Roothan equations*)
- 6) Calculate new density matrix  $P_{new}$  from C = RC'.
- 7) If  $\mathbf{P}_{new} \approx \mathbf{P}$  within some given convergence criteria, then algorithm is done. Otherwise, repeat starting at step (3) using  $\mathbf{P}_{new}$  as the new guess density matrix

**Algorithm 1.** Self-consistent field (SCF) method for solving Roothan's equations to calculate the Hartree-Fock energy using a specified basis set.

Assuming that the molecular specification constitutes a quantum system that can be stably defined a singular Slater determinant, the SCF process summarized as Algorithm 1 will converge to the correct Hartree-Fock energy, bounded from below variationally by the Hartree-Fock limit ( $E_{HF}$  calculated using an infinite basis set). Practically, calculation of **G** is the most computationally expensive part of the HF procedure; for a basis set of K spin orbitals, the number of unique two-electron integrals is O(K<sup>4</sup>/8).

## 4) Perturbative Corrections to HF Theory

The previous section outlines a method for calculating the energy of an arbitrary molecule in an *ab initio* fashion. For many common molecular systems (e.g. in a singlet electronic state), the Hartree-Fock approximation is a good one – the common rule of thumb is that HF treats "99%" of a molecule's total energy.<sup>17</sup> However, the approximation is only good in the local vicinity of the minimum of a bound molecular state, and in regions away from the minimum the HF approximation can often fail. This is not surprising, given HF assumes the molecular state can be well approximated by a single Slater determinant. But even for a system as simple as H<sub>2</sub>, HF theory does not even predict dissociation correctly, as the Slater determinants corresponding to a homolytically cleaved bond are not considered in the HF calculation. This can be alleviated by using a method called *unrestricted Hartree-Fock*, where the assumption that the number of electrons of  $\alpha$  spin is equal to the number with  $\beta$  spin is relaxed, but the inherent mean-field approximation associated with HF results in poor dissociation energies for many molecules.

The mean-field approximation also gives poor results for systems where van der Waals (or dispersive) interactions are important, particularly in non-covalently bound molecular complexes. Dispersive interactions typically arise from correlated motion of electrons in one monomer to those in the other; however, HF treats these motions in an averaged fashion and, as a result, cannot accurately represent this kind of correlated motion. As such, HF often poorly determines the binding geometry in a non-covalent system and can sometimes predict no binding at all.<sup>18</sup>

However, for well-behaved closed-shell molecular systems, this correlated motion is typically very small with respect to the HF energy of a molecular system (e.g.  $E_{corr}/E_{HF}$ < 0.01), but for non-covalent systems this correlation energy can be a large fraction of the total intermolecular binding energy (which in itself is typically a small fraction of the total energy of the molecular system). Therefore electron correlation can often be treated perturbatively. The most common form of perturbation theory used for HF to treat electron correlation is called n<sup>th</sup> order Møller–Plesset perturbation theory, or MPn theory.<sup>11,19</sup> Fortunately, MPn theory can give excellent results with respect to experiment even when evaluated to second order (MP2), which is the first non-trivial order of the MPn perturbation series. In short, MPn theory rewrites the molecular Hamiltonian into two parts,  $\hat{H} = \hat{H}^{(0)} + \hat{V}$  where

$$\hat{H}^{(0)} = \sum_{i}^{N} h(i) + v^{HF}(i)$$
$$\hat{V} = \sum_{i < j}^{N} \hat{r}_{ij}^{-1} - v^{HF}(i)$$

such that the perturbation V is the difference of real electron-electron interaction energy and what is calculated by the mean-field approximation.

Applying perturbation theory up to second order and simplifying, one finds that the energetic terms of the perturbation series are

$$E_0^{(0)} + E_0^{(1)} = \sum_a \epsilon_i - \frac{1}{2} \sum_a \sum_b \langle ab | ab \rangle - \langle ab | ba \rangle = E_{HF}$$
$$E_0^{(2)} = \sum_{a < b; r < s} \frac{|\langle ab | rs \rangle|^2}{\epsilon_a + \epsilon_b - \epsilon_r - \epsilon_s}$$

where a, b are occupied orbitals and r, s are virtual (unoccupied) orbitals.

The interpretation of these equations is quite simple. MP1 is equivalent to Hartree-Fock theory, and MP2 corrects Hartree-Fock by calculating the sum of the amplitudes between all double combinations of occupied and virtual orbitals. This is precisely what one would expect where the "real" wavefunction of a ground-state system is mostly a single Slater determinant with no excitations, and a small contribution of Slater determinants where energetically-favorable excitations into nearby virtual orbitals of certain electrons can occur. Note that since MP2 only consists of a two-particle operator, then only "double" (2-electron) excitations can occur in this approximation, as Brillouin's theorem guarantees no contributions from single electron excitations. Higher orders of MPn theory can potentially treat the contributions arising from Slater determinants where triple or higher excitations can occur, but in practice MP3 and better do not often improve the results as MP(n) is generally not a convergent perturbative method.<sup>20</sup> As such, more careful treatment of electron correlation requires methodology like *coupled cluster* (CC) theory<sup>21,22</sup>, which can often treat contributions arising from triple excitations or more in an accurate and convergent fashion. Since MP2 scales as  $O(N^5)$ , and MP3 scales the same as CC theory accounting for single, double and triple excitations (called CCSD(T)) at  $O(N^7)$ , MP3+ methods are typically never used.

# 5) Density Functional Theory

An alternative method for calculating the energy of a molecular system is through density functional theory, which instead uses the probability density representation of the electronic wavefunction instead of a Slater determinant. If one can represent a given Nelectron system with a total electronic wavefunction of 3N degrees of freedom as  $\Psi(r_1, r_2, ..., r_n)$ , then one can write the total three-dimensional probability density  $n(\mathbf{r})$  as the integral

$$n(\mathbf{r}) = N \int d^3 r_2 \dots \int d^3 r_N \Psi^*(\mathbf{r}, r_2, \dots, r_n) \Psi(\mathbf{r}, r_2, \dots, r_n)$$

However, for a given density of N electrons, there are an infinite number of possible wavefunctions (under U(1) symmetry, e.g. by application of constant phase factor  $e^{i\theta}$ ). This initially seems like an untenable position, but the Hohenberg-Kohn (H-K) theorems provide an escape.<sup>23</sup> The first H-K theorem guarantees that the properties of a multi-electron system can be *uniquely* determined by a 3 dimensional electron density; the

second H-K theorem guarantees that there exists a *functional*  $F[n(\mathbf{r})]$  such that minimization of this functional reproduces the ground-state energy of the equivalent multi-electron wavefunction. More specifically, for input density *n* for a system with N electrons and nuclear potential  $v(\mathbf{r})$ , one is guaranteed to have a function F[n] such that

$$E_0 = \min_{n \to N} F[n] + \int d^3 \vec{r} \, n(\vec{r}) v(\vec{r})$$

If one compares this expression for the ground state energy to the wavefunction-based expression then the "universal" functional F[n] can be written as

$$F[n] = \underbrace{T[n]}_{kinetic \ energy} + \underbrace{J[n]}_{Coloumb} + \underbrace{E_{corr}[n]}_{functional}$$

Both H-K theorems guarantee that this universal functional exists. However, there is no prescription to the actual *form* of this functional, so approximations must be made in order to do calculations. This is somewhat the opposite problem of HF theory; in HF theory, the form of the energy-generating functional (or operator) is exactly known, but the multi-dimensional nature of the wavefunction makes the solution impossible to calculate exactly. In DFT, however, the electronic structure of the ground state is known explicitly, but the functional used to calculate the energy cannot generally be known exactly.

The classic way of representing the functional in chemistry is through use of the Kohn-Sham (K-S) equations<sup>24</sup>, which treat the electronic density as a set of singleelectron Slater orbitals  $\varphi_i$ , normalized such that  $\sum_i |\phi_i(\vec{r})|^2 = n[\vec{r}]$ . The K-S equations then show that the energy functional F[n] can be rewritten as

$$F[n] = \int_{elec-nuc functional} \frac{Z_I}{r_I} \sum_{l} \frac{Z_I}{r_I} - \frac{1}{2} \sum_{i}^{N} \langle \phi_i | \nabla^2 | \phi_i \rangle + \frac{1}{2} \iint_{e-e \ Coloumb \ functional} \frac{n[\vec{r}_1]n[\vec{r}_2]}{r_{12}}$$

$$+ E_{XC}[n]$$

Assuming a finite basis, this functional can be rewritten into a matrix eigenvalue optimization problem, just like the HF equations transformed with Roothan's equations. However, the mysterious exchange-correlation functional  $E_{XC}$  is in general not analytic (unlike the exchange integral in HF), so typically a latticed numerical integration scheme is used to evaluate  $E_{XC}$ .

In practice, one must empirically derive a form of  $E_{XC}$  as no analytical form is known, except for the uniform electron gas. The K-S equations guarantee that  $E_{XC}$  should only be a function of the electron density (local density approximation, or LDA, functional), but many functionals also include dependence on values such as  $\nabla n$ (generalized-gradient approximation, or GGA, functionals), spin (local-spin density approximation, or LSDA, functionals), or even with a weighted contribution of the exact exchange calculated using HF theory (hybrid functionals).<sup>7</sup>

For molecular calculations, the most common type of functional used is a hybrid functional. For the ubiquitous hybrid functional B3LYP,  $E_{XC}$  is a mix of LDA, GGA and Hartree-Fock exact exchange, where the weights between the different types of exchange energies are determined by training with a set of atomic data such as atomization energies, ionization potentials, and total atomic energies.<sup>25</sup> More recent functionals such as the Minnesota family of meta-GGA functionals (M06-2X being the most common) improve on hybrid functionals such as B3LYP by including dependence on  $\nabla^2 n[\vec{r}]$  and training the functional using a large variety of experimental atomic and experimental data.<sup>26,27</sup> Although this in effect suggests that DFT is more of an empirical theory than an *ab initio* one, systematic studies of modern DFT functionals show that DFT-derived properties are often as accurate than those predicted using *ab initio* theory (such as MP2) at a significant fraction of the computational cost, as DFT scales formally as  $O(N^4)$  and can be reduced to  $O(N^2)$  with additional approximations.<sup>28</sup>

One relevant limitation of DFT is its poor treatment of long-range correlation, especially important in non-covalently bound molecular systems. This is due to the fact that correlation is treated *locally* by the exchange-correlation functionals typically used. In general, these local functionals cannot reproduce the R<sup>-6</sup> asymptotic form of London dispersion. To treat this discrepancy, a number of approaches have been undertaken to improve the prediction of long-range behavior. The most common is to include an additional, empirical, dispersion correction to the functional of the form

 $E_{disp} = -\sum_{i,j} \frac{c_6^{ij}}{r_{ij}^6}$  where the c<sub>ij</sub> coefficients are trained using experimental data.<sup>29</sup> The most advanced and recent of these empirical methods is the –D3 correction of Grimme<sup>30,31</sup>, which has a dispersion correction consisting of London-type dispersion modified with an empirical damping function:

$$E_{disp}^{D3} = -\sum_{n=6,8} \sum_{A,B} \frac{C_n^{AB}}{(R_{AB})^n} \frac{1}{1 + 6(\frac{R_{AB}}{s_{r,n}R_0^{AB}})^{-\alpha_n}}$$

where  $R_{AB}$  is the internuclear distance between two atoms of the target molecule,  $R_0^{AB}$  is the van der Waals cutoff radii, and  $s_{r,n}$  and  $\alpha_n$  are empirical coefficients for scaling the damping function. Other approaches include implementing corrections using MP2 theory to treat correlation (so-called "double hybrid" functionals)<sup>32</sup>, as well as functionals that treat correlation by calculating the C<sub>6</sub> London dispersion coefficients by calculating them using an explicitly non-local treatment of density in  $E_{XC}$  ("van der Waals density functionals", or vdW-DF)<sup>33,34</sup>. However, these methods are computationally expensive relative to standard DFT calculations, and in the case of vdW-DF, only treat pure van der Waals interactions effectively (and not the subtle balance between dispersion and electrostatics generally seen in most non-covalent molecular complexes).

#### 6) Basis Sets

The previous sections have summarized a number of methods to calculate the energy of any molecular species; however, there is no prescription to how one should represent the basis set approximating the set of spatial orbitals used to construct a reference Slater determinant or density. There are primarily two goals one must balance in order to specify an appropriate basis set; one, that the functions used in the basis set accurately depict actual atomic orbitals; and two, that calculation of the one and two-electron integrals in the Fock operator is efficient.

In practice, these two goals end up being somewhat contradictory. The natural choice to represent the spatial orbitals of an arbitrary atom is to expand them in a basis set of hydrogenic orbitals called Slater-type orbitals<sup>35</sup>  $\phi$ <sup>STO</sup>, which have the form,

$$\phi^{STO}(r,L,\xi) = Nx^a y^b z^c e^{-\xi r}$$

where L = a + b + c controls the angular momentum part of the orbital, N is a normalization constant, and  $\xi$  is a fitted parameter dependent on the atom being represented. STOs are hydrogenic in nature, and contracted sums of STOs often converge to an accurate representation of an arbitrary atomic orbital quickly.<sup>36</sup>

Modern basis sets use contracted sums of Gaussians in lieu of Slater-type orbitals.<sup>37</sup> The immediate disadvantage to using Gaussians to model orbitals is the incorrect  $e^{-r^2}$  asymptotic behavior, which is typically corrected by adding additional Gaussian functions to the basis set that are *diffuse*, which retards the Gaussian decay at distances far from the nucleus to better model the Slater-type  $e^{-r}$  behavior. A typical contracted Gaussian type orbital (CGTO) has the form,

$$\emptyset^{CGTO}(\{\xi\}, r) = \sum_{i=1}^{M} d_i N x^a y^b z^c e^{-\xi_i r^2}$$

where M is the length of the contraction, a + b + c = L is the angular momentum part of the orbital, and  $d_i$  and  $\xi_i$  are atom-dependent contraction parameters. The coefficient sets {d} and {  $\xi$ } are typically calculated variationally by minimizing the *ab initio* SCF energy of individual atoms with respect to the coefficients.

The construction of a CGTO atomic orbital occurs by building up Gaussian *1s*type lobes. For instance, a *3s* orbital of an atom is purely built up by a contracted sum of *1s* CTGOs. A *2p* orbital is built by placing two *1s*-type lobes in the correct orientation in space; for example, the CTGO representing  $2p_x$  has the form  $Nxe^{-\xi r^2}$ .

A single atomic orbital requires a sum of a large number of contracted Gaussians (generally > 3) in order to represented with chemical accuracy.<sup>38,39</sup> However, this

complication is readily counteracted due to the closed form of Gaussian integrals, so computational time to calculate the two-electron integrals in the Fock matrix using CTGOs is significantly shorter than for STOs. This is enabled by the Gaussian product theorem, which enables reduction of the four-center integral [ij|kl] to a two-center integral [J|K]. To explore this, consider the two-electron integral defined earlier in equation (5), with Gaussian orbitals as the chosen spatial basis set. Each product of orbitals sandwiching the  $r_{12}^{-1}$  operator can be rewritten as a new Gaussian function on a third center,

$$\phi_{i}^{*}(\xi_{i}, r - r_{A})\phi_{j}(\xi_{j}, r - r_{B}) = K_{IJ}\phi_{ij}'(\xi_{i} + \xi_{j}, r - r_{C})$$

where  $K_{IJ}$  is a constant dependent on the contraction coefficients  $\xi$  and the internuclear distance  $|r_A - r_B|$ , and  $r_C$  is a new weighted centroid of  $r_A$  and  $r_B$ . Therefore, the four-center integral can always be deconvoluted as

$$[ij|kl] = K_{IJ}K_{KL} \int dr_1 dr_2 \, \phi'(\xi_i + \xi_j, r_1 - r_C) \frac{1}{r_{12}} \phi(\xi_k + \xi_l, r_2 - r_C)$$

Since practical basis sets representing atomic orbitals contain sums of multiple Gaussians per orbital, multiple steps of the product theorem might be required to reduce the integral to two centers.

The most ubiquitous basis set used for quantum chemistry calculations is John Pople's *split-valence basis sets*, typically noted as a *X-YZg* basis.<sup>40</sup> In this notation, X corresponds to the number of Gaussian functions used to represent a core atomic orbital; Y+Z is the number of Gaussians used for the valence orbitals. For instance, consider a carbon atom in the 6-31G basis. This implies that the 1s orbital is represented by 6 Gaussian functions, and the 2s and 2p orbitals are represented in hybrid fashion as a sum of 3 + 1 Gaussian functions, for a total of 10 Gaussians per carbon atom. In atoms like carbon, Pople basis sets typically represent the valence s/p orbitals as "SP" hybrids in order to reduce complexity. Typically, the valence orbital represented by Z has a much smaller  $\xi$  coefficient than the Y valence orbitals, in order to account for Slater-type probability density far away from the nucleus, which is essential for behavior that occurs on the "edges" of the atomic probability density, such as intermolecular interactions or anionic states where the lone electron is typically held far away from the core.

The 6-31G basis is typically called a *split-valence, double-\zeta basis set.*<sup>37</sup> The splitvalence refers to the use of only 1 contracted function of Gaussians for each core orbital, and a larger basis size for the valence. The double-  $\zeta$  refers the use of two sets of contracted Gaussians to represent the valence, e.g. the 3+1 motif in 6-31G. Consequently, better performance can often be obtained by increasing the basis set size to a triple- $\zeta$ split-valence basis, such as 6-311G.

Polarization and additional diffuse functions are also commonly added in a Pople basis set. For instance, the 6-31+G basis set contains an additional set of *s* and *p*-type Gaussians with small  $\zeta$  coefficients to improve long-range interactions for non-hydrogen atoms, and the 6-311++G basis includes these improvements for hydrogens as well. The 6-31G(d) includes a *d*-orbital Gaussian for mixing with the valence in non-hydrogen atoms in order to improve polarization, and 6-31G(d,p) also includes *p*-orbital character to hydrogen atoms. These common scale up to including up to 3 *d*-orbitals and 1 f orbital for heavy atoms, and 3 *p*- and 1 *d*- for hydrogens (e.g. 6-31G(3df,3pd)). Not surprisingly, these two kinds of improvements can be mixed, to make basis sets as large as 6-311++G(3df,3pd), which corresponds to a carbon basis set size of 16 Gaussian functions.<sup>41,42</sup>

When Pople basis sets are applied to calculations that correct for electron correlation – e.g. MPn, CCSD(T), etc – the convergence towards the infinite basis set limit is often not achieved.<sup>43</sup> Pople basis sets were constructed primarily to extrapolate towards the Hartree-Fock limit, but when correlated methods are applied larger Pople basis sets can sometimes lead to poorer results than expected, or smaller basis sets might garner unpredictably better results due to error cancelation.<sup>44</sup>

To combat this, a number of basis sets have been developed in order to maximize the effectiveness of post-Hartree Fock calculations. In many of these basis sets, the coefficients on the CGTOs are optimized using calculations that explicitly treat electron correlation in atoms, whereas Pople-style basis sets merely derive their parameterization using the SCF method. The most common of these correlation-consistent basis sets is Dunning's (*aug*)-*cc-pVnZ* series<sup>45</sup>, where n = D(ouble), T(riple), Q(uadruple), ... is the specification for the  $\zeta$  characteristic of the basis set. The *cc-pVnZ* series was designed specifically to ensure smooth convergence to the complete basis set (CBS) limit<sup>46</sup>, for example the energy bounded from below by, for instance, a second-order perturbation treatment of electron correlation with an infinite basis set. Initially the *cc-pVnZ* basis sets were designed to converge quickly for *frozen-core* calculations, where correlation is only explicitly treated between electrons in the valence shell, but more recent developments from Dunning and coworkers have improved the core-valence correlation convergence of these basis sets.<sup>47</sup> The use of CBS-extrapolated calculations using Dunning's basis sets and the

CCSD(T) level of theory is typically considered the "gold standard" of quantum chemistry calculations, with accuracies of approximately 0.1% for correlation energies of atoms and small molecules<sup>48</sup> and 1.5% for non-covalent interaction energies.<sup>49</sup>

# References

- (1) Dirac, P. A. M. Proc. R. Soc. Lond. Ser. Contain. Pap. Math. Phys. Character **1929**, *123*, 714.
- (2) Mulliken, R. S. In *Nobel Lectures, Chemistry 1963-1970*; Elsevier: Amsterdam, 1972.
- (3) Hoffmann, R.; Woodward, R. B. Acc. Chem. Res. 1968, 1, 17.
- (4) Gimarc, B. M. Acc. Chem. Res. 1974, 7, 384.

(5) Szabo, A. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Courier Dover Publications, 1996.

(6) Cook, D. B. *Handbook of Computational Quantum Chemistry*; Courier Dover Publications, 2012.

- (7) Becke, A. D. J. Chem. Phys. **2014**, 140, 18A301.
- (8) Simons, J. J. Phys. Chem. **1991**, 95, 1017.
- (9) Slater, J. Rev. Mod. Phys. 1963, 35, 484.
- (10) Roothaan, C. Rev. Mod. Phys. 1951, 23, 69.
- (11) Møller, C.; Plesset, M. S. Phys. Rev. 1934, 46, 618.
- (12) Seiler, R. *Helvetica Phys. Acta* **1973**, *46*, 230.
- (13) Sutcliffe, B. T.; Woolley, R. G. J. Chem. Phys. 2012, 137, 22A544.
- (14) Sutcliffe, B. T.; Woolley, R. G. J. Chem. Phys. 2014, 140, 037101.
- (15) Khoromskij, B. N. ArXiv14084053 Math 2014.
- (16) Hall, G. G. Proc. R. Soc. Lond. Math. Phys. Eng. Sci. 1951, 205, 541.
- (17) Raghavachari, K.; Anderson, J. B. J. Phys. Chem. 1996, 100, 12960.

- (18) Johnson, E. R.; Becke, A. D. J. Chem. Phys. 2006, 124, 174104.
- (19) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. Chem. Phys. Lett. 1988, 153, 503.

(20) Leininger, M. L.; Allen, W. D.; Iii, H. F. S.; Sherrill, C. D. J. Chem. Phys. 2000, 112, 9213.

- (21) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. J. Chem. Phys. 1987, 87, 5968.
- (22) Scuseria, G. E.; Iii, H. F. S. J. Chem. Phys. 1989, 90, 3700.
- (23) Hohenberg, P.; Kohn, W. Phys. Rev. 1964, 136, B864.
- (24) Kohn, W.; Sham, L. J. Phys. Rev. 1965, 140, A1133.
- (25) Becke, A. D. J. Chem. Phys. 1993, 98, 1372.
- (26) Zhao, Y.; Truhlar, D. G. Theor. Chem. Acc. 2008, 120, 215.
- (27) Peverati, R.; Truhlar, D. G. *ArXiv12120944 Cond-Mat Physicsphysics Physicsquant-Ph* **2012**.
- (28) Goerigk, L.; Grimme, S. Phys. Chem. Chem. Phys. 2011, 13, 6670.
- (29) Burns, L. A.; Mayagoitia, Á. V.-; Sumpter, B. G.; Sherrill, C. D. J. Chem. Phys. **2011**, *134*, 084107.
- (30) Grimme, S. **2011**, *1*, 221.
- (31) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. J. Chem. Phys. 2010, 132, 154104.
- (32) Biczysko, M.; Panek, P.; Scalmani, G.; Bloino, J.; Barone, V. J. Chem. Theory Comput. **2010**, *6*, 2115.
- (33) Lee, K.; Murray, É. D.; Kong, L.; Lundqvist, B. I.; Langreth, D. C. *Phys. Rev. B* **2010**, *82*, 081101.
- (34) Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. *Phys. Rev. Lett.* **2004**, *92*, 246401.
- (35) Guseinov, I. I. J. Phys. B At. Mol. Phys. 1970, 3, 1399.
- (36) Chong, D. P.; Van Lenthe, E.; Van Gisbergen, S.; Baerends, E. J. *J. Comput. Chem.* **2004**, *25*, 1030.
- (37) Hehre, W. J.; Stewart, R. F.; Pople, J. A. J. Chem. Phys. 1969, 51, 2657.
- (38) Davidson, E. R.; Feller, D. Chem. Rev. 1986, 86, 681.
- (39) Almlof, J.; Helgaker, T.; Taylor, P. R. J. Phys. Chem. 1988, 92, 3029.
- (40) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. J. Chem. Phys. 1980, 72, 650.
- (41) Feller, D. J. Comput. Chem. 1996, 17, 1571.

(42) Schuchardt, K. L.; Didier, B. T.; Elsethagen, T.; Sun, L.; Gurumoorthi, V.; Chase, J.; Li, J.; Windus, T. L. *J. Chem. Inf. Model.* **2007**, *47*, 1045.

- (43) Dunning, T. H. J. Chem. Phys. 1989, 90, 1007.
- (44) Feller, D.; Peterson, K. A. Chem. Phys. Lett. 2006, 430, 459.
- (45) Woon, D. E.; Jr, T. H. D. J. Chem. Phys. 1995, 103, 4572.
- (46) Petersson, G. A.; Braunstein, M. J. Chem. Phys. 1985, 83, 5129.
- (47) Peterson, K. A.; Jr, T. H. D. J. Chem. Phys. 2002, 117, 10548.
- (48) Ranasinghe, D. S.; Petersson, G. A. J. Chem. Phys. 2013, 138, 144104.
- (49) Řezáč, J.; Hobza, P. J. Chem. Theory Comput. 2013, 9, 2151.

# **III.** A Computational Approach to Intermolecular Interactions

#### 1) Introduction

Since a number of experimental studies in this dissertation relate to noncovalently bound molecular complexes, it is necessary to provide a brief, qualitative overview of intermolecular interactions, in particular those in a molecular dimer. Typically, many-body effects are treated as expansions on two-body interactions, so description of the intermolecular forces within a dimer should be sufficient for understanding the fundamentals of non-covalent interactions. Although the only true physical observable in an intermolecular interaction is the binding energy between two or more monomers,

$$E_{\text{binding}}(A-B) = E_{AB} - E_A - E_B \tag{1}$$

multiple approaches can be undertaken in order to decompose the binding energy into set of chemically meaningful terms. Historically, one of the original treatments is Fritz London's perturbation theory formulation, where the monomers are treated as noninteracting (zeroth-order), and the perturbing interaction potential is replaced with a multipole expansion.<sup>1</sup>

In more modern calculations, using perturbative corrections to Hartree-Fock theory, such as Møller–Plesset (MPn) perturbation theory<sup>2,3</sup> or coupled-cluster<sup>4,5</sup> calculations, one can calculate intermolecular binding energies with excellent accuracy. This kind of traditional interaction energy calculation requires calculation of all three terms on the right side of equation (1). However, in the calculation of E<sub>AB</sub>, the basis set used is a coupled sum of the monomer orbitals ("dimer-centered") while  $E_A$  and  $E_B$  are calculated using monomer-centered orbitals. Favorable overlap between each set of monomer orbitals in the dimer-centered calculation can lead to overestimation of the binding energy that is not accounted for in the monomer energy calcuations. In order to correct for this, a common method called *counterpoise correction* is used, where the binding energy is corrected using monomer energies in the dimer-centered basis set.<sup>6</sup>

However, these modern methods only return the physical observable, the total interaction energy. However, especially in the context of London's theory of intermolecular interactions, one would expect that this observable can be partitioned into a set of chemically meaningful terms, such as the Coloumb and exchange interactions between the two monomers, and attractive inductive or dispersive effects. After all, the interactions seen in noble gas dimers are physically very different than the hydrogen bonding interaction in  $(H_2O)_2$ , so it would be very meaningful to quantify the specific differences between these two kinds of intermolecular interactions.

In particular, symmetry adapted perturbation theory (SAPT) was developed in order to directly analyze a chemically relevant partitioning of a dimer's interaction energy. Instead of returning only the total interaction energy, SAPT is partitioned such that energetic values for electrostatic, exchange, induction and dispersion interactions can be directly calculated, as seen in Equation 2.

$$E_{SAPT0} = E_{elst}^{(1)} + E_{exch}^{(1)} + E_{ind}^{(2)} + E_{disp}^{(2) 2}$$
(2)

 $<sup>^{2}</sup>$  The superscripts (1)/(2) refer to the order of the perturbation series in the interaction operator at which these terms are first non-zero.

By using this formalism as a basis for a discussion on intermolecular interactions, one can correlate common motifs in intermolecular interactions in a computationally revealing one, as a set of couplings and interactions between the orbital spaces of each monomer. This is a somewhat different approach than that of London<sup>1</sup>, where dispersion and electrostatics are treated as extensions of classical theory. Since this dissertation focuses on the computational aspects of molecular science, it is advantageous to describe intermolecular forces as a set of responses born out in a target system's set of occupied and virtual (unoccupied) orbitals.

Much of this discussion is adapted from Szalewicz and coworkers' excellent review on SAPT<sup>7</sup>, as well as the references from early literature made in that review article. I will also use the diagrammatic construction of the SAPT partition, developed initially by Szalewicz<sup>8,9</sup> and coworkers and later described wonderfully by Parrish and Sherill<sup>10</sup>, in order to describe the computational formalism for each of the intermolecular forces. In order to simplify the discussion, SAPT theory that is zeroth-order in the monomer basis (SAPT0) will only be discussed. This implies that each monomer is treated with only Hartree-Fock theory (e.g. the calculated intramolecular potentials are mean-field), and only the interactions between each monomer are treated perturbatively to higher order. Otherwise, couplings between intramolecular and intermolecular excitations can occur in the occupied and virtual orbital spaces of both monomers, which significantly complicated the mathematics. Additionally, discussion of the SAPT expansion itself will be limited to second-order, as the terms arising in second-order from the expansion are enough to describe the most important fundamental aspects of intermolecular interactions.

Α	В	Definition			
А	В	Set of all monomer nuclei			
а	b	Occupied orbitals of monomer			
m	n	Virtual (unocc.) orbitals of monomer			
Matrix Element		Definition			
$V_a^A$		Nuclear-electron Coloumb integral			
$S_b^a$		One-electron overlap integral; <b>[a b]</b>			
$v^{ab}_{mn}$		Two-electron integral; [ab mn]			
$V^A_{ab}$		Electrostatic potential of monomer A; = $V_{bm}^{A} + 2v_{ba}^{ma} = \left[a \left  \frac{Z_{a}}{r_{aA}} \right  m\right] + [am ba]$			

**Table 1.** Table of symbols used in this section, for a dimer with monomers A and B. Capital letters refer to monomer nuclei, and lower-case variables refer to electronic orbitals. Einstein summation notation is assumed where variables are repeated in expressions. X/x and Y/y correspond to arbitrary but unique monomers. To prevent confusion,  $V^{\alpha}_{\beta}$  will always refer to the electrostatic potential of monomer  $\alpha$  only when  $\beta$  is

a product of two occupied orbital sets.

## 2) Elements of the SAPTO Partition

## a) Electrostatics

The primary contributor to an intermolecular potential is its electrostatic interaction, which is succinctly defined as the total sum of Coloumb forces between the electrons and nuclei of each monomer. The calculable expression for this term<sup>3</sup> is

$$E_{elst} = \iint d\vec{r}_1 \ d\vec{r}_2 \ \rho_A(\vec{r}_1) \frac{1}{r_{12}} \rho_B(\vec{r}_2) = 2V_a^B + 2V_b^A + 4v_{bb}^{aa} + \frac{Z_A Z_B}{r_{AB}}$$
(3)

In lieu of considering the individual electrostatic contributions from each occupied orbital (e.g. electrons) of each monomer, which would be a multiple summation over all combinations of occupied orbitals and nuclei, an identical result can be found by integrating over the charge densities  $\rho$  of monomers A and B. Consequently, the matrix elements arising from this integration can be interpreted as the total attractive interactions between occupied orbitals and nuclei of each monomer  $(2V_a^B + 2V_b^A)$ , the standard four-index, two-electron repulsion integrals  $v_{bb}^{aa}$ , and the total nuclear-nuclear repulsion  $\frac{Z_A Z_B}{r_{AB}}$ .

<sup>&</sup>lt;sup>3</sup> Einstein summation notation is used on the right side of the expression, where the terms are summed over the set of all nuclei (A/B) and occupied orbitals (a/b) in each monomer. This applies to all equations in this section.

## b) Exchange

Exchange is a purely quantum mechanical effect, arising from the exchange symmetry of identical particles. In this discussion only the exchange effect between two electrons localized in a set of two spin orbitals is considered – e.g. matrix elements of the form  $\langle i | \hat{O} | j \rangle$  – where only values where the two orbitals are spin-matched are non-zero, as the "spin overlap" part of the overlap integral,  $\langle \alpha(i) | \beta(j) \rangle$ , ( $\alpha \neq \beta$ ), is zero by definition. Since the only non-zero contributions to the exchange energy are between orbitals containing electrons of the same spin, exchange is purely a repulsive effect by Pauli exclusion.

To evaluate the exchange energy computationally, the typical approach is to split the calculation into three electronic response terms, e.g.  $A \rightarrow B$ ,  $A \leftarrow B$ , and  $A \leftrightarrow B$ . This is expressed in the following equation:

$$E_{exch} = -2S_a^b S_b^m V_{am}^B - 2S_b^a S_a^n V_{bn}^A - 2S_a^n S_b^m v_{mn}^{ab}$$
(4)

The first two terms correspond to exchange responses  $A \leftarrow B$  and  $A \rightarrow B$ , respectively. For instance, in the  $A \rightarrow B$  term  $-2S_b^a S_a^n V_{bn}^A$ , the exchange interaction correspond to non-zero interactions between the occupied orbitals of monomers A and B ( $S_b^a$ ), minor but nonzero overlap between the unoccupied orbitals of B and the occupied orbitals of A ( $S_a^n$ ) and the electrostatic potential parameterized by a two-electron integral. Effectively, this corresponds to an induction between a monomer's electrons and the other body; for instance, the A $\rightarrow$ B term implies that exchange of electrons in monomer A are occurring through Pauli-excluded overlap with orbitals in monomer B, induced by the total electrostatic field generated by B. Although this approximation is done in first order in the SAPT partition, the exchange term is actually second-order in the perturbation expansion of the polarization operator<sup>9</sup>, so one must also consider concerted two-particle exchange ("doubles" excitation) in our first order expression. Physically it quantifies, to leading-order, concerted exchange between two electrons. Computationally, this does not complicate matters as four-index matrix elements such as  $v_{ab}^{mb}$  are calculated directly in the Hartree-Fock process, which treats electrostatic and exchange effects exactly through calculation of these four-index integrals.

#### c) Induction

In the SAPT formalism, induction is a very similar interaction to that of exchange. However, as noted in the previous section, exchange arises through an Pauli-excluded overlap interaction between, in the B $\rightarrow$ A interaction, the occupied orbitals of monomer A and the virtual orbitals of B, induced by the electromagnetic field of monomer B. However, in B $\rightarrow$ A, induction manifests itself as a single self-excitations in the occupied orbitals of monomer A into its own virtual orbitals, induced by the field created by monomer B. For the B $\rightarrow$ A inductive effect, this is represented as,

$$E_{ind}^{(2)} = 2x_{am}^{B}V_{am}^{B'}$$
(5)

where  $x_{am}^{B}$  is the single excitation amplitude, which can be defined as:

$$x_{am}^{B} = \frac{V_{am}^{B}}{E_{a} - E_{m}}, \text{ where } V_{am}^{B} = \sum_{nuclei B} \iint dr_{1} dr_{2} \psi_{a}(r_{1}) \psi_{m}(r_{1}) \frac{1}{r_{12}} \rho_{B}(r_{2})$$
(6)

However, this evaluation typically overestimates the induction energy, as it does not account for Pauli exclusion arising from the electrons in monomer B (accounted for in  $\rho_B$  in equation 6). In order to correct for this, an additional term is calculated, the induction-exchange ( $E_{exch-ind}$ ) energy, that also must be calculated. This term has a similar form to  $E_{ind}$ , except that the electrostatic potential term  $V_{ar}^{B'}$  is replaced with a modified potential that accounts for exchange between occupied orbitals of both monomers, but with field-induced singles excitations within the orbital space of the induced monomer.

