CHIROPTICAL SPECTROSCOPIC STUDIES AND METHODS FOR NATURAL PRODUCTS AND SMALL SYNTHETIC MOLECULES

By

Jordan Lee Johnson

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Chemistry

May 31, 2022

Nashville, Tennessee

Approved:

Prasad L. Polavarapu, Ph.D.

Michael P. Stone, Ph.D.

Lauren E. Buchanan, Ph.D.

Anita Mahadevan-Jansen, Ph.D.

For the poor.

ACKNOWLEDGMENTS

I would like to thank my colleagues and friends I have made throughout my journey here at Vanderbilt. They have been an invaluable source of support and an often needed distraction.

Thanks to my family and all the amazing friends who supported me in my decision to come to Vanderbilt, and their encouragement to press on when the work became tedious and tiring. A major thanks to my oldest and closest friend Matt, without whom my success at Vanderbilt would not have been possible.

I would also like to express my sincere gratitude to my advisor Prof. Prasad Polavarapu for his guidance in the lab and his support in allowing me to explore my personal scientific curiosity. He has provided me many wonderful opportunities, like my multiple trips to Poland to gain firsthand experience in using an ROA spectrometer, or the conference in Pisa, Italy, where I was able to share my research and practice my public speaking with a group of some of the brightest minds I have had the pleasure of meeting. Thanks are due to Dr Malgorzata Baranska for allowing me such an extensive use of her lab and ROA spectrometer, and to Dr Grzegorz Zajac who was a fantastic mentor and friend through out those trips. I cannot forget to thank my good friend Lukasz as well; he showed me the best of Krakow and gave me clothes to wear to work when I arrived for my second trip on a Sunday night but my luggage did not. I would also like to thank the rest of my committee, Dr Lauren Buchanan, Dr Anita Mahadevan-Jansen, and Dr Michael Stone, for taking their time to guide me throughout this process.

I would like to acknowledge the funding I have received through the NSF and Vanderbilt University, as well as the Mitchum E Warren Jr Foundation for their support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AC	Absolute Configuration
CC	Coupled Cluster
CD	Circular Dichroism
CE	Cotton Effect
CID	Circular Intensity Difference
CPL	Circularly Polarized Luminescence
DCP	Dual Circular Polarization
DF	Dissymmetry Factor
DFT	Density Functional Theory
EA	Electronic Absorption
ECD	Electronic Circular Dichroism
EDF	Electronic Dissymmetry Factor
FPMD	First Principle Molecular Dynamics
ICP	Incident Circular Polarization
IR	Infrared
KK	Kramers-Kronig
MD	Molecular Dynamics
NMR	Nuclear Magnetic Resonance
NVP	Nuclear Velocity Perturbation
OR	Optical Rotation
ORD	Optical Rotatory Dispersion
PCM	Polarization Continuum Model
QC	Quantum Chemical
RDF	Radial Distribution Function
RMSD	Root Mean Square Deviation
ROA	Raman Optical Activity
SCP	Scattered Circular Polarization
SOR	Specific Optical Rotation
SSO	Spectral Similarity Overlap
VA	Vibrational Absorption
VCD	Vibrational Circular Dichroism
VDF	Vibrational Dissymmetry Factor
VOA	Vibrational Optical Activity
VROA	Vibrational Raman Optical Activity

CHAPTER 1

Introduction

1.1 Statement of Motivation

When I was seven years old, I remember joining a group of kids from my neighborhood for my first game of backyard baseball. One of the boys asked me if I needed a right or left-handed baseball mitt. Although I often still confused my left from my right, I knew from school I was right-handed and asked for a right-handed mitt. Of course, if you catch the ball with your gloved right hand, you are forced to throw the ball with you free left hand. I quickly tried to switch the mitt to my left hand. To my immense surprise, the mitt simply would not fit. My seven year old self would continue to struggle differentiating between left and right, but I became aware then that the difference is not as meaningless as I had believed.

The property I was encountering was chirality; hands are familiar chiral objects and the word chiral comes from the Greek word for hand. Lord Kelvin first brought the term to the scientific world in 1893 at the second Robert Boyle Lecture. He said "I call any geometrical figure, or group of points, 'chiral', and say that it has chirality if its image in a plane mirror, ideally realized, cannot be brought to coincide with itself." Chirality permeates through modern science, from the spin of elementary particles to all biological life. In physics, chirality or handedness describes the orientation of a particle's spin in relation to its momentum. When spin of a particle is in the same direction as its momentum, the particle is called right-handed, and when spin and motion are in opposite directions, the particle is called left-handed.[1] All massive particles should exist in both left and right-handed forms (and even convert between the two states).