### d) Dispersion

Dispersion is unique in the zeroth-order SAPT partition in that it manifests itself purely in second order, in that calculations in the Hartree-Fock limit explicitly do not treat the effect. Dispersion arises as an attractive effect caused by correlated one-electron excitations between each of the monomer's occupied and virtual orbital spaces, e.g.  $a \rightarrow m$  and  $b \rightarrow n$ . In the context of second-order Møller-Plesset perturbation theory (MP2), the energy of these excitations can be accounted for in the dimer-centered frame of reference (e.g. the dimer is treated as a singular molecular object) arise in the standard MP2 energy calculation:

$$E_{MP2}^{(2)} = \sum_{a < b; m < n} \frac{|v_{ab}^{mn}|^2}{(\epsilon_a + \epsilon_b) - (\epsilon_m + \epsilon_n)}$$
(7)

where, as usual, m/n and a/b are arbitrary virtual and occupied 1-electron orbitals, respectively.

However, in the SAPT partition, as each monomer is treated as a separable part of the system, these cross-monomer excitations are "mean-field" to any other excitations occurring between monomers. For instance, consider the example where excitations corresponding to both  $[a \rightarrow m / b \rightarrow n]$  and  $[a' \rightarrow m' / b' \rightarrow n']$  have non-zero amplitudes, implying the two electron repulsion matrix elements  $v_{ab}^{mn}$  and  $v_{a'b'}^{m'n'}$  are non-zero. In the SAPTO approximation, these are completely uncoupled, in that crossterms of the form  $v_{ab'}^{mn'}$  (or  $v_{a'b}^{m'n} / v_{ab'}^{m'n}$ , for instance) are assumed to be zero. This assumption can lead to overbinding between the two monomers, since the amplitude of a specific intra-monomer excitation can be reduced/quenched by a non-zero amplitude arising in one of these ignored cross-terms. This is contrary to the dimer-centered MP2 formalism, where these terms are all considered in the calculation of the overall electron correlation energy. This drawback of SAPT0 can be improved by using Møller-Plessettype perturbative corrections to the dispersion contribution, or doing the SAPT calculation to a higher leading order, but these are computationally costly and beyond the scope of this summary.<sup>11</sup>

In order to evaluate the dispersion energy, the amplitudes and matrix elements of the two-electron excitations are merely summed over, for a total dispersion energy of

$$E_{disp}^{(2)} = 4 \sum_{a,b} \frac{v_{ab}^{mn} v_{mn}^{ab}}{(\epsilon_a + \epsilon_b) - (\epsilon_m + \epsilon_n)}$$
(8)

Like induction, one has to account explicitly for the fact that this term does not account for penetration of a monomer into the Pauli exclusion zone of the other, so an additional dispersion-exchange energy correction is typically calculated in order to explicitly account for this repulsive effect.

#### 3) SAPTO Results for Simple Models

SAPT0 is applied to a select set of in-house optimized structures and compared to literature results, in order to illustrate the ease and simplicity of analysis that SAPT0 provides for an experimental chemist interested in the energetic profile of a target non-covalent complex. Unless otherwise noted, all studied systems were initially optimized using the MP2/6-311++g(d,p) level of theory, and then subsequently reoptimized using B3LYP-D3/aug-cc-pVTZ. All SAPT0 calculations are done using David Sherrill and coworkers' implementation of SAPT0 with the jun-cc-pVDZ basis set, which has been well studied in the literature as a reliable level of theory for qualitative analysis of non-covalent interactions without significant computational cost.<sup>11</sup> All calculations were performed using the PSI4 beta 5 quantum chemistry package, which is open source and freely available online.<sup>12</sup>

Table 1 summarizes the SAPT0 interaction energies for a variety of benchmark noncovalent complexes, as well as a selection of intermolecular parameters determined by a selection of common quantum chemical methodologies. These complexes range from those with intermolecular interactions that are highly electrostatic in nature --  $(H_2O)_2$  for instance – to others that are almost entirely dispersive in nature, such as  $(CH_4)_2$  and the T-shaped benzene dimer. In general, the SAPT0 results are in good qualitative agreement with accepted literature values for binding energies.

		SAPT0/						
System	<b>∆E</b> elst	$\Delta E_{exch}$	$\Delta E_{ind}$	$\Delta E_{disp}$	$\Delta E_{total}$	Lit. Value	Lit methodology	
(benzene) <sub>2</sub> [T-								
shaped]	-2.14	4.27	-0.63	-4.29	-2.78	-2.74	CCSD(T)/CBS <sup>13</sup>	
benzene … CH4	-0.95	2.30	-0.27	-2.17	-1.09	-1.47	CCSD(T)/CBS <sup>5</sup>	
$(CH_4)_2 (D_{3h})$	-0.19	0.69	-0.03	-0.61	-0.14	-0.54	CCSD(T)/CBS <sup>5</sup>	
(phenol) <sub>2</sub>	-9.91	10.54	-3.12	-4.89	-7.37	-6.64	CCSD(T)/CBS <sup>14</sup>	
(H <sub>2</sub> O) <sub>2</sub>	-8.88	7.14	-2.15	-1.28	-5.17	-5.02	Expt + $ab$ initio $D_e$ corrections <sup>15</sup>	
Structural parameters [6-311++g(d,p)]								
System	HF	<i>MP2</i>	B3LYP	B3LYP-D3	M06-2X	Lit. Value	Lit methodology	
(H2O)2							<b>MB-ERS</b> / $r_s^{16}$	
r[O … O] /Å	3.00	2.91	2.90	2.91	2.88	2.98(1)		
180° - θ[O-H ··· O]	0.9	3.2	5.0	7.4	6.1	7(6)		
(CH4)2 [D3h]							CCSD(T)/CBS <sup>17</sup>	
$r[C \cdots C] / \mathring{A}$	5.18	4.01	7.05	3.64	3.84	3.63		
(benzene) <sub>2</sub> [T-								
shaped]							FTMW / $r_s^{18}$	
$r[H \cdots \pi] / \mathring{A}$	3.29	2.31	3.10	2.43	2.43	2.50		
benzene ··· CH4							CCSD(T)/aug-cc-pVQZ <sup>5</sup>	
$r[H \cdots \pi] / \mathring{A}$	3.84	2.65	4.31	2.65	2.58	2.70		
(phenol) <sub>2</sub>							CP-FTMW / r <sub>0</sub> (Chapter 3)	
$r[O \cdots O] / \mathring{A}$	2.98	2.83	2.88	2.86	2.83	2.83(2)		
hinge angle / °	104.5	48.2	110.5	57.0	59.6	64(1)		

**Table 1.** SAPT0 and structural results for a selection of molecular complexes. All calculations were performed with identical convergence criteria of  $10^{-6}$  Ha for  $\Delta E$ ,  $3x10^{-4}$  Ha/bohr for force, and  $1.2x10^{-3}$  bohr for geometric displacement.

The methane dimer is a benchmark example of a very weakly bound, dispersiondominated molecular dimer. With a predicted binding energy of  $< 200 \text{ cm}^{-1}$ , the methane dimer has yet to be detected experimentally. However, it provides a suitable benchmark structure to gauge the efficacy of theoretical methodology since the interaction requires treatment of the intermolecular potential to, at the very least, second order. The SAPTO results are indicative of this, as  $\Delta E_{disp}$  dominates the attractive part of the intermolecular interaction energy.

Structurally, the dispersion-bound nature of the methane dimer is illustrated through the C  $\cdots$  C distances for HF and B3LYP, which are 5.18 and 7.05 Å respectively. These values reflect structures are within the convergence criteria for energy, but not in terms of displacement. This suggests that a minimum is found in the potential energy surface, but it is extraordinarily shallow so that the geometric optimization itself oscillates around the minimum. Tightening of the convergence criteria could improve the HF and B3LYP determinations, but the poor result is illustrative of the importance of proper treatment of displaying proper optimization convergence. M06-2X also shows similar results. MP2 is slightly poorer, but this is a typical result as MP2 has been shown to poorly estimate binding energies and geometries for dispersion-bound clusters compared to higher level *ab initio* methods such as CCSD(T).<sup>19–21</sup>

Replacing a methane with the  $\pi$  system of a benzene ring, a trend begins to appear where the addition of additional electrostatic attraction improve the relevant structural parameters from HF and B3LYP. However in these cases the SAPTO attraction due to dispersion is still twice the magnitude of the electrostatic attraction, and as such the HF and B3LYP intermolecular distances are off by roughly 0.7 and >1 Å for (benzene)<sub>2</sub> and benzene…methane, respectively.

On the far end of the interaction spectrum is  $(H_2O)_2$ , where the attractive force is almost exclusively electrostatic ( $\Delta E_{elst} + \Delta E_{ind}$ ) in nature. In this case, nearly all methodologies used result in a satisfactory parameterization of the water hydrogen bond. This is consistent with the SAPTO formalism, where dispersion largely arises as an interaction that first appears at second order. As such, the simplest methodologies can represent the electrostatic hydrogen bond with relatively good accuracy. However, in cases such as (phenol)<sub>2</sub>, where the intermolecular interaction has competition between electrostatic (hydrogen bond) and dispersive ( $\pi$ - $\pi$  stacking) forces, the simplest theoretical methods fail outright. A more detailed discussion of the competition between dispersion and electrostatics and the associated structural phenomena can be found in the study of phenol dimer and trimer in Chapter 3.

SAPT0 is a useful qualitative tool for decoding intermolecular interactions of a dimer. The SAPT0 calculation time is miniscule (no calculation took more than 5 minutes of wall time on a desktop computer) and provide results that are in qualitative agreement with accepted literature values. For experimentalists interested in molecular structure, SAPT0 can be a valuable tool for not only analyzing intermolecular interactions in observed dimers, but also as a means of feedback for "tuning" the theoretical methodology used for structure prediction in order to optimize the balance between speed and proper treatment of intermolecular interactions. Although the method here is only applicable to a two-body molecular complex, expansion of the SAPT partition can be expanded into a many-body calculation. John Herbert's group at the Ohio State University has published promising results for creating a many-body polarization operator<sup>22</sup> that lends itself to caluclable many-body SAPT partitions with low computational cost.<sup>23</sup> The so-called XSAPT fragment method scales as O(n) where *n* is the number of monomers, which is promising for expanding this kind of intermolecular exploration to larger complexes. They can also be combined with dispersion-corrected density functionals in order to fully optimize the computational cost of the fragmentation method.<sup>24</sup> However, these methods are not yet available in common quantum chemistry suites, so their application for routine use is still a future, but promising, goal.

# References

- (1) London, F. *Trans. Faraday Soc.* **1937**, *33* (0), 8b.
- (2) Møller, C.; Plesset, M. S. Phys. Rev. 1934, 46 (7), 618.
- (3) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* **1988**, *153* (6), 503.
- (4) Scuseria, G. E.; Iii, H. F. S. J. Chem. Phys. 1989, 90 (7), 3700.
- (5) Sherrill, C. D.; Takatani, T.; Hohenstein, E. G. J. Phys. Chem. A **2009**, *113* (38), 10146.
- (6) Boys, S. F.; Bernardi, F. Mol. Phys. **1970**, 19, 553.
- (7) Jeziorski, B.; Moszynski, R.; Szalewicz, K. Chem. Rev. 1994, 94 (7), 1887.
- (8) Rybak, S.; Jeziorski, B.; Szalewicz, K. J. Chem. Phys. 1991, 95 (9), 6576.
- (9) Moszynski, R.; Jeziorski, B.; Szalewicz, K. J. Chem. Phys. 1994, 100 (2), 1312.
- (10) Parrish, R. M.; Sherrill, C. D. J. Chem. Phys. 2014, 141 (4), 044115.

(11) Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D. J. Chem. *Phys.* **2014**, *140* (9), 094106.

(12) Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. A.; Fermann, J. T.; Mintz, B. J.; Burns, L. A.; Wilke, J. J.; Abrams, M. L.; Russ, N. J.; Leininger, M. L.; Janssen, C. L.; Seidl, E. T.; Allen, W. D.; Schaefer, H. F.; King, R. A.; Valeev, E. F.; Sherrill, C. D.; Crawford, T. D. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2* (4), 556.

(13) Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. J. Am. Chem. Soc. **2002**, *124* (36), 10887.

(14) Kołaski, M.; Kumar, A.; Singh, N. J.; Kim, K. S. *Phys. Chem. Chem. Phys.* **2010**, *13* (3), 991.

(15) Leforestier, C.; Szalewicz, K.; Avoird, A. van der. J. Chem. Phys. **2012**, *137* (1), 014305.

(16) Dyke, T. R.; Mack, K. M.; Muenter, J. S. J. Chem. Phys. 1977, 66 (2), 498.

(17) Li, A. H.-T.; Chao, S. D. J. Mol. Struct. THEOCHEM 2009, 897 (1–3), 90.

(18) Schnell, M.; Erlekam, U.; Bunker, P. R.; von Helden, G.; Grabow, J.-U.; Meijer, G.; van der Avoird, A. *Angew. Chem. Int. Ed.* **2013**, *52* (19), 5180.

(19) Jureka, P.; Sponer, J.; Cerny, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8* (17), 1985.

(20) Tkatchenko, A.; Jr, R. A. D.; Head-Gordon, M.; Scheffler, M. J. Chem. Phys. **2009**, *131* (9), 094106.

(21) Cybulski, S. M.; Lytle, M. L. J. Chem. Phys. 2007, 127 (14), 141102.

(22) Herbert, J. M.; Jacobson, L. D.; Lao, K. U.; Rohrdanz, M. A. *Phys. Chem. Chem. Phys.* **2012**, *14* (21), 7679.

- (23) Lao, K. U.; Herbert, J. M. J. Phys. Chem. A 2015, 119 (2), 235.
- (24) Lao, K. U.; Herbert, J. M. J. Phys. Chem. Lett. 2012, 3 (22), 3241.

# **IV.** Chirped-Pulse Fourier Transform Microwave Spectroscopy

#### 1) Historical Background

Microwave (MW) spectroscopy had its first major developments after World War II, thanks to the significant research and development on RADAR technology performed during wartime. After the invention of the klystron amplifier by the Varian brothers and subsequent commercialization by companies such as Varian and defense contract Raytheon, the first generation of laboratory microwave spectrometers became available.<sup>1</sup>

One of the most successful designs of the first generation was the Stark modulation absorption spectrometer of Hughes and Wilson<sup>2,3</sup>, which used the Stark effect to frequency-modulate a target molecular transition to be on-resonance with a klystron amplifier in order to maximize the absorption. In Wilson's design, the cell was a long piece of evacuated K-band (20-40 GHz) waveguide filled with a sample gas at room temperature and low pressure (ca. 10 mTorr). In practice the K-band region is not ideal for measurement of rotational transitions at room temperature, as the rotational partition function for polyatomic molecules scale as  $T^{1.5}$ , and the peak of the Boltzmann distribution for a small polyatomic molecule such as methanol is >100 GHz. However, the technique was highly successful for molecular detection<sup>4–6</sup> and observation of effects such as internal rotation<sup>7</sup>, tunneling effects<sup>8</sup> and nuclear quadrupole coupling.<sup>9,10</sup> In fact, in the early 1970s, Hewlett-Packard (unsuccessfully) marketed a commercialized version of Hughes and Wilson's design with a more modern backwards-wave oscillator (BWO) amplifier<sup>11</sup>, a design that is still used by some microwave labs.<sup>12–14</sup> The development of pulsed-jet molecular beams by Smalley, Wharton and Levy<sup>15,16</sup> enabled a means to improve the inherent limitations of static-gas spectrometers in the cm-wave (2-30 GHz) region by cooling the gas to a rotational temperature of <10 K using a supersonic adiabatic expansion into a vacuum chamber evacuated to a very low pressure (ca. 10<sup>-6</sup> torr). In lieu of absorption measurements, molecular excitations could be probed by using a short pulse of resonant microwave radiation in order to establish a coherence, which is then allowed to relax and emit a coherent free induction decay (FID). The FID is then heterodyned and collected in the time-domain and then Fourier transformed in a similar fashion to FT-NMR.<sup>17</sup> For a typical microwave transition excited coherently in the pulsed jet, the Doppler lattice relaxation time scale (*T*<sub>2</sub>) is generally ca. 5-10 µs so a high duty cycle, along with high frequency resolution, can easily be achieved. Additionally, with a properly phase-stable receiver, this signal can be averaged in the time-domain and achieve an additional signal to noise (S/N) ratio improvement equivalent to the square root of the number of averages.

The most influential instrument design exploiting the pulsed-jet beam for MW spectroscopy is Balle and Flygare's Fabry–Pérot cavity pulsed-jet spectrometer, first published in 1981.<sup>18,19</sup> The Balle-Flygare FTMW spectrometer applies a similar technique as described above, but adds a highly reflective microwave cavity ( $Q > 10^4$ ) tuned to the target molecular resonance in order to achieve a significant sensitivity boost.

The Balle-Flygare design has significant advantages over the first generation of microwave spectrometers. One, the FTMW technique is background-free, as only coherent emission from the resonant molecules in the jet is observed. As such, saturation effects from the incident radiation or high number density are a non-issue as they are in

absorption techniques. In theory, the "N-emitter" effect guarantees that, for sample molecular beams containing N molecules where the collision relaxation rate  $T_p$  is slower than  $T_2$ , the intensity of the emitted decay should scale as N. In practice, this is largely defined by the pumping rate of the vacuum system, in order to minimize the chamber pressure when each sample injection pulse occurs. Additionally, the incident polarization pulse (typically a square-wave pulse) is tuned in order to generate a " $\pi/2$ " coherence, which in a two-state Rabi problem implies an ideal coherence of  $1/\sqrt{2}$  (|*lower*) + |*upper*)).<sup>20</sup>

Second, the pulsed-jet sample introduction cools the sample to near 1 K, which optimizes the Boltzmann peak to near the cm-wave frequencies used in the experiment. This also reduces the number of populated vibrational states (except for some largeamplitude motions) to near-zero, significantly simplifying the observed spectra.

The pulsed-jet cooling introduces a simple method to form non-covalently bound molecular clusters in significant abundance. Coupled with the high sensitivity of the multi-pass cavity, the high frequency resolution of the technique enabled detection of direct molecular structure with 0.1 Å or better structural resolution for molecular clusters. As such, Balle-Flygare FTMW spectrometers revolutionized the study of molecular clusters, and have seen much success over the last 35 years.<sup>21–23</sup>

However, the most significant limitation of the Balle-Flygare technique is the cavity itself; in order to maintain a high Q factor for molecular emission, the excitation bandwidth must stay very small, generally on the order of 1 MHz. The experiment is therefore bottle-necked in time by the need to tune the physical cavity to stay in

resonance with the incident light. This ultimately limits the applicability of the Balle-Flygare technique for large molecular systems or complex mixtures which often result in spectrum with high densities of transitions arising from multiple conformers, isotopologues and molecular clusters. Since many of these transitions can be detected in a single band, identification of a target spectrum from the "weeds" can be difficult or require a long frequency scan to acquire a sufficiently large set of transitions. Additionally, since the mode quality will vary from band to band in a Balle-Flygare cavity, relative intensities of molecular transitions are often unreliable. For decoupling target rotational spectra out of dense clumps of resolved transitions, the lack of reliable relative intensity information can make identification difficult.

#### 2) The Chirped-Pulse FTMW Technique

## a) Introduction

Less than a decade before the publication of the Balle-Flygare cavity, McGurk et al. outlined a *fast passage* technique using a short microwave pulse with a fast linear frequency sweep to excite a large bandwidth of molecular transitions within the frequency range of the sweep.<sup>24,25</sup> By sweeping at a rate much faster than the dephasing time of the molecular coherence, a large swath of molecular transitions can be coherently excited simultaneously. If the sweep rate over a resonance is such that the Rabi "flip angle" for each two-level transition is small, then the resultant polarization (or correspondingly the intensity of the emission) will scale linearly with the power of the incident pulse or inversely proportional to the sweep rate. This enables the experimenter to have full control over the optimization of the detected intensity of a target molecular spectrum or set of transitions by tuning the bandwidth, or the power spectral density of the excitation pulse (W/Hz), and the sweep rate, or how long the polarization pulse's power "stays" near-resonant with a molecular transition.

However, at the time of McGurk and coworkers' study, the technology for generating linear frequency sweeps with high bandwidth and power was limited. Thankfully, modern computing technology has enabled the production of arbitrary waveform generators (AWG) with high sampling rates (Nyquist frequencies of greater than 10 GHz) and bit depth, and high-speed digitizers for detection of both heterodyned (for high frequency) or directly detected molecular emission. The development of chirped-pulse Fourier transform microwave (CP-FTMW) spectroscopy at the University of Virginia exploits these developments directly.<sup>26</sup> By using an AWG to generate a short linear frequency sweep over a target bandwidth and a modern oscilloscope to detect the molecular emission, large bandwidths of molecular transitions can be probed simultaneously. In fact, the bandwidth of the pulse is largely limited by the availability of amplification sources, which are usually limited to specific standardized bands due to engineering limitations, and the sampling rate of the AWG/digitizer. The latest generation of AWGs offer sampling rates up to 65 GS/s with 20 GHz of bandwidth (Agilent M8195A) and oscilloscopes with 100 GS/s and >30 GHz of bandwidth (Tektronix DPO7000 series).

Since the large bandwidth of the CP-FTMW pulses prevent the use of cavity methods to enhance signals, high power sources must be used to provide a sufficient amount of polarization to a pulsed-jet sample. For instance, the 2-8 GHz spectrometer at UVa<sup>27</sup> uses a 600 W peak power traveling wave tube (TWT) amplifier, coupled with a high gain/high directionality broadcast antenna, to provide a polarization pulse of sufficient intensity. Since coherent emission power will scale as  $N^2$ , multiple pulsed-jet nozzles are used in order to improve the signal by a factor of N (since the *electric field*, not the power, is the detected quantity). Additionally, since all acquisitions are coherently averaged, the signal to noise ratio improves as  $(time)^{1/2}$ . Therefore, by using N nozzles for sample injection, the required time to reach a desired signal to noise ratio reduces as  $N^2$ .

# b) Principle of Operation

The typical CP-FTMW experiment consists of a vacuum chamber at high vacuum (ca.  $10^{-6}$  torr) with one or more pulsed jet nozzles arranged perpendicularly to the polarization axis of the experiment. Two high-gain directional microwave horn antennas designed for the frequency range chosen for the experiment are placed at opposite ends of the chamber. For the instrument design at UVa for the 2-8 and 6-18 GHz arrangements, 5 pulsed nozzles separated at ca. 30 cm is optimal, but this reduces to 4 and 3 for 18-26 and 26-40 GHz, respectively.

The chirped pulse is a linear frequency sweep with starting frequency  $v_0$  and sweep rate  $\dot{v}$  and has the functional form

$$S(t) = \cos(v_0 t + \frac{1}{2}\dot{v}t^2)$$

where the bandwidth  $\Delta v = \dot{v} t_{pulse}$ .

For measurements at cm-wave, the chirped pulse is typically on the order of 1-4  $\mu$ s. Since the *T*<sub>2</sub> dephasing time is typically on the order of 10  $\mu$ s and the emission begins

immediately after the pulse dissipates, a full single CP-FTMW acquisition can occur within 50  $\mu$ s. Since each gas pulse lasts approximately 700  $\mu$ s, multiple acquisitions can occur during a single gas pulse. The UVa spectrometer uses the Fast Frame averaging capability of the Tektronix high-performance oscilloscopes to exploit the length of the gas pulse, acquiring and averaging up to 10 coherent acquisitions per sample injection cycle. The overall repetition rate of this sequence depends largely on the processing speed of the digitizer and the pumping rate of the vacuum system. For a typical 2-8 GHz measurement at UVa using the 100 Gs/s Tektronix oscilloscope, a record length of 40  $\mu$ s at a 25 GS/s sampling rate for 8 acquisition frames limits the overall repetition rate to about 5 Hz , depending on backing pressure used. At 6-18 GHz, a similar repetition rate can be achieved by reducing to 20  $\mu$ s at 50 GS/s for 10 acquisition frames. However, this rate is largely a function of the speed of the digitizer; the previous generation 50 GS/s oscilloscope used at UVa could only achieve a repetition rate of about 1 Hz.

The assumption of the "small angle", or weak pulse, limit of CP-FTMW spectroscopy introduces a large amount of experimental flexibility. For instance, as high-powered broadband TWTs are very expensive, the sensitivity loss due to loss of incident power can be overcome by reducing the bandwidth of the pulse. In the weak pulse limit, the detected intensity of a transition scales as  $(\Delta v)^{-1/2}$  and as  $E_{pulse} \approx P_{input}^{0.5}$ , so the losses caused by reducing the input power by a factor of two can be made up by reducing the bandwidth of the chirped pulse by a factor of four.<sup>28–30</sup>

#### c) Instrument Design

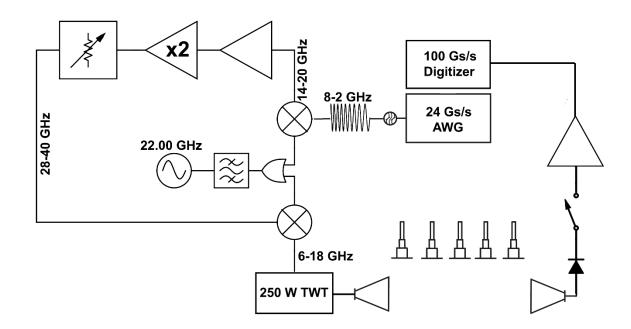
In short, a basic CP-FTMW experiment has three primary stages. The first is the excitation stage, where a chirped pulse is generated and amplified at the appropriate frequencies. The second is the interaction region, where the chirp pulse polarizes the pulsed-jet sample. And third is the detection stage, where the resultant molecular emission is amplified and/or heterodyned for digitization in the time-domain.

For the 2-8 GHz band, the chirped pulse is directly generated from the 24 Gs/s AWG (Tektronix 7122B dual-channel, 12 Gs/s per channel with interleaving). Since the AWG used can also generate spurious content that is at integer multiples of the output frequency, a 12 GHz low-pass cavity filter (K&L Microwave) is used to retain only the desired pulse, which is output by the AWG at 0 dBm. This pulse is then amplified to 400 W RMS / 600 W max power with a pulsed traveling wave tube (TWT) microwave amplifier (Applied Systems Engineering 167S/C). The amplified pulse is then broadcasted using a high gain, 2-18 GHz microwave horn antenna (Q-Par Angus WBH 2-18-NHG) and allowed to interact with the sample from up to 5 pulsed nozzles (Parker General Valve Series 9). Each pulsed nozzle is separated by ca. 25 cm and are perpendicular to the propagation of the chirped pulse. The molecular emission is then detected using another microwave horn antenna of the same make.

Since the amplified chirped pulse is significant enough to cause damage to the components used, a high power PIN diode limiter and PIN diode SPST switch (ATM Microwave S1517D, 20 ns switching time) are placed immediately after the receiver horn in order to protect the rest of the system from the high powered polarization pulse. Once the chirp is allowed to dissipate, the switch is opened to allow the molecular FID to pass, which is then amplified by a low-noise, high-gain amplifier (Narda-Miteq 1-20 GHz

LNA, 2.2 dB noise figure). The amplified FID is then directly detected and digitized on a 100 Gs/s, 33 GHz bandwidth Tektronix oscilloscope (Tektronix DPO73304D).

For the 6-18 GHz band, since the chirped pulse cannot be directly generated by the 24 Gs/s AWG, a slightly more complicated circuit must be used for the excitation stage. First, a chirped pulse is generated from 8-2 GHz (reverse chirp). This pulse is then filtered using a 12 GHz low-pass filtered and then upconverted to 14-20 GHz by mixing with the signal of a 22.00 GHz phase-locked resonant dielectric oscillator (PDRO). The pulse is then amplified by a general purpose microwave pre-amplifier (Wright Tech 10-20 GHz, 30 dBm gain) and doubled to 28-40 GHz by an active doubler (Marki ADA1020, +20 dBM output). This intermediate high frequency pulse is attenuated to ca. 0 dBm using a variable attenuator, and then mixed and downconverted with the same 22.00 GHz PDRO to reach the target 6-18 GHz pulse. This pulse is then amplified using a 250 W RMS, 450 W max TWT amplifier (Applied Systems Engineering 167X/KU) and is broadcasted and detected in the interaction region using either the 2-18 GHz Q-Par Angus broadband horns or a specialized set of 6-18 GHz horns (ATM Microwave 650-442-C1). Figure 1 shows a schematic of 6-18 GHz setup just described. In order to ensure phase stability for coherent averaging throughout the entire experiment, a 10 MHz rubidium oscillator is used to lock all components with clock inputs, such as the digitizer, AWG and any single frequency source (e.g. PDROs).



**Figure 1**. Schematic of the 6-18 GHz CP-FTMW instrument at UVa. In the case of the 2-8 GHz instrument, the 2-8 GHz pulse is directly inputted into the TWT from the AWG with no intermediate mixing circuit. The remainder of the circuit (interaction and detection) is identical. The 10 MHz rubidium oscillator is not shown.

#### d) Data Reduction and Analysis Tools

Since the collected data is a coherently averaged time-domain molecular signal, it must be Fourier transformed in order to extract the relevant frequency information. The typical work-up process at UVa is application of a discrete fast Fourier transform (FFT) routine written using the numpy and scipy<sup>31</sup> FFTPACK libraries for Python 2.7<sup>32</sup>, which extends a set of core numerical analysis routines written natively in C for optimal performance. Using Cython<sup>33</sup>, improved performance can be obtained by statically typing the data structures used to process the time-domain signal by compiling the code in a standard C compiler. For 2-8 GHz spectra measured at a record length of 40 µs, this FFT process takes less than 500 ms without the use of Cython. With Cython, the processing time can be reduced by more than order of magnitude. This allows the user to balance between the ease of use and flexibility of Python with the speed and power of C.

The time-domain average is zero-padded to twice the length, which improves the frequency spacing of the FT by a factor of two. For a 40 µs time-domain average, this results in a FT frequency spacing of 12.5 kHz. For more precise measurement of center frequencies of specific molecular frequencies, the spectrum is typically peak-picked after being interpolated to 1-2 kHz using a cubic spline implemented in scipy, or by picking single transitions and fitting them to a Gaussian lineshape function. In general, these two approaches result in very similar center frequency measurements. As such, the cubic spline is typically used for routine analysis as it can be applied to a complete broadband spectrum quickly. In contrast, application of a Gaussian lineshape function requires a full least-squares optimization for every transition in the data set.

Additionally, in order to improve baseline sensitivity of the Fourier transform, a Kaiser-Bessel window function is applied to the time-domain data. While the Kaiser-Bessel window tends to broaden the FWHM of a given transition, the sensitivity gained along the wings is improved significantly, which can be essential for detection of isotopologues and other weak species with transitions near strong parent transitions. For more information on the benefits of window functions in CP-FTMW spectroscopy, see the discussion in Chapter 3.

Analysis of a pure rotational spectrum at UVa generally follows a two-step process. Initial inspection of the broadband spectrum is generally performed using David Plusquellic's JB95 fitting program.<sup>34</sup> JB95 is amazingly useful for CP-FTMW spectra, as it allows one to overlay a predicted spectrum on the experimental data and use a series of scroll bars that change the values of the dipole moment projections and rotational constants (up to quartic distortion) on the fly using a linear interpolation routine. If a potential combination of fit lines is found, then JB95 can fit the rotational spectrum to a specified set of rigid rotor and/or distortion constants, and then display the RMS error and updated prediction associated with this interim fit.

This allows for fast visual inspection of a CP-FTMW spectrum, as user-supplied changes in the predicted spectrum are immediately reflected across the entire bandwidth of the data set. With the accurate relative intensities provided by CP-FTMW spectroscopy, this often results in extremely fast visual identification of experimental spectra. If a target spectrum cannot be identified in the experimental data, perhaps due to spectral congestion or poor predictions, AUTOFIT can be used to check all possible

combinations of experimental transitions to search for a fit consistent with the theoretical predictions. A detailed description of AUTOFIT can be found in Chapter 2.

If an experimental rotational spectrum is identified, then publication-quality fits are typically generated using Kisiel's AABS spectrum viewer coupled with Pickett's SPCAT and SPFIT (CALPGM)<sup>35</sup> for spectral prediction and least-squares fitting, respectively. Typically, this combination results in fits with RMS errors better than 10 kHz, especially for near-rigid molecules at 2-8 GHz.

CALPGM is an extremely flexible program suite for fitting rotational spectra, as it provides a complete set of parameters for fitting perturbations and effects that arise from electronic and spin angular momentum as well as vibrational states. As such, spectra complicated due to nuclear quadrupole coupling from one or more nuclei, or vibrationally-excited spectra perturbed by Coriolis coupling or Fermi resonances can all be fit to better than experimental accuracy for practically any frequency region in the microwave, millimeter, or sub-millimeter ranges, with a program-set hard limit of J  $\leq$  369 / 999 for SPCAT/SPFIT. However, CALPGM is relatively hard to use for spectra complicated by internal rotation, and other programs such as Isabelle Kleiner's BELGI<sup>12,36,37</sup>, Peter Groner's *ERHAM*, <sup>38</sup> or Hartwig and Driezler's XIAM<sup>38</sup> are most commonly used to quantify internal rotation effects. More details on fitting internal rotation spectra can be found in Chapter 4.

#### 3) Applications of CP-FTMW Spectroscopy

The following chapters in this dissertation will summarize a number of studies of molecules and molecular complexes that illustrate the high sensitivity and applicability of the CP-FTMW technique for structural determination. Chapter 2 first describes the use of automated spectral fitting techniques via the program AUTOFIT. In general, this chapter describes the typical protocol for broadband spectrum analysis at UVa, including a study of quantum chemical methodologies for relevant for structure prediction and relative energy orderings. In particular, the focus will be on accurate predictions with computational complexity feasible for a standard desktop workstation. Using only a systematic quantum chemical study of the 1-hexanal conformational landscape, AUTOFIT enabled the full and automated assignment of all conformers detectable within the sensitivity limits of the acquired 6-40 GHz broadband spectrum. By comparison of the relative intensities of each conformer's pure rotational spectrum to the theoretical relative energy ordering, an approximate "conformational temperature" for the expansion, or equivalently an approximate energetic upper bound for detectable thermally populated conformers of 1-hexanal, can be determined.

Chapter 3 changes the focus to intermolecular interactions with the structural study of phenol dimer and trimer. Due to the large sizes of these non-covalent complexes, the 2-8 GHz instrument at UVa is particularly suited for structural studies as the band is particularly suited for the Boltzmann peak of pure rotational spectra of large complexes.<sup>27</sup> Phenol dimer in particular features a number of challenges due to the fine interplay of dispersive ( $\pi$ - $\pi$  stacking) and electrostatic-dominated (hydrogen bonding) interactions as well as large amplitude motion. The high structural precision arising from the substitution

and least-squares ground state structure determination of the phenol dimer allows for a case-study on the quantum chemical treatment of intermolecular interactions in structural prediction and provides significant improvement over previous experimental studies of the structure of phenol dimer.

Chapter 4, a structural study of the cluster of sevoflurane, a common inhalant anesthetic, and benzene, follows similar themes of the phenol dimer and trimer study. The 2-8 GHz instrument is used to determine the structure of sevoflurane...benzene with an unparalleled data set of over 40 isotopologues, the majority of which are detected in natural abundance. This enables a least-squares fit ground state structure determination with nearly every independent structural parameter floated in the fit. Additionally, the sevoflurane...benzene spectrum features the first detection of a high barrier six-fold internal rotation splitting arising from the hindered rotation of the benzene top about the sevoflurane frame. A full analysis of the internal rotation effect is presented, and a fit to experimental accuracy can be obtained using BELGI. Chapter 5 follows the sevoflurane...benzene study with the detection of the sevoflurane dimer. The dimer contains 24 heavy atoms and is one of the largest clustered studied to date by rotational spectroscopy. Thanks to the sensitivity afforded by the 2-8 GHz spectrometer, most heavy atom isotopologues were detected in natural abundance, enabling substitution (r<sub>s</sub>) and least-squares ground state  $(r_0)$  structure determinations.

Finally, Chapter 6 will discuss the structure determination of the tetramer of carbonyl sulfide (OCS). This section is a part of the study performed at UVa in collaboration with Robert McKellar that encompasses the structure determination of two OCS trimers and the tetramer using both theoretical methods and IR/microwave

spectroscopy.<sup>39</sup> The potential energy surface for OCS aggregations is complex and very shallow, and as such theoretical searches offer a large number of candidate structures very closely spaced in relative energy. In fact, a previous theoretical study of OCS clusters discussed four specific isomers of the tetramer – none being consistent with the observed experimental structure! As such, this study is a particular example of using broadband rotational spectroscopy to "guide" theoretical methodology. With the assumption that the monomer structure is unchanged upon complexation, a brute-force method can be used to calculate all possible structures arising from the data set used. Since the Kraitchman equations can constrain all atoms with experimental isotopic information to sub-Å precision and the experimental rotational constants of the parent species give a precise constraint for the overall structure, in theory only one of the possible structures can be consistent with the experimental data set. In practice, imaginary Kraitchman coordinates for some of the atoms cause some ambiguity and increases the number of consistent candidate structures.

These five studies reflect the versatility and power of the CP-FTMW technique for structure determination. However, they only represent a small subset of the output using the UVa CP-FTMW instrument. Previous work over the past 4 years has included the detection and structural elucidation of  $(H_2O)_{6-10}$ <sup>27,40,41</sup>, CH···· $\pi$  interaction studies<sup>42,43</sup>, clusters of small molecules with water<sup>41,44</sup>, routine and efficient structure determination of small silicon-containing organics<sup>45–48</sup>, and chirality detection<sup>49</sup>, amongst others. This prolific output is fully thanks to the high sensitivity and speed of the CP-FTMW technique, and further optimizations should increase the viability of detection of large molecules (>20 heavy atoms) and higher order molecular clusters.

# References

(1) Caryotakis, G. High Power Klystrons: Theory and Practice at the Stanford Linear Accelerator Center, Part I, 2005.

- (2) Hughes, R. H.; Wilson, E. B. *Phys. Rev.* 1947, 71, 562.
- (3) Jr, K. B. M.; Hughes, R. H.; Jr, E. B. W. *Rev. Sci. Instrum.* 1949, *20*, 821.
- (4) Martins, J. F.; Wilson Jr., E. B. J. Mol. Spectrosc. 1968, 26, 410.
- (5) Wilson, E. B. *Science* 1968, *162*, 59.
- (6) Wilson, E. B.; Smith, Z. Acc. Chem. Res. 1987, 20, 257.
- (7) Jr, R. E. N.; Jr, E. B. W. J. Chem. Phys. 1957, 26, 1057.

(8) Baughcum, S. L.; Smith, Z.; Wilson, E. B.; Duerst, R. W. J. Am. Chem. Soc. 1984, 106, 2260.

(9) Kivelson, D.; Jr, E. B. W.; Jr, D. R. L. J. Chem. Phys. 1960, 32, 205.

(10) Kurland, R. J.; Jr, E. B. W. J. Chem. Phys. 1957, 27, 585.

(11) Harrington, H. W.; Hearn, J. R.; Rauskolb, R. F. *Hewlett-Packard Journal*. June 1971, pp 1–16.

(12) Tudorie, M.; Kleiner, I.; Hougen, J. T.; Melandri, S.; Sutikdja, L. W.; Stahl, W. *J. Mol. Spectrosc.* 2011, *269*, 211.

(13) Nguyen, H. V. L.; Kleiner, I.; Shipman, S. T.; Mae, Y.; Hirose, K.; Hatanaka, S.; Kobayashi, K. *J. Mol. Spectrosc.* 2014, 299, 17.

(14) Samdal, S.; Grønås, T.; Møllendal, H.; Guillemin, J.-C. J. Phys. Chem. A 2014, 118, 1413.

- (15) Smalley, R. E.; Wharton, L.; Levy, D. H. J. Chem. Phys. 1975, 63, 4977.
- (16) Smalley, R. E.; Wharton, L.; Levy, D. H. Acc. Chem. Res. 1977, 10, 139.
- (17) Ekkers, J.; Flygare, W. H. Rev. Sci. Instrum. 1976, 47, 448.
- (18) Balle, T. J.; Flygare, W. H. Rev. Sci. Instrum. 1981, 52, 33.

(19) Campbell, E. J.; Buxton, L. W.; Balle, T. J.; Keenan, M. R.; Flygare, W. H. J. *Chem. Phys.* 1981, 74, 829.

(20) Grabow, J.-U. In *Handbook of High Resolution Spectroscopy*; John Wiley & Sons, Ltd, 2011; Vol. 1, pp 723–799.

(21) Legon, A. C. Annu. Rev. Phys. Chem. 1983, 34, 275.

(22) Leopold, K. R.; Fraser, G. T.; Novick, S. E.; Klemperer, W. *Chem. Rev.* 1994, *94*, 1807.

(23) Novick, S. Bibliography of Rotational Spectra of Weakly Bound Complexes,2015.

(24) McGurk, J. C.; Mäder, H.; Hofmann, R. T.; Schmalz, T. G.; Flygare, W. H. J. *Chem. Phys.* 1974, *61*, 3759.

(25) McGurk, J. C.; Schmalz, T. G.; Flygare, W. H. J. Chem. Phys. 1974, 60, 4181.

(26) Brown, G. G.; Dian, B. C.; Douglass, K. O.; Geyer, S. M.; Shipman, S. T.; Pate, B. H. *Rev. Sci. Instrum.* 2008, *79*, 053103.

(27) Pérez, C.; Lobsiger, S.; Seifert, N. A.; Zaleski, D. P.; Temelso, B.; Shields, G. C.; Kisiel, Z.; Pate, B. H. *Chem. Phys. Lett.* 2013, *571*, 1.

(28) Grubbs, G. S.; Dewberry, C. T.; Etchison, K. C.; Kerr, K. E.; Cooke, S. A. *Rev. Sci. Instrum.* 2007, *78*, 096106.

(29) Shipman, S. T.; Pate, B. H. In *Handbook of High-resolution Spectroscopy*; John Wiley & Sons, Ltd, 2011.

(30) Obenchain, D. A.; Elliott, A. A.; Steber, A. L.; Peebles, R. A.; Peebles, S. A.; Wurrey, C. J.; Guirgis, G. A. *J. Mol. Spectrosc.* 2010, *261*, 35.

(31) Van der Walt, S.; Colbert, S. C.; Varoquaux, G. Comput. Sci. Eng. 2011, 13, 22.

(32) Oliphant, T. E. Comput. Sci. Eng. 2007, 9, 10.

(33) Behnel, S.; Bradshaw, R.; Citro, C.; Dalcin, L.; Seljebotn, D. S.; Smith, K. *Comput. Sci. Eng.* 2011, *13*, 31.

(34) Plusquellic, D. F.; Suenram, R. D.; Maté, B.; Jensen, J. O.; Samuels, A. C. J. *Chem. Phys.* 2001, *115*, 3057.

(35) Pickett, H. M. J. Mol. Spectrosc. 1991, 148, 371.

(36) Kleiner, I.; Hougen, J. T. J. Chem. Phys. 2003, 119, 5505.

(37) Kleiner, I. J. Mol. Spectrosc. 2010, 260, 1.

(38) Hartwig, H.; Dreizler, H. Z Naturforsch 1996, 51a, 923.

(39) Evangelisti, L.; Perez, C.; Seifert, N. A.; Pate, B. H.; Dehghany, M.; Moazzen-Ahmadi, N.; McKellar, A. R. W. J. Chem. Phys. 2015, 142, 104309.

(40) Pérez, C.; Muckle, M. T.; Zaleski, D. P.; Seifert, N. A.; Temelso, B.; Shields, G. C.; Kisiel, Z.; Pate, B. H. *Science* 2012, *336*, 897.

(41) Pérez, C.; Zaleski, D. P.; Seifert, N. A.; Temelso, B.; Shields, G. C.; Kisiel, Z.; Pate, B. H. *Angew. Chem. Int. Ed Engl.* 2014, *53*, 14368.

(42) Ulrich, N. W.; Songer, T. S.; Peebles, R. A.; Peebles, S. A.; Seifert, N. A.; Pérez, C.; Pate, B. H. *Phys. Chem. Chem. Phys.* 2013, *15*, 18148.

(43) Ulrich, N. W.; Seifert, N. A.; Dorris, R. E.; Peebles, R. A.; Pate, B. H.; Peebles, S. A. *Phys. Chem. Chem. Phys.* 2014, *16*, 8886.

(44) Kisiel, Z.; Lesarri, A.; Neill, J. L.; Muckle, M. T.; Pate, B. H. *Phys. Chem. Chem. Phys.* 2011, *13*, 13912.

(45) Guirgis, G. A.; Klaassen, J. J.; Pate, B. H.; Seifert, N. A.; Darkhalil, I. D.; Deodhar, B. S.; Wyatt, J. K.; Dukes, H. W.; Kruger, M.; Durig, J. R. *J. Mol. Struct.* 2013, *1049*, 400.

(46) Guirgis, G. A.; Overby, J. S.; Palmer, M. H.; Peebles, R. A.; Peebles, S. A.; Elmuti, L. F.; Obenchain, D. A.; Pate, B. H.; Seifert, N. A. *J. Phys. Chem. A* 2012, *116*, 7822.

(47) Pate, B. H.; Seifert, N. A.; Guirgis, G. A.; Deodhar, B. S.; Klaassen, J. J.; Darkhalil, I. D.; Crow, J. A.; Wyatt, J. K.; Dukes, H. W.; Durig, J. R. *Chem. Phys.* 2013, 416, 33.

(48) Durig, J. R.; Guirgis, G. A.; Sawant, D. K.; Seifert, N. A.; Deodhar, B. S.; Pate,
B. H.; Panikar, S. S.; Groner, P.; Overby, J. S.; Askarian, S. M. *Chem. Phys.* 2014, 445, 68.

(49) Lobsiger, S.; Perez, C.; Evangelisti, L.; Lehmann, K. K.; Pate, B. H. J. Phys. Chem. Lett. 2014, 6, 196.

# **Chapter 2: AUTOFIT, an Automated Fitting Tool for Broadband Rotational Spectra, and Applications to 1-Hexanal**

(*The writing presented here has been peer-reviewed and published in the Journal of Molecular Spectroscopy, with the following citation:* N. A. Seifert, I. A. Finneran, C. Perez, D. P. Zaleski, J. L. Neill, A. L. Steber, R. D. Suenram, A. Lesarri, S. T. Shipman, B. H. Pate, *J. Mol. Spectrosc.* **2015**, *312*, 13-21.)

# 1) Introduction

The development of chirped-pulse Fourier transform microwave spectroscopy (CP-FTMW) has enabled broadband rotational spectroscopy measurements with high sensitivity.<sup>1,2</sup> However, the increase in sensitivity produces dense spectra that can achieve average line densities of over 1 MHz<sup>-1</sup>.<sup>3–6</sup> With line densities this high, traditional (e.g. manual) spectrum fitting techniques can be inefficient in part because visual patterns associated with distinct combinations of rotational transitions<sup>7</sup> can be difficult to identify at high line densities. These challenges are also encountered in room temperature gas samples where rotational transitions from thermally populated vibrationally excited states can cause significant congestion.<sup>8–10</sup>

One advantage of rotational spectroscopy, especially for jet cooled samples of large molecules, is that the rotational spectrum can often be quantitatively fit to the Watson asymmetric top Hamiltonian.<sup>11</sup> Even for molecules with internal rotors, like a methyl rotor, one of the symmetry species can be fit to this Hamiltonian. Furthermore, at low temperature the contribution from centrifugal distortion to the transition frequencies is small and the spectrum can be fit to good precision using just the rigid rotor Hamiltonian. Therefore, an approximate fit to the spectrum can be obtained from the starting point of fitting three transitions to the rigid rotor Hamiltonian to obtain the rotational constants *A*, *B*, and *C*. The validity of this spectral fit can be assessed by predicting a small set of additional transitions and checking for their presence. The AUTOFIT program described here implements this simple data analysis algorithm to fit all possible combinations of three transitions (called a 'triple') within a given search window. The search windows are established using *ab initio* estimates of the rotational spectroscopy parameters. Therefore, AUTOFIT is designed to test whether a candidate molecule is present in the broadband spectrum.

Previously published methods of automated spectral fitting have found success for rovibronic spectra using heuristic models such as genetic algorithms<sup>12–14</sup> or spectrum cross-correlation<sup>15</sup> to assign rotational structure automatically. Although these algorithms can determine rotational constants without the use of an initial parameter set of quantum numbers, spectra are fit using both frequency and intensity information. For rotational spectroscopy, the spectra of all species present in the sample overlap in the measurement spectral range and pose significant difficulty for fitting algorithms that use intensity information and implicitly assume a single spectrum is present.

The performance of AUTOFIT for analyzing complex mixtures is illustrated by the conformational analysis of 1-hexanal. The conformational flexibility of its aliphatic backbone leads to multiple populated conformers even in a cold molecular beam. In addition, the dynamic range of the measurement is high enough that isotopologues in natural abundance can be detected for the four lowest energy conformers. As a result, a complex sample mixture is generated from a single chemical sample. Success with similar molecular systems has been achieved using cavity FTMW and CP-FTMW techniques, with multiple conformers being detected for conformationally flexible molecules such as the related 1-heptanal<sup>16</sup>, 2-hexanol<sup>17</sup>, the long-chain alkenes 1-pentene through 1-octene<sup>18–20</sup>, cyclic species such as 15-crown-5 ether<sup>21</sup>, and numerous other molecules.<sup>22–26</sup>

# 2) Methodology

# a) Experimental

All measurements were performed with the CP-FTMW spectrometer at the University of Virginia. The spectroscopy was performed using 3 different arrangements in order to collect broadband data in the 6-18 GHz, 18-26, and 26-40 GHz bands. The details of the 6-18 GHz <sup>27</sup> and 26-40 GHz <sup>28</sup> spectrometers have been published previously, and also can be found in Chapter 1, Part IV of this dissertation. The full details of the 18-26 GHz arrangement are presented below, and a schematic of the microwave circuit can be found in Figure 1.

An initial 1 µs linear frequency sweep is generated from 10-1.5 GHz using a 24 GS/s arbitrary waveform generator (AWG, Tektronix AWG7122B). This sweep is then filtered using a 12.2 GHz low-pass filter in order to remove higher frequency components generated by mixing between the chirp and the clock frequency of the AWG. The chirp is then upconverted using a triple-balanced mixer (Miteq TB0440LW1) with a 28 GHz local oscillator (Microwave Dynamics, PLO-4000-28.00) that is filtered with a bandpass filter (K&L 3C62-28000/T100-K/K) in order to improve LO purity.

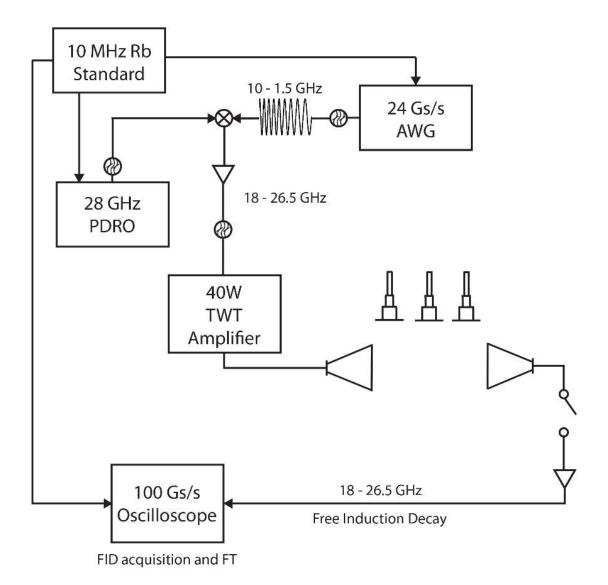


Figure 1. Schematic of the 18-26 GHz CP-FTMW spectrometer.

The lower sideband of this mixing stage, 18 – 26 GHz is attenuated to approximately 0.3 mW (-5.2 dBm) by passing through an 18-26.5 GHz preamplifier (Miteq AMF-3F-18002650-25-8P) and variable attenuator. The pulse is filtered using an 18-26.5 GHz bandpass filter (K&L 9IR05-22350/T8500-K/K) and amplified using a 40W 18-26.5 traveling wave tube amplifier (Applied Systems Engineering 187K).

The excitation pulse is transmitted through a vacuum chamber using 20 dBi gain microwave horn antennas (Advanced Technical Materials PNR 42-442-6), where it interacts with the pulsed jet expansion of three Parker General Valve Series 9 pulsed valves (model # 009-0181-050-2) separated by 20 cm. The receiver is protected from the high power TWT pulse by a PIN diode limiter (Aeroflex, ACLM-4765C36) and an SPST switch (American Microwave Corp., SWCH1K-DC40-SK). The molecular free induction decay (FID) is then amplified using an 18-40 GHz low-noise (3.3 dB noise figure), 48 dBm gain amplifier (Miteq, JS44-18004000-33-8P) and directly digitized on a 100 GS/s oscilloscope (Tektronix DPO73304D). 10 FIDs are collected on each sample injection cycle. A 10 MHz rubidium standard is used to synchronize and phase-lock all frequency generators in order to enable stable time-domain signal averaging.

Stark effect measurements were performed using a 7.5 – 18.5 GHz CP-FTMW spectrometer<sup>29</sup> with a Stark cage designed by Emilsson et al. <sup>30</sup> Due to limitations arising from the design of the Stark cage, only one pulsed valve could be used, and a total of 600 000 averages were collected at a field strength of 155 V/cm. 3,3,3-trifluoropropyne ( $\mu$  = 2.317 D) was used to calibrate the electric field strength. <sup>31</sup>

A 0.2% gas mixture of 1-hexanal (98%, Aldrich) in neon (>99%, GTS Welco/Praxair) was used. The gas sample was then pulsed into the vacuum chamber using 1 atm backing pressure. For the 6-18 GHz measurement a 50 GS/s sampling rate and a record length of 20  $\mu$ s were used (point spacing of 24 kHz), and 960,000 FIDs were averaged in the time-domain to achieve a Fourier transform with a dynamic range of approximately 5000:1. The 18-26 GHz data was digitized directly at 100 GS/s, and consists of a time-domain average of 1.9 million FIDs at a 10  $\mu$ s record length (point spacing of 50 kHz), resulting in an approximate dynamic range of 3500:1. Finally, the 26-40 GHz data is an average of 1.1 million FIDs with a sampling rate of 50 GS/s and a record length of 10  $\mu$ s (point spacing of 47.8 kHz), with a resultant dynamic range of 800:1. The repetition rate of the experiments were 1.2 Hz, 3 Hz and .77 Hz, for the 6-18, 18-26 and 26-40 GHz experiments, respectively.

## b) Quantum Chemistry

All possible conformers of 1-hexanal (a total of 36 distinct structures) were optimized using B3LYP <sup>32</sup> with the standard Pople basis set 6-311++g(d,p). <sup>33</sup> The B3LYP structures were then re-optimized using the M06-2X/6-311++g(d,p) level of theory. <sup>34</sup> Additionally, all experimentally observed conformers were re-optimized using the MP2 <sup>35</sup> and B3LYP-D3 <sup>36,37</sup> methods. The B3LYP and M06-2X calculations were performed using GAUSSIAN 09 <sup>38</sup> with default convergence criteria (10<sup>-6</sup> Hartree), and the MP2 and B3LYP-D3 calculations were performed using Psi4 beta 5 <sup>39</sup> with identical convergence criteria. Unless otherwise noted, references to calculated values with a defined methodology (e.g. M06-2X) use a 6-311++g(d,p) basis set. In order to speed up calculations, the MP2 calculations were performed using density fitting with an auxiliary aug-cc-pVTZ-RI basis set. <sup>40</sup> The B3LYP-D3/6-311++g(d,p) calculations were performed using density fitting with the auxiliary aug-cc-pVTZ-JKFIT basis set. In addition, a B3LYP-D3/aug-cc-pVTZ (aVTZ) <sup>41</sup> set of calculations were performed in order to compare B3LYP-D3 improvements by increasing the basis set size.

#### c) AUTOFIT

AUTOFIT automates the process of assigning pure rotational spectra in broadband microwave data. By providing the program with a predicted geometry and set of dipole moments predicted with quantum chemistry and experimental parameters such as frequency range, intensity bounds and rotational temperature, AUTOFIT will search an input microwave data set for an experimental spectrum consistent with the predicted parameters. If an experimental spectrum associated with the target system is both found and has a sufficient signal/noise ratio, AUTOFIT can then use the inputted geometry to predict and search for isotopically-substituted spectra for the purposes of structural determination or confirmation.

Rotational spectrum prediction and fitting is done by interfacing with the SPFIT and SPCAT programs of the CALPGM suite. <sup>42</sup> AUTOFIT functions as an automated generator of SPCAT and SPFIT input files and an analyzer of the corresponding output files. First, an initial SPCAT prediction is generated from the input geometry and dipole moments. Three predicted transitions ("fitting" transitions) are then chosen with the criteria that they are sufficiently intense and linearly independent, in order to minimize the uncertainties on the rotational constant determination. AUTOFIT can generate a set of fitting transitions automatically using the independence and intensity metrics, but the ability to manually choose is also available to the user. An additional set of transitions ("scoring" transitions) are also chosen by the user and later used to verify a potential experimental spectral fit using the three fitting transitions.

AUTOFIT will then automatically fit every possible combination of observed spectral lines (called a "triple") within three search windows corresponding to the three fitting transitions. The size of the search windows are set by the error associated with rotational constant prediction of the chosen quantum chemical methodology. For instance, if the user-chosen triplet of fit transitions has initial predicted frequencies of 10, 12 and 15 GHz, and a 500 MHz search window is applied to each, then AUTOFIT will check every triple of observed spectral lines between 9.75-10.25, 11.75-12.25 and 14.75-15.25 GHz, respectively. AUTOFIT can automatically provide the user with appropriate search window sizes for each fitting transition, but the option to manually override the window sizes is also available.

After the algorithm assigns the fitting transitions to an experimental triple in SPFIT, the spectrum is then repredicted using the least-squares fit rotational constants. The previously chosen set of scoring transitions are then forward predicted and are evaluated by comparing the predicted frequencies to the nearest observed transitions in the spectrum. A total RMS fit error is then calculated using the residuals of the N additional transitions ("scoring" transitions) with the center frequencies of the nearest observed experimental lines. The determined rotational constants and the microwave RMS error of the (N+3) line fit are recorded by AUTOFIT in an output file. A simplified flowchart of this AUTOFIT process can be found in Figure 2.

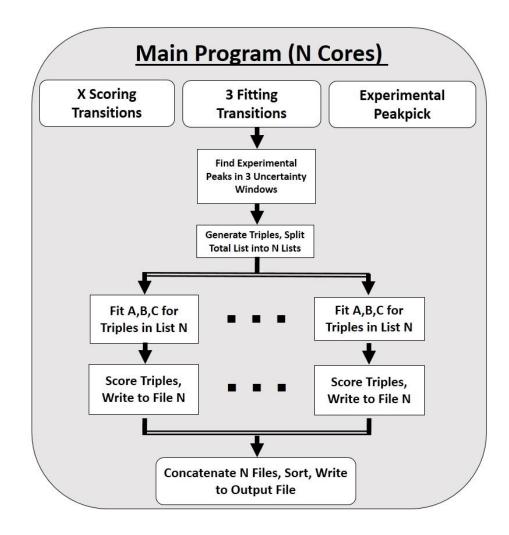


Figure 2. Simplified flowchart of the main AUTOFIT process.

One limitation to AUTOFIT is that distortion constants are considered as fixed initial values, such as those obtained from quantum chemistry estimates. AUTOFIT does not directly float distortion when processing a triple. Only the triple is used to fit the spectrum, so there is insufficient data to determine more than the three rigid rotor rotational constants. Although this can lead to inflations in the RMS error of the fit, it prevents determination of unphysical rotational constants by biasing the fit with distortion constants. Since all the AUTOFIT results treat distortion as a fixed quantity, the final results of an AUTOFIT search generally follow the trend of the best results having the lowest RMS error. Despite this, AUTOFIT results for spectra at low frequency (e.g. cmwave) are not significantly reduced in quality by the lack of distortion constants in the fit, and satisfactory results can generally be achieved by inputting only the rigid rotor parameters. However, AUTOFIT does have the capability of taking a user-selected result after an AUTOFIT run has completed and automatically fitting user-specified distortion constants to the result.

AUTOFIT is specifically tuned to do spectral searches in low-frequency (e.g. 2-40 GHz) microwave data using supersonic expansions. This limits the value of both J and  $K_a/K_c$  for transitions observed in a broadband spectrum, minimizing the negative impact of fixing distortion constants to *ab initio* values or to zero. Fixing all quartic distortion constants to zero, the resultant fit RMS errors are within 100 kHz. Although these fits are by no means quantitative, an RMS error of this magnitude implies predicted frequencies are generally within the typical linewidth in the CP-FTMW experiment (ca. 120 kHz). Inclusion of *ab initio* values of the quartic distortion constants would likely improve the initial AUTOFIT results. However, by revising these distortion-less fits with a set of fitted

quartic distortion constants, the average RMS error reduces to <10 kHz, verifying that the AUTOFIT search results are physical.

AUTOFIT is also able to do isotopologue searches using scaled experimental rotational constants. By using scaled predicted isotopologue rotational constants and thereby assuming the theoretical structure approximates the actual structure (within a few percent error in A, B and C), the search windows are dramatically reduced in size with respect to those of the parent species and a full set of possible isotopologue fits can be generated in a few minutes.

Since AUTOFIT is a brute force algorithm, additional approaches need to be considered in order to counteract the O(N<sup>3</sup>) scaling for minimizing run time on a typical desktop computer. A number of simple approaches to reducing runtime are available in the script. These include using intensity bounds for experimental lines, and search window sizes for each of the three chosen transitions, which can be important when dealing with transitions that have different statistical dependences on the rotational constants. For instance, the compounded error in a typical  $\mu_a$ -type (*a*-type) transition will be lower than a pure  $\mu_b$ -type (*b*-type) transition since the *a*-type transition will likely only depend mostly on *B* and *C*, whereas the *b*-type transition will also depend on *A*. Therefore, placing a smaller search window on the *a*-type transitions will reduce the number of required triples to check without losing search quality.

AUTOFIT was programmed exclusively in Python 2.7. Typical speeds for AUTOFIT on a modern, multicore desktop computer range from 35-50 Hz per thread, so an eight-threaded Intel i7 CPU can achieve speeds of up to 400 Hz. Since a typical "worst-case" AUTOFIT search is around 16-20 million triples, most AUTOFIT jobs can be completed

97

overnight on a typical desktop computer. AUTOFIT is open source and is freely available from the Pate lab website <sup>43</sup> along with download links to all required dependencies.

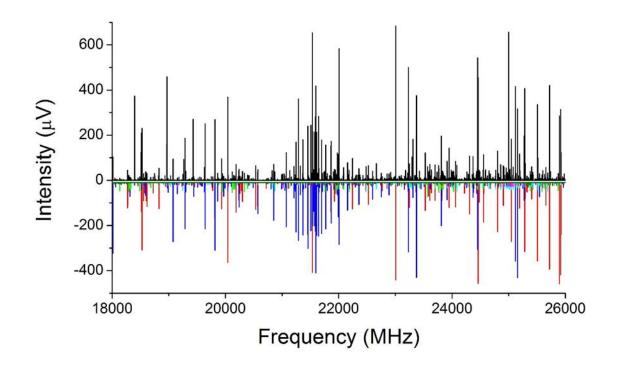
# 3) Results

There are 38 distinct conformations for 1-hexanal. Each of the three aliphatic dihedrals in 1-hexanal can only take on three possible conformations  $^{44}$  – ca. 180° (anti, or "a") or ca.  $\pm 60^{\circ}$  (gauche, or "g<sup>±</sup>"). Additionally, there is the carbonyl dihedral ( $\psi$ [O-C<sup> $\alpha$ </sup>-C<sup> $\beta$ </sup>-C<sup> $\gamma$ </sup>]), which has an global minimum eclipsing arrangement of ca. 0°. Therefore, there are a maximum of 3<sup>3</sup> possible conformers of carbonyl-eclipsing 1-hexanal conformers, but by accounting for mirror symmetry this set of conformers reduces to 14 unique structures.

Of these 14 possible conformers, 4 are of significantly higher energy due to compounding gauche interactions, and their energies range from 500-800 cm<sup>-1</sup> above the global minimum aaa-e conformer, so it is unlikely these will be detected in a molecular beam experiment. This leaves a total of 10 possible eclipsed-carbonyl conformers likely detectable. There are an additional 4 gauche carbonyl (where  $\psi$ [O-C<sup> $\alpha$ </sup>-C<sup> $\beta$ </sup>-C<sup> $\gamma$ </sup>]  $\approx \pm 120^{\circ}$ ) rotamers calculated to be within 500 cm<sup>-1</sup> of the global minimum aaa-e conformer, for a total of 14 candidate conformers.

A total of 12 conformers were detected in the CP-FTMW spectrum using AUTOFIT. A visual summary of the CP-FTMW spectrum, along with predicted intensities and frequencies of the 12 assigned conformers, can be found in Figure 3. The experimental rotational parameters and dipole moments of the 12 identified 1-hexanal conformers can be found in Table 1. Only conformers 1-6 were sufficiently intense to perform Stark effect measurements. The detected conformers include all four of the predicted lowest-energy gauche-carbonyl rotamers, as well as eight conformers with an eclipsing carbonyl group. The remaining two conformers, ggg-e and ag<sup>-</sup>g<sup>+</sup>-e, were not found after extensive AUTOFIT searches with window sizes corresponding to 5% error in rotational constants, as well as manual searches. Final fits were performed using the CALPGM program suite. The theoretical relative energies and rotational spectroscopy parameters are given in Table 2.

The two most intense conformer spectra are strong enough to detect <sup>18</sup>O isotopologues in natural abundance (ca. 5000:1 S/N). Conformers 3 and 4 are slightly weaker and only <sup>13</sup>C isotopologues are detectable (ca. 300:1 S/N). With this data, full heavy atom Kraitchman structures <sup>45,46</sup> were determined for conformers 1 and 2, as well as for the carbon backbones of 3 and 4. Figure 4 shows a comparison between the Kraitchman results and the B3LYP-D3/aVTZ structures of conformer 1-4, and a comparison of the experimental heavy atom back bone structural parameters can be found in Table 3.



**Figure 3.** The 18-26 GHz spectrum of 1-hexanal (top, black), with predictions of the 12 assigned conformers (multiple colors, bottom). Specifically, blue transitions correspond to conformer 1 (aaa-e), red to conformer 2 (aag-e), and green to conformer 3 (aga-e).

Conformers with weaker intensity are additionally shown with differently colored transitions. Relative intensities have been scaled using the same averaged scale factors used in Figure 5. The rotational temperature chosen for the predictions was 3.6 K, which best reproduces the spectral intensities for all 12 conformers.

	1 (aaa-e)	2 (aag-e)	3 (aga-e)	4 (gaa-e)	
A / MHz	9769.63045(68)	5399.89397(20)	8975.4925(15)	5995.20639(45)	
B / MHz	868.845880(92)	1143.248678(62)	933.43862(21)	1046.73937(11)	
C / MHz	818.518770(92)	1028.990827(75)	898.08347(22)	945.97194(10)	
⊿J / kHz	0.04627(10)	0.30363(15)	0.09094(23)	0.19870(29)	
$\Delta_{JK}$ /					
kHz	-0.8843(19)	-1.6494(11)	-2.5930(39)	-2.0086(20)	
$\Delta_K / \mathrm{kHz}$	24.193(85)	14.5675(39)	72.48(30)	18.060(33)	
$\delta_J$ / kHz	0.004957(39)	0.064466(49)	-0.00256(11)	0.049058(68)	
$\delta_K / \mathrm{kHz}$	0.101(31)	1.2275(65)	-3.005(97)	1.996(17)	
$P_{cc}$ / u					
$Å^2$	7.98273(5)	22.25264(2)	17.49621(9)	16.43335(4)	
$ \mu_a $ / D	1.274(3)	0.515(2)	1.918(8)	0.983(13)	
$ \mu_b $ / D	2.288(2)	2.292(5)	1.651(6)	2.370(10)	
$ \mu_c $ / D		1.012(7)	0.877(7)	0.715(15)	
Ν	270	323	237	259	
$\sigma$ / kHz	8.5	8.4	8.7	9.1	

	5 (g <sup>+</sup> ag <sup>-</sup> -e)	6 (g <sup>+</sup> ag <sup>+</sup> -e)	7 (aag <sup>+</sup> -g <sup>-</sup> )	8 (ag <sup>+</sup> g <sup>+</sup> -e)	
A / MHz	6116.50288(72)	4667.6175(11)	5455.60826(59)	4827.91033(52)	
B / MHz	1167.30037(22)	1336.76933(37)	1055.75603(11)	1240.84154(21)	
C / MHz	1059.54145(19)	1166.12535(37)	937.27598(10)	1159.91361(32)	
⊿J / kHz ⊿ <sub>JK</sub> /	0.28619(89)	0.8711(14)	0.50355(30)	1.00661(96)	
kHz	-2.5335(68)	-5.4158(58)	-6.1640(22)	-9.3475(83)	
$\Delta_K / \mathrm{kHz}$	17.889(62)	20.716(59)	41.404(41)	40.213(60)	
$\delta_J$ / kHz	0.03319(26)	0.25912(45)	0.118903(82)	0.16839(34)	
$\delta_K / \mathrm{kHz}$	0.772(41)	3.324(26)	2.323(11)	1.620(82)	
$P_{cc}$ / u Å <sup>2</sup>	19.29667(6)	26.4752(2)	16.06209(6)	38.13096(7)	
$ \mu_a $ / D	0.0461(22)	0.581(7)			
$ \mu_b $ / D	2.251(10)	2.469(8)			
$ \mu_c $ / D	0.833(15)	0.19(4)			
N σ/kHz	143 8.5	142 9.2	201 8.8	121 8.2	

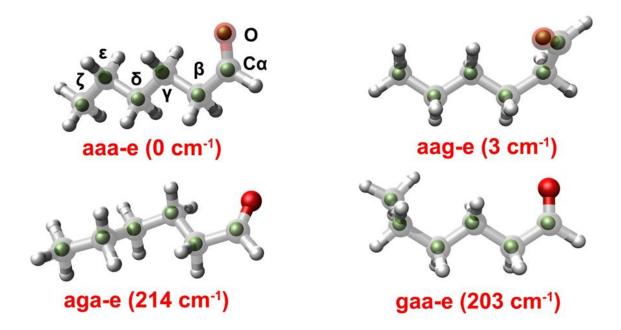
	9 (aaa-g)	10 (gga-e)	11 (ag+g+-g-)	12 (g <sup>-</sup> ag <sup>-</sup> -g <sup>+</sup> )	
A / MHz	12969.7951(84)	5461.5546(17)	3961.10569(51)	5439.6534(17)	
B / MHz	775.68441(42)	1149.07728(24)	1400.19125(31)	1098.76209(40)	
C / MHz	763.26201(48)	1102.89081(28)	1158.63001(28)	1050.33986(35)	
$\Delta_J/\mathrm{kHz}$	0.07102(46)	0.51927(90)	2.4272(17)	0.6234(11)	
$\Delta_{JK}$ /					
kHz	-7.2775(77)	-4.2407(87)	-10.948(11)	-8.88272(80)	
$\Delta K / kHz$	[0]	37.03(15)	25.561(22)	[0]	
$\delta_J$ / kHz	[0]	0.09360(39)	0.70955(80)	0.1283(11)	
$\delta_K / \mathrm{kHz}$	[0]	[0]	4.647(53)	[0]	
$P_{cc}$ / u Å <sup>2</sup>	14.1810(3)	38.29749(8)	26.16716	26.16716(7)	
$ \mu_a $ / D					
$ \mu_b $ / D					
$ \mu_c $ / D					
Ν	60	87	97	54	
$\sigma$ / kHz	9.6	9.7	7.9	9.0	

Table 1. Experimentally determined parameters for the 12 detected conformers of hexanal. For conformers 1-6, the dipole moments presented are experimentally determined via the Stark effect. For conformers 7-12, there was insufficient sensitivity for the determination of the dipole moments.

					l			I		
		$\Delta E$ (c	cm <sup>-1</sup> )		A (MHz)	В	С	$\mu_a$ (D)	$\mu_b$	$\mu_c$
Conformer	M06-2X	RI- MP2	D3	D3/aVTZ			D3/aVTZ			
1 (aaa-e)	86	203	0	0	9863.36	859.74	810.43	1.38	- 2.37	0.02
2 (aag-e)	0	0	11	3	5394.94	1133.00	1019.39	0.52	2.38	1.03
3 (aga-e)	192	203	199	214	9064.72	922.74	887.46	2.03	- 1.62	0.98
4 (gaa-e)	195	279	191	203	5998.64	1036.25	935.87	1.04	2.45	0.70
5 (g+age)	157	182	217	239	6145.47	1154.87	1047.52	-0.04	2.41	0.71
6 (g+ag+-e)	156	157	197	219	4692.13	1313.60	1154.07	0.31	- 2.59	0.04
7 (aag+-g-)	289	132	210	221	5503.54	1037.15	924.15	1.81	2.38	0.63
8 (ag+g+-e)	218	189	240	271	4675.12	1257.95	1168.09	0.55	1.82	1.81
	201	220	2 10	220	12000 42	771 40		2.07	-	1 40
9 (aaa-g)	386	228	240	229	12889.43	771.49	759.46	-2.87	0.11	1.40
10 (gga-e)	193	233	316	357	5431.14	1140.01	1093.33	-1.92	0.25	2.05
11 (ag+g+-g-)	289	141	336	369	3932.57	1388.51	1149.33	-0.74	2.87	0.04
12 (g-agg+)	415	279	401	434	5514.80	1076.17	1033.67	2.30	1.84	0.83

Table 2. Theoretical energies and rotational constants for the 12 assigned conformers of hexanal. Except for D3/aVTZ (i.e. B3LYP-

D3/aug-cc-pVTZ), the basis set used is 6-311++g(d,p).