In chemistry, any molecule that is not superimposable on its mirror image is said to be chiral. Chiral molecules have a stereogenic element (e.g. stereogenic center, stereogenic axis and/or planar chirality) from which chirality arises. A stereogenic center would be an atom such as a carbon in sp^3 hybridization with four distinct groups attached to it. A stereogenic axis often comes from sterically restricted rotation about a bond leading to axial chirality. Common examples are biaryl compounds wherein the rotation about the aryl-aryl bond is restricted, such as in BINAP and BINOL. Planar chirality generally arises from compounds with a planar ring connected to an ansa or handle. If the ring cannot be swiveled through the hoop, then the compound is considered to posses planar chirality. Molecules with a curved skeleton wherein the terminal rings of the compound twist in opposite direction are said to posses helical chirality due to the resulting development of a helical twist. DNA is an example of a molecule possessing helical chirality that arises from the stacking of nucleic acid base pairs. All chiral molecules belong to the point groups C_n, D_n, T, O, and I.

Most of the compounds that comprise living organisms possess chirality, like DNA and proteins, giving any chiral substance a different response from its mirror image, referred to as its enantiomer. Many medications, supplements, insecticides, herbicides, fungicides, pesticides, fragrances, and foods are chiral or contain chiral compounds. As with the right-handed mitt issue I encountered at seven, a chiral compound tends to interact with another chiral molecule differently than its enantiomer would. This was made evident by the tragic case of thalidomide[2] wherein the *R* enantiomer was safe to consume, but the *S* enantiomer led to birth defects and deaths in thousands of children. By knowing the effects (positive, negative, or neutral) of a chiral compound on a system and the compounds structure, it becomes possible to recognize how other similarly structured compounds could bring about similar effects to lesser or greater degrees. More so, it brings about the possibility of constructing a stereo-selective synthesis for producing only the desired enantiomer. Knowing the chirality of a compound and its effects also allows for the possibility of isolating enantiomers from one another through methods such as chiral column chromatography or recognizing when chiral separation is not even necessary due to inert properties of the non-desired enantiomer. To study the interactions of any chiral molecule within a biological system, it is therefore important to know and understand molecular chirality. But a pair of enantiomers have, as far as we can measure, the same physical properties and can only be distinguished by reference to a known chiral system, which could be light or the olfactory receptors in your nose.[3] Since it is fairly easy to construct instruments that emit and measure light, that is the route that modern chemists, physicists, and biologists have taken to study molecular chirality. Let us now dive into the world of measurements and predictions of the interaction of light with chiral systems: chiroptical spectroscopy.

The power to discriminate the two enantiomers of a chiral compound is not enough however, at least for useful and safe pharmaceutical research, as one needs to know the three dimensional structure exactly, called the Absolute Configuration (AC). This is due in large part to drug discovery often involving modifications to compounds known to have certain effects. Knowing precisely which enantiomer of a drug is providing a certain effect allows for isolation and modification or removal of that compound alone. As chiral drug compounds often have more than one stereogenic element, leading to 2^n diastereomers where n is the number of stereogenic elements, there is a strong possibility of one or more of the diastereomers to have unwanted effects. Even without general side effects, extraneous drug-like compounds in medication will lead to additional stress on the liver. Too provide a certain desired effect from a drug without general side effects or unnecessary liver stress, it is vital to know the AC of the drugs being researched for pharmacological purposes. If we know both the AC of a particular drug and how it interacts with its active site, then purposeful modification of the compound can be done to create a compound with desired properties (such as a stronger or weaker binding affinity) without altering the 3D structure that allows for the interaction to take place. Fortunately, it seems that the interactions of light with chiral substances can be reasonably modeled using modern quantum

chemical techniques.[4,5] With the combination of chiroptical techniques and quantum chemical (QC) predictions of the exact 3D structure that gives rise to the chiroptical properties, the chiroptical properties can be informative on how a molecular system is changing.[4,6,7,8]

The currently available spectroscopic tools in the exploration of chirality for molecular systems include optical rotatory dispersion (ORD), electronic circular dichroism (ECD), vibrational circular dichroism (VCD), vibrational Raman optical activity (ROA), and circularly polarized luminescence (CPL). Each has its own advantages and disadvantages such as usable solvents, concentration ranges, collection times, sensitivities, etc. The techniques that will be used in this work will be detailed in the following sections and chapters.