**Figure 4.** B3LYP-D3/aVTZ structures of the four lowest energy conformers of 1hexanal, overlaying the experimentally-determined Kraitchman coordinates (green spheres). The labeling scheme used throughout the paper (O,  $C\alpha$ - $\zeta$ ) is shown on the *aaa-e* conformer.

	1 (aaa-e)	2 (aag-e)	3 (aga-e)	4 (gaa-e)	
Lengths / Å					
Ο - Cα	1.2078(51)	1.2105(55)			
Cα - Cβ	1.5112(24)	1.5032(46)	1.5256(25)	1.5146(80)	
Cβ - Cγ	1.596(12)	1.5575(65)	1.6213(55)	1.5507(78)	
Ογ - Οδ	1.420(14)	1.4782(68)	1.4847(82)	1.4731(81)	
Cδ - Cε	1.5327(42)	1.5436(34)	1.5309(48)	1.5401(42)	
Cε - Cζ	1.532(20)	1.5254(59)	1.5254(45)	1.5193(72)	
Angles / deg					
Ο - Cα - Cβ	124.58(29)	124.81(26)			
Cα - Cβ - Cγ	116.80(46)	115.02(22)	113.24(36)	115.51(66)	
Cβ - Cγ - Cδ	112.90(61)	112.30(30)	109.81(28)	112.53(88)	
Cγ - Cδ - Cε	111.26(58)	111.78(45)	115.87(26)	112.97(52)	
Cδ - Cε - Cζ	112.8(16)	113.33(40)	112.87(61)	114.06(24)	
Dihedrals /					
deg					
Ο - Cα - Cβ -					
Cγ	-3.8(76)	6.62(64)			
Cα - Cβ - Cγ - Cδ	179 2(17)	71 72(50)	178 60(21)	170 68(67)	
Cδ Cβ - Cγ - Cδ -	178.2(47)	71.72(50)	178.69(31)	179.68(67)	
Cρ - Cγ - Co - Cε	179.8(38)	-178.68(46)	66.25(39)	177.44(51)	
Cγ - Cδ - Cε -	172.0(30)	170.00(40)	00.25(57)	1,1,1,1,1)	
Cζ	-179.4(48)	-178.74(48)	172.08(69)	64.90(55)	

**Table 3** List of experimental Kraitchman ( $r_s$ ) internal parameters for the four conformerswith assigned isotopologues. Oxygen-containing parameters for conformers 3 and 4 areleft blank, as no  $^{18}$ O isotopologues were assigned.

Proper assignment of isotopologues by AUTOFIT can additionally be supported by the near identity of the planar moment  $P_{cc}$  for each of the assigned species of a given conformer. For a species like conformer 1 (aaa-e) where the heavy atom backbone is planar,  $P_{cc}$  provides additional assurance that the isotopologues found by AUTOFIT are correctly associated with the targeted parent species. From the values of  $P_{cc}$  tabulated in Table 4 for conformer 1, all isotopologues exhibit near-identical values; however, the <sup>18</sup>O isotopologue displays a slightly reduced value for  $P_{cc}$ , likely due to a small reduction in zero-point out-of-plane motion of the carbonyl moiety associated with the increased mass of the oxygen atom.

## 4) Discussion

#### a) 1-Hexanal Conformational Energies

Experimental detection of the mixture of 1-hexanal conformers in conjunction with the accurate relative intensity information provided by CP-FTMW spectroscopy enable determination of an experimental relative energy ordering. This poses a number of challenges, as deconvolution of the relative intensity information only provides experimental relative abundances. The abundance ordering must then be compared to relative energies provided by quantum chemical calculations. Therefore, a selection of four theoretical methodologies, presented in Table 2, were chosen to generate sets of relative energies for each of the 12 assigned conformers, and were directly compared to the experimental abundance ordering.

	A / MHz	<i>B</i> / MHz	C / MHz	$P_{cc}$ / u Å <sup>2</sup>
Parent	9769.63045(68)	868.845880(92)	818.518770(92)	7.98273(5)
<sup>13</sup> Ca	9750.86955(63)	858.080623(82)	808.829889(73)	7.98314(4)
<sup>13</sup> Cβ	9626.80436(61)	866.314556(67)	815.254427(70)	7.98014(3)
<sup>13</sup> Cγ	9763.83782(59)	868.825621(88)	818.465184(76)	7.98464(4)
<sup>13</sup> Cδ	9732.46478(63)	866.819546(74)	816.460549(80)	7.98312(5)
<sup>13</sup> Cε	9703.87446(66)	861.112525(98)	811.197212(84)	7.98352(4)
<sup>13</sup> Cζ	9769.34761(58)	849.117232(66)	800.985873(75)	7.98327(4)
<sup>18</sup> O	9511.5885(16)	843.55541(34)	794.27108(19)	7.9793(1)

Table 4 Rotational constants and out-of-plane planar moments  $(P_{cc})$  for conformer 1

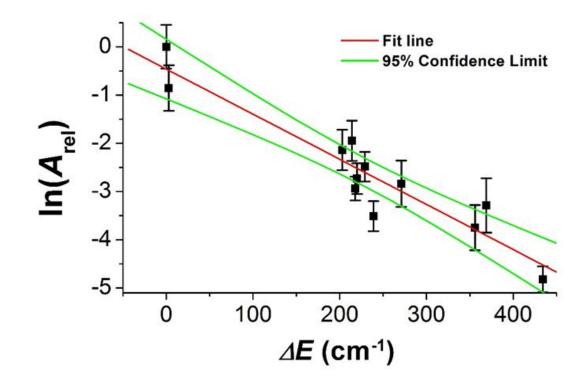
(aaa-e) and observed isotopologues.

Some discrepancies exist in the M06-2X or B3LYP relative energy rankings. Although the two most intense assigned conformers, 1 and 2, correlate to the two lowest energy conformers for both levels of theory, the rest of the ordering is inconsistent with experimental observations. For example, the undetected candidate conformers, ggg-e and ag<sup>-</sup>g<sup>+</sup>-e, have energies intermediate in the overall rankings. Additionally, conformers 3 (aga-e) and 4 (gaa-e) are significantly higher in energy in the initial DFT search than their relative intensities would suggest, falling in the middle of the table of twelve.

In order to further explore these discrepancies, higher level optimizations were performed on the 14 candidate conformers using the B3LYP-D3/aVTZ level of theory. Comparing these dispersion-corrected structures to the experimental results, a more consistent energetic profile is seen, as shown in Table 2. Conformers 1-4 correlate to the four lowest energy structures, and the two undetected candidates can be found with relative energies comparable to conformer 12, which has the highest relative energy of all the candidate structures. Although the energies of 3-9 are similar, the experimental intensity gap is not significant as many of these conformers' strongest transitions have intensities just below the <sup>13</sup>C detection limit.

The experimental relative abundance ordering can be determined by comparing the scale factors between the SPCAT-calculated and experimental intensities for a given conformer. Assuming thermodynamic equilibrium, a linear relationship should exist between the natural logarithm of the normalized scale factor and the relative energy of the conformer ( $\Delta E$ , cm<sup>-1</sup>). As shown in Figure 5, a linear relationship between the B3LYP-D3/aVTZ relative energies and experimental relative populations is observed for the conformers of 1hexanal. Calculated intensities use experimental dipole moments where available (conformers 1-6), and B3LYP-D3/aVTZ theoretical values otherwise. Comparison of the predicted and experimental dipoles for conformers 1-6 are in good agreement, so use of the DFT-derived dipole moments for the other conformers is appropriate. The rotational temperature for the predicted spectrum was set to 3.6 K, which minimizes the error in the predicted intensities. Additionally, since all but conformer 1 (aaa-e) have double-well minima (due to mirror symmetry), the scale factors for conformers 2-12 were divided by 2 to account for the degeneracy.

One caveat to this comparison is that thermal equilibrium is not necessarily a good approximation in a pulsed jet experiment, as conformational interconversion can potentially occur across barriers during the expansion process.<sup>47,48</sup> However, the linear relationship seen in Figure 5 suggests that non-equilibrium effects are not particularly significant for 1-hexanal. Use of B3LYP-D3 with the smaller 6-311++g(d,p) basis does not degrade the correlation ( $R^2 = 0.92$ ). However, when MP2 or M06-2X energies are used instead, much poorer correlations are seen, with  $R^2$  values of 0.81 and 0.53 respectively.



**Figure 5**. A plot of the normalized natural logarithm of the relative abundance ( $A_{rel}$ ) versus the B3LYP-D3/aVTZ relative energies for the 12 observed 1-hexanal conformers. The data fits to a line of the following form:  $\ln(A_{rel}) = -0.36(24) - 0.0106(10) \Delta E$  (cm<sup>-1</sup>), with an  $R^2 = 0.91$ .

#### b) Decreasing AUTOFIT Run Times

The most effective approach to reducing run time is to improve the accuracy of the predicted rotational constants by consulting higher levels of theory. However, it is important to consider the inherent tradeoff between computational accuracy and time. For instance, using methods such as CCSD(T) with a large basis set would certainly improve the accuracy of predictions <sup>49</sup>, but the time it takes to optimize a structure using such high levels of theory competes with the time it would take to complete an AUTOFIT search with predictions made with a less accurate method.

A recent benchmark study by Grimme and Steinmetz discusses directly the viability of multiple theoretical methods, both density functional and *ab initio*, in determining accurate rotational constant predictions. <sup>50</sup> They show that dispersion-corrected density functional theory such as B3LYP-D3 can compete directly with MP2 in predicting rotational constants of molecular species with dispersive interactions. Since B3LYP is a naturally less computationally intensive methodology than MP2, this result is highly attractive for improving AUTOFIT run times and structural accuracy. However, Grimme and Steinmetz use a large, quadruple- $\zeta$  basis set (def2-QZVP) for their optimizations and also calculate anharmonic corrections in order to directly compare the ground-state experimental rotational constants to the calculated ground-state, vibrationally averaged structures. In order to gauge the validity of these results in the context of using AUTOFIT on the desktop, similar methodologies have been applied with smaller basis sets (e.g. 6-311++g(d,p)) and no anharmonic corrections. As the benchmark set of molecules, the 12 experimentally assigned conformers of 1-hexanal are used.

The approximate triples count for a standard AUTOFIT run was estimated using a simple search window model. A tabulated summary of results can be found in Table 5. Three types of standard AUTOFIT searches were considered. First is a pure  $\mu_a$ -type spectrum, where the frequencies of the transitions are primarily dependent on just the values of *B* and *C*. Second, a pure  $\mu_b$ -type spectrum where the frequencies are dependent on all three rotational constants, and third, a  $\mu_a/\mu_b$  hybrid spectrum.  $\mu_c$ -type spectra were not considered, since  $\mu_c$ -type spectra depend primarily on just *A* and *B* and are thusly similar to  $\mu_a$ -type spectra in terms of modeling their AUTOFIT search characteristics. The uncertainties in the theoretical rotational constants are set to 3 times the standard deviation of the *ab initio* determinations (the "SD" column in Table 5). Assuming the uncertainties are Gaussian distributed in nature, this implies a 99.7% certainty on the search window. A full description of this model can be found in the Appendix at the end of this chapter.

In general, the 1-hexanal results support the conclusions made by Grimme and Steinmetz. Application of a dispersion corrected functional significantly improves the rotational constant prediction, with nearly a factor of two improvement in signed RMS error using B3LYP-D3 compared to B3LYP, MP2 and M06-2X. These results can be seen in Table 5, where AUTOFIT search speeds are reduced by nearly a factor of 20 for B3LYP-D3/6-311++g(d,p) compared to standard B3LYP. Increasing the basis set size further to aVTZ improves the B3LYP-D3 prediction RMS error by another 25%, but the additional up-front computational time might not be worthwhile when compared to the additional search time associated with a smaller basis set.

	% signed error in rotational constant predictions		Estimated AUTOFIT run time (min) [rel speedup wrt. B3LYP/6- 311++g(d,p)]			
Method	Mean	SD	RMS	μ <sub>a</sub> -type fit	$\mu_b$ -type fit	μ <sub>a</sub> /μ <sub>b</sub> hybrid fit
B3LYP/6-311++g(d,p)	1.23	2.47	2.75	250	3390	915
M06-2X/6-311++g(d,p)	-1.14	2.37	2.63	225 [11%]	3420 [-1%]	825 [11%]
MP2/6-311++g(d,p) B3LYP-D3/6-	-0.46	2.40	2.41	105 [138%]	1100 [205%]	490 [86%]
311++g(d,p)	0.66	1.05	1.24	14 [1708%]	179 [1792%]	132 [592%]
B3LYP-D3/aVTZ	0.30	0.90	0.95	7 [3596%]	52 [6464%]	85 [989%]

**Table 5** Model AUTOFIT run times and rotational constant prediction errors for the 12assigned hexanal conformers. Assumed AUTOFIT speed is 50 triples/s.

## 5) Conclusion

CP-FTMW rotational spectra of 12 conformers of 1-hexanal, all <sup>13</sup>C isotopologues of conformers 1-4, and the <sup>18</sup>O isotopologues of conformers 1-2 have been assigned using automated fitting techniques. Dipole moments have additionally been determined for conformers 1-6. The fitted results are in good agreement with theoretical predictions, with absolute RMS rotational constant errors as low as 0.88% for the B3LYP-D3/aug-cc-pVTZ level of theory. Based on comparisons between theoretical energies and experimental relative intensities, the full set of potentially detectable 1-hexanal conformers have been assigned at the sensitivity limit for the experiment. With the substantial recent improvements to the efficiency and accuracy of density functional theory methodologies, the feedback loop between experiment and theory is becoming more efficient. AUTOFIT exploits these improvements directly, and can consequently be a helpful tool for rotational spectroscopy as a routine chemical technique for structure determination and mixture analytics.

# References

(1) Pérez, C.; Lobsiger, S.; Seifert, N. A.; Zaleski, D. P.; Temelso, B.; Shields, G. C.; Kisiel, Z.; Pate, B. H. *Chem. Phys. Lett.* **2013**, *571*, 1.

Seifert, N. A.; Zaleski, D. P.; Pérez, C.; Neill, J. L.; Pate, B. H.; Vallejo-López,
M.; Lesarri, A.; Cocinero, E. J.; Castaño, F.; Kleiner, I. *Angew. Chem.* 2014, *126*, 3274.

(3) Seifert, N. A.; Steber, A. L.; Neill, J. L.; Pérez, C.; Zaleski, D. P.; Pate, B. H.; Lesarri, A. *Phys. Chem. Chem. Phys.* **2013**, *15*, 11468.

(4) Fortman, S. M.; McMillan, J. P.; Neese, C. F.; Randall, S. K.; Remijan, A. J.; Wilson, T. L.; De Lucia, F. C. *J. Mol. Spectrosc.* **2012**, *280*, 11.

(5) Stephens, S. L.; Walker, N. R.; Legon, A. C. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20736.

(6) Stephens, S. L.; Walker, N. R.; Legon, A. C. *Phys. Chem. Chem. Phys.* **2011**, *13*, 21093.

(7) Cooke, S. A.; Ohring, P. J. Spectrosc. 2012, 2013.

(8) Finneran, I. A.; Shipman, S. T.; Widicus Weaver, S. L. *J. Mol. Spectrosc.* **2012**, 280, 27.

(9) Kroll, J. A.; Shipman, S. T.; Widicus Weaver, S. L. J. Mol. Spectrosc. 2014, 295,
52.

(10) Reinhold, B.; Finneran, I. A.; Shipman, S. T. J. Mol. Spectrosc. 2011, 270, 89.

(11) Watson, J. K. G. Mol. Phys. 1968, 15, 479.

(12) Hageman, J. A.; Wehrens, R.; Gelder, R. de; Meerts, W. L.; Buydens, L. M. C. J. *Chem. Phys.* **2000**, *113*, 7955.

(13) Leo Meerts, W.; Schmitt, M. Int. Rev. Phys. Chem. 2006, 25, 353.

(14) Meerts, W. L.; Schmitt, M.; Groenenboom, G. C. Can. J. Chem. 2004, 82, 804.

(15) Helm, R. M.; Vogel, H.-P.; Neusser, H. J. Chem. Phys. Lett. 1997, 270, 285.

(16) Fisher, J. M.; Xu, L.-H.; Suenram, R. .; Pate, B.; Douglass, K. J. Mol. Struct. **2006**, *795*, 143.

(17) Tubergen, M. J.; Conrad, A. R.; Chavez III, R. E.; Hwang, I.; Suenram, R. D.; Pajski, J. J.; Pate, B. H. *J. Mol. Spectrosc.* **2008**, *251*, 330.

(18) Fraser, G. T.; Xu, L.-H.; Suenram, R. D.; Lugez, C. L. J. Chem. Phys. **2000**, *112*, 6209.

(19) Fraser, G. T.; Suenram, R. D.; Lugez, C. L. J. Phys. Chem. A 2000, 104, 1141.

(20) Fraser, G. T.; Suenram, R. D.; Lugez, C. L. J. Phys. Chem. A 2001, 105, 9859.

(21) Gámez, F.; Martínez-Haya, B.; Blanco, S.; López, J. C.; Alonso, J. L. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12912.

(22) Laurie, V. W. Acc. Chem. Res. 1970, 3, 331.

(23) Alonso, J. L.; Pérez, C.; Sanz, M. E.; López, J. C.; Blanco, S. *Phys. Chem. Chem. Phys.* **2009**, *11*, 617.

(24) Blanco, S.; López, J. C.; Mata, S.; Alonso, J. L. *Angew. Chem. Int. Ed.* **2010**, *49*, 9187.

(25) Peña, I.; Cocinero, E. J.; Cabezas, C.; Lesarri, A.; Mata, S.; Écija, P.; Daly, A. M.; Cimas, Á.; Bermúdez, C.; Basterretxea, F. J.; Blanco, S.; Fernández, J. A.; López, J. C.; Castaño, F.; Alonso, J. L. *Angew. Chem.* **2013**, *125*, 12056.

(26) Shipman, S. T.; Neill, J. L.; Suenram, R. D.; Muckle, M. T.; Pate, B. H. J. Phys. Chem. Lett. **2011**, *2*, 443.

(27) Shipman, S. T.; Pate, B. H. In *Handbook of High-resolution Spectroscopy*; John Wiley & Sons, Ltd, 2011.

(28) Zaleski, D. P.; Neill, J. L.; Muckle, M. T.; Seifert, N. A.; Brandon Carroll, P.; Widicus Weaver, S. L.; Pate, B. H. *J. Mol. Spectrosc.* **2012**, *280*, 68.

(29) Brown, G. G.; Dian, B. C.; Douglass, K. O.; Geyer, S. M.; Shipman, S. T.; Pate, B. H. *Rev. Sci. Instrum.* **2008**, *79*, 053103.

(30) Emilsson, T.; Gutowsky, H. S.; Oliveira, G. de; Dykstra, C. E. J. Chem. Phys. **2000**, *112*, 1287.

(31) Kasten, W.; Dreizler, H. Z Naturforsch 1984, 39, 1003.

- (32) Becke, A. D. J. Chem. Phys. 1993, 98, 5648.
- (33) Frisch, M. J.; Pople, J. A.; Binkley, J. S. J. Chem. Phys. 1984, 80, 3265.
- (34) Zhao, Y.; Truhlar, D. G. Theor. Chem. Acc. 2008, 120, 215.
- (35) Werner, H.-J.; Manby, F. R.; Knowles, P. J. J. Chem. Phys. 2003, 118, 8149.
- (36) Grimme, S. **2011**, *1*, 221.

(37) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. J. Chem. Phys. 2010, 132, 154104.

(38) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.;
(Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji,
H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.;
Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida,
M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.;

Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.;
Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision A.1*; Gaussian, Inc.: Wallingford, CT, 2009.

(39) Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. A.; Fermann, J. T.; Mintz, B. J.; Burns, L. A.; Wilke, J. J.; Abrams, M. L.; Russ, N. J.; Leininger, M. L.; Janssen, C. L.; Seidl, E. T.; Allen, W. D.; Schaefer, H. F.; King, R. A.; Valeev, E. F.; Sherrill, C. D.; Crawford, T. D. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 556.

- (40) Weigend, F.; Köhn, A.; Hättig, C. J. Chem. Phys. 2002, 116, 3175.
- (41) Dunning, T. H. J. Chem. Phys. 1989, 90, 1007.
- (42) Pickett, H. M. J. Mol. Spectrosc. 1991, 148, 371.
- (43) Shipman, S. T.; Finneran, I. A.; Seifert, N. A. Autofit.

(44) Carey, F. A.; Sundberg, R. J. Advanced Organic Chemistry, Part A: Structure and *Mechanisms*, 5th edition.; Springer, 2008.

- (45) Costain, C. C. J. Chem. Phys. 2004, 29, 864.
- (46) Kraitchman, J. Am. J. Phys. 1953, 21, 17.

(47) Ruoff, R. S.; Klots, T. D.; Emilsson, T.; Gutowsky, H. S. J. Chem. Phys. **1990**, 93, 3142.

- (48) Campbell, E. J.; Buxton, L. W.; Balle, T. J.; Keenan, M. R.; Flygare, W. H. J. *Chem. Phys.* **1981**, *74*, 829.
- (49) Puzzarini, C. Phys. Chem. Chem. Phys. 2013, 15, 6595.
- (50) Grimme, S.; Steinmetz, M. Phys. Chem. Chem. Phys. 2013, 15, 16031.

# **Chapter 3: Phenol Dimer and Trimer – Structures from Broadband Rotational Spectroscopy**

(This chapter is adapted from a published article that has been peer-reviewed and published in Physical Chemistry Chemical Physics, with the following citation: N. A. Seifert, A. L. Steber, J. L. Neill, C. Perez, D. P. Zaleski, B. H. Pate, A. Lesarri, *Phys. Chem. Chem. Phys.* **2013**, *15*, 11468.)

# 1) Introduction

The structure of the phenol dimer presents a challenge to computational chemistry because the hydrogen bond interaction of the hydroxyl group and the dispersion interaction of the benzene ring prefer different geometries.<sup>1–3</sup> Since dispersion is ignored in the Hartree-Fock approximation<sup>4</sup> and traditional DFT functionals treat it incompletely, accurate predictions are immediately limited to post-Hartree-Fock methods such as MP2 or coupled cluster, or dispersion-modified density functional theory.<sup>5–9</sup> However, results for phenol dimer using these advanced levels of theory give inconsistent results in the treatment of dispersion.<sup>3</sup> It is therefore important for an accurate experimental structure to be determined in order to not only accurately gauge the efficacy of these methods<sup>10,11</sup>, but also to provide an experimental perspective on an interplay of competing non-covalent interactions that play an essential role in everything from biochemical processes<sup>12,13</sup> to catalysis<sup>14,15</sup> and supramolecular chemistry.<sup>16</sup>

This intriguing interplay of non-covalent interactions in phenol dimer has not gone unnoticed, as a number of theoretical and spectroscopic studies have been published over the past few decades, beginning with the multiphoton ionization studies of Fuke and Kaya.<sup>17,18</sup> Phenol dimer is a member of the JSCH-2005 benchmark, which was created for the purpose of benchmarking the efficacy of new and existing theoretical methods in

treating noncovalent interactions in biologically-relevant systems.<sup>19</sup> Experimental structural studies have also been previously published and performed with a variety of spectroscopic methods including rotational coherence spectroscopy (RCS)<sup>20,21</sup>, IR-UV double resonance spectroscopy<sup>22,23</sup>, and high-resolution UV absorption spectroscopy.<sup>24</sup> The RCS measurements have provided rotational constants for the phenol dimer and important features of the molecular structure with comparison to *ab initio* predictions have also been discussed.<sup>1</sup> Subsequent high-resolution UV spectroscopy studies determined the molecular rotational constants to higher precision.<sup>25</sup> Also, the spectra of several, commercially available isotopic species were analyzed to provide more detail about the structure of phenol dimer.<sup>24</sup>

Although these methods have shed considerable light on the structure of phenol dimer, they lack the high resolution and structural precision afforded by high resolution microwave spectroscopy. In this study, the use of chirped-pulse Fourier transform microwave (CP-FTMW) spectroscopy<sup>26</sup> in the 2-8 GHz band<sup>27,28</sup> enables the determination of a full heavy atom (e.g. carbons and oxygens) experimental structure of phenol dimer, which is in excellent agreement with *ab initio* results. This structure, as well as the least-squares fitted, ground state r<sub>0</sub> structure determined in this study, resolves a number of the inconsistencies between previous experimental structures<sup>21,24</sup> of phenol dimer and the *ab initio* results, including the hydrogen bonding distance and the relative orientations of the two benzene rings due to dispersive interactions. However, that study was limited in scope compared to the results presented here, and the data collected and analyzed for this study is completely independent of the data used in the previous microwave study.

In addition, the first direct structural determination of phenol trimer is reported, which has a symmetric, barrel-like structure. Although the trimer was previously detected and correctly assigned to a barrel structure in the aforementioned IR-UV double resonance study<sup>22</sup>, no direct structural information could be obtained from the O-H stretch region vibrational band data. Like phenol dimer, the trimer features an interplay between dispersion and hydrogen bonding. The trimer suffers some of the inconsistencies between varying levels of theory like the dimer, though the constraint of  $C_3$  symmetry reduces the number of independent degrees of freedom, so potential differences between the observed and calculated structure that can reduce the quality of the spectroscopic predictions are less prevalent than in the dimer. However, the correct choice of theory produces a good trimer prediction, and the theoretical results reported here are in good agreement with experimental results.

## 2) Experimental

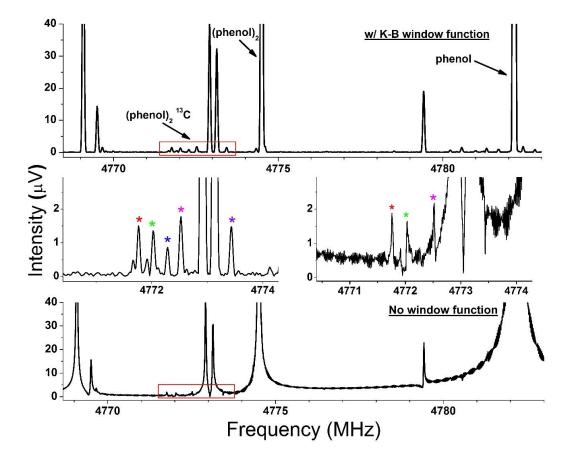
The broadband rotational spectrum of the phenol clusters was measured in the 2-8 GHz frequency range using the CP-FTMW spectrometer at UVa. As the description of this instrument can be found in Chapter 1, Section IV, only experimental information specific to study will be described.

Phenol is loaded into the reservoir of the pulsed nozzles and heated to  $85^{\circ}$ C. Neon carrier gas at a pressure of 1.67 atm is used for the pulsed jet expansion with a pulse duration of 700 µs. Each nozzle pulse consumes an estimated 10 nmol of phenol. Following gas sample injection into the spectrometer vacuum chamber, a series of 8 back-to-back measurements of the broadband rotational spectrum are performed. These

measurements use a 4 µs duration chirped-pulse with a linear frequency sweep from 2-8 GHz, with a digitizer record length of 40 µs at a sampling rate of 25 Gs/s. Gas sample injection repetition rates of 8 Hz (64 Hz FID signal averaging) are achieved. The 2-8 GHz spectrum of phenol used in this analysis contained a total of 9.2 M FID averages. The frequency domain spectrum is obtained from this averaged FID by a discrete fast Fourier transform following the application of a Kaiser-Bessel window function to give improved baseline resolution in the spectrum.

### 3) **Results**

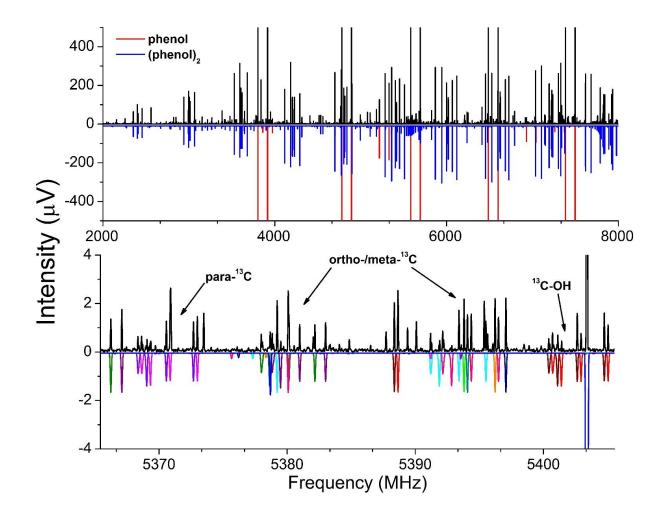
A section of the 2-8 GHz broadband rotational spectrum of the pulsed-jet phenol sample is shown in Fig. 1. This figure also shows the importance in applying the Kaiser-Bessel window function to improve the baseline resolution so that weak transitions, from natural abundance isotopic species for example, can be detected in the presence of the strong phenol and phenol dimer rotational transitions. The ability to shape the rotational spectrum line shape is a special feature of broadband rotational spectroscopy. In comparison, the narrow bandwidth of a cavity-FTMW spectrometer results in the polarization of molecules "in the wings" of the inhomogeneous residual Doppler distribution making it impossible to minimize the baseline resolution of very strong transitions.<sup>30</sup> Although CP-FTMW spectrometers have lower resolution when defined as the full-width at half-maximum, their superior baseline resolution yields a higher useable dynamic range in the measurement.



**Figure 1**. A 20 MHz section of the 2-8 GHz CP-FTMW spectrum of phenol. Use of the Kaiser-Bessel window function (top spectrum) greatly improves the baseline resolution, enabling observation of multiple transitions of the <sup>13</sup>C isotopologues of phenol dimer (starred, center) that are poorly resolved or not resolved at all in the raw spectrum (bottom).

The strongest transitions in the spectrum come from the phenol monomer and have about 30000:1 signal-to-noise ratio. The spectrum for the phenol dimer is a factor of 10 weaker. The measurement sensitivity, therefore, is sufficient to observe the singly <sup>13</sup>C and <sup>18</sup>O isotopologues of the phenol dimer in natural abundance. The full 2-8 GHz spectrum, as well as an expanded region showing the <sup>13</sup>C isotopic spectra in natural abundance, is shown in Fig. 2. A summary of the rotational spectrum analysis of the phenol dimer and the 14 heavy-atom isotopologues using the A-reduction of the Watson Hamiltonian<sup>31</sup> is presented in Table 1. The centrifugal distortion constants for the isotopologues were fixed to the normal isotopic species fit values (fitting these constants for each isotopologues produced nearly the same values within the confidence interval of the parameter estimation).

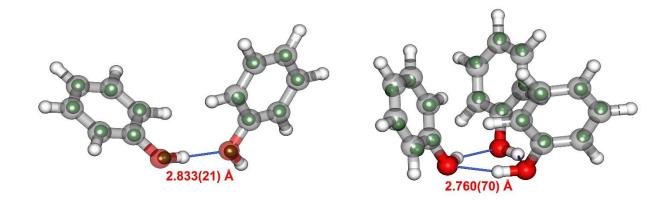
The phenol trimer is observed in the spectrum as an oblate symmetric top species, as seen in Fig. 3. With the improved spectrometer performance, the <sup>13</sup>C isotopologues of phenol trimer can be observed in natural abundance (~1% abundance). Since the phenol trimer is C<sub>3</sub> symmetric, the intensities of the isotopologues were tripled due to symmetry. For ease of analysis, the prolate I<sup>r</sup> representation was chosen for assignment of the isotopologues. Since the isotopic substitutions are off the C<sub>3</sub> symmetry axis, isotopologues of phenol trimer in the prolate basis show pure  $\mu_c$ -type asymmetric top spectra.



**Figure 2.** The 2-8 GHz spectrum of phenol, with 2 K predictions of phenol monomer (top, red) and phenol dimer (top, blue). The full dynamic range of the phenol monomer transitions has been truncated in order to emphasize the phenol dimer spectrum. The <sup>13</sup>C isotopologues of phenol dimer (bottom) are also compared to predictions (bottom, multiple colors). As shown, the isotopologues can be seen as spatially resolved in frequency space, where the substitutions farthest from the center of mass are the most redshifted from the parent transition (bottom, strong blue transition at ca. 5403 MHz). <sup>18</sup>O isotopologues are not shown.

	Parent	Donor	A (MHz)	В	С	Ν	σ (kHz)
A (MHz)	1415.32747(14)	1- <sup>13</sup> C	1413.0983(14)	312.52409(32)	287.16539(32)	120	8.08
В	313.368020(41)	2- <sup>13</sup> C	1412.74240(48)	312.75594(19)	287.52814(20)	136	5.77
С	287.961282(38)	3- <sup>13</sup> C	1403.8469(63)	312.00616(35)	286.73261(34)	111	8.25
		4- <sup>13</sup> C	1406.0358(10)	310.68675(28)	285.40748(28)	129	7.98
Δ」 (kHz)	0.372930(76)	5- <sup>13</sup> C	1413.37840(97)	310.05088(26)	285.13627(28)	133	8.75
$\Delta_{JK}$	-3.92349(48)	6- <sup>13</sup> C	1411.85779(88)	311.08467(29)	286.08891(29)	113	6.98
$\Delta_{K}$	13.0664(27)	1- <sup>18</sup> 0	1387.525(25)	312.54360(80)	286.32269(76)	60	10.4
δJ	0.057377(17)	Acceptor					
δκ	0.6928(19)	1- <sup>13</sup> C	1413.23440(69)	312.50304(16)	287.15363(16)	123	5.53
		2- <sup>13</sup> C	1412.24751(35)	312.72811(22)	287.52476(22)	133	5.38
Ν	481	3- <sup>13</sup> C	1403.92180(82)	311.90897(22)	286.72685(23)	129	7.11
σ (kHz)	6.39	4- <sup>13</sup> C	1405.66540(96)	310.71737(19)	285.40780(19)	123	6.79
		5- <sup>13</sup> C	1413.02009(74)	310.10090(26)	285.17718(26)	121	6.41
		6- <sup>13</sup> C	1411.56899(64)	311.12121(17)	286.14677(18)	136	6.84
		2- <sup>18</sup> O	1388.945(42)	312.70639(67)	286.47370(65)	69	7.02

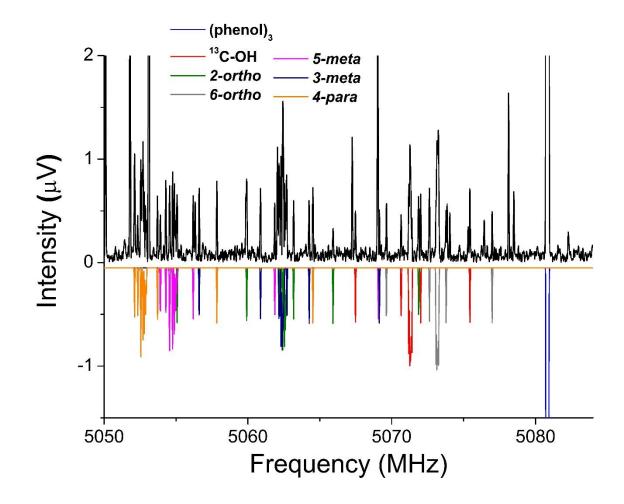
**Table 1.** Spectroscopic parameters for phenol dimer and its 14 heavy atom isotopologues. N is the total number of transitions contained in the fit, and  $\sigma$  is the RMS average of each fit. Distortion for all isotopologues is held fixed to the reported parent species values.