1.2 Statement of Research Plan

The original goal of the project was to determine the AC of natural products and small synthetic molecules through direct QC prediction and comparison of chiroptical spectra. After a year of investigating various compounds with chiroptical spectroscopy, I realized that more robust methods, techniques, or experimental planning were necessary to unequivocally determine the AC or other structural information for many compounds. I built upon previous works to better model various chiral structures and system in order to successfully reproduce chiroptical spectra. This included finding a way to efficiently and systematically solvate specific chiral compounds in order to reproduce chiroptical properties and confidently assign the AC of the structure. Almost counter intuitively, this also included investigating and showing cases in which one can avoid explicit solvation altogether despite what one's chemical intuition might expect and successfully predict chiroptical properties. I wanted to delve into the quantum mechanical formalisms of these chiroptical properties, and in doing so managed to show the importance of a specific quantum mechanical term in Raman optical activity. I wanted to determine if chiroptical properties, specifically the vibrational optical activities, of compounds could be used as unique identifiers, which, for compounds of the same chemical formula appears to be viable. These procedures and practices that are employed in the following chapters clearly demonstrate their own value in teasing out structural details that may be missed or otherwise indiscernible, while often reducing the computational cost required.

1.3 Research Methods

Herein I describe the overarching research methods that I have used in my time here at Vanderbilt. I provide the fundamentals of the four chiroptical methods that I have utilized in my investigations, and those of several common analytical methods that I made use of for my projects. These methods again are overarching, providing the basis for my research. These methods have been used extensively with a variety of compounds and

careful planning to elucidate structural details as well as pertinent information to successfully reproducing chiroptical spectra for a variety of compounds in a variety of circumstances.

1.3.1 Optical Rotatory Dispersion (ORD)

Discovered by Biot in 1812, Optical Rotation (OR or α) is the rotation of the plane of polarization of electromagnetic radiation as it passes through a medium.[4,9] If the light rotates clockwise as it approaches the observer, the sample is called dextrorotatory and OR is defined to be positive, while negative or counterclockwise rotations come from levorotatory samples, which is a distinction used to label all chiral compounds as (+) or (-). The characteristic specific optical rotation (SOR or $[\alpha]$) is a widely used method for differentiating between chiral molecules, and is generally written as,

$$[\alpha]_{\lambda}^{T} = \frac{\alpha}{C * l} \tag{1.1}$$

where α is the observed rotation in degrees, λ is the wavelength of light in nm, T is the temperature, C is the concentration in $g\ mL^{-1}$, and l is the path length in dm. Specific optical rotation should be independent of concentration, but in some cases it has been found to vary slightly with concentration.[10] The standard units of SOR are $deg\ mL\ g^{-1}\ dm^{-1}$. Specific optical rotation changes with the wavelength of light, which is an effect called Optical Rotatory Dispersion (ORD). Optical rotatory dispersion was first calculated within the static limit in 1997, but these calculations are only valid far from electronic transitions.[11] Optical rotatory dispersion can be calculated using linear response theory for any given wavelength and gives accurate results in most cases with Density Functional Theory (DFT) but the gold standard for prediction is Coupled Cluster (CC) theory.[12,13] ORD is calculated from the imaginary part of the electric dipole-magnetic dipole polarizability (called the G' tensor), which complicates the calculation of ORD (and other chiroptical properties) by the origin dependence of the magnetic dipole moment operator.[12] To obtain origin independent results, Gauge-Invariant Atomic Orbitals or the velocity gauge must be used.[14]

Optical rotation is observed because left and right handed circularly polarized light travel at different speeds in a chiral medium. Optical rotation is related to the imaginary components of refractive index. Since linearly polarized light can be written as the sum of equal amounts of right and left circularly polarized light, a difference in speed manifests as a change in the polarization angle. The difference in the real components of refractive index is related to difference in absorption of left and right circularly polarized light also called Circular Dichroism (CD). The real and imaginary parts of the complex refractive index of light, as a real complex pair, are interchangeable through the Kramers-Kronig (KK) transformations. The relationship of

the real and imaginary parts of a complex linear response function $f(\omega)$ are related by:[15]

$$Re[f(\omega)] = \frac{2}{\pi}\int_0^\infty \frac{\mu Im[f(\mu)]}{\mu^2-\omega^2}\,d\mu \tag{1.2}$$

$$Im[f(\omega)] = \frac{-2\omega}{\pi}\int_0^\infty \frac{\mu Re[f(\mu)]}{\mu^2-\omega^2}\,d\mu \tag{1.3}$$

where *Re* stands for the real part, *Im* stands for the imaginary part, and μ is a variable used for integration. The form of the KK relations makes ORD a long ranged effect when compared to CD and can be measured for chiral compounds for which the ECD cannot be measured due to instrumental restrictions. The ORD spectrum can be measured routinely for any sample or solvent as long as some light passes through the sample at the wavelengths of interest. At an absorption transition, OR becomes sigmoidal and changes sign. Rotations as small as 0.003 degrees can be measured and there is no instrumental restriction on how large the angle of rotation can be. In our lab, the Autopol IV polarimeter (Rudolph Research Analytical, Flanders, NJ, USA) is used to measure optical rotation at six discreet wavelengths (633, 589, 546, 435, 405, and 365 nm).

1.3.2 Electronic Circular Dichroism (ECD)

Circular Dichroism (CD) is the differential absorption of left and right handed circularly polarized light. This phenomenon was discovered by Aime Cotton in 1895.[16] The differential molar absorptivity ($\Delta\varepsilon$) is defined by,

$$\Delta\varepsilon = \varepsilon_L - \varepsilon_R = \frac{(A_L - A_R)}{C*l} = \frac{\Delta A}{C*l} \tag{1.4}$$

where l is the path length in cm, or the length that the light beam passes through the sample, A_L is the absorption of left circularly polarized light, A_R is the absorption of right circularly polarized light, and C is the concentration in $mol\ L^{-1}$. Electronic circular dichroism (ECD) measures the differential absorption of left and right handed circularly polarized light for electronic transitions observed in the UV-Visible spectral region. ECD is different from Electronic Absorption (EA) in that it can be positive or negative. The strength of a given electronic transition for ECD can be calculated by the relation[17]

$$R_k = Im\left[\left\langle i\middle|\mu\middle|k\right\rangle\left\langle k\middle|m\middle|i\right\rangle\right] = Im\left[\mu_{ik}\bullet m_{ki}\right] \tag{1.5}$$

where the transition goes from state $|i\rangle$ to state $|k\rangle$, and μ and m are the electric and magnetic dipole moment operators respectively. For the absorption process, D_k is called the dipole transition strength and determines

the integrated absorption band intensity. D_k is defined as:

$$D_k = \left[\left\langle i|\mu|k\right\rangle\left\langle k|\mu|i\right\rangle\right] = \left[\mu_{ik} \bullet \mu_{ki}\right]. \tag{1.6}$$

For electronic transitions, theoretical spectral simulations are normally carried out with Gaussian spectral intensity distribution, given as:

$$G_k(x) = y_k^0 e^{-\left(\frac{x-x_k^0}{w_G}\right)^2} \tag{1.7}$$

where x_k^0 is the wavelength at the center of k^{th} band, w_G represents the half-width at 1/e of the band maximum.

1.3.3 Vibrational Circular Dichroism (VCD)

Vibrational transitions can absorb left and right handed light differently giving rise to VCD. The rotational strength R_k and dipole transition strength D_k for a given vibrational transition are defined in the same general terms as provided for electronic transitions in equations 1.6 and 1.7. However, the calculated rotational and dipole transition strengths correspond to the integrated molar absorption coefficients determined experimentally and are given by the relations,

$$R_k = \frac{0.23x10^{-38}}{\bar{\nu}_k}\int \Delta\varepsilon(\bar{\nu})d\bar{\nu} \tag{1.8}$$

$$D_k = \frac{0.92x10^{-38}}{\bar{\nu}_k}\int \varepsilon(\bar{\nu})d\bar{\nu} \tag{1.9}$$

where $\bar{\nu}_k$ is the transition frequency in cm^{-1}.