**Figure 3.** Structures of the phenol dimer and trimer. In both cases, the shown frameworks are the  $r_0$  and  $r_m^{(1)}$  structures for phenol dimer and trimer respectively, derived from M06-2X/6-311++g(d,p) geometries and experimental data. The green inner spheres represent the Kraitchman coordinates derived directly from the experimental data. The labeled O···O lengths are taken from the  $r_0$  and  $r_m^{(1)}$  geometries.

However, since the asymmetry generated by a single <sup>13</sup>C substitution is slight, isotopologues with substitutions near the symmetry axis generate small asymmetry splittings. Since assignments for these isotopologues include transitions with a relatively high J quantum number ( $J \le 14$ ), the region where transitions with a high K<sub>c</sub> appear for a given J tend to be quite dense. Since these high K<sub>c</sub> transitions are essential in determining the distortion for a given species, floating distortion was only possible for the parasubstituted and a meta-substituted isotopologue, where the asymmetry is relatively high. A visual summary of the phenol trimer spectrum can be seen in Fig. 4.

Two other symmetric top species were identified in the spectrum with rotational constants similar to the normal species of phenol trimer. These spectra are about 30 times weaker than the normal species phenol trimer. Because a plausible phenol trimer isomer that is a symmetric top with nearly the same rotational constant of the dominant species cannot be proposed, it is believed that these spectra come from vibrationally excited states of the assigned trimer species that are populated in the pulsed molecular beam. The harmonic vibrational calculation at the M06-2X/6-311++g(d,p) level of theory<sup>32,33</sup>, gives 6 vibrational normal modes with frequencies below 100 cm<sup>-1</sup>. A summary of the spectral analysis of all phenol trimer species in provided in Table 2. All calculations done for this study were performed exclusively with the GAUSSIAN 09, Rev A.02 suite of programs.<sup>34</sup>



**Figure 4.** A section of the 2-8 GHz spectrum representing the J = 9-8 transition of phenol trimer (top), with 2 K predictions of the parent species and the 6 <sup>13</sup>C isotopologues (bottom). As seen, the asymmetry generated by the <sup>13</sup>C substitution is slight, and leads to dense pile-ups of high K<sub>c</sub>,  $\mu_c$ -type transitions.

	Parent	1- <sup>13</sup> C	6- <sup>13</sup> C	2- <sup>13</sup> C	5- <sup>13</sup> C	3- <sup>13</sup> C	4- <sup>13</sup> C
A(MHz)		282.19749(14)	282.21553(17)	282.26789(23)	282.23405(52)	281.96181(31)	281.89683(66)
B (MHz)	282.280790(56)	281.30995(14)	280.33503(17)	281.44815(23)	279.47608(53)	280.56222(30)	279.57290(65)
C(MHz)		187.33(14)	187.069(50)	187.36(21)	186.897(44)	187.34(14)	186.824(66)
$\Delta_{J}$ (kHz)	0.10300(19)	[0.0315]	[0.0315]	[0.0315]	0.0297(10)	[0.0315]	0.0334(12)
$\Delta_{JK}$	-0.13600(28)	[0.0752]	[0.0752]	[0.0752]	0.0779(26)	[0.0752]	0.725(34)
$\Delta_{K}$		[0]	[0]	[0]	[0]	[0]	[0]
δJ		[0.03249]	[0.03249]	[0.03249]	0.03227(56)	[0.03249]	0.03270(65)
δκ		[0]	[0]	[0]	[0]	[0]	[0]
Ν	20	63	87	53	108	53	100
σ (kHz)	1.37	5.96	7.11	9.10	9.33	9.49	8.62

**Table 2.** Spectroscopic parameters for phenol trimer and its 6  ${}^{13}$ C isotopologues. Distortion was fit for the most abundant species, aswell as for  ${}^{13}$ C isotopologues 4 and 6. Distortion parameters for the other  ${}^{13}$ C isotopologues were fixed to the average of isotopologues4 and 6.

## 4) Discussion

The analysis of the rotational spectra of the isotopologues can be used to generate experimental structures of the phenol dimer and trimer. The rotational constants for each isotopologue can be converted into the position of the substituted atom in the principal axis system using the method first described by Kraitchman<sup>35</sup> and analyzed by Costain.<sup>36</sup> This method builds the experimental structure atom-by-atom and requires no direct input from theoretical calculations. However, the analysis provides only the magnitudes of the principal axis coordinates. In this work, there are theoretical structures available that have calculated rotational constants that are in good agreement with experiment and, therefore, are likely to be good approximations of the experimental structures. These theoretical structures have been used to guide the selection of the signs of the coordinates for each atom in the "substitution structure." Once a close match between the experimental and theoretical heavy-atom structure is obtained, it is possible to produce a "refined" structure that fits structural parameters directly to the observed isotopic moments-of-inertia using a limited set of constraints from the theoretical model.

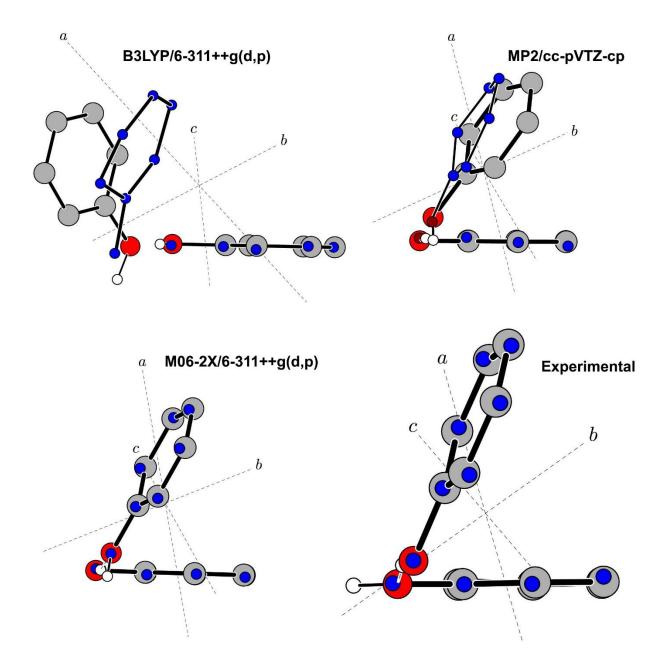
## a) Phenol Dimer

The experimental heavy atom substitution structure of the phenol dimer, showing the positions of the twelve distinct carbon atoms and the two oxygen atoms, is shown in Fig. 3. The phenol dimer structure is nearly symmetric (only the free O-H reduces the symmetry) and the carbon atoms appear in pairs with nearly identical atom magnitudes of the coordinate positions. In this case, there is the possibility that atom pairs have not been properly assigned. However, the effect of changing the positions has negligible effect on the experimental substitution structure. The phenyl ring structure from this analysis shows small deviations from planarity. In a few cases of structure determination of rigid, aromatic molecules by rotational spectroscopy, more planar structures have been observed. It is likely that the effects of large amplitude vibrational motions are responsible for the deviations from planarity observed for phenol dimer.

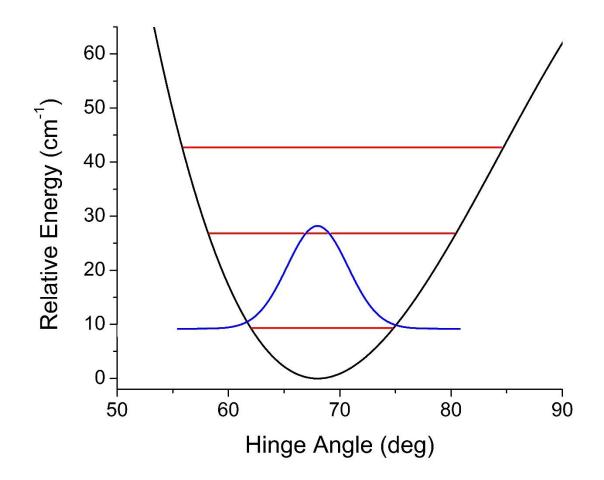
A key structural parameter for the phenol dimer is the "hinge" angle between the two phenol rings. This aspect of the geometry reflects the balance between the intermolecular hydrogen bond and the dispersion interaction of the aromatic rings. The experimental values for the hinge angle form the current work, a previous rotationallyresolved UV spectroscopic study<sup>24</sup>, and several computational chemistry models are given in Table 3. As discussed before, different computational models predict significantly different equilibrium geometries for the phenol dimer.<sup>3</sup> In particular, the incomplete treatment of dispersion in density functional theory leads to a structure with a large hinge angle that is dominated by the geometry preference of the hydrogen bond. Functionals with exchange-correlation energy approximations that have improved treatment of dispersion interactions, like M06-2X, give equilibrium structures that better match the experimental structure. The equilibrium structures from MP2/6-311++g(d,p)also give good approximations to the geometry that was observed in the experiment. However, as the basis set size increases, the MP2 theoretical geometries give angles that are smaller than experiment suggesting that the stabilization energy due to dispersion is overestimated.<sup>37</sup> Similar problems are seen in the CCSD calculations with Pople basis sets, though in this case CCSD actually underestimates the dispersion, with larger hinge angles than experiment. Hybrid functionals like M06-2X give results that appear to

stabilize as one increases basis set size. This is consistent with previous theoretical results, where high level dispersion modified DFT calculations, such as the RI-DFT-D/aQZ' results by Kolář and Hobza, give excellent results compared to experiment.<sup>3</sup> A visual comparison of three different levels of theory with experiment can be found in Fig. 5.

One important assumption behind the comparisons between the substitution structure and theoretical equilibrium geometries is that the hinge angle is well-defined in the dimer. The "hinge motion" of the dimer correlates to the lowest frequency normalmode vibration with a harmonic frequency of 28 cm<sup>-1</sup> calculated at the M06-2X/6-311++g(d,p) level. This is high compared to the measured frequency<sup>38</sup> of 9 cm<sup>-1</sup> but other calculations obtain similar results, though it appears the application of counterpoise correction enables a more accurate estimation of this mode's vibrational energy.<sup>25</sup> The one-dimensional potential energy of this motion was estimated by performing energy minimization at a series of fixed hinge angles allowing geometry optimization of all other structure parameters. This potential is shown in Fig. 6 and is reasonably harmonic. The approximate probability distribution of the hinge angle in the zero-point energy level, calculated in the harmonic limit, is shown in the figure. The calculated quantum mechanical fluctuation in the harmonic limit is about  $\pm 3.9^{\circ}$  (1 $\sigma$ ), which is smaller than the variation of hinge angles seen in the calculations tabulated in Table 3.



**Figure 5**. Comparison of theoretical and an experimental structure of phenol dimer with the  $r_s$  (Kraitchman) results. The small blue circles represent the  $r_s$  coordinates, and the larger connected frameworks represent the labeled *ab initio* method used, except for the Experimental entry (which uses the  $r_0$  structure as the framework).



**Figure 6.** Phenol dimer "hinge" potential energy surface, calculated at the M06-2X/6-311g(d,p) level of theory. The total *harmonic* probability distribution for the ground state (blue) suggests a hinge angle of  $68.0(39)^\circ$ . The ground state and the first two excited vibrational states (red) are shown at their calculated anharmonic frequencies.

	A (MHz)	B (MHz)	C (MHz)	Hinge Angle
MP2/6-311g(d,p)	1165.2	410.3	363.3	49.0
MP2/6-311++g(d,p)	1365.2	325.8	305.7	69.5
MP2/cc-PVTZ-cp <sup>19</sup>	1286.0	356.1	317.6	64.4
CCSD/6-311g(d,p)	1545.2	282.4	275.6	88.7
CCSD/6-311++g(d,p)	1459.6	308.9	286.8	79.5
M06-2X/6-311g(d,p)	1352.7	342.1	318.3	68.2
M06-2X/6-311++g(d,p)	1382.5	336.6	306.4	59.6
M06-2X/6- 311++g(df,pd)	1382.9	338.4	307.6	59.6
B3LYP/6-31G(d,p)	1800.0	250.8	243.7	111.9
B3LYP/6-311++g(d,p)	1946.3	231.8	229.6	110.5
RI-DFT-D/aQz' <sup>3</sup>	1399.7	318.5	292.5	61
CP-FTMW (r <sub>0</sub> , this study)	1415.3275(11)	313.36802(32)	287.96128(30)	62.3(14)
UV-VIS (r <sub>0</sub> , Schmitt et al.)	1416.99(39)	313.51(1)	288.11(1)	63.0

**Table 3**. Phenol dimer rotational constants and hinge angles from various levels of theory.

The *ab initio* potential energy surface data was fit to a cubic function of the form  $a\theta^3 + b\theta^2$  (a = -0.004472 cm<sup>-1</sup>, b = 0.232 cm<sup>-1</sup>, R<sup>2</sup> = 0.99). The cubic corrections were made using standard second order perturbation theory. The reduced mass for the normal mode associated with this hinge angle is 5.5868 amu, which leads to a harmonic force constant of 0.255 N·m<sup>-1</sup>.

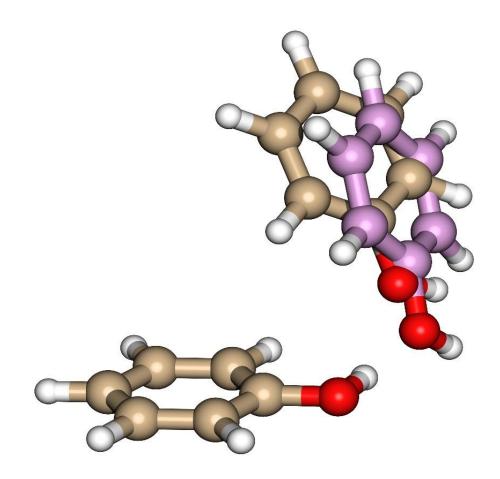
Finally, a refined  $r_0$  structure has been calculated using STRFIT<sup>39</sup> that combines the experimental isotopic data with some constraints from the M06-2X/6-311++g(d,p) theoretical structure. The bond lengths and angles within each phenol unit are fixed to their theoretical values (and these are essentially identical to the geometry of the phenol monomer). This constrains the heavy atoms in each phenol unit to be planar. The torsional angle of the –OH group for each phenol molecule is also constrained to the theoretical value. The hydroxyl hydrogen atom in each phenol is close to the plane of its heavy atoms (dihedral angles of 0.633° and 2.336° are calculated for the donor and acceptor rings, respectively). Therefore, the constraints from theory are close to assuming that the phenol monomer geometry is unchanged by complexation.

The appropriate fitted structural parameters are those that determine the intermolecular structure of the cluster. These include the hinge angle, the O…O distance, and the angles and dihedrals that define the angles between the two planes of the phenol rings. With the simple assumption of a static monomeric structure, the required number of degrees of freedom to fit a good structure is reduced to 8, which is small compared to the number of experimental structural parameters obtained in this study (45, or 3 rotational constants from the parent species and each of the 14 assigned isotopologues). Therefore, the fitted structure can be well-determined precisely, and in fact the precision

of the  $r_0$  fit presented here is limited by the RMS error generated by inaccuracies in the *ab initio* structure. A separate  $r_0$  fit with *ab initio* constraints taken from a MP2/cc-pVTZcp benchmark structure<sup>19</sup> shows similar intermolecular parameters to that of the M06-2Xderived  $r_0$  structure, albeit with poorer determination (since the MP2 geometry is more inaccurate with respect to experimental positions). The MP2-derived  $r_0$  hinge angle, 56.5(37)°, is slightly smaller than the angle determined with the M06-2X geometry, 62.3(14)°, so there is some spread in the determination of this parameter. However, all other intermolecular parameters appear consistent between the two determinations. In addition, comparison of the monomer parameters between the M06-2X and MP2 reveals no significant difference (within the confidence interval for the  $r_0$  determination) between the two levels of theory.

In general, the large amplitude anharmonic hinge motion does not significantly affect the ground state structural determination. Any significant zero-point vibrational deformation of the ground state structure should reveal itself in a discrepancy between the  $r_0$  and  $r_s$  geometries, since the  $r_0$  fit structure does not account for zero-point vibrational motion, whereas the  $r_s$  geometry accounts for some with an additive constant factor to the ground state moments of inertia.<sup>36</sup> Therefore, if a significant zero-point contribution to the ground state molecular structure exists, it should reveal itself in a direct comparison of the  $r_0$  and  $r_s$  geometries. However, Fig. 5 (bottom right panel) reveals that the agreement between these two experimental geometries is excellent. This guarantees that there is no significant contribution to the ground state geometry from the anharmonic hinge potential. The refined phenol dimer structure is shown in Fig. 3 (left) and Fig. 5 (bottom right) and the fitted  $r_0$  structural parameters are reported in Table 4.

The results of this r<sub>0</sub> fit clear up a number of inconsistencies between previous experimental results and ab initio predictions. For instance, previous studies had assumed a hydrogen bonding distance of over 3 angstroms.<sup>1,24</sup> However, ab initio results suggest that the hydrogen bond is actually more water-dimer like, with an O…O distance of about 2.8 Å, a prediction consistent with the r<sub>0</sub> distance of 2.833(21) Å. Microwave measurements of the water dimer report an O…O distance of 2.98(1) Å.<sup>40</sup> Additionally, the previous results have overestimated the inter-ring effects, where a more pronounced C-H… $\pi$  effect is visible between the hydrogens of the acceptor ring and the  $\pi$  system of the donor. A comparison of the new r<sub>0</sub> with previously published results<sup>24</sup> can be found in Fig. 7.



**Figure 7.** Comparison of the new  $r_0$  structure (purple) with the previously published  $r_0$  structure by Schmitt et al. (brown). The donor rings of both structures are overlapping in order to emphasize the differences in the intermolecular geometry ("new" donor ring not shown).

Parameter	M06-2X/6- 311++g(d,p)	MP2/cc-pVTZ- cp	r₀, M06-2X/6- 311++g(d,p)	r <sub>0</sub> , MP2/cc- pVTZ-cp	UV study, r₀
r(H O)	1.889	1.873	1.8730(226)	1.879(38)	2.354(49)
r(O O)	2.833	2.833	2.8326(211)	2.830(36)	3.221(25)
<(07-H7-08)	168.05	170.50	170.5(21)	166.2(37)	150.6(18)
<(C9-O8-H7)	118.88	122.48	122.5(10)	125.8(18)	138.6(15)
t(08-H7-07-C1)	85.07	128.97	75.5(59)	87(12)	109.6(45)
t(C9-08-H7-07)	-23.41	-41.22	-27.7(47)	-35(10)	-26.5(46)
t(C10-C9-O8-H7)	14.18	14.86	10.6(17)	14.8(44)	-1.0(19)
t(C1-O7-O8-C9) (hinge)	59.61	48.17	64.0(13)	56.5(37)	63.0 <sup>4</sup>

**Table 4.** Fitted r<sub>0</sub> intermolecular parameters for phenol dimer for two sets of *ab initio* constraints (M06-2X/6-311++g(d,p) and MP2/cc-PVTZ-cp), with comparisons to the M06-2X/6-311++g(d,p) *ab initio* structure used and a previous high resolution UV study.<sup>24</sup> The MP2/cc-pVTZ-cp values come from benchmark calculations found in Ref. 19. Some parameters fitted in this study (such as C-O bond lengths) are omitted from this table.

 $^4$  Value taken from  $r_s$  parameters reported in the UV study [24], since no  $r_0$  value was reported.

#### b) Phenol Trimer

The experimental carbon atom structure of the trimer obtained from a Kraitchman analysis is shown in Fig. 3 where it is compared to the refined experimental structure obtained from constraining the M06-2X/6-311++g(d,p) *ab initio* structure. Because the parent species is a symmetric top, the Kraitchman analysis provides only two structure parameters for each <sup>13</sup>C atom: the coordinate of the atom on the symmetry axis and the distance of the atom from the symmetry axis. The relative angle of rotation around the symmetry axis between two <sup>13</sup>C atoms was determined by requiring the distance of each carbon atom to the selected reference atom (the <sup>13</sup>C-O position) to equal the C-C distance in the phenol monomer (the theoretical structures show that the geometry of the phenyl group is essentially unchanged by complexation). These distance constraints determine the relative positions of the carbon atoms in a single phenyl ring, the other two positions are generated by  $(2\pi/3)$  rotations of the coordinate positions around the symmetry axis to generate the symmetric top structure.

Theoretical results are generally consistent across varying levels of theory, unlike phenol dimer. Use of diffuse functions in the Pople basis set appears to slightly improve predictions with respect to experiment, but not to the dramatic extent seen in the dimer. MP2 results agree well with the M06-2X/6-311++g(d,p) theoretical structure reported in this study. As with the dimer, the dispersion plays a crucial role in determining the complexation geometry, and use of B3LYP causes the barrel-like structure to open up into an *asymmetric* umbrella structure. These results, with comparison to the experimental rotational constants, can be found in Table 5.

	M06-2X/6-	
Parameter	311++g(d,p)	r <sub>m</sub> (1)
r(H … O) / Å	1.940	1.895(86)
r(O … O)	2.811	2.760(70)
r(C1 - C1')	4.010	3.967(83)
<(0 - 0' - 0'') / °	60.0	60.03(73)
<(O - H ··· O')	147.3	147.1(16)
<(C1 - O'' O')	-23.4	117.9(14)
<(C1'' - O'' - O)	117.9	114.0(16)
<(C1 ··· C1' ···		
C1'')	60.0	60.00(75)
t(C1-O-O'-C1')	-6.42	-5.9(28)
t(O-O'-C1'-C6')	85.2	85.2(33)
c <sub>aa</sub> = c <sub>bb</sub> / amu Å <sup>2</sup>		1.140(85)

**Table 5**. Fitted  $r_m^{(1)}$  intermolecular parameters for phenol trimer, with comparisons to the M06-2X/6-311++g(d,p). Due to the C<sub>3</sub> symmetry, these parameters are symmetric upon cyclic interchange of atoms  $A \rightarrow A' \rightarrow A''$  for all atoms A contained in the parameter.

A refined  $r_m^{(1)}$  structure using constraints from the M06-2X/6-311++g(d,p) theoretical structure is shown in Fig. 3. Like the phenol dimer analysis, the constraints from theory fixed the geometry of each phenol rings where, also for the trimer, the calculated phenol geometry in the cluster is essentially the same as the phenol monomer.  $r_m^{(1)}$  was chosen over  $r_0$  due to the lack of experimental <sup>18</sup>O positions. Without these positions, r<sub>0</sub> determination of oxygen-containing structural parameters was poorer than  $r_m^{(1)}$ . This is due to the fact that the  $r_m^{(1)}$  model functions very similarly to Kraitchman, except that the vibrational corrections, which are significant for atoms near the principal axes, are fit with three mass-dependent c<sub>ii</sub> coefficients.<sup>41</sup> Kraitchman's equations, however, assume a constant mass-independent vibrational correction to the ground-state moments of inertia.<sup>36</sup> In the case of a symmetric top like phenol trimer,  $c_{aa} = c_{bb}$ , and  $c_{cc}$ cannot be determined due to the  $C_3$  symmetry about the c-axis (in the prolate basis), so only one  $c_{ii}$  coefficient was fit. The fitted value for  $c_{aa} = c_{bb}$ , as well as the fitted intermolecular parameters for phenol trimer, can be found in Table 6. The experimental O····O distance is slightly shortened compared to phenol dimer, at 2.760(70) Å. This is similar to observations of water trimer versus dimer, where the trimer  $O \cdots O$  distance is shortened from 2.98 Å to 2.85 Å due to stabilizing three body effects.<sup>42</sup> Additionally, a C-H  $\cdots \pi$  interaction is apparent, where the ortho hydrogen of a donor ring coordinates with the 5-6 bond in the acceptor ring. This creates a roughly perpendicular arrangement of the rings in the barrel (the dihedral  $O \cdots O' - C_1' - C_6'$  is  $85.2(33)^\circ$  in the  $r_m^{(1)}$  structure). In this case, it appears that parallel  $\pi$  stacking is not optimal for stable complexation, as the rings try to associate in as perpendicular of a fashion without sacrificing the monomer planarity and hence the water trimer-like hydrogen bonding motif.

	A (MHz)	B (MHz)	C (MHz)
M06-2X/6-311g(d,p)	199.494	292.600	
M06-2X/6-311++g(d,p)	198.084	290.540	
MP2/6-311g(d,p)	205.635	296.244	
MP2/6-311++g(d,p)	211.380	300.117	
B3LYP/6-311++g(d,p)	240.6	222.786	130.919
Experiment	[188]	282.280790(56)	

**Table 6.** Phenol trimer rotational constants in the oblate basis for various levels of theory,compared to experiment. The experimental A constant is assumed based on the value of

C in the isotopologues (prolate basis), since it cannot be fit experimentally.

A recent study by Yang and Waller on cooperative effects in gas-phase homotrimers reveals common threads between the results presented here for phenol trimer and the benchmark molecules used in their study, and sheds additional light on the efficacy of theoretical methods in treating the cooperative effects between dispersion and hydrogen bonding.<sup>43</sup> Yang and Waller note that use of a limited basis set in an RI-MP2 calculation can lead to significant differences when extrapolating to the complete basis set (CBS) limit in three-body interaction energies, sometimes up to 5% (for methanol trimer). However, as seen with our success with M06-2X, they note that dispersionmodified DFT calculations tend to give the best balance between computational cost and accuracy. Similar results were found for the interplay between hydrogen bonding and interring dispersion, where direct parallel  $\pi$  stacking is an inefficient complexation geometry for molecules that contain hydrogen bonding interactions.

## 5) Conclusion

In this study, the structures for phenol dimer and trimer have been determined through the use of chirped-pulse FTMW spectroscopy in the 2-8 GHz band. For the dimer, the r<sub>s</sub> and least-squares fit r<sub>0</sub> geometries are improved over previously reported experimental geometries when compared to the best available theoretical structures. The dimeric structure is found to represent a case where an interplay between dispersion and hydrogen bonding play an essential role in fixing the complexation geometry, with a water dimer-like O…O distance of 2.833(21) Å, and a small dihedral between the rings of 62.3(14)°. However, both DFT and MP2 calculations struggle in determining an accurate structure. Instead, functionals like M06-2X and dispersion-modified DFT tend to give the most accurate results. For phenol trimer, a C<sub>3</sub> symmetric barrel-like helical structure is found, again with excellent agreement with the best available theoretical structures. The determined least-squares fit  $r_m^{(1)}$  structure the three body interactions strengthen the hydrogen bonding in the trimer, with a shortened O…O distance of 2.760(60) Å, similar to that of water trimer.

With the new improvements to the 2-8 GHz CP-FTMW instrument design, structure determination of gas-phase clusters with over ten heavy atoms is becoming routine. In this regime of large gas-phase clusters, the complex interactions often involved in determining the actual global minimum cluster geometry is often a difficult problem for standard *ab initio* calculations to solve accurately and consistently. Therefore, it is important to supplement the theoretical methods with accurate experimental structures such as those for phenol dimer and trimer, in order to fine tune the accuracy of *ab initio* structures both for theory and for experimental predictions. Experimentally, this is especially important for large species with many isotopologues that are often weak in intensity and closely spaced together in a dense spectral region (like both species studied in this report), where a good theoretical geometry is required to make accurate scaled predictions for isotopologue rotational constants.

## References

(1) Hobza, P.; Riehn, C.; Weichert, A.; Brutschy, B. *Chem. Phys.* **2002**, *283* (1–2), 331.

(2) Parthasarathi, R.; Subramanian, V.; Sathyamurthy, N. J. Phys. Chem. A 2005, 109
(5), 843.

(3) Kolář, M.; Hobza, P. J. Phys. Chem. A 2007, 111 (26), 5851.

(4) Kitaura, K.; Morokuma, K. Int. J. Quantum Chem. 1976, 10 (2), 325.

(5) Wu, X.; Vargas, M. C.; Nayak, S.; Lotrich, V.; Scoles, G. J. Chem. Phys. 2001, 115 (19), 8748.

(6) Klimeš, J.; Michaelides, A. J. Chem. Phys. 2012, 137 (12), 120901.

(7) Strømsheim, M. D.; Kumar, N.; Coriani, S.; Sagvolden, E.; Teale, A. M.; Helgaker, T. *J. Chem. Phys.* **2011**, *135* (19), 194109.

(8) Peverati, R.; Truhlar, D. G. *ArXiv12120944 Cond-Mat Physicsphysics Physicsquant-Ph* **2012**.

(9) Burns, L. A.; Mayagoitia, Á. V.-; Sumpter, B. G.; Sherrill, C. D. J. Chem. Phys. **2011**, *134* (8), 084107.

(10) Riley, K. E.; Pitoňák, M.; Jurečka, P.; Hobza, P. Chem. Rev. 2010, 110 (9), 5023.

(11) Müller-Dethlefs, K.; Hobza, P. Chem. Rev. 2000, 100 (1), 143.

(12) Riley, K. E.; Hobza, P. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2011, 1 (1), 3.

(13) Wang, W.; Donini, O.; Reyes, C. M.; Kollman, P. A. Annu. Rev. Biophys. Biomol. Struct. 2001, 30 (1), 211.

(14) Knowles, R. R.; Jacobsen, E. N. Proc. Natl. Acad. Sci. 2010, 107 (48), 20678.

(15) Tang, W.; Johnston, S.; Iggo, J. A.; Berry, N. G.; Phelan, M.; Lian, L.; Bacsa, J.; Xiao, J. *Angew. Chem. Int. Ed.* **2013**, *52* (6), 1668.

(16) Lehn, J.-M. Chem. Soc. Rev. 2007, 36 (2), 151.

(17) Fuke, K.; Kaya, K. Chem. Phys. Lett. 1982, 91 (4), 311.

(18) Fuke, K.; Kaya, K. Chem. Phys. Lett. 1983, 94 (1), 97.

(19) Jureka, P.; Sponer, J.; Cerny, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8* (17), 1985.

(20) Connell, L. L.; Ohline, S. M.; Joireman, P. W.; Corcoran, T. C.; Felker, P. M. J. *Chem. Phys.* **1992**, *96* (4), 2585.

- (21) Weichert, A.; Riehn, C.; Brutschy, B. J. Phys. Chem. A 2001, 105 (23), 5679.
- (22) Ebata, T.; Watanabe, T.; Mikami, N. J. Phys. Chem. 1995, 99 (16), 5761.
- (23) Vaden, T. D.; Lisy, J. M. J. Chem. Phys. 2006, 124 (21), 214315.

(24) Schmitt, M.; Böhm, M.; Ratzer, C.; Krügler, D.; Kleinermanns, K.; Kalkman, I.; Berden, G.; Meerts, W. L. *ChemPhysChem* **2006**, *7* (6), 1241.

(25) Brause, R.; Santa, M.; Schmitt, M.; Kleinermanns, K. *ChemPhysChem* 2007, 8(9), 1394.

Brown, G. G.; Dian, B. C.; Douglass, K. O.; Geyer, S. M.; Shipman, S. T.; Pate,
B. H. *Rev. Sci. Instrum.* **2008**, *79* (5), 053103.

(27) Neill, J. L.; Shipman, S. T.; Alvarez-Valtierra, L.; Lesarri, A.; Kisiel, Z.; Pate, B. H. *J. Mol. Spectrosc.* **2011**, *269* (1), 21.

(28) Pérez, C.; Lobsiger, S.; Seifert, N. A.; Zaleski, D. P.; Temelso, B.; Shields, G. C.; Kisiel, Z.; Pate, B. H. *Chem. Phys. Lett.* **2013**, *571*, 1.

(29) Steber, A. L.; Neill, J. L.; Zaleski, D. P.; Pate, B. H.; Lesarri, A.; Bird, R. G.; Vaquero-Vara, V.; Pratt, D. W. *Faraday Discuss.* **2011**, *150*, 227.

- (30) Campbell, E. J.; Lovas, F. J. Rev. Sci. Instrum. 1993, 64 (8), 2173.
- (31) Watson, J. K. G. Mol. Phys. 1968, 15 (5), 479.
- (32) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120* (1-3), 215.
- (33) Zhao, Y.; Truhlar, D. G. Acc. Chem. Res. 2008, 41 (2), 157.

(34) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.;
Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji,
H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.;
Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida,
M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.;
Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.;
Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J.
C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.;
Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.;
Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.;
Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.;
Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.;
Fox, D. J. *Gaussian 09, Revision A.1*; Gaussian, Inc.: Wallingford, CT, 2009.

- (35) Kraitchman, J. Am. J. Phys. **1953**, 21 (1), 17.
- (36) Costain, C. C. J. Chem. Phys. 1958, 29 (4), 864.

(37) Valdés, H.; Klusák, V.; Pitoňák, M.; Exner, O.; Starý, I.; Hobza, P.; Rulíšek, L. J. *Comput. Chem.* **2008**, *29* (6), 861.

(38) Schmitt, M.; Henrichs, U.; Müller, H.; Kleinermanns, K. J. Chem. Phys. 1995, 103 (23), 9918.

- (39) Kisiel, Z. J. Mol. Spectrosc. 2003, 218 (1), 58.
- (40) Dyke, T. R.; Mack, K. M.; Muenter, J. S. J. Chem. Phys. 1977, 66 (2), 498.
- (41) Watson, J. K. G.; Roytburg, A.; Ulrich, W. J. Mol. Spectrosc. **1999**, 196 (1), 102.
- (42) Keutsch, F. N.; Cruzan, J. D.; Saykally, R. J. Chem. Rev. 2003, 103 (7), 2533.
- (43) Yang, J.; Waller, M. P. J. Phys. Chem. A 2013, 117 (1), 174.

# **Chapter 4: The Structure and Internal Rotation Dynamics of Sevoflurane…Benzene**

(Portions of this research have been peer-reviewed and published in Angewandte Chemie International Edition, with the following citation: N. A. Seifert, D. P. Zaleski, C. Perez, J. L. Neill, B. H. Pate, M. Vallejo-Lopez, A. Lesarri, E. J. Cocinero, F. Castaño, I. Kleiner, *Angew. Chemie. Int. Ed.*, **2014**, *126*, 3274.)

# 1) Introduction

To demonstrate the unparalleled sensitivity of the current generation of CP-FTMW instrumentation, the assignment and experimental structural determination of the 27-atom sevoflurane...benzene complex using CP-FTMW spectroscopy is reported. The complex of sevoflurane, a commonly-used anesthetic with the formula (CF<sub>3</sub>)<sub>2</sub>CHOCH<sub>2</sub>F, with benzene has been recently studied by Dom *et al.*, who reported the observation of two isomers in gas-phase and xenon cryosolution IR spectroscopy.<sup>1</sup> The most abundant isomer showed vibrational bands consistent with a complex forming primarily from a C-H… $\pi$  interaction between the isopropyl hydrogen of sevoflurane and the benzene ring, whereas the weaker isomer's bands were consistent with a combination of C-H… $\pi$  and C-H…F interactions between the ring and the fluoromethyl group. However, due to the nature of the experiment, conclusive structural determination was not made and was limited to a comparison of the vibrational band assignments and experimental complexation enthalpies with *ab initio* calculations.

The structural assignment from CP-FTMW spectroscopy confirms the existence of the most abundant isomer seen by Dom and coworkers, but the second isomer was not detected in the cold supersonic expansion of the CP-FTMW experiment. Along with the parent species of the detected isomer, all 12 <sup>2</sup>D and <sup>13</sup>C benzene-associated isotopologues have been assigned, as well as 33 of the 60 possible  ${}^{13}C/{}^{2}D$  double isotopologues. With this data, a full least-squares fit r<sub>0</sub> structure was assigned for the observed complex in excellent agreement with the reported theoretical geometries. The experimental results are additionally complemented by a comprehensive conformational analysis, exploring the predictive capacity of different *ab initio* and DFT methods.

This complex shows a novel form of high barrier six-fold internal rotation, where the benzene acts as a C<sub>6</sub>-symmetric top about the C<sub>1</sub> frame of sevoflurane. This case is different than previously studied cases of six-fold internal rotation in benzene clusters<sup>2–7</sup>, where the six-fold symmetry typically arises from a linear or C<sub>2v</sub> top (such as water) rotating about the benzene frame. In this case, the inherently low energetic barrier to internal rotation in the non-covalent complexes causes analysis to be restricted to the low barrier limit or free rotation.<sup>8</sup> However, in sevoflurane...benzene a clear six-fold internal rotation in the high barrier limit is observed. This is due to the relatively small rotational constant of benzene top along its C<sub>6</sub> symmetry axis, which causes the lowest-lying torsional levels to be bound well below the barrier. Fortunately, this kind of six-fold internal rotation can be reduced into a two-state C<sub>3</sub> top problem, which enables the fitting of the rotational spectrum in the torsional ground state to below experimental uncertainty with the program BELGI-C<sub>1</sub>.<sup>9,10</sup>

#### 2) Experimental

## a) Spectroscopy

The experiments were conducted with the CP-FTMW spectrometer at the University of Virginia. In all cases, both sevoflurane (brand name Ultane, Abbott Laboratories) and benzene (99%, Aldrich) were mixed in a tank and diluted to 0.2% each by pressurizing with *ca*. 6 atm of the desired backing gas (either He or Ne). Additionally, for measurement of all <sup>2</sup>D single and <sup>13</sup>C/<sup>2</sup>D double isotopologue spectra, a mixture of 0.2% benzene-d<sub>1</sub> (98%, Aldrich) and 0.2% sevoflurane was used.