The CD for vibrational transitions is approximately 10,000 times weaker than the Vibrational Absorption (VA) of unpolarized light, meaning there are inherent problems with noise when measuring VCD. Due to the weakness of the VCD signal and other instrumental difficulties, measurements of VCD were not performed until 1974.[18] The removal of the linear birefringence signal[19] has improved the reliability of VCD measurements in newer instruments,[2] which can reliably measure from 4000-800 cm^{-1}. This range allows for the observation of a large number of vibrational transitions, which can be compared to Quantum Chemical (QC) calculations. In our lab, the ChiralIR spectrometer (BioTools, Jupiter, FL, USA) is used for the measurement of VCD spectra.

There are practical limitations on VCD measurements. Measurements of VCD are most easily made in the solution phase, but nearly all solvents will absorb light in the infrared (IR) region. Deuterated solvents and short path lengths (50-200 μm) must be used to minimize solvent absorbance. To maximize the absorbance from the sample, high concentrations must be used which can lead to aggregation effects. Also the sample

may not dissolve at high concentrations in all solvents. If hydrogen bonding solvents are necessary, then the vibrations of the molecule will be perturbed by hydrogen bonds with the solvent. The solvent interactions must be accounted for in comparisons with QC calculations, which can alter the results significantly.[20] VCD measurements on finely dispersed solid particles and films can be made, but these require a rotating sample holder to remove linear dichroism and the effects of strain on the sample.[21]

In spectral simulations, the VA bands are generally represented by Lorentzian band shapes given by:[22]

$$L_k(x) = \gamma_k^0 \frac{\gamma_L^2}{\gamma_L^2 + \left(x - x_k^0\right)^2} \tag{1.10}$$

where γ_L is the half-width at half-maximum for the Lorentzian band and γ_k^0 is the peak y-value at x_k^0.

1.3.4 Raman Optical Activity (ROA)

Raman optical activity (ROA) is the measurement of differential intensity of Raman scattering between right and left circularly polarized light.[23,24] ROA is supported by both rotational[25,26] as well as vibrational[23,24] transitions of chiral molecules. ROA associated with vibrational transitions, referred to as vibrational Raman optical activity (VROA), is the most common form that is being widely investigated. The experimental vibrational Raman and ROA spectral intensities depend on various experimental parameters, including the incident laser intensity, and therefore cannot be quantitatively compared with the intensities calculated using QC methods. However, dimensionless circular intensity difference (CID), the ratio of VROA to corresponding Raman, eliminates the dependence on experimental parameters and permits quantitative comparisons in the analysis of QC predictions.

VROA originates[23,24] from the products of the normal coordinate derivatives of three polarizability tensors, the electric-dipole—electric-dipole polarizability (α), the electric-dipole—magnetic-dipole polarizability (***G'***), and the electric-dipole—electric-quadrupole polarizability (***A***). The VROA intensities are governed by a combination of three different tensor product invariants: (a) the product of the means of the $\frac{d\alpha_{\alpha\beta}}{dQ_k}$ and $\frac{dG'_{\alpha\beta}}{dQ_k}$ tensors (designated to as $\omega^{-1}\bar{\alpha}_k\bar{G}'_k$, where ω is the circular frequency of incident laser light), (b) the anisotropy of the product $\frac{d\alpha_{\alpha\beta}}{dQ_k}\frac{dG'_{\alpha\beta}}{dQ_k}$ tensor (designated as $\omega^{-1}\gamma_k^2$), and (c) the anisotropy of the product $\frac{d\alpha_{\alpha\beta}}{dQ_k}\frac{dA_{\alpha\beta\gamma}}{dQ_k}$ tensor (designated as $\omega^{-1}\delta_k^2$). The notation used here follows that from Polavarapu[23,24] and the units for these tensor product invariants, $\omega^{-1}\bar{\alpha}_k\bar{G}'_k$, $\omega^{-1}\gamma_k^2$, $\omega^{-1}\delta_k^2$, are A^5/amu.[24]