The initial spectroscopy was performed in the 2-8<sup>11</sup> and 6.5-18.5 GHz<sup>12</sup> bands utilizing various experimental setups. The original spectrum was recorded from 6.5-18.5 GHz with a typical nozzle arrangement perpendicular to the polarization pulse, for a total of *ca*.  $5x10^5$  averages. However, the natural Doppler linewidth at this frequency (*ca*. 120 kHz), coupled with the short free induction decay (FID) record lengths of 20 µs (in neon) and 10 µs (in helium), is too broad to resolve the internal rotation splittings.

A 2-8 GHz spectrum was obtained in a coaxial arrangement, where the expansion is approximately collinear with the polarization pulse. Since the Doppler distribution profile is significantly narrowed via this method, and a collinear expansion increases the Doppler dephasing time, a longer FID record length of 100  $\mu$ s was possible, increasing the overall frequency resolution. With a total of 4 FIDs per pulse at a pulse rate of 0.4 Hz, a total averaged spectrum of *ca*. 10<sup>5</sup> averages was obtained. Linewidths of approximately 25 kHz were achieved, which enabled resolved measurement of the internal rotation splittings. Since the average direction of the expansion is in the direction of the microwave propagation, a slight Doppler shift is detectable in the measured frequencies. By comparing the frequencies of sevoflurane transitions between the coaxial and the perpendicular experiments, an average frequency ratio was determined for both the neon and helium expansions. This ratio was then applied to each spectrum in order to correct for the Doppler shifting in the coaxial spectrum. For the neon expansion that ratio was determined to be 0.99999827; while for the helium expansion that ratio was determined to be 0.99999545.

Another spectrum of the complex in neon backing gas was recorded later using newly upgraded components for the 2-8 GHz band, described in detail previously in Chapter 1, Section IV. With a FID record length time of 40  $\mu$ s, an average linewidth of approximately 60 kHz was achieved, which is sufficient for resolving the internal rotation splittings of the parent species. Approximately  $1.1 \times 10^7$  FIDs were obtained and averaged in the time domain. For the sevoflurane/benzene-d<sub>1</sub> measurements, the experiment was performed in identical fashion, and a total of  $6.2 \times 10^6$  FIDs were collected and averaged together in the time domain.

Unless otherwise noted, the analysis in this report is derived primarily from the spectra obtained on the recently upgraded CP-FTMW instrument, which includes the most recent improvements to the 2-8 GHz arrangement. These new spectra, of sevoflurane/benzene and sevoflurane/benzene-d<sub>1</sub>, are significantly improved in terms of dynamic range and sensitivity compared to the original spectra and provide the dynamic range required to observe <sup>13</sup>C isotopologues of the complex.

#### b) Computational

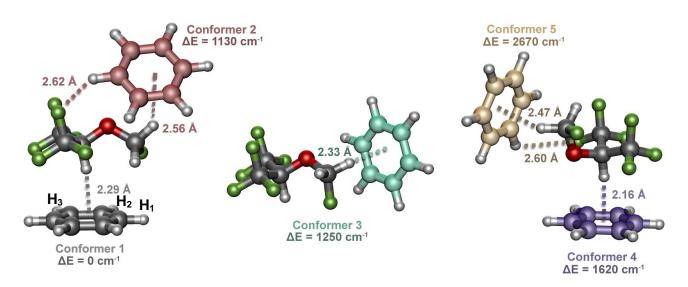
The initial conformational screening of sevoflurane-benzene used the Merck Molecular Force Field (MMFFs) combined with Monte-Carlo and Large-Scale/Low-Modes search algorithms, as implemented in Macromodel.<sup>13</sup> The set of 95 lowest-lying structures were later fully reoptimized using several ab initio  $(MP2)^{14}$  and DFT methods  $(B3LYP and M06-2X)^{15}$ , converging into 26 different structures within an energy window of 43 kJ mol-1 (6-311++g(d,p) basis set was used for all methods). The rovibrational analysis was based the harmonic approximation.

Binding energies were corrected for basis-set superposition errors using the counterpoise procedure.<sup>16</sup> All ab initio and DFT predictions were calculated using the GAUSSIAN 09 suite.<sup>17</sup> Zeroth-order symmetry adapted perturbation theory (SAPT0) calculations were also performed using the jun-cc-pVDZ basis (aug-cc-pVDZ with no diffuse functions on hydrogen and no diffuse *f* functions on heavier atoms in the basis) using the Psi4 quantum chemistry package.<sup>18</sup> In order to speed up calculations, density fitted SAPT0 was used, with the aug-cc-pVDZ-JKFIT auxiliary basis set used for the SCF calculations and aug-cc-pVDZ-RI for the SAPT part of the calculations.<sup>19,20</sup>

## 3) **Results**

A visual summary of the initial conformational screening for the sevoflurane...benzene can be found in Figure 1. A full summary of relevant energetic and rotational parameters for the five lowest energy conformers can be found in Table 1. The five lowest energy conformers shown in Figure 1 and Table 1 contain a variety of hydrogen bonding effects. Conformers such as 2 and 5 both show C-H…F and C-H… $\pi$  interactions, but it appears that the single C-H… $\pi$  interaction of Conformer 1 is the most energetically favorable complexation mode by 13.5 kJ·mol<sup>-1</sup> (MP2). Labeling for the benzene positions in Conformer 1 begins at the 1- position (written H<sub>1</sub> or 1-C, depending on which atom is being discussed), which is the hydrogen eclipsed by the fluoromethyl group, and proceeds clockwise (perspective is opposite the sevoflurane-bound face of the ring).

Conformer 1 shows an intriguing topology for the benzene not seen in the other conformers, where eclipsing one hydrogen with the fluoromethyl group allows for the other 5 hydrogens to be optimally staggered between the sevoflurane substituents. The C-F…H-C bonds in Conformer 1 are longer than those in Conformers 2 and 5, ranging from 2.9 Å (OCH<sub>2</sub>F…H<sub>1</sub>) to 3.5 Å (CF<sub>3</sub>…H<sub>3/5</sub>), but given the staggered arrangement of the fluorines with the benzene hydrogens there is likely some stabilizing contribution from these weak hydrogen bonds to the overall complexation geometry.



**Figure 1.** Structures of the five lowest energy conformers of the sevoflurane...benzene complex. For conformers 1/2 and 4/5, sevoflurane monomers are shown as overlapping, since they share the same conformation. Structures and relative energies shown are those calculated using the MP2/6-311++g(d,p) level of theory.

	Theory ( MP2 / M06-2X / B3LYP )				
	Conf. 1	Conf. 2	Conf. 3	Conf. 4	Conf. 5
A / MHz	508/ 515 / 500	650 / 670 / 636	663 / 698 /636	543 / 557 / 538	685 / 697 / 667
<i>B</i> / MHz	376/ 379 / 319	264 / 263 / 209	256 / 253 / 211	358 / 355 / 305	273 / 279 / 228
C / MHz	353/ 358 / 302	235 / 235 / 191	234 / 233 / 193	325 / 318 / 281	229 / 234 / 196
$ \mu_a $ / <b>D</b>	2.2/ 2.3 / 2.1	1.9 / 2.3 / 2.1	1.1 / 1.3 / 2.1	1.1 / 1.0 / 1.1	1.6 / 1.8 / 1.6
$ \mu_b  / \mathbf{D}$	0.5/ 0.3 / 0.2	0.6  /  0.4  /  0.8	1.3 / 1.7 / 0.8	0.8 / 0.8 / 0.9	1.1 / 1.0 / 1.1
$ \boldsymbol{\mu}_c  / \mathbf{D}$	1.6/ 1.6 / 1.7	1.3 / 1.3 / 1.2	1.3 / 1.0 / 1.2	1.3 / 1.3 / 1.3	0.5  /  0.6  /  0.6
$ \mu_{TOT}  / \mathbf{D}$	2.7/ 2.8 / 2.7	2.4 / 2.6 / 2.6	2.2 / 2.3 / 2.6	1.8 / 1.8 / 1.9	2.0 / 2.1 / 2.1
D <sub>J</sub> / kHz	0.02 / 0.02 /	0.02 / 0.01 /	0.02 / 0.01 /	0.02 / 0.02	0.008 / 0.007 /
	0.07	0.1	0.07	/0.04	0.05
$D_{JK}$ / kHz	0.004 / 0.01 / 0.2	0.2 / 0.2 / - 0.3	0.1 / 0.2 / - 0.03	0.08 / 0.1 / 0.3	0.1 / 0.1 / 0.04
$D_K / \mathrm{kHz}$	0.04 / 0.04 / 0.3	0.1 / 0.1 / 0.4	0.8 / 0.2 / 0.1	-0.07 / -0.1 / -0.2	0.2 / 0.1 / -0.02
$d_1$ / Hz	-2.8/ -2.8/ - 14.4	1.4 / -1.6 / - 5.6	1.0 / 0.7 / -3.4	-2.6 / -3.3 / - 7.0	-0.8 / -0.5 / -7.0
$d_2$ / Hz	0.2 / 0.2/ -0.7	0.6 / -0.1 /- 0.2	0.3 / 0.2 / -0.2	-0.3 / -0.2 / - 0.3	0.0 / 0.1 / -0.5
⊿G / kJ·mol <sup>-1</sup>	0.0 / 0.0 / 0.0	8.6 / 4.5 / - 1.5	7.9 / 10.2 / 0.1	19.0 / 13.8 / 17.0	31.0 / 22.7 / 17.8
∆ (E+ZPE) / kJ∙mol <sup>-1</sup>	0.0 / 0.0 / 0.0	13.5/ 10.3/ 2.2	14.9/ 13.5/ 2.2	19.4/ 17.0/ 15.5	31.9/ 24.6/ 18.5
E <sub>d</sub> / kJ·mol <sup>-1</sup>	-17.8 / -31.1 /- 20.5	-12.8 / -20.9 / -5.0	-11.9 / -19.1 / -5.3	-17.4 / -26.1/ -5.6	-12.4 / -20.6 / - 4.4

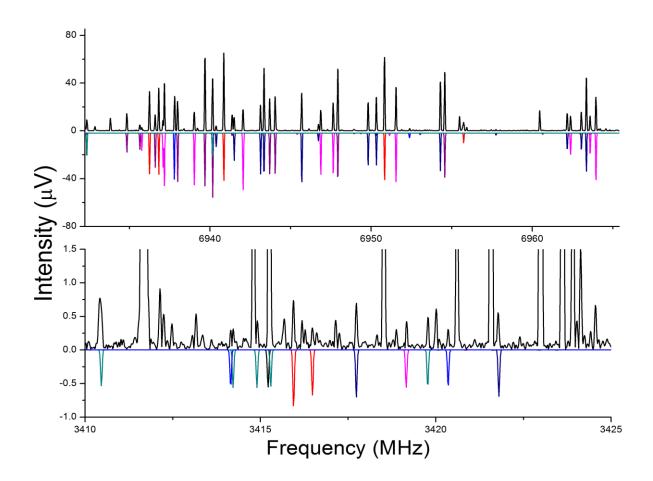
**Table 1.** Rotational constants and energetics of the five lowest energy conformers found in the conformational screening, which are shown in Figure 1. The basis set used is 6-311++g(d,p) in all cases.

As stated earlier, the vibrational results published by Dom and coworkers present evidence for two species, one similar to Conformer 1 where the single isopropyl C-H $\cdots\pi$ interaction dominates, and one similar to Conformer 2, where the primary hydrogen bonding arises from interactions between the fluoromethyl hydrogens and the benzene ring and the fluorine with one of the benzene hydrogens. Based on their calculations of complexation energies, they predict a preference for the isopropyl-bonded complex by a factor of 3.75, which they claim to be underestimated since the vibrational band intensity comparison suggests a population ratio of over 15:1. Though the pulsed jet in this expansion is colder than the methods used by the previous IR study, the energetic environment within the expansion can enable a buildup of population of higher energy conformers and species<sup>21</sup>, especially with the light inert carrier gases such as neon and helium as used in this experiment.<sup>22-24</sup> Therefore, spectral inspection and automated fitting searches, using AUTOFIT, were undertaken in order to find higher energy isomers. However, only one sevoflurane-benzene isomer was observed, consistent with Conformer 1 on the basis of rotational constants alone. An overview of the 2-8 GHz broadband rotational spectrum of a mixture of sevoflurane and benzene is shown in Figure 2.

The strongest transitions in the spectrum arise from sevoflurane monomer, which have a dynamic range of approximately 20 000:1. The sevoflurane...benzene parent species is approximately a factor of 50 weaker than the sevoflurane monomer transitions, but an average signal to noise ratio of *ca*. 400-600:1 is more than sufficient to resolve all relevant heavy atom single isotopologues (<sup>13</sup>C, <sup>18</sup>O). However, the sevoflurane...benzene spectrum is complicated by internal rotation splittings arising from the rotation of the benzene monomer about the sevoflurane frame. A single rotational transition is therefore

split into four components, which correspond to the 4 irreducible representations (A, B,  $E_1, E_2$ ) of the C<sub>6</sub> point group.

Due to the internal rotation, all possible isotopologues where the substitution arises in the sevoflurane monomer are additionally complicated by the splitting. However, substitutions on the benzene break the C<sub>6</sub> symmetry, and the resulting isotopologues show simple asymmetric top rigid rotor spectra. The six benzene-<sup>13</sup>C isotopologues were detected in natural abundance and assigned using the standard Watson A-reduction Hamiltonian.<sup>25–27</sup> In order to enable easy assignment of the sevoflurane-derived isotopologues, benzene-d<sub>1</sub> is used in order to force the symmetry breaking.



**Figure 2.** Overview of the sevoflurane····benzene-d<sub>1</sub> spectrum (black). The six deuterated complexes (multiple colors) are shown in the 30 MHz in the top panel, and are found at approximately a factor of 20-50 times weaker than the monomer. The bottom panel shows a small 15 MHz portion, roughly a factor of 100 times weaker in intensity than the transitions of the top panel, of the spectrum with predicted transitions for some of the 33 assigned  ${}^{13}C/{}^{2}H$  isotopologues (multiple colors). The colors of the  ${}^{13}C/{}^{2}H$  isotopologue predictions correspond directly to the colors of the  ${}^{2}H$  predicted spectra in the top panel, although in the bottom panel the specific carbon that is substituted varies from transition to transition.

In the sevoflurane/benzene-d<sub>1</sub> spectrum, similar intensities are seen for the sevoflurane monomer and the sevoflurane...benzene-d<sub>1</sub> complexes. Although the average transition intensity is boosted by roughly a factor of 4 by breaking the C<sub>6</sub> symmetry, it is additionally split six times between the six unique d<sub>1</sub> isotopologues. Despite this, these isotopologues are generally intense enough (e.g. SNR  $\geq$  300:1) to assign <sup>13</sup>C/<sup>2</sup>D double isotopologues in natural abundance, of which there are 60 possible.

Each hydrogen position on the ring appear to not be equivalent in the supersonic expansion, as the intensities vary from position to position. For instance, the spectrum associated with deuterium substitution at the 1- position (defined as the hydrogen eclipsing the fluoromethyl group) is sufficiently weak that transitions associated with  ${}^{13}C$ isotopologues of the 1-<sup>2</sup>D isotopologue are below the noise. This variation in intensity, coupled with the extremely high spectral density of transitions at the SNR associated with these  ${}^{13}C$  isotopologues (e.g. SNR < 10:1) can lead to abnormally weak isotopologue transitions, blends with stronger transitions and confusion-limited pile-ups with other weak lines. Despite this limitation, 33 of the 60 possible  ${}^{13}C/{}^{2}D$  isotopologues were assigned and every carbon has at least one isotopologue assignment associated with it. All assigned isotopologues and the parent species have their distortion constants fixed to the fitted values of the  $5^{-2}$ D isotopologue. This choice of isotopologue is arbitrary, but floating distortion for other similarly intense isotopologues (e.g.  $2^{-2}D$  and  $4^{-2}D$ ) results in statistically similar distortion constants. A summary of the fitted rotational constants for the parent species (specifically, the A and B symmetry states, whose spectra can be fit with a Watson Hamiltonian) and all assigned <sup>13</sup>C and <sup>2</sup>D isotopologues can be found in Table 2.

Species	A /MHz	B /MHz	C /MHz	Δ <sub>J</sub> /kHz	Δ <sub>JK</sub> /kHz	Δ <sub>K</sub> /kHz	Ν	RMS /kHz
Parent (A)	508.42070(40)	358.82931(13)	338.32685(13)	[0.03490]	[0.0550]	[-0.0630]	77	6.04
Parent (B)	507.74250(60)	358.82020(12)	338.30995(12)	"	••	"	107	6.41
Parent (avg)	508.08160(50)	358.82476(13)	338.31840(13)	"	**	"		
5- <sup>2</sup> D	505.704895(85)	356.433250(86)	335.438320(92)	0.03490(44)	0.0550(14)	0.0630(16)	225	6.91
1- <sup>2</sup> D	505.425300(95)	355.461685(38)	335.793300(41)	"	"	"	175	6.24
2- <sup>2</sup> D	504.157430(60)	357.392938(38)	335.338172(35)	••	"	"	254	7.13
3- <sup>2</sup> D	503.901760(58)	356.776863(34)	336.551764(34)	"	"	"	247	6.37
4- <sup>2</sup> D	504.480480(54)	355.830667(33)	337.133881(31)	"	"	"	272	6.57
6- <sup>2</sup> D	506.345460(81)	355.463814(29)	335.477671(28)	.,	"	"	188	4.81
5- <sup>13</sup> C	506.7452(77)	356.63733(29)	336.95711(22)	[0.03490]	[0.0550]	[-0.0630]	49	6.15
1- <sup>13</sup> C	507.7542(70)	356.32317(20)	336.13907(22)	"	"	"	57	5.23
2- <sup>13</sup> C	507.2520(41)	356.44752(14)	336.24013(15)	"	"	"	52	3.53
3- <sup>13</sup> C	506.5724(49)	357.20783(16)	336.21528(17)	"	"	"	52	3.74
4- <sup>13</sup> C	506.3589(71)	357.04374(21)	336.76461(21)	"	"	"	59	5.96
6- <sup>13</sup> C	507.3754(94)	356.67671(27)	336.29174(30)	"	"	"	56	6.09

**Table 2.** Fitted rotational constants for the parent species (A and B symmetry states) and the benzene associated <sup>13</sup>C and <sup>2</sup>D isotopologues. Distortion is held fixed to the values fit for the 5-<sup>2</sup>D isotopologue. Experimental uncertainty is approximately 10 kHz. Fit constants for the other 33 assigned double <sup>13</sup>C/<sup>2</sup>D isotopologues can be found in the published study by Seifert *et al.*, cited at the beginning of this chapter.

## 4) Discussion

#### a) Structure

The analysis of the rotational spectra of a large number of isotopologues of the sevoflurane···benzene complex enables the determination of an experimental structure, including parameters for all heavy atoms and the six hydrogens on the benzene ring. Since Kraitchman's equations are limited to single substitutions, a least squares fitting procedure to determine the ground-state  $r_0$  structure is more appropriate given the abundance of double isotopic substitution-associated rotational constants. Additionally, application of the Kraitchman coordinates derived from the sevoflurane monomer isotopologues measurements reported by Lesarri et al.<sup>28</sup>, and subsequent rotation into the inertial frame of the sevoflurane···benzene complex, shows excellent agreement with the *ab initio* complex structure. Therefore, for the purposes of the least-squares  $r_0$  fit, the assumption that the global minimum sevoflurane monomer structure does not change upon complexation is valid.

A requirement of an accurate least squares structure determination is an overabundance of experimental moments of inertia with respect to the independent degrees of inertial freedom in the molecule. Excluding the three hydrogens on the sevoflurane monomer, there are 3N-6 = 66 independent degrees of freedom. With one parent species,  $6^{13}C$  and  $6^{2}D$  benzene isotopologues, and  $33^{13}C/^{2}D$  doubles assigned, there are a total of 138 experimental ground-state moments of inertia (3 rotational constants per species). Assuming  $D_{6h}$  symmetry for the benzene moiety also reduces the number of independent degrees of freedom. Using only *ab initio* constraints from the

M06-2X/6-311++g(d,p) structure on the positions of the three hydrogens in sevoflurane and the F-C-F bond angles of the isopropyl group, which have a statistically insignificant dependence on the experimental moment of inertia information available, a full  $r_0$ structure of the sevoflurane····benzene complex was determined using STRFIT.<sup>29</sup> In fact, use of the parent species constants is not necessary and statistically identical results are achieved using only the isotopic data. A comparison between the M06-2X/6-311++g(d,p) structure and the  $r_0$  results can be found in Figure 3.

A potential issue that could bias the  $r_0$  results is use of an *ab initio* structure for constraints whose rotational constants most closely approximates the experimental rotational constants of the parent species instead of the structure with the most accurate treatment of the energetics associated with complexation. This was a potential issue in the work on (phenol)<sub>2</sub> (see Chapter 3), where the M06-2X/6-311++g(d,p) appeared to approximate the structure better than the higher level counterpoise-corrected MP2/cc-pVTZ benchmark structure.<sup>30</sup> However, in that study, similar  $r_0$  results were found with constraints from either level of theory. Likewise in this study, use of the MP2/6-311++g(d,p) structure results shows little difference in  $r_0$  parameter determinations compared to M06-2X.

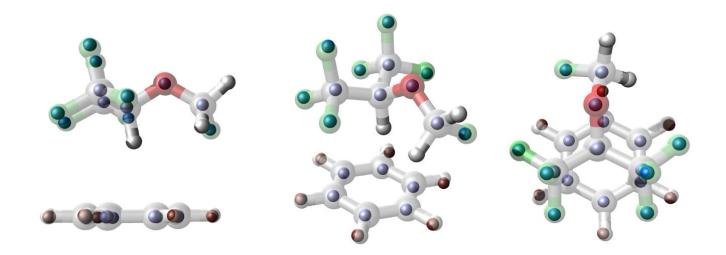


Figure 3. Comparison of M06-2X/6-311++g(d,p) structure (transparent molecular frame) with the experimentally-determined  $r_0$  structure (small colored spheres).

The only striking difference between the MP2 and M06-2X calculations and the  $r_s/r_0$  experimental structures is a slight rotation in the benzene orientation. This can be interpreted as a small deviation of the ground state structure from the equilibrium structure due to the low-lying benzene torsional mode (v = 14 cm<sup>-1</sup>). Further anharmonic corrections to the theoretical structure would likely alleviate this discrepancy.

Since dispersion plays an important energetic role in C-H··· $\pi$  interactions<sup>31</sup>, B3LYP is generally a poor choice<sup>30</sup> for accurate spectroscopic predictions of complexes like sevoflurane···benzene, which can be seen by the poor agreement with experimental rotational constants. However, considering polarity effects, the isopropyl hydrogen is rather Lewis acidic, as it is flanked by three strong electron withdrawing groups (two perfluoro and one fluoromethyl). The electrostatic contribution due to the interaction of this relatively acidic hydrogen and the Lewis basic benzene  $\pi$  cloud likely activates the C-H··· $\pi$  interaction with respect to typical C-H··· $\pi$  model systems<sup>31,32</sup> where the effect is mostly dispersive in nature.

The r<sub>0</sub> experimental separation between the isopropyl carbon and the benzene center of mass (C···· $\pi$ ) is 3.495(11) Å. This is comparable to *ab initio* distances for simple CX<sub>3</sub>-H···benzene systems, which a previous dispersion-corrected DFT study (PBE-D/aug-cc-pVTZ) report ranging from 3.69 Å (X = H) to 3.39 Å (X = F).<sup>33</sup> Comparable results are also found for acetylene···benzene, where CCSD(T) values in the CBS limit calculates the C··· $\pi$  distance as 3.4 Å, a value consistent with experimental results.<sup>34,35</sup> The acetylene···benzene system is a special case compared to the halomethanes, since the acetylene hydrogen is quite acidic (pK<sub>a</sub> (gas) = 25)<sup>36</sup> with respect to other simple halogenated hydrocarbons, such as fluoroform (pK<sub>a</sub> (gas) = 30.5).<sup>37</sup> Both experimental<sup>34</sup> and theoretical<sup>38</sup> studies show that electrostatics are dominant in the noncovalent interaction due to the acidic nature of the acetylene hydrogen. A predicted  $pK_a(H_2O)$  value for sevoflurane derived from a cheminformatic model for medicinal chemistry<sup>39</sup> is around 15.7, so a similar energetic profile for its benzene cluster is possible.

The  $r_0$  distances for the C-H···F distances are consistent with theoretical calculations, ranging from 3.095(14) Å for the OCH<sub>2</sub>F····H<sub>1</sub> distance to 3.293(12) (short) and 3.531(15) Å (long) for the CF<sub>3</sub>···H interactions. Since these range anywhere from 0.5 to nearly 1 Å longer than the fluorine hydrogen bonding distances seen in the other conformers, it is likely that this interaction mode plays a small role in stabilizing the complex. This could indicate that these weak fluorine hydrogen bonds are a contributor to determining the relative orientation (staggered) of the benzene ring with respect to sevoflurane.

Application of the  $r_m^{(1)}$  model results in nearly statistically equivalent results to that of the  $r_0$  structure. Although the  $r_m^{(1)}$  model fits an additional correction to the ground state moments of inertia equal to  $c_\alpha(I_\alpha)^{1/2}$  for each of the fitted experimental moments of inertia, the values for the three  $c_\alpha$  constants along each of the principal axes are statistically equivalent, with values of  $c_a = 0.49(24)$ ,  $c_b = 0.66(19)$ ,  $c_c = 0.57(19)$  amu<sup>1/2</sup> Å. Due to these additional fitted parameters, the fluorine-related parameters that were fit in the  $r_0$  data had to be fixed to *ab initio* values in order to achieve even some statistical certainty on the  $c_\alpha$  parameters, which suggests that the sevoflurane···benzene structural data is relatively invariant to vibrational corrections in leading order  $I_\alpha^{1/2}$ . The near equivalency of the three  $c_\alpha$  coefficients suggest that the vibrational deviations from the  $r_0$ model are roughly isotropic in nature, similar to the assumption in Kraitchman's equations. Despite this, the deviations between the  $r_0$  and  $r_m^{(1)}$  determined parameters are small with respect to the standard error of the fit, so any real meaningful comparisons between the  $r_0$  and  $r_m^{(1)}$  structures cannot be made. Additionally, increasing the fit order to  $r_m^{(2)}$  was unsuccessful, as the three  $d_\alpha$  coefficients could not be fit to statistical certainty.

However, it is clear that dispersion still plays a role in determining the complexation geometry in sevoflurane…benzene, as the B3LYP determination of the C… $\pi$  length is much longer, at 3.9 Å. In addition, the B3LYP calculation predicts the ring as slightly tilted with respect to the C-H bond, with a tilt of 98°, whereas MP2 and M06-2X values, 92.3° and 91.9° respectively, agree with the r<sub>0</sub> value of 92.65(43)°. To make matters worse for the B3LYP results, the benzene ring is actually rotated away by 25° from eclipsing the 1- hydrogen with the fluoromethyl group as found in the experimental structure. Thus, optimal staggering of the benzene hydrogens about the sevoflurane frame in the global minimum structure is dependent on the proper treatment of the dispersion contribution to the C-H…  $\pi$  interaction.

This structural analysis is corroborated by the results of a SAPT0 calculation, which suggests a close balance between dispersion and electrostatics. The net SAPT0 electrostatic interaction energy is -8.83 kcal·mol<sup>-1</sup>, whereas the net dispersion energy is -11.2 kcal·mol<sup>-1</sup>. Adding in the net repulsive exchange energy (14.2 kcal·mol<sup>-1</sup>) and induction terms (-2.47 kcal·mol<sup>-1</sup>) results in a net SAPT0 interaction energy of -5.97 kcal·mol<sup>-1</sup>. This is in good agreement with the complete basis set (CBS)-extrapolated MP2/aug-cc-pVXZ complexation energy of -7.62 kcal·mol<sup>-1</sup> reported by Dom et al. The discrepencies between the SAPT0 and the MP2(CBS) results can likely be explained by the fact that the intermolecular interactions are only treated at the lowest order in SAPT0; higher order perturbation theory could resolve some of the discrepancies between these two values.

In the context of other benzene-containing complexes studied using SAPT techniques, sevoflurane-benzene shares similar energetics with acetylene-benzene, as predicted previously on the basis of C… $\pi$  distances discussed previously. A recent DFT-SAPT/aug-cc-pVDZ study of the acetylene-benzene complex reports an electrostatic component of -1.75 kcal·mol<sup>-1</sup> and a dispersion interaction of -2.2 kcal·mol<sup>-1</sup>. <sup>40</sup> Although this is a weaker interaction overall (total SAPT interaction energy of -2.3 kcal·mol<sup>-1</sup>) than seen sevoflurane-benzene, the relative proportions of electrostatics to dispersion are quite similar. This is in contrast to the archetypal dispersion-dominated benzene complexes, such as benzene dimer, which has electrostatic and dispersion interaction energies of -0.48 and -5.68 kcal·mol<sup>-1</sup>, respectively.<sup>41</sup>

## b) Internal Rotation

The internal rotation dynamics of the benzene on the  $C_1$  frame of sevoflurane is unusual when compared to other examples of six-fold internal rotation in the literature. Typically, the six-fold symmetry arises in previously studied systems via two common cases. The first is having a  $C_3$  top (typically a methyl group) on a two-fold symmetric frame. Classic examples of this include toluene <sup>42,43</sup> and nitromethane.<sup>44,45</sup> Here, the barriers are low due to the increased symmetry of the top with respect to typical highbarrier methyl internal rotation, so the spectra appear similar to that of a rigid symmetric or asymmetric rotor, with the torsional motion treated as a perturbation. The second common form of six-fold internal rotation has also been seen in simple benzene-containing clusters, such as those with  $SO_2^{-2}$ ,  $H_2O^{-3-6}$  and  $N_2$ .<sup>46</sup> Note that the six-fold internal rotation arises in these complexes from the planarity of the "top" species rotating about the C<sub>6</sub> frame of benzene. Since the energies of the torsional states are defined by the effective rotational constant of the top, which is large for such small species, these complexes often appear to have effective free rotation or a very low barrier. Additionally, since the interaction characteristics are mostly dispersive in nature, the actual barriers to internal rotation are small, cementing the spectroscopy into the low-barrier or free rotor limits. In these cases, the |J, K, m> basis is often appropriate where *m* specifies the torsional level of the rotor.<sup>47–49</sup>

In all of these cases, the coupling between the angular momenta of the top and the frame quickly becomes significant due to the low barrier and often requires high-order perturbative corrections in order to achieve quantitative fits of the experimental data. Additionally, the internal rotation axis of these previously mentioned systems are typically collinear or close to collinear with the overall rotation symmetry axis. This leads to non-negligible Coriolis-type coupling terms between the internal and overall angular momenta, which can additionally complicate the fitting of the spectrum and cause issues in convergence when determining a quantitative fit.<sup>50–52</sup>

The issues noted above are not applicable in the case of sevoflurane...benzene, as the six-fold internal rotation arises from the  $C_6$  symmetry of the benzene top about the  $C_1$ frame of sevoflurane. Here, the reduced rotational top of the top, F, should be close to the C constant of benzene (ca. 2850 MHz), though not exactly since the rotor axis is off-axis from the principal axes of the overall inertial frame. Assuming a harmonic rotor potential, the torsional energies are given by  $E = hv(v_t + \frac{1}{2})$ , where  $v = (36FV_6)^{1/2}$ . An MP2/6-311++g(d,p) calculation predicts the barrier height V<sub>6</sub> to be 48 cm<sup>-1</sup>. A quick calculation shows that the ground and first excited torsional states lie at 7 and 21 cm<sup>-1</sup> respectively, which are well below the barrier. Therefore, the analysis can be grounded firmly in the high barrier limit. In this case, it is expected the ground state rotational spectrum would be split into four components as observed in the experimental spectrum, with each components corresponding to one of the four irreducible representations of the C<sub>6</sub> point group (A, B, E<sub>1</sub> and E<sub>2</sub>). Additionally, due to the internal rotation axis being off-axis from the principal axes for overall rotation, the Coriolis-like coupling between overall and internal rotation should be small, so a perturbative fit of the internal rotation should converge easily in similar fashion to that of high-barrier methyl rotor calculation.

To properly fit this spectrum, the fitting program BELGI-C<sub>1</sub> is used. The effectiveness of BELGI for fitting internal rotation has been well established, especially for low barrier problems where principal axis methods converge poorly.<sup>52</sup> Unlike the previously mentioned approach in the principal axis frame, BELGI rotates the system into the *rho-axis*. This has the benefit of removing the Coriolis-type cross terms between overall rotation and internal rotation in all but the rho axis direction, which is close (but not exactly in the case of a near-prolate asymmetric top) to the benzene rotor axis. However, since the system is not in the principal axis frame, cross terms arise in the operators governing overall rotation, so cross terms such as the anticommutator  $\{P_a, P_b\}$  (and/or  $\{P_b, P_c\}$  or  $\{P_a, P_c\}$  for a C<sub>1</sub> symmetric complex) are required to fit the spectrum.

Since BELGI-C<sub>1</sub> is a program developed for C<sub>3</sub> rotors on a C<sub>1</sub> frame, a slightly modified approach is required to fit the C<sub>6</sub> internal rotation. the C<sub>6</sub> internal rotation can be decoupled into a problem solvable by BELGI as follows. Given a rigid rotor state  $J'_{Ka'Kc'}$ , the torsional symmetry is defined by an additional set of symmetry labels (v<sub>C3</sub>,  $\sigma$ ), where (0, 0) and (1,0) define the A and B states, or (0, ±1) and (1, ±1) define the E<sub>1</sub> and E<sub>2</sub> states, respectively. Therefore, the rotational transitions have the additional selection rules  $\Delta v_{C3} = \Delta \sigma = 0$ . Note that the label v<sub>C3</sub> only specifies the symmetry doublet within in a given single C<sub>6</sub> torsional state. Therefore, v<sub>C3</sub> is binary in that it can only take values modulo 2; a v<sub>C3</sub> = 2 would imply that the system is in the v<sub>C3</sub> = 0 state of the first excited six-fold torsional state, where v<sub>t</sub> = 1.

This construction allows us to treat the six-fold rotation as a two-state problem with C<sub>3</sub>-like symmetry. This is allowed mathematically since C<sub>6</sub> is isomorphic to C<sub>2</sub>  $\oplus$ C<sub>3</sub>. This is easy to see, as multiplication of the A and E irreducible representations of C<sub>3</sub> by the representations A and B of C<sub>2</sub> will give four new irreducible representations (A **x** A ; E **x** A) = (A ; E<sub>1</sub>) and (A **x** B ; E **x** B) = (B ; E<sub>2</sub>) in the product point group, which are identical to those of the C<sub>6</sub> point group.

With the assumption that  $V_3 = 0$ , the splittings can be naively determined by fitting two classes of constants. One is the constant that determines the splittings between the components of the A/E<sub>1</sub> or B/E<sub>2</sub> doublets, which in the high-barrier case is V<sub>6</sub>, the barrier height. Since in this approach the splittings between the two C<sub>3</sub>-like doublets are tied to a change in torsional excitation in the C<sub>3</sub> basis, the splittings can be determined to first-order by fitting the torsion-rotation constants. One is  $\rho$ , which connects energy levels in the torsional manifold with  $\Delta K = 0$ . The other are the two anticommutators, D<sub>ab</sub> and  $D_{ac}$ , which connect energy levels with  $\Delta K = 1$ . By fitting these three constants, one can model the experimental splittings to first order. Thankfully, since the spectrum is grounded firmly in the high barrier case, it is not necessary to measure the actual torsional excited states ( $v_t > 0$ ) in order to determine the barrier, so determination of the energies of the higher torsional states is trivial within the uncertainty generated by the assumption that the actual torsional potential energy surface is defined only by  $V_6$ . A visual summary of this internal rotation scheme can be found in Figure 4, where the experimentally-determined splittings between torsional levels and symmetry states for the J = 0 state are shown.