VROA measurements have been conducted with incident circularly polarized light, which are referred to as incident circular polarization (ICP) measurements.[27] The initial ICP measurements,[27] performed in the 90° right angle scattering geometry with scattered light polarization parallel to the scattering plane, were referred to as depolarized ICP VROA. Polarized ICP VROA, monitoring scattered light polarization per-

pendicular to the scattering plane, were also measured in subsequent years.[28] However, improved signal-to-noise was realized for ICP measurements in the 180° backscattering geometry.[29] In a different approach, VROA measurements can be conducted with linearly polarized incident light, and monitoring the circular polarization components in the scattered light, which are referred to as scattered circular polarization (SCP) measurements.[30] SCP measurements are normally conducted in 180° backscattering geometry. When incident laser light is circularly polarized and scattered light is analyzed for circular polarizations, the resulting measurements are referred to as dual circular polarization (DCP) measurements.[31,32] DCP measurements are normally conducted in 180° backscattering geometry. The DCP setup with incident and scattered circular polarizations in-phase is referred to as DCP_I and as DCP_{II} when they are out of phase. A limited number of VROA measurements were also reported with these DCP setups. Whereas the ICP measurements performed in 90° right angle depolarized scattering geometry,[27,28] and in 180° backscattering geometry,[29] were common in the early stages of development, the SCP VROA measurements in 180° backscattering geometry with unpolarized incident light are practiced currently. While a graduate student in Dr. Polavarapu's lab, I was able to make many ROA measurements in collaboration with Dr. Malgorzata Baranska and Dr. Grzegorz Zajac. Dr. Baranska's group at Jagiellonian University, Krakow, Poland, have a commercially available ChiralRAMAN spectrometer (BioTools, Jupiter, FL, USA) for ROA measurements. This spectrometer utilizes SCP VROA measurements in 180° backscattering geometry with unpolarized incident light.

VROA, and Raman, intensities for kth vibration are given, respectively, by Equations 1.12 and 1.13:[23,24]

$$(I_\alpha^\gamma - I_\beta^\delta)_k = K \frac{(\bar{\nu} \mp \bar{\nu}_k)^4}{\bar{\nu}_k(\pm 1 \mp e^{\frac{\mp hc\bar{\nu}_k}{bT}})} P_k, \tag{1.11}$$

$$(I_\alpha^\gamma + I_\beta^\delta)_k = K \frac{(\bar{\nu} \mp \bar{\nu}_k)^4}{\bar{\nu}_k(\pm 1 \mp e^{\frac{\mp hc\bar{\nu}_k}{bT}})} S_k, \tag{1.12}$$

where $I_\alpha^\gamma - I_\beta^\delta$ represents the VROA, and $I_\alpha^\gamma + I_\beta^\delta$ the corresponding Raman, the subscripts on $\boldsymbol{I}$ represent the polarization of the scattered light, and the superscripts represent the polarization of the incident light, $\bar{\nu}$ is the incident laser wavenumber, $\bar{\nu}_k$ is the wavenumber of the kth vibration, b is Boltzmann's constant, c is the speed of light, h is Planck's constant, T is temperature, K is a constant, P_k is the scattering geometry dependent Raman circular intensity difference activity, and S_k is the scattering geometry dependent Raman activity for kth vibration. P_k and S_k are defined in the following general Equations 1.14 and 1.15, respectively:

$$P_k = \frac{2\pi}{\lambda}(a_3\omega^{-1}\bar{\alpha}_k\bar{G'}_k + a_4\omega^{-1}\gamma_k^2 + a_5\omega^{-1}\delta_k^2), \tag{1.13}$$

$$S_k = (a_1\bar{\alpha}_k^2 + a_2\beta_k^2), \tag{1.14}$$

with λ as the incident laser light wavelength; $a_1 - a_5$ are scattering geometry and polarization dependent constants.; $\bar{\alpha}_k^2$is the mean of electric-dipole polarizability derivative tensor and β_k^2 its anisotropy, for kth vibration. The units for $\bar{\alpha}_k^2$, β_k^2, P_k, and S_k are A^4/amu.[24] The ratio of Equations 1.14 and 1.15 gives the dimensionless CID, $\Delta_k = \frac{P_k}{S_k}$. The Raman activity S_k is determined solely by β_k^2 in depolarized ICP right angle scattering, as well as in DCP_I backscattering, geometry, whereas that in SCP, as well as ICP, backscattering and polarized ICP measurements are determined by both $\bar{\alpha}_k^2$ and β_k^2. In spectral simulations, the vibrational Raman bands are generally represented by Lorentzian band shapes as with VA.