Since the top is non-covalently (and weakly) bound to the frame, an additional variety of terms that account for nonrigidity in both the overall and internal rotation are required in order to fit the spectrum to better than experimental uncertainty. Accounting for this nonrigidity, an excellent fit using 14 floated parameters was achieved with an RMS error of 6.3 kHz for the A/B states and 9.7 kHz for the E<sub>1</sub> and E<sub>2</sub> states. The V<sub>6</sub> barrier is well determined at 32.8688(27) cm<sup>-1</sup>. The direction cosines between the *a* principal axis and the direction of the benzene top are in good agreement with the theoretical of Conformer 1, which lends additional spectroscopic confirmation that Conformer 1 is the correct experimental structure. A full listing of the fit parameters can be found in Table 3.

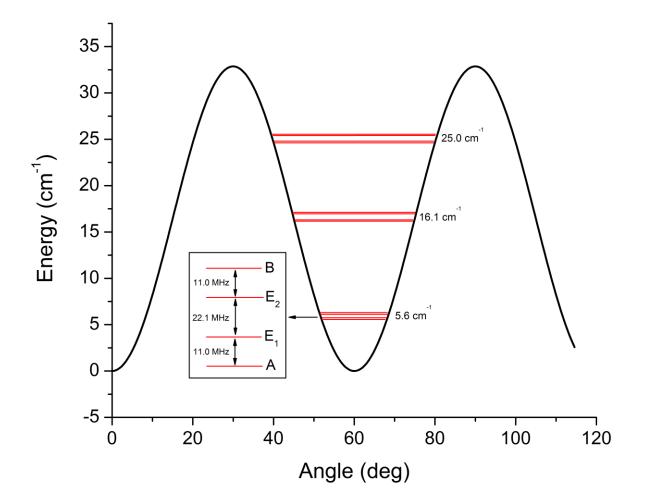


Figure 4. A schematic of the six-fold potential energy surface with a barrier of 32.8687(27) cm<sup>-1</sup>. Additionally, the experimentally determined symmetry splittings for the  $|v_t, J, K\rangle = |0, 0, 0\rangle$  state are also listed (small box, bottom-left).

Parameter		Operator	Value
T.1.1	Systematic		
Label used in Program	Label	D 2	500 604(1.5)
OA (MHz)	A "	$P_a^2$	509.684(1.5)
$PAM \ value^5$		$J_a^2$	515.778(3.1)
B (MHz)	В	$P_b^2$	365.02(12)
PAM value	"	$J_b^2$	361.05(16)
C (MHz)	С	$P_c^2$	338.242(84)
PAM value	"	$J_c^2$	336.11(16)
$V_{6}$ (cm <sup>-1</sup> )	$V_6$	$(1 - \cos(6\alpha))/2$	32.8688(27)
RHORHO (unitless)	ρ	$P_{\alpha}P_{a}$	0.170557(54)
F (MHz)	F	$P_{\alpha}^{2}$	[3452.484]
DELTA (MHz)	$F_{ab}$	$\mathbf{P}_{\alpha}^{2}\{\mathbf{P}_{a},\mathbf{P}_{b}\}^{6}$	-0.0306(44)
DAB (MHz)	$D_{ab}$	$\{\mathbf{P}_{a},\mathbf{P}_{b}\}$	25.40(24)
DACI (MHz)	$D_{ac}$	$\{\mathbf{P}_{a},\mathbf{P}_{c}\}$	17.945(81)
DJ (kHz)	$\Delta_{ m J}$	$-\mathbf{P}^4$	0.0117(23)
DJK (kHz)	$\Delta_{ m JK}$	$-\mathbf{P}^{2}\mathbf{P}_{a}^{2}$	0.648(85)
DK (kHz)	$\Delta_{\mathrm{K}}$	$-\mathbf{P}_{a}^{4}$	-7.85(81)
C11 (MHz)	V <sub>6bc</sub>	$(1-\cos(6\alpha))(P_b^2-P_c^2)$	-13.45(40)
BK2 (MHz)	$V_{6K}$	$(1-\cos(6\alpha))(P_a^2)$	-13.2(48)
AK2 (MHz)	$F_K$	$P_{\alpha}^{2}P_{a}^{2}$	-0.234(28)
N (A/B states) <sup>7</sup>			178
σ <sub>AB</sub> (kHz) <sup>8</sup>			6.3
N (E states)			123
σe (kHz)			9.7
N (total)			301
σ (kHz)			7.7
Direction Cosines		Experiment	M06-2X
$\cos(\lambda_a)^9$		0.961	0.951
$\cos(\lambda_b)$		-0.214	-0.230
$\cos(\lambda_c)$		0.176	0.208

Table 3. Fitted rotational constants and internal rotation parameters for the BELGI-C1 fit of the parent species.

<sup>&</sup>lt;sup>5</sup> Value for A, B, and C rotated into the principal axis after diagonalizing the moment of inertia tensor formed by the diagonal A, B and C rotational constants and the off-diagonal elements Dab and Daci in the rho-axis system.

<sup>&</sup>lt;sup>6</sup> The anticommutator  $\{A,B\}$  is defined as  $\{A,B\} = AB + BA$ <sup>7</sup> Number of assigned transitions for the given symmetry states

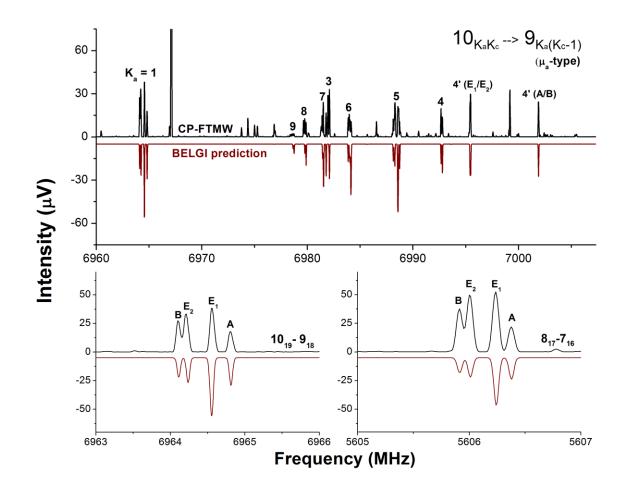
<sup>&</sup>lt;sup>8</sup> RMS average of all (obs-calc) frequency errors for the stated set of transitions

<sup>&</sup>lt;sup>9</sup> Angle between the unit vector in the **a** direction in the principal axis frame and the internal rotor axis. Ab*initio* values are derived from the M06-2X/6-311++g(d,p) geometry.

Additionally, due to spin statistics, the relative intensities of each symmetry component are not as simple as would be expected for a two-state C<sub>3</sub> fit, where the intensities between  $\sigma = 0$  and  $\sigma = \pm 1$  are 1:1, so appropriate relative intensities in the experimental spectrum would additionally verify the high-barrier case postulated here. The spin statistics for the C<sub>6</sub> point group are 10 A : 14 B : 22 E<sub>1</sub> : 18 E<sub>2</sub>.<sup>65</sup> These are consistent with the general relative intensity trends seen in the experimental spectrum and are within the expected variance caused by experimental inhomogeneities in the CP-FTMW experiment. Figure 5 shows an overview of the spectrum complicated by the sixfold internal rotation, with comparisons to the fit parameters and predicted relative intensities calculated by BELGI.

# 5) Conclusion

The rotational spectra of sevoflurane...benzene, as well as 12 single and 33 double isotopologues, have been assigned through the use of CP-FTMW spectroscopy in the 2-8 GHz band. The spectrum of the parent isotopic species is complicated by high-barrier, six-fold internal rotation of the benzene around the sevoflurane frame. All four symmetry states of this species were fit simultaneously with BELGI-C<sub>1</sub> to an excellent overall RMS error of 7.7 kHz.



**Figure 5.** Part of a single  $\mu_a$ -type progression for J' = 10 for the sevoflurane…benzene parent species spectrum (top panel). The predictions shown are generated with BELGI, using the internal rotation fitting process described in this study. The bottom two panes

show characteristic transitions with fully resolved internal rotation splittings.

Due to the abundance of experimental isotopic information, a full heavy atom (plus benzene hydrogens) r0 structure was determined in excellent agreement with the ab initio global minimum structure found from the initial molecular mechanics conformer screening. In the case of the observed conformer, the experimentally observed global minimum structure of the sevoflurane monomer is preserved upon complexation. Like the primary conformer in the previous vibrational study, the structure has a C-H··· $\pi$  interaction between the benzene ring and the relatively Lewis acidic isopropyl hydrogen on sevoflurane as its primary complexation mode. The r<sub>0</sub> C··· $\pi$  distance is 3.495(11) Å, which is comparable to other C-H··· $\pi$  complexes with relatively acidic donor hydrogens, such as acetylene···benzene. The next lowest energy conformer ( $\Delta E_{MP2} = 1130$  cm<sup>-1</sup>), where the complexation occurs with the fluoromethyl hydrogens, is not observed in the microwave spectrum.

New advances in high-speed electronics have enabled significant improvements to broadband rotational spectroscopy techniques that have led to unparalleled levels of sensitivity for molecules and molecular clusters of increasingly large sizes. With the ability to structurally probe increasingly larger and more complex molecular clusters, the use of a theoretical methodology with a balance between accuracy and computational cost and portability is crucial for efficient and accurate spectroscopic predictions. Therefore, structural studies for complexes such as sevoflurane…benzene, where the interplay between electrostatics and dispersion are crucial for determining the complexation geometry, are essential in benchmarking and determining an appropriate theoretical procedure for efficient evaluation and prediction for broadband rotational spectroscopy.

## References

(1) Dom, J. J.; van der Veken, B. J.; Michielsen, B.; Jacobs, S.; Xue, Z.; Hesse, S.; Loritz, H.-M.; Suhm, M. A.; Herrebout, W. A. *Phys. Chem. Chem. Phys. PCCP* **2011**, *13*, 14142.

(2) Taleb-Bendiab, A.; Ii, K. W. H.; Kuczkowski, R. L. J. Chem. Phys. **1992**, 97, 2996.

(3) Gotch, A. J.; Garrett, A. W.; Severance, D. L.; Zwier, T. S. *Chem. Phys. Lett.* **1991**, *178*, 121.

(4) Gutowsky, H. S.; Emilsson, T.; Arunan, E. J. Chem. Phys. 1993, 99, 4883.

(5) Emilsson, T.; Gutowsky, H. S.; Oliveira, G. de; Dykstra, C. E. J. Chem. Phys. **2000**, *112*, 1287.

(6) Suzuki, S.; Green, P. G.; Bumgarner, R. E.; Dasgupta, S.; Goddard, W. A.; Blake, G. A. *Science* **1992**, *257*, 942.

(7) Rodham, D. A.; Suzuki, S.; Suenram, R. D.; Lovas, F. J.; Dasgupta, S.; Goddard, W. A.; Blake, G. A. *Nature* **1993**, *362*, 735.

(8) Lin, C. C.; Swalen, J. D. *Rev. Mod. Phys.* **1959**, *31*, 841.

(9) Kleiner, I.; Hougen, J. T. J. Chem. Phys. 2003, 119, 5505.

(10) Lavrich, R. J.; Walker, A. R. H.; Plusquellic, D. F.; Kleiner, I.; Suenram, R. D.; Hougen, J. T.; Fraser, G. T. *J. Chem. Phys.* **2003**, *119*, 5497.

(11) Neill, J. L.; Shipman, S. T.; Alvarez-Valtierra, L.; Lesarri, A.; Kisiel, Z.; Pate, B. H. *J. Mol. Spectrosc.* **2011**, *269*, 21.

(12) Brown, G. G.; Dian, B. C.; Douglass, K. O.; Geyer, S. M.; Shipman, S. T.; Pate, B. H. *Rev. Sci. Instrum.* **2008**, *79*, 053103.

(13) *Maestro, version 9.2; Macromodel, version 9.5.207*; Schrödinger, Inc.: New York, 2010.

(14) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. Chem. Phys. Lett. 1988, 153, 503.

(15) Peverati, R.; Truhlar, D. G. *ArXiv12120944 Cond-Mat Physicsphysics Physicsquant-Ph* **2012**.

(16) Boys, S. F.; Bernardi, F. Mol. Phys. 1970, 19, 553.

(17) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.;
Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji,
H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.;
Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida,

M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.;
Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.;
Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.;
Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.;
Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.;
Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.;
Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.;
Fox, D. J. *Gaussian 09, Revision A.1*; Gaussian, Inc.: Wallingford, CT, 2009.

(18) Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. A.; Fermann, J. T.; Mintz, B. J.; Burns, L. A.; Wilke, J. J.; Abrams, M. L.; Russ, N. J.; Leininger, M. L.; Janssen, C. L.; Seidl, E. T.; Allen, W. D.; Schaefer, H. F.; King, R. A.; Valeev, E. F.; Sherrill, C. D.; Crawford, T. D. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 556.

(19) Burns, L. A.; Mayagoitia, Á. V.-; Sumpter, B. G.; Sherrill, C. D. J. Chem. Phys. **2011**, *134*, 084107.

(20) Hohenstein, E. G.; Parrish, R. M.; Sherrill, C. D.; Turney, J. M.; Iii, H. F. S. J. *Chem. Phys.* **2011**, *135*, 174107.

(21) Ruoff, R. S.; Klots, T. D.; Emilsson, T.; Gutowsky, H. S. J. Chem. Phys. **1990**, 93, 3142.

(22) Cocinero, E. J.; Lesarri, A.; Écija, P.; Basterretxea, F. J.; Grabow, J.-U.; Fernández, J. A.; Castaño, F. *Angew. Chem. Int. Ed.* **2012**, *51*, 3119.

(23) Alonso, J. L.; Pérez, C.; Sanz, M. E.; López, J. C.; Blanco, S. *Phys. Chem. Chem. Phys.* **2009**, *11*, 617.

(24) Shipman, S. T.; Neill, J. L.; Suenram, R. D.; Muckle, M. T.; Pate, B. H. J. Phys. Chem. Lett. **2011**, *2*, 443.

(25) Watson, J. K. G. J. Chem. Phys. 1967, 46, 1935.

(26) Watson, J. K. G. Mol. Phys. 1968, 15, 479.

(27) Watson, J. K. G. In *Handbook of High-resolution Spectroscopy*; John Wiley & Sons, Ltd, 2011.

(28) Lesarri, A.; Vega-Toribio, A.; Suenram, R. D.; Brugh, D. J.; Grabow, J.-U. *Phys. Chem. Chem. Phys. PCCP* **2010**, *12*, 9624.

(29) Kisiel, Z. J. Mol. Spectrosc. 2003, 218, 58.

(30) Jureka, P.; Sponer, J.; Cerny, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985.

(31) Tsuzuki, S.; Fujii, A. Phys. Chem. Chem. Phys. 2008, 10, 2584.

(32) Klimeš, J.; Michaelides, A. J. Chem. Phys. 2012, 137, 120901.

(33) Shibasaki, K.; Fujii, A.; Mikami, N.; Tsuzuki, S. J. Phys. Chem. A **2007**, 111, 753.

(34) Shibasaki, K.; Fujii, A.; Mikami, N.; Tsuzuki, S. J. Phys. Chem. A **2006**, *110*, 4397.

(35) Dey, R. C.; Seal, P.; Chakrabarti, S. J. Phys. Chem. A 2009, 113, 10113.

(36) Fujii, A.; Morita, S.; Miyazaki, M.; Ebata, T.; Mikami, N. J. Phys. Chem. A **2004**, *108*, 2652.

(37) Bégué, J.-P.; Bonnet-Delpon, D. *Bioorganic and Medicinal Chemistry of Fluorine*; John Wiley & Sons, 2008.

(38) Nishio, M.; Hirota, M.; Umezawa, Y. *The CH/\pi Interaction: Evidence, Nature, and Consequences*; John Wiley & Sons, 1998.

(39) Csizmadia, F.; Tsantili-Kakoulidou, A.; Panderi, I.; Darvas, F. J. Pharm. Sci. **1997**, *86*, 865.

(40) Tekin, A.; Jansen, G. Phys. Chem. Chem. Phys. 2007, 9, 1680.

(41) Sherrill, C. D. Acc. Chem. Res. 2013, 46, 1020.

(42) Rudolph, P. H.; Dreizler, H.; Jaeschke, A.; Wendling, P. Z. Naturforsch **1967**, 22a, 940.

(43) Ilyushin, V. V.; Kisiel, Z.; Pszczółkowski, L.; Mäder, H.; Hougen, J. T. *J. Mol. Spectrosc.* **2010**, *259*, 26.

(44) Tannenbaum, E.; Johnson, R. D.; Myers, R. J.; Gwinn, W. D. J. Chem. Phys. **1954**, 22, 949.

(45) Rohart, F. J. Mol. Spectrosc. 1975, 57, 301.

(46) Ohshima, Y.; Kohguchi, H.; Endo, Y. Chem. Phys. Lett. 1991, 184, 21.

(47) Fraser, G. T.; Lovas, F. J.; Suenram, R. D.; Jr, D. D. N.; Klemperer, W. J. Chem. *Phys.* **1986**, *84*, 5983.

(48) Kroto, H. W. *Molecular Rotation Spectra*; Dover Publications, 2003.

(49) Gordy, W.; Cook, R. L. Microwave molecular spectra; Wiley, 1984.

(50) Laurie, V. W. J. Chem. Phys. 1959, 31, 1500.

(51) Herschbach, D. R. J. Chem. Phys. 1959, 31, 91.

(52) Kleiner, I. J. Mol. Spectrosc. **2010**, 260, 1.

# **Chapter 5: The Structure of the Sevoflurane Dimer**

(A portion of the writing has been adapted from a manuscript currently in the peer-review and submission process, with citation: N. A. Seifert, C. Perez, J. L. Neill, B. H. Pate, M. Vallejo-López, A. Lesarri, E. J. Cocinero, F. Castaño, *Phys. Chem. Chem. Phys., accepted.* Explicit permission has been granted from the other principal co-author, A. Lesarri, for adaption of this manuscript)

## 1) Introduction

The results presented in Chapter 4 on the cluster of sevoflurane and benzene show the amenability of the CP-FTMW technique for detecting clusters of large molecules with structural determination capabilities unparalleled by other gas-phase techniques. With 18 heavy atoms, sevoflurane...benzene is one of the largest non-covalent complexes assigned by microwave spectroscopy, and certainly the only with a spectral data set as large as 45 isotopologues. However, the available sensitivity and bandwidth of the UVa CP-FTMW instrument can also cause significant spectral congestion due to multiple interloping species, including other complexes and sample impurities, so detection of species via visual inspection can often be difficult.<sup>1</sup> The major bottleneck in discovery of unassigned complexes is generating candidate structures via computational chemistry, which requires a thorough and systematic sampling of interaction topologies through molecular dynamics (MD) as well as high levels of theory to optimize the minima found in the MD search.<sup>2,3</sup> Modeling of these large complexes require accurate treatment of multiple competing forms of intermolecular interactions, including moderate/weak hydrogen bonding<sup>4,5</sup> and electrostatic/dispersive forces<sup>6–8</sup>.

One such cluster found during the study of sevoflurane...benzene via visual detection of distinct "clumps" of  $\mu_a$ -type transitions and subsequent theoretical structure

searches is the cluster of sevoflurane with itself, a 24 heavy atom complex. The sevoflurane dimer appears to be largely stabilized by a variety of C-H…X interactions (X = O, F) similar to the cluster with benzene. Despite sevoflurane having no chiral centers, hindered internal rotation of the terminal fluoromethyl group gives rise to transient axial chirality.<sup>9,10</sup> As such, formation of the sevoflurane dimer results in two diastereomers of homo- (e.g. RR or SS) or hetero- (RS or SR) chirality. The IR spectrum of sevoflurane had previously shown blue-shifted vibrational bands tentatively assigned to sevoflurane oligomers, but no structural information on these species was available and stereoisomerism was ignored.<sup>11</sup> This study therefore constitutes the first detailed spectroscopic study of the sevoflurane dimer.

Previous rotational studies have examined the formation of some small chiral intermolecular clusters between neutral molecules, mostly alcohol dimers<sup>12–15</sup> or alcohol clusters involving other chiral or achiral partners.<sup>16–18</sup> These studies have proved useful to observe the formation of homo- and heterochiral diastereoisomers, the conformational preferences, the nature of the chiral discriminating forces (mainly conventional O-H…O hydrogen bonding in aliphatic alcohols), and the performance of alternative techniques and ab initio calculations.

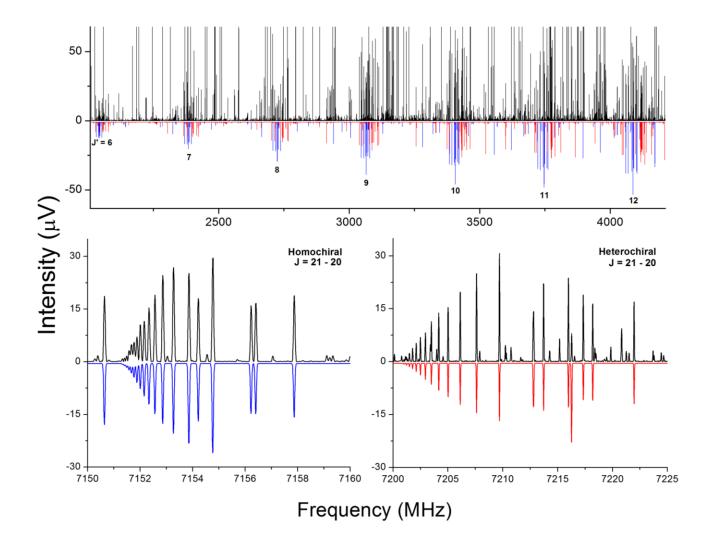
#### 2) Experimental & Theoretical

The chirped-pulse Fourier transform microwave (CP-FTMW) spectrum of sevoflurane dimer was performed using the 2-8 GHz CP-FTMW spectrometer at the University of Virginia. Details of this instrumental setup in the context of structure determination have been presented previously in Chapter 1, Section IV so only details specific to this experiment are presented here.

The sample was generated by mixing sevoflurane vapor (brand name Ultane, Abbott Laboratories) with ca. 6 atm of neon (GTS Welco) for an approximate concentration of 0.2%. The spectrum was acquired by pulsing the sample gas with a backing pressure of ca. 1 atm through 5 pulsed nozzles arranged perpendicularly to the polarization pulse, at a repetition rate of 3.3 Hz. During each valve injection cycle, 4 free induction decays (FIDs) of the jet-cooled sampled are measured at a total record length of 40  $\mu$ s, which are coherently averaged in the time-domain. The final spectrum consists of an averaged spectrum of approximately 9.1 million FIDs. All isotopologues measured in this experiment were observed in natural abundance. A portion of this 2-8 GHz spectrum is shown in Figure 1.

The 2-8 GHz spectrum was comprised of ca. 9,600 rotational transitions with signal-to-noise ratios over 4:1 (1.6 MHz<sup>-1</sup>). The experiment was assisted by different theoretical calculations for the prediction of the dimer geometries and molecular properties, which combined a molecular mechanics conformational search with high-level *ab initio* reoptimization and vibrational frequency calculations, as in sevoflurane····benzene.

The initial conformational screening of the sevoflurane dimer used the Merck Molecular Force Field (MMFFs) combined with Monte-Carlo and Large-Scale/Low-Modes search algorithms, as implemented in Macromodel.<sup>19</sup> Theoretical predictions for the structure of both detected diasteromers were calculated with GAUSSIAN 09<sup>20</sup> using a variety of methods. Initial optimizations were performed using MP2<sup>21</sup> and M06-2X<sup>22</sup> with the Pople 6-311++g(d,p) basis set.<sup>23</sup> Binding energies were corrected for basis set superposition error (BSSE) by using Boys and Bernardi's counterpoise correction.<sup>24</sup> Further reoptimizations were performed using B3LYP<sup>25</sup> with Grimme's –D3 empirical dispersion corrections<sup>26</sup> using the 6-311++g(d,p) basis set. Symmetry-adapted perturbation theory (SAPT) calculations of the intermolecular interaction energies were performed using the zeroth-order SAPT (SAPT0)<sup>27</sup> approximation with a truncated *augcc-pVDZ (jun-cc-pVDZ)* basis set.<sup>28</sup> A detailed explanation of the SAPT0 process can be found in Chapter 1, Section III. The B3LYP-D3 and SAPT0 calculations were both performed using the Psi4 (beta 5) quantum chemistry suite.<sup>29</sup>



**Figure 1**. The broadband MW spectrum of the sevoflurane dimer. The top trace shows a 2 GHz spectral section showing the J=6 to J=12 series transitions for the dimer. The bottom panel illustrates typical  $\mu_a$ -type progressions for the two homo and heterochiral species. The experimental spectrum (black) is compared with simulations (blue, red) using the fit rotational parameters. The dynamic range of the top spectrum was truncated to show the intensity region relevant to the dimer.

#### 3) Results & Discussion

Using the initial molecular modeling results, two different species were subsequently detected in the spectrum and identified as the homochiral and heterochiral species of the sevoflurane dimer. The detection of a very large set of experimental transitions (700-1000) for each cluster assured the internal consistency of a Watson's semirigid-rotor Hamiltonian<sup>30,31</sup> fit and led to accurate determination of the rotational constants and some quartic centrifugal distortion parameters. The experimental results are compared with the theoretical predictions in Table 1. The conformational assignment was unequivocally confirmed by the detection of eighteen additional isotopologues in natural abundance. All possible <sup>13</sup>C species (ca. 1%) were detected for both homo- and heterochiral clusters, while the two weaker <sup>18</sup>O isotopologues (ca. 0.2%) could be observed only for the homochiral form.

The abundance of isotopic data allowed an accurate determination of the cluster structure using both substitution  $(r_s)^{32,33}$  and effective  $(r_0)$  methods.<sup>34</sup> The experimental (vibrationally-averaged) structures are not directly comparable to the equilibrium *ab initio* structures, but the differences are often within the uncertainty of the derived structural parameters. In the sevoflurane dimer the substitution coordinates are in excellent agreement with the MP2 *ab initio* structure, with a total (rms) coordinate deviation of 0.067 Å and 0.183 Å, respectively, for the hetero- and homochiral species.

		Homochiral			
	Experimental		MP2	M06-2X	B3LYP-D3
	CP-FTMW			6-311++g(d,p)	
A (MHz) <sup>[a]</sup>	307.789308(39) <sup>[b]</sup>	A (MHz) <sup>[a]</sup>	308.11 (0.15%) <sup>[c]</sup>	314.05 (2.0%)	304.78 (0.99%)
В	172.119904(23)	В	181.76 (3.1%)	185.89 (4.5%)	176.89 (2.7%)
С	168.437022(23)	С	172.83 (1.4%)	182.49 (4.6%)	173.07 (2.7%)
⊿, (kHz)	0.011420(29)	⊿, (kHz)	0.00616	0.0224	0.0185 <sup>[d]</sup>
$\Delta_{JK}$	0.029193(43)	$\Delta_{JK}$	0.0185	0.0340	0.0220
$\Delta_{\kappa}$	-0.035955(67)	$\Delta_{\kappa}$	-0.0153	-0.0283	-0.0217
$(\mu_a/\mu_b)^2$	16.	$(\mu_a/\mu_b)^2$	11.	2.86	2.92
$(\mu_a/\mu_c)^2$	0.3	$(\mu_a/\mu_c)^2$	0.33	0.35	0.34
		$\mu_a$ (D)	-1.86	-1.86	-1.71
$N_{\text{lines}}$	1051	$\mu_b$	0.55	1.10	1.00
RMS (kHz)	8.19	$\mu_c$	3.26	3.15	2.95
		$\mu_{total}$	3.79	3.82	3.55
		Heterochiral			
	Experimental		MP2	M06-2X	B3LYP-D3
	CP-FTMW			6-311++g(d,p)	
A (MHz)	304.70027(32)	A (MHz) <sup>[a]</sup>	306.56 (0.59%)	309.30 (1.5%)	301.15 (1.2%)
В	175.56391(12)	В	183.87 (4.7%)	188.37 (7.3%)	179.12 (2.0%)
С	167.25459(13)	С	174.85 (4.5%)	177.53 (6.1%)	171.40 (2.4%)
⊿,(kHz)	0.00850(24)	⊿, (kHz)	0.00548	0.0110	0.0122
$\Delta_{JK}$	0.04428(62)	$\Delta_{JK}$	0.0271	0.0429	0.0794
$\Delta_{\kappa}$	0.0487(13)	$\Delta_{\kappa}$	0.0247	0.0349	0.0712
$(\mu_a/\mu_b)^2$	3.2	$(\mu_a/\mu_b)^2$	2.2	3.4	1.5
$(\mu_a/\mu_c)^2$		$(\mu_a/\mu_c)^2$	> 100	> 100	41.
		$\mu_{a}$ (D)	2.03	2.53	1.67
N <sub>lines</sub>	726	$\mu_b$	1.36	1.37	1.37
RMS (kHz)	7.25	$\mu_c$	0.02	-0.14	-0.26
		$\mu_{total}$	2.44	2.88	2.17

**Table 1.** Rotational parameters for the homo- and heterochiral diastereoisomers

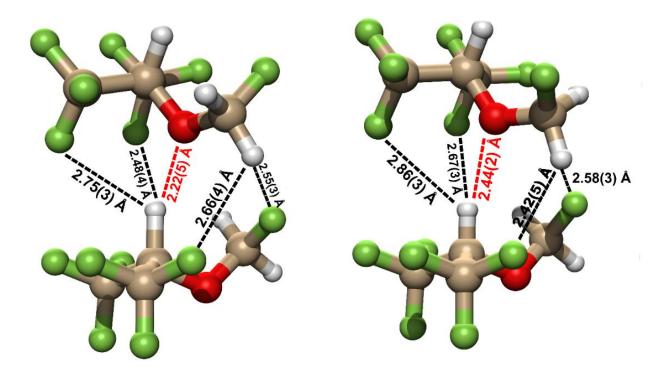
 of the sevoflurane dimer, and comparison with the theoretical predictions.

The larger deviations for the homochiral dimer are due to imaginary coordinates corresponding to near coincidental positions of several atoms and the inertial axes, a well-known issue of the Kraitchman structural method. Figure 2 shows the effective structures of both diastereoisomers and the most relevant intermolecular parameters calculated by a least-squares fit to the moments of inertia,<sup>34</sup> and Table 2 summarizes the essential intermolecular parameters with comparison to the theoretical structures. In this fit all bond parameters containing purely carbon atoms (and oxygen for the homochiral dimer) have been floated, including an intermolecular connectivity between neighboring perfluoro carbons in each of the sevoflurane subunits. All other internal parameters were constrained to the MP2/6-311++g(d,p) structure. A comparison of the effective structure parameters of the two subunits of each dimer to that of bare sevoflurane shows an excellent agreement (rms error of 0.025 Å), which confirms the usual assumption that the monomer geometry is not distorted upon complexation through moderate/weak hydrogen bond forces.

Special attention was paid to the calculation of spectral intensities in order to estimate conformational abundances and, eventually, a preference for one of the two homo- or heterochiral diastereoisomers. The rotational temperature in the jet was first determined by a least-squares fit minimizing the residuals of the intensity ratio between the experimental and predicted values, which resulted in a rotational temperature of 1.32(24) K (including transitions from both disastereoisomers).

	Homochiral			Heterochiral				
Intermolecular		MP2	M06-2X	B3LYP-D3		MP2	M06-2X	B3LYP-D3
distances / Å	<i>r</i> <sub>0</sub>	6-311++g(d,p)			<i>r</i> <sub>0</sub>	6-311++g(d,p)		
[syn-perfluoro F] ···· H	2.86(3)	2.70	2.48	2.69	2.48(4)	2.50	2.40	2.54
[anti-perfluoro F] ···· H	2.67(3)	2.49	2.37	2.56	2.75(3)	2.74	2.54	2.86
О … Н	2.44(2)	2.23	2.28	2.28	2.22(5)	2.21	2.20	2.25
[fluoromethoxy H] [fluoromethoxy F]	2.58(3)	2.64	2.71	2.62	2.55(3)	2.48	2.35	2.40
[fluoromethoxy H] ···· [syn-perfluoro F]	2.42(5)	2.54	2.38	2.50	2.66(4)	2.63	2.57	2.59

**Table 2**. Selected intermolecular parameters of the observed sevoflurane dimers, with comparison to theoretical structures. The parameters used here correlate with those illustrated in Figure 2, and the notation for the linkages are [acceptor atom] ··· [donor atom].



**Figure 2.** The effective ( $r_0$ ) experimental structures derived for the heterochiral (left) and homochiral (right) species of the sevoflurane dimer.

Since the intensity profile of the rotational spectrum increases with the square of the electric dipole moment, errors in the relative proportions of the theoretical dipole components may cause some uncertainties. To mitigate this, only the common  $\mu_a$ -type transitions were used, including all assigned transitions with  $K_a \leq 20$ . The experimental intensity ratio between the homochiral and the heterochiral dimers was then estimated as 1.3 : 1. Assuming the difference in entropy is negligible between the two dimers, this value is in fair agreement with the MP2 counterpoise-corrected energy difference of 22.9 cm<sup>-1</sup>, which corresponds to a statistical mixture of 1.12 : 1 at 298 K.

The structural data obtained for the sevoflurane dimer confirms that the cluster is held together by a combination of weak hydrogen bonds, with both C-H···F and C-H···O intermolecular linkages. The primary hydrogen bond operates through the isopropyl C-H bond of one of the monomer subunits, acting as proton donor to the oxygen atom in the second monomer. Comparable C-H···F interactions are also found in between the hydrogen in the acceptor's fluoromethoxy group that points closest to the donor fluorines. The  $r_0$ -determined C-H···F linkages range from 2.42(3) Å to 2.86(3) Å, which fall on both sides of the average C-H···F distance of 2.6(1) Å detected in crystallographic studies.<sup>35</sup> The  $\angle$ CH···F bonding angles range from ca. 120° in the acceptor fluoromethoxy linkages, to 143° in those containing the donor isopropyl hydrogen. The average angle seen in crystallographic studies falls again somewhere in between (135°). Comparatively, the C-H···O bonding distance and angle seen in both the hetero and homochiral dimers exhibit values of 2.22(5) Å / 146° and 2.44(2) Å / 151(1)°, respectively (vs. 2.401(16) Å in sevoflurane···benzene). These values fall in line with the typical C-H···O interaction in crystallographic studies, with typical mean interaction distances of 2.4 Å and angles of  $140^{\circ}$ .<sup>35</sup>

Since the only significant contributor to an energetic difference between the diastereomers would be found in the intermolecular binding potential, a zeroth-order symmetry adapted perturbation theory (SAPT) calculation was performed to decompose the energetic contributions to the intermolecular binding forces. At this level of theory, the total SAPT0 binding energy difference is 6.8 cm<sup>-1</sup>. Recent benchmark studies of non-covalent interactions suggest that the mean absolute error (MAE) on counterpoise-corrected (cp) binding energy calculations using a Dunning triple- $\zeta$  (aug-cc-pVTZ) basis set is 0.126 kcal/mol (44 cm<sup>-1</sup>),<sup>36</sup> which acts as a lower bound for the determination of the counterpoise-corrected calculations. In this study older generation triple- $\zeta$  Pople basis sets were used to reduce the computational costs. However, assuming similar benchmark errors, the SAPT(0)/jun-cc-pVDZ binding energy value of 6.8 cm<sup>-1</sup>.

The intermolecular bonding characteristics in the sevoflurane dimers have been compared quantitatively with related sevoflurane clusters using the SAPT(0) energy decomposition, with the results tabulated in Table 3. Comparison of the SAPT(0) results for both dimers with those of the sevoflurane...benzene complex reveal surprisingly similar characteristics between both types of complexes. In particular, the binding energies are very similar (-5.97 kcal mol<sup>-1</sup> for the benzene-containing complex).

SAPT(0)/jun-cc-pVDZ (kcal mol <sup>-1</sup> )								
Complex	$\Delta E_{elst}$	$\Delta E_{ind}$	$\Delta E_{exch}$	$\Delta E_{disp}$	$\Delta E_{tot}$			
(sevo) <sub>2</sub> , homochiral	-7.65	-1.76	9.22	-5.86	-6.05			
(sevo) <sub>2</sub> , heterochiral	-7.52	-1.80	8.96	-5.70	-6.06			
sevoflurane…benzene	-8.83	-2.50	14.20	-8.84	-5.97			
sevoflurane…H <sub>2</sub> O	-10.42	-2.03	7.61	-2.37	-7.21			
$(H_2O)_2$	-8.84	-2.15	7.03	-1.27	-5.22			
benzeneacetylene	-2.90	-1.06	4.34	-3.32	-2.95			
benzene…methane	-0.99	-0.30	2.46	-2.24	-1.07			

**Table 3.** Energy decompositions (SAPT(0)/jun-cc-pVDZ) for a selection of sevoflurane clusters, together with three benchmark examples for electro-static-dominated (water dimer), dispersion-dominated (benzene...methane) and mixed (benzene...acetylene) intermolecular interactions.

Like sevoflurane...benzene, the sevoflurane dimer exhibits an intermediate mix of electrostatic and dispersive interactions, consistent with the bond distance/angle analysis presented previously. In this regime, the intermolecular binding between sevoflurane monomers falls between the energetic characteristics of a classical hydrogen bond such as the water dimer and that of a dispersion-dominated interaction such as that in benzene...methane. The sevoflurane...H<sub>2</sub>O complex has a different C-H...O interaction motif than the sevoflurane dimer. The water-containing complex has an intermolecular binding that is largely electrostatic in character and the C-H...O linkage is much shorter (2.13 Å) and more linear (162°), as reflected in the dominant  $\Delta E_{elst}$  term from the SAPT(0) analysis.

## 4) Conclusion

The combination of a large set of rotational data from the parent and eighteen different isotopologues and comprehensive *ab initio* calculations fully specified the molecular structure, conformational abundances and intermolecular binding effects, providing unprecedented details compared to previous molecular studies of weakly-bound intermolecular complexes. No stereoselectivity is observed in the dimer, as the two homo- and heterochiral diastereoisomers are formed in practically equal (1.1:1) proportions. The similar topology of the hydrogen bonds in the two diastereoisomers, based on a primary C-H…O link assisted by weaker C-H…F contacts, is probably on the origin of this similarity.

This study lends additional evidence to the versatility of CP-FTMW for structure determination and studies of non-covalently bound molecular complexes. The availability of experimental structural information of species with greater than 20 heavy atoms opens up significant potential for use of CP-FTMW for studies of more complex molecules, such as those relevant for medicinal applications, and their interactions with mimics of biologically-relevant binding substrates. In particular, these kinds of studies provide a precise and stringent test of the treatment of intermolecular interactions by quantum chemistry by providing an experimental means of directly probing the most favorable complexation geometries in systems with competing interactions.

## References

(1) Pérez, C.; Zaleski, D. P.; Seifert, N. A.; Temelso, B.; Shields, G. C.; Kisiel, Z.; Pate, B. H. *Angew. Chem. Int. Ed Engl.* **2014**, *53*, 14368.

(2) Zehnacker, A. Int. Rev. Phys. Chem. 2014, 33, 151.

(3) Wales, D. *Energy Landscapes: Applications to Clusters, Biomolecules and Glasses*; Cambridge University Press, 2003.

(4) Steiner, T. Angew. Chem. Int. Ed Engl. 2002, 41, 49.

(5) Steiner, T.; Desiraju, G. R. *Chem. Commun.* **1998**, 891.

(6) Müller-Dethlefs, K.; Hobza, P. Chem. Rev. 2000, 100, 143.

(7) Jureka, P.; Sponer, J.; Cerny, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985.

(8) Řezáč, J.; Hobza, P. J. Chem. Theory Comput. 2013, 9, 2151.

(9) Zehnacker, A.; Suhm, M. A. Angew. Chem. Int. Ed Engl. 2008, 47, 6970.

(10) Zask, A.; Murphy, J.; Ellestad, G. A. Chirality 2013, 25, 265.

(11) Dom, J. J.; van der Veken, B. J.; Michielsen, B.; Jacobs, S.; Xue, Z.; Hesse, S.; Loritz, H.-M.; Suhm, M. A.; Herrebout, W. A. *Phys. Chem. Chem. Phys. PCCP* **2011**, *13*, 14142.

(12) Thomas, J.; Jäger, W.; Xu, Y. Angew. Chem. Int. Ed. 2014, 53, 7277.

(13) Thomas, J.; Xu, Y. J. Phys. Chem. Lett. 2014, 5, 1850.

(14) Liu, X.; Borho, N.; Xu, Y. Chem. – Eur. J. 2009, 15, 270.

(15) Snow, M. S.; Howard, B. J.; Evangelisti, L.; Caminati, W. J. Phys. Chem. A 2011, 115, 47.

(16) Scharge, T.; Häber, T.; Suhm, M. A. Phys. Chem. Chem. Phys. 2006, 8, 4664.

(17) Scharge, T.; Cézard, C.; Zielke, P.; Schütz, A.; Emmeluth, C.; Suhm, M. A. *Phys. Chem. Chem. Phys.* **2007**, *9*, 4472.

(18) Scharge, T.; Luckhaus, D.; Suhm, M. A. Chem. Phys. 2008, 346, 167.

(19) *Maestro, version 9.2; Macromodel, version 9.5.207*; Schrödinger, Inc.: New York, 2010.

(20) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.;
Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji,
H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.;

Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision A.1*; Gaussian, Inc.: Wallingford, CT, 2009.

- (21) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. Chem. Phys. Lett. 1988, 153, 503.
- (22) Zhao, Y.; Truhlar, D. G. Theor. Chem. Acc. 2008, 120, 215.
- (23) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. J. Chem. Phys. 1980, 72, 650.
- (24) Boys, S. F.; Bernardi, F. Mol. Phys. 1970, 19, 553.
- (25) Becke, A. D. J. Chem. Phys. 1993, 98, 5648.
- (26) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. J. Chem. Phys. 2010, 132, 154104.
- (27) Rybak, S.; Jeziorski, B.; Szalewicz, K. J. Chem. Phys. 1991, 95, 6576.

(28) Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D. J. Chem. *Phys.* **2014**, *140*, 094106.

(29) Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. A.; Fermann, J. T.; Mintz, B. J.; Burns, L. A.; Wilke, J. J.; Abrams, M. L.; Russ, N. J.; Leininger, M. L.; Janssen, C. L.; Seidl, E. T.; Allen, W. D.; Schaefer, H. F.; King, R. A.; Valeev, E. F.; Sherrill, C. D.; Crawford, T. D. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 556.

- (30) Watson, J. K. G. J. Chem. Phys. 1967, 46, 1935.
- (31) Watson, J. K. G. Mol. Phys. **1968**, 15, 479.
- (32) Kraitchman, J. Am. J. Phys. 2005, 21, 17.
- (33) Costain, C. C. J. Chem. Phys. 2004, 29, 864.
- (34) Kisiel, Z. J. Mol. Spectrosc. 2003, 218, 58.
- (35) Schneider, H.-J. Chem. Sci. 2012, 3, 1381.
- (36) Goerigk, L.; Grimme, S. Phys. Chem. Chem. Phys. 2011, 13, 6670.

# **Chapter 6: Structure Determination of the Carbonyl Sulfide (OCS) Tetramer**

## 1) Introduction

As the previous chapters have demonstrated, CP-FTMW spectroscopy provides an extremely sensitive method for absolute structure determination of molecules and molecular complexes. However, determination using microwave spectroscopy generally requires accurate theoretical structure predictions in order to guide the analysis, both in terms of narrowing the search in a broadband acquisition for the target rotational spectrum and for interpretation of structural results from isotopic data such as for the signs of atomic coordinates derived from Kraitchman's equations.<sup>1,2</sup> This is only a requirement of convenience; proper constraint and/or *a priori* assumptions on the structural parameters of a molecule or molecular complex can enable determination of the experimental structure without access to accurate theoretical predictions. For noncovalently bound molecular complexes, if the absolute structure of the monomers is known and one assumes that the monomeric geometry does not distort upon complexation, then only the intermolecular parameters need be determined.

In this chapter, the structure determination of the OCS tetramer is presented as a demonstration of "theory-blind" experimental structure determination. These results arose from a recently published collaborative study at UVa on OCS clusters with Robert McKellar and coworkers that included this work on tetramer as well as complete substitution structures of two OCS trimers.<sup>3</sup> The study on OCS trimers, while interesting,

consists of a full set of experimental isotopic data and good agreement with theory, unlike the tetramer analysis described here.

The tetramer presents a particular challenge to theory due to the extremely shallow potential energy surface and the additional complexity arising from the addition of a fourth monomer. In the trimer, a barrel-like structure is formed<sup>4,5</sup> – well predicted by ab initio calculations and pairwise potential energy calculations using a fully correlated CCSD(T) parametrized (OCS)<sub>2</sub> potential energy surface.<sup>6</sup> In the tetramer, it is not energetically favorable for the fourth monomer to add a fourth vertex to the trimer barrel.<sup>7</sup> Instead, the monomer must add to the barrel externally in a  $(OCS)_3 + OCS$  type fashion. This is easier said than done: considering all possible combinations on a 14dimension potential energy surface for (OCS)<sub>4</sub> using the pairwise potential, over 20 isomers are found within 100 cm<sup>-1</sup> of the global minimum. Higher-order optimization of these isomers using *ab initio* or density functional techniques prove difficult as the optimization routines were found often to "bounce" out of the shallow minima, even with an ultrafine integration grid. In fact, a recent *ab initio* study by Sahu et al.<sup>7</sup> on OCS clusters presented four low-energy tetramers out of a sample of 30 trial structures using dispersion-corrected DFT calculations. After comparison of these structures with the observed tetramer in the CP-FTMW spectrum, none were consistent with the experimental structure!

#### 2) Experimental

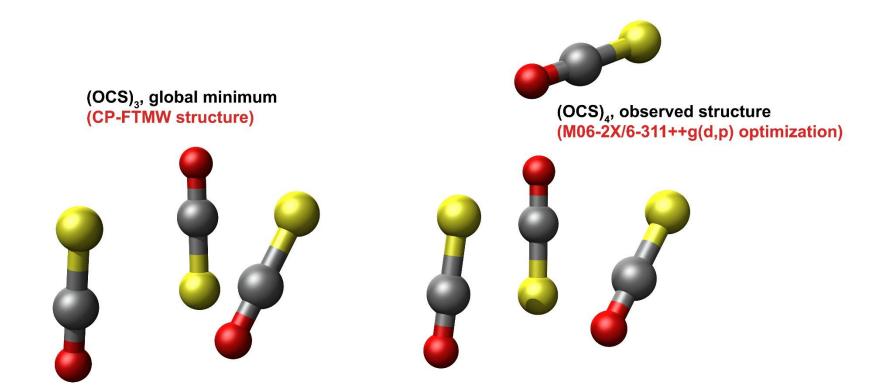
The rotational spectrum of (OCS)<sup>4</sup> was detecting using the CP-FTMW instrument at UVa in two reduced bandwidth segments of 3-6 and 6-9 GHz. The 3-6 GHz segment was obtained using the 2-8 GHz arrangement of the UVa instrument and the 6-9 GHz segment used the 6-18 GHz arrangement, both setup identically as described in Chapter 1, Section IV. These reduced bandwidth segments were used since the OCS cluster signals were optimal in the 3-9 GHz region, and use of two reduced bandwidth segments, assuming the small angle limit, increase the intensity of the spectrum by a factor of 2<sup>1/2</sup> and 2 relative to full bandwidth measurements for the 3-6 and 6-9 GHz segments respectively. The 3-6 GHz segment was coherently averaged with a 40 µs record length and at 25 Gs/s for a total averaged free induction decay (FID) consisting of 7.8 million averages, or approximately 82 hours at a repetition rate of 26.4 FIDs/second. The 6-9 GHz segment was recorded with a 20 µs length and at 50 Gs/s for a total number of 8.5 million averages, or approximately 72 hours at a rate of 33 FIDs/second.

The sample used was a 1% mixture of OCS (Aldrich, >99%) in neon, pulsed into the interaction region through five pulsed nozzles with an approximate backing pressure of 3.5 atm, which promotes production of molecular clusters.

#### 3) Spectroscopic Results

Since the more intense of the two observed OCS trimers was detected at a high dynamic range of over 2000:1, the spectrum of OCS tetramer would likely be observable at a dynamic range amenable for detection of isotopic species in natural abundance. Preliminary constants for an OCS tetramer were previously published in a review paper by Moazzen-Ahmadi and McKellar<sup>8</sup>, and input of these constants into AUTOFIT resulted in a detection of a rotational spectrum with constants consistent with one of the predicted (OCS)<sub>4</sub> structures. By input of this predicted geometry into AUTOFIT, a search for the isotopologues of this tetramer resulted in detections of eight spectra with intensities consistent with all 4 <sup>13</sup>C (1% rel. <sup>12</sup>C) and <sup>34</sup>S (4.5% rel. <sup>32</sup>S) isotopologues of the tetramer.

Application of Kraitchman's equations<sup>1</sup> to the rotational constants of all nine detected species generated a set of unsigned Cartesian coordinates for the four oxygen and four sulfur atoms of the tetramer. However, comparison of these coordinates to the geometries calculated using the pairwise potential and from the *ab initio* study from Sahu et al.<sup>7</sup> revealed discrepancies between the experimental and theoretical structures, as the experimental structure suggested a different orientation than any of the theoretical structures with rotational constants close in magnitude to the experimental determination. Comparison of this consistent theoretical structure of (OCS)<sub>4</sub> to barrel structure of (OCS)<sub>3</sub> can be found in Figure 1. A table of rotational constants for (OCS)<sub>4</sub> and its isotopologues can be found in Table 1, and the determined Kraitchman coordinates for the eight detected isotopologues are tabulated in Table 2.



**Figure 1.** Observed structure of (OCS)<sub>4</sub>, with comparison to the global minimum structure of (OCS)<sub>3</sub> determined using CP-FTMW spectroscopy.

	A / MHz	<i>B</i> / MHz	C / MHz	N lines	RMS / kHz
Р	611.32965(80)	315.42238(33)	308.46549 32)	235	3.3
<b>S</b> 1	604.89928(58)	314.46752(24)	307.30546(20)	133	8.9
S2	603.74823(58)	313.09109(30)	306.55920(28)	134	7.4
<b>S</b> 3	611.31131(61)	309.52360(28)	302.82690(24)	125	7.6
<b>S</b> 4	605.56494(58)	314.82650(33)	306.50326(30)	125	6.6
C1	610.50(18)	313.8306(77)	306.72909(73)	29	8.3
C2	609.43(18)	314.01217(79)	306.74085(81)	26	6.9
C3	608.06(15)	314.6652(49)	308.35212(53)	34	7.6
C4	607.66(13)	314.75645(49)	307.96093(42)	33	8.4

**Table 1.** Experimental rotational constants for the parent species of  $(OCS)_4$  – labeled "P" – and the eight detected <sup>13</sup>C and <sup>34</sup>S isotopologues. Fitted Distortion constants of the substituted species were fixed to the normal species values:  $D_J = 0.05417(93)$  kHz;  $D_{JK} = 0.1694(32)$  kHz;  $D_K = 0.0881(136)$  kHz;  $d_1 = -1.59(69)$  Hz;  $d_2 = -0.87(32)$  Hz.

	a	b	c
<b>S</b> 1	0.7516(20)	1.5141(10)	1.4674(10)
S2	1.7106(9)	1.3527(11)	1.8583(8)
<b>S</b> 3	3.9259(4)	[0.0]	0.118(13)
S4	1.1887(13)	1.9655(8)	0.3564(44)
C1	2.852(21)	1.071(58)	[0.0]
C2	2.628(24)	1.515(42)	0.54(12)
C3	0.07(69)	0.730(62)	1.979(26)
C4	0.732(62)	1.407(30)	1.739(28)

**Table 2.** Kraitchman coordinates for the eight detected isotopologues of  $(OCS)_4$ .Imaginary coordinates are fixed to zero, as in the /b/ coordinate for S3.

#### 4) Structure Determination

At the time the experiment was undertaken, the theoretical tetramer structure in Figure 1 predicted a trimer barrel with two oxygens and one sulfur pointing towards the fourth monomer with the sulfur end towards the barrel – hereby called the "finger" (see Figure 2 for this structure). Although the rotational constants for this structure were in qualitative agreement with the experimental spectrum, the derived Kraitchman coordinates for the carbon and sulfur atoms suggested this was not the appropriate structure. Reorientation of the trimer barrel so that two sulfurs and one oxygen pointed towards the "finger" somewhat improved the agreement, but due to the position of the center of mass with respect to the center of the barrel, transforming the barrel from O-O-S towards the finger to O-S-S did not significantly affect the rotational constants.

Therefore, theoretical input to the structure determination problem was not reliable in the traditional sense of directly comparing unsigned Kraitchman coordinates to a signed theoretical geometry. Also, due to the lack of oxygen isotopic information, a standard least-squares r<sub>0</sub> structure determination was not immediately feasible.

However, if one assumes that the OCS monomer structure does not distort upon complexation, then only the intermolecular arrangement of the four monomers need be determined. As the experimental OCS monomer structure is well-known to precision greater than that afforded by this experiment<sup>9</sup>, one can constrain unique pairs of carbon and sulfur Kraitchman coordinates to the experimental monomer C-S distances. Since the experiment guarantees that (OCS)<sub>4</sub> is  $C_1$  symmetric – all carbons and sulfurs have unique rotational spectra – there should be 4 unique pairs of carbons and sulfurs that give Cartesian distances close to the 1.565 Å experimental C-S monomer bond length. However, since the experimental coordinates are unsigned, one must permute through all relative sign combinations for each carbon/sulfur pair.

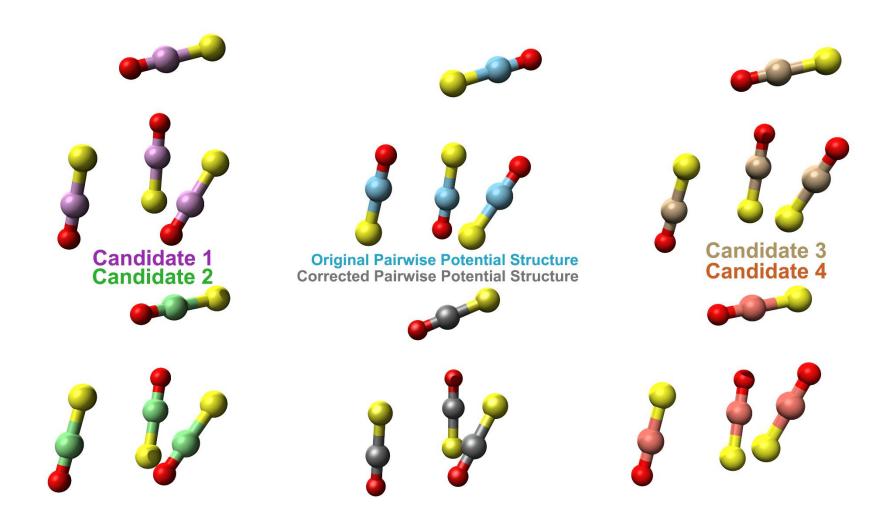
Consider a pair of carbon and sulfur atoms,  $(C_i; S_j)$  with Kraitchman coordinates  $(a_i, b_i, c_i; a_j, b_j, c_j)$ . The eight relative sign permutations for this pair are (+++; +++), (-++; +++), (+-+; +++), (+-+; +++), (--+; +++), (---; +++). Note that since these are relative signs, there is a pair-wise symmetry; e.g. (---; +++) is the same as (+++, ---).

Permuting through these signs and the coordinates in Table 2, three pairings generate realistic C-S bond lengths: (S1; C4) = (-++; +++); (S2; C3) = (+++; +++); (S4; C2) = (+++,+++). The fourth pairing (S3; C1) has an additional ambiguity since the */b/* coordinate for S3 and the */c/* coordinate for C1 are imaginary and fixed to zero, so there are four permutations that generate a physical C-S bond length: (+++; +++), (+-+; +++); (+++, +++), (+-+; +++); (+++, +++), (+-+; +++); (+++, +++), (+-+; +++); (+++, +++), (+-+; +++); (+++, +++), (+-+; +++); (+++, +++), (+-+; +++); (+++, +++), (+-+; +++); (+++, +++), (+-+; +++); (+++, +++), (+-+; +++); (+++, +++), (+-+; +++); (+++, +++), (+-+; +++); (+++, +++), (+-+; +++); (+++, +++), (+-+; +++); (+++, +++), (+-+; +++); (+++, +++), (+-+; +++); (+++, +++), (+-+; +++); (+++, +++), (+-+; +++); (+++, +++), (+-+; +++).

By fixing one of these pairs to a specified octant of  $\mathbb{R}^3$ , the first OCS monomer can be built by placing the C and S atoms at their octants specified by their consistent relative signage, and adding an oxygen atom at 1.156 Å from the carbon atom and with a linear  $\theta$ (O-C-S). The next monomer to add can be in any of the eight octants, but the relative positions of the C and S atoms must be consistent with the relative signage determined above. To compute all possible structures of the OCS tetramer, a Python script was written to brute-force through all possible arrangements of four monomers built up from the C and S atoms with distance and sign constraints described earlier. The procedure for the Python script is as follows:

- Fix the first OCS monomer using the octants specified by the relative signage of its C and S atom pairings. For instance, for (S3; C1) the carbon is fixed to the (+++) octant and therefore the sulfur can be in any of the following four octants (+++), (+-+), (++-) and (+--). Then, add the oxygen atom at the appropriate position for this monomer.
- For each of these possibilities, add OCS monomer #2 using the same steps as part (1). There are eight possible positions for monomer #2 corresponding to each of the eight octants, since the relative signs between two pairs of C and S coordinates are unknown.
- 3) Continue for the other two monomers. As such, there are  $8^3 = 512$  possible structures for each (S3; C1) sign combination. Since there are four possibilities for (S3; C1), the total number of structures that must be analyzed is 512 x 4 = 2048 possible geometries.

Using this routine, there are *four* unique candidate structures that most closely match the experimental rotational constants for the parent species. This is consistent with the four-fold ambiguity of the (S3; C1) imaginary coordinates. The structures of these four candidates are visualized in Figure 2.

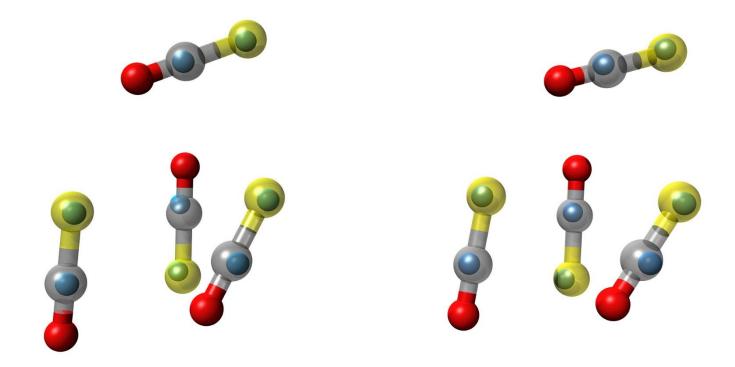


**Figure 2.** Candidate structures from brute-force structure determination routine, with comparison to the original (incorrect) pairwise potential structure and the corrected theoretical structure. Note the difference in barrel orientations between candidates 1/2 and 3/4.

In order to determine which of the four most closely matches the experimental structure, two metrics were used. One metric is the average deviation between the predicted rotational constants for the <sup>34</sup>S and <sup>13</sup>C isotopologues of the candidate structures with their experimental rotational constants. The structure in the greatest agreement should have the lowest deviation in isotopologue rotational constants (labeled  $\sigma_{fit}$ , measured in MHz). The second is the candidate's structure deviation from the experimental principal axes. Since the structural procedure is adding both oxygen atoms and monomers at independent positions relative to each other, there should be small changes in the position of the center of mass and Euler angles due to these additions with respect to the coordinate system defined by the experimental Kraitchman determination. As such, once the candidate structure is built it must be recentered to the center-of-mass frame and re-rotated to diagonalize its moment of inertia tensor. Therefore, the best-fit candidate structure should have a rotation matrix close to the identity matrix and the smallest center of mass shift ( $\Delta_{COM}$ , measured in Å). Rotational constants for all four candidate structures, as well as values for  $\sigma_{fit}$  and  $\Delta_{COM}$ , can be found in Table 3. Comparison of the Kraitchman structure with the pairwise potential structure and the M06-2X/6-311++g(d,p) reoptimized theoretical structures can be found in Figure 3.

Candidate	1	2	3	4	Experiment
A / MHz	614.9842	613.9524	621.9364	618.8134	611.32965(80)
В	317.6737	317.4038	318.302	318.1013	315.42238(33)
С	309.6543	309.4475	308.2825	309.9487	308.46549(32)
σfit (MHz)	0.217	0.254	0.378	0.285	
$\Delta$ сом (Å)	0.013	0.0078	0.052	0.057	

**Table 3.** Rotational constants and best-fit metrics for the 4 candidate structures, shown inFigure 2, with comparison to the experimentally determined rotational constants.



**Figure 3.** Comparison of the corrected pair potential structure (left, ball & stick model) and the M06-2X/6-311++g(d,p) reoptimized structure (right, ball & stick model) with the experimental Kraitchman coordinates for C and S atoms (small blue spheres).

Table 3 shows that candidate structures #1 and #2 show a distinct advantage in rotational constant agreement and both metrics over structures #3 and #4. Although it is hard to distinguish, both numerically and visually, between structures 1 and 2, the evaluation fully suggests that the orientation of the barrel in the tetramer is S-S-O pointing towards the "finger". With this experimental result in hand, it was then discovered that the pairwise potential implementation had a small bug that lead to a preference of the O-O-S barrel. Upon correction of this bug, the pairwise calculations reverted to an S-S-O orientation preferentially. Reoptimization of this corrected pairwise potential structure using M06-2X/6-311++g(d,p) improved the structure determination with respect to the experimental sulfur and carbon Kraitchman positions, as can be seen in Figure 3.

### 5) Conclusions

Using CP-FTMW spectroscopy, the structure of one isomer of (OCS)<sub>4</sub> has been determined. Although only the <sup>13</sup>C and <sup>34</sup>S isotopologues were observed, assumption of the well-determined OCS monomer structure was used to generate a best-fit candidate structure that is consistent with the experimental rotational constants of the tetramer. The structure consists of the barrel-like OCS trimer structure with an additional monomer added to one end of the barrel. This structure was not predicted in a previous theoretical study by Sahu *et al.* and was in initial disagreement with the pairwise potential structures provided by McKellar and coworkers. It is important to note that the observed (OCS)<sub>4</sub> structure is not necessarily the global minimum, but rather the only structure that survives

in the supersonic expansion of the CP-FTMW experiment.<sup>10</sup> Further experimentation is required in order to confirm its status as the global minimum, such as a measurement with argon as the buffer gas, as used in the experimental determination of the global minimum of the water hexamer.<sup>11</sup>

Above all else, the (OCS)<sub>4</sub> structure determination is a demonstration of the importance of experiment for studies of non-covalent complexes. With the structural resolution afforded by the sensitivity of the CP-FTMW technique, precise measurement of structures of global minimum or closely lying isomers of molecular complexes can be performed. This is important in the light of the previous theoretical work done on (OCS)<sub>4</sub>, where the dispersion-corrected DFT study by Sahu *et al.*<sup>7</sup> generated *no* structures consistent with the observed (OCS)<sub>4</sub> structure, and the pairwise potential unknowingly giving incorrect results.

With the fast evolution of theoretical chemistry techniques<sup>12,13</sup>, both in speed and accuracy, it suffices to say that for structural applications, experiment plays an increasingly less important role.<sup>14</sup> However, in situations where the potential energy surface is complex and/or extremely shallow, theory requires guidance from experiment to gauge its own predictive abilities.

## References

(1) Kraitchman, J. Am. J. Phys. 2005, 21 (1), 17.

(2) Harmony, M. D.; Laurie, V. W.; Kuczkowski, R. L.; Schwendeman, R. H.; Ramsay, D. A.; Lovas, F. J.; Lafferty, W. J.; Maki, A. G. *J. Phys. Chem. Ref. Data* **1979**, 8 (3), 619.

(3) Evangelisti, L.; Perez, C.; Seifert, N. A.; Pate, B. H.; Dehghany, M.; Moazzen-Ahmadi, N.; McKellar, A. R. W. J. Chem. Phys. **2015**, *142* (10), 104309.

(4) Peebles, R. A.; Kuczkowski, R. L. J. Phys. Chem. A 1999, 103 (32), 6344.

(5) Valdés, H.; Sordo, J. A. J. Phys. Chem. A 2003, 107 (39), 7845.

(6) Brown, J.; Wang, X.-G.; Dawes, R.; Jr, T. C. J. Chem. Phys. **2012**, *136* (13), 134306.

(7) Sahu, N.; Singh, G.; Gadre, S. R. J. Phys. Chem. A 2013, 117 (42), 10964.

(8) Moazzen-Ahmadi, N.; McKellar, A. R. W. Int. Rev. Phys. Chem. 2013, 32 (4),
611.

(9) Watson, J. K. G.; Roytburg, A.; Ulrich, W. J. Mol. Spectrosc. **1999**, 196 (1), 102.

(10) Ruoff, R. S.; Klots, T. D.; Emilsson, T.; Gutowsky, H. S. J. Chem. Phys. **1990**, *93* (5), 3142.

(11) Pérez, C.; Muckle, M. T.; Zaleski, D. P.; Seifert, N. A.; Temelso, B.; Shields, G. C.; Kisiel, Z.; Pate, B. H. *Science* 2012, *336* (6083), 897.

(12) Gavroglu, K.; Simões, A. *Neither Physics nor Chemistry: A History of Quantum Chemistry*; The MIT Press, 2011.

(13) Lanyon, B. P.; Whitfield, J. D.; Gillett, G. G.; Goggin, M. E.; Almeida, M. P.; Kassal, I.; Biamonte, J. D.; Mohseni, M.; Powell, B. J.; Barbieri, M.; Aspuru-Guzik, A.; White, A. G. *Nat. Chem.* **2010**, *2* (2), 106.

(14) Puzzarini, C. Phys. Chem. Chem. Phys. 2013, 15 (18), 6595.

# Conclusions

## 1) Summary

This dissertation has presented a selection of studies illustrating the amenability of the CP-FTMW technique for structure determination of gas-phase molecules and molecular complexes. The inherently broadband nature of the chirped pulse, coupled with the scalability of sensitivity due to tuning the excitation pulse and/or adding additional pulsed jet sources, readily extends microwave spectroscopy into the regime where 20 heavy atoms is a "large" molecular system, instead of the ~10 heavy atoms regime commonly seen in studies during the Balle-Flygare cavity era of microwave spectroscopy.<sup>1,2</sup> Although some sensitivity is lost with CP-FTMW in the context of frequency resolution/lineshapes or single acquisition detection limits<sup>3</sup>, the ability to interrogate multiple GHz of bandwidth significantly improves the efficiency of microwave techniques for large molecule studies.

Using a combination of modern computational techniques, including quantum chemistry and automated spectroscopic assignment programs such as AUTOFIT, rotational spectroscopy is now useful for routine structure determination and molecular identification, and provides an experimental framework to do traditional microwave-based structure studies at a significantly reduced cost of both time and user effort. Use of AUTOFIT enables fast assignment of isotopologues and weakly intense species, allowing for studies of systems sevoflurane-benzene with isotopic data sets of unparalleled size.

## 2) Future Developments

As this dissertation has shown, the CP-FTMW technique, especially in the 2-8 GHz region, is especially powerful for studies of large molecules.<sup>4</sup> The technique shows incredible sensitivity for complexes of volatile species. However, the issue of sample introduction becomes significant as the size of the molecule increases. Large organic molecules with 20 or more heavy atoms typically have high melting points and low vapor pressures. Techniques such as laser ablation have been successful at detection of molecules with low volatility<sup>5,6</sup>, but the laser ablation technique typically limits the experiment to only one pulsed jet nozzle. Recently at UVa, the CP-FTMW instrument has been upgraded with three heated reservoir nozzles each with independent temperature control. The reservoirs can be filled with a solid or high-boiling liquid sample and stably heated to temperatures nearing 200° C. Initial results on single molecule targets have been promising, and it is likely this technique can be coupled with a seeded gas of a volatile substance in order to promote formation of mixed clusters. This heated nozzle setup will also be amenable for studies of mixtures of diastereomers of complex organic molecules; where standard analytical techniques have issues distinguishing between diastereomers of organic molecules with multiple chiral centers<sup>7,8</sup>, the high resolution afforded by microwave spectroscopy is more than sufficient for resolving distinct spectra arising from small structural changes between diastereomers.

A common rule of thumb used here at UVa is that a "well-assigned" CP-FTMW spectrum typically only has 50% of its total observed transitions assigned. What are the carriers of the other, unassigned half? There is some indication some of these might arise due to spurious signals from intermodulation of intense molecular transitions, or mixing

between clock spurs from the digitizer with molecular signals.<sup>9</sup> However, spurious signals surely do not make up the entirety of the undetected half. As such, the detection of higher-order molecular clusters might not even require further developments in instrumental techniques; rather, better theoretical methodology and data reduction techniques might unlock larger detections in the broadband data sets already obtained.

Finally, on issue that arises as microwave spectroscopy approaches and passes the 20 heavy atom regime is the accuracy of quantum chemical methodology. However, the studies in this dissertation have shown that modern theoretical techniques not only has strong predictive power, but sufficient levels of accuracy can be achieved without significant computational power outside of that available on a standard modern desktop computer. Dispersion corrected density functional theory<sup>10</sup>, coupled with modern basis set approximations such as density fitting<sup>11,12</sup>, can determine structures of large, noncovalent complexes to accuracy appropriate for spectroscopic searches at a computational cost far less than modern Moller-Plesset or Coupled Cluster techniques.<sup>13</sup> Continuing development of the fragment-based method<sup>14</sup>, where the computational cost scales not as the number of canonical orbitals but rather the number of fragments (or monomers), will likely decrease the computational cost for all desired levels of theoretical accuracy. Since analysis of rotational spectroscopic results is inexorably linked to the development and evolution of theoretical methodology, the continuing progress in detection of larger and more complex chemical species is inevitable.

## References

(1) Caminati, W. In *Handbook of High-resolution Spectroscopy*; John Wiley & Sons, Ltd, 2011.

(2) Novick, S. Bibliography of Rotational Spectra of Weakly Bound Complexes,2015.

(3) Brown, G. G.; Dian, B. C.; Douglass, K. O.; Geyer, S. M.; Shipman, S. T.; Pate, B. H. *Rev. Sci. Instrum.* **2008**, *79* (5), 053103.

(4) Pérez, C.; Lobsiger, S.; Seifert, N. A.; Zaleski, D. P.; Temelso, B.; Shields, G. C.; Kisiel, Z.; Pate, B. H. *Chem. Phys. Lett.* **2013**, *571*, 1.

(5) Grubbs, G. S.; Dewberry, C. T.; Etchison, K. C.; Kerr, K. E.; Cooke, S. A. *Rev. Sci. Instrum.* **2007**, *78* (9), 096106.

(6) Lesarri, A.; Mata, S.; López, J. C.; Alonso, J. L. *Rev. Sci. Instrum.* **2003**, *74* (11), 4799.

(7) Nguyen, L. A.; He, H.; Pham-Huy, C. Int. J. Biomed. Sci. IJBS 2006, 2 (2), 85.

(8) Labuta, J.; Ishihara, S.; Šikorský, T.; Futera, Z.; Shundo, A.; Hanyková, L.; Burda, J. V.; Ariga, K.; Hill, J. P. *Nat. Commun.* **2013**, *4*.

(9) Muckle, M. T.; Zaleski, D. P.; Steber, A.; Harris, B.; Pate, B. H. 2012.

(10) Grimme, S.; Steinmetz, M. Phys. Chem. Chem. Phys. 2013, 15 (38), 16031.

(11) Sodt, A.; Subotnik, J. E.; Head-Gordon, M. J. Chem. Phys. 2006, 125 (19), 194109.

- (12) Werner, H.-J.; Manby, F. R.; Knowles, P. J. J. Chem. Phys. 2003, 118 (18), 8149.
- (13) Becke, A. D. J. Chem. Phys. 2014, 140 (18), 18A301.
- (14) Richard, R. M.; Lao, K. U.; Herbert, J. M. Acc. Chem. Res. 2014, 47 (9), 2828.