1.3.5 Quantum Chemistry (QC)

Quantum chemistry is the application of quantum mechanics to chemical systems. The Schrödinger equation is used in QC for understanding the dynamics of molecules and their electronic structures. Density functional theory (DFT) is a method for computationally modeling quantum mechanics that is used by physicists, chemists, and material scientists. For most typical organic molecules, their structure and properties can be calculated reliably using DFT.[15] Density functional theory can provide accurate results in most cases by empirically accounting for electron correlation, and by fitting parameters to experimental data.[33] The density functional of choice in our lab is typically B3LYP, which includes Hartree-Fock character as well as Becke's exchange[34] and correlation parameterized by Lee, Yang, and Parr.[35] DFT methods can be used to calculate the properties of molecular systems up to a several hundred atoms. In some cases[36] DFT does not properly model the electronic structure of the molecule and Coupled Cluster (CC) methods are necessary,[5,37,38,39,40,41,42] but CC methods are practically limited to molecules roughly 30 atoms or less with standard computational resources.

To calculate the properties of a molecule, first all possible conformations or different three dimensional arrangements of a given structure obtained by rotating about molecular bonds, are found by a search algorithm. Then the structures of each of the conformations are optimized, or the bond lengths, bond angles, and dihedral angles are altered until the minimum energy is found. From the optimum energies of the conformers, the ones with the lowest energies in a certain threshold are selected as the dominant conformations, and any conformations with energies higher than a given cutoff are eliminated. The cutoff depends upon the relative accuracy of the method used ($\sim$2kcal/mol for DFT methods).[43] All conformations that have energies within the cutoff are used in the calculation of molecular properties. To obtain the conformational average from selected spectra, the individual spectra from the relevant conformers are Boltzmann weighted according to

the formula,

$$P_i = \frac{N_i}{N_{tot}} = \frac{e^{-E_i/RT}}{\sum_a e^{-E_a/RT}} \tag{1.15}$$

where E_i is the energy of a given conformation, R is the gas constant, and T is the temperature.

The effects of the solvent on the solute can be accounted for quickly by the use of continuum models, with the model of choice being the Polarization Continuum Model (PCM).[44] In the PCM, the solute is surrounded by a dielectric medium with a spacing between the molecule and the dielectric medium given by an empirical force field. PCM can provide a reasonable model for solvation by less polar solvents but can fail to model hydrogen bonding solvent effects.[45]

1.3.6 Analytical Methods

Chiroptical spectra can be analyzed in many ways. When one has only an experimental spectrum, the traditional approach of correlating band positions with specific transitions known from the literature can be useful. For example, a VCD band at $\sim$1800 cm^{-1} is indicative of the compound containing a carbonyl group, as C=O stretching vibration is generally the only vibrational mode at $\sim$1800 cm^{-1}. The inclusion of QC predicted spectra for each diastereomer of the compound of interest allows for better structural discrimination. By qualitatively comparing experimental chiroptical spectra to the corresponding spectra predicted for each diastereomer, one may be able to identify the AC of the compound. This comparison is usually done band-by-band between experimental and predicted spectra to validate or discriminate between the signs and relative positions of the bands. While a qualitative match between experimental and predicted spectra are necessary for reliable interpretation, quantitative comparisons are desirable for more objective and reliable interpretations of the spectra and their comparison. The following subsections introduce a few of the major quantitative methods I take advantage of when analyzing chiroptical spectra.

1.3.6.1 Dissymmetry Factor and Circular Intensity Difference

Dissymmetry Factor (DF or g), also known as anisotropy factor, is the ratio between CD and corresponding absorption given by:

$$g = \frac{\Delta\varepsilon}{\varepsilon} = \frac{\Delta A}{A} = \frac{4R_k}{D_k}. \tag{1.16}$$

Vibrational Dissymmetry Factor (VDF) is obtained as the ratio of VCD to VA. Electronic Dissymmetry Factor (EDF) is obtained as the ratio of ECD to EA. Circular Intensity Difference (CID or Δ_k) is the ratio of ROA to corresponding Raman and given by:

$$\Delta_k = \frac{I_\alpha^\gamma - I_\beta^\delta}{I_\alpha^\gamma + I_\beta^\delta} = \frac{P_k}{S_k}. \tag{1.17}$$

The errors in concentration and path length do not influence the EDF and VDF spectra when both absorption and CD spectra are measured simultaneously for the same sample and using the same instrument. Similarly, CID eliminates the dependence on experimental parameters and permits quantitative comparisons in the analysis of QC predictions. If the CD or ROA baseline is not horizontal, then the corresponding DF or CID spectra cannot be used. For this reason, it is important to have a good-quality experimental CD or ROA spectrum with a horizontal baseline before proceeding to calculate the DF or CID spectrum.

Calculation of DF or CID spectra incorporates two features: (a) to help remove the amplification of noise in the regions where absorbance or Raman is nearly zero and the associated CD or ROA signal is only noise, a baseline tolerance value for absorption or Raman spectra is incorporated; and (b) the reliability level of experimental CD or ROA signals, determined by the experimental reliability criterion (previously referred to as the experimental robustness criterion[46], based on the experimental $\Delta A/A$ or P_k/S_k values, is introduced. The choice for the experimental reliability criterion can vary from instrument to instrument and for individual research groups. In measuring the VCD spectra the absolute values of VDF less than 10^{-5} are not routinely measurable and calculated VDF magnitudes, using the criterion of Gobi and Magyarfalvi,[47] less than $4x10^{-5}$ may not be reliable.[48,49] Thus, the use of $|g| > 4 * 10^{-5}$ consistently for both experimental and calculated VDF spectra would satisfy the robustness criteria for VDFs. In the case of ROA, the experimental CID magnitudes are roughly similar to those of VDFs, so the threshold of 10^{-5} is used for analyzing the experimental CIDs. Since there is a factor of 4 difference in the definitions of VDF and CID, a threshold of 10^{-5} is used for both experimental and calculated CIDs. In the case of ECD, the experimental EDFs are 1 to 2 orders of magnitude larger than VDFs due to good the signal-to-noise ratio of ECD spectra in general. Thus for calculating the EDF spectra, I do not impose the experimental reliability criterion since ECD spectra generally have a good signal-to-noise ratio. The use of 40 ppm (i.e., $\Delta A/A = 4 \times 10^{-5}$) and 10 ppm as the threshold for reliability for VDF and CID respectively results in blanking out the spectral regions with VDF less than 40 ppm or CID less than 10 ppm, leading to spectral discontinuities in the VDF or CID spectra. To avoid spectral discontinuities for VDF less than 40 ppm or CID less than 10 ppm, I have generally used 0 ppm for the robustness criterion.

1.3.6.2 Spectral Similarity Overlap

Spectral Similarity Overlap (SSO) is a quantitative comparison between spectra, generally between experimental and QC predicted spectra. Spectral Similarity Overlap is calculated using the *Sim* function suggested by Shen et al.[50] This *Sim* function is written as[51]

$$SimXXX(\sigma) = \frac{\int f_{XXX}(\sigma : x) h_{XXX}(x) dx}{\int f_{XXX}(\sigma : x) f_{XXX}(\sigma : x) dx + \int h_{XXX}(x) h_{XXX}(x) dx - |\int f_{XXX}(\sigma : x) h_{XXX}(x) dx|}. \quad (1.18)$$

In this equation, h_{XXX}(x) represents the experimental spectrum as a function of the running index for x-axis values; subscript XXX represents the type of spectrum (XXX = VA, VCD, VDF, EA, ECD, EDF, Raman, ROA, or CID); $f_{XXX}(\sigma$:x) represents the simulated predicted spectrum obtained after scaling the predicted transition frequencies/wavelengths with a scale factor σ; *Sim*XXX(σ) represents the numerical measure of SSO between experimental and predicted spectra as a function of σ for the type of spectrum XXX; the integrals run over a continuous spectral region of interest. The display of *Sim*XXX(σ) as a function of σ is referred to as the SSO plot.

For similarity analysis, VA, VCD, VDF, EA, ECD, and EDF spectra are normalized individually using square root of the sum of squared intensities.[52,53] The range for *Sim*ECD, *Sim*VCD, *Sim*EDF, and *Sim*VDF values is -1 to +1. A value of +1 indicates perfect agreement of experimental spectra with predicted spectra for the AC used in calculations. A value of -1 indicates perfect agreement of experimental spectra with predicted spectra for the AC opposite to the one used for calculations. The *Sim*VA and *Sim*EA values, which have a range of 0-1, do not relay any information on AC. While separate evaluations of QC predictions of EA and ECD or VA and VCD spectral intensities may appear satisfactory, EDF and VDF (ratio of ECD/VCD to EA/VA) spectra are more challenging for QC predictions. For this purpose, it has been advocated for the quantitative analysis of experimental and predicted EDF and VDF spectra also for verifying the agreement between experimental and predicted ECD and VCD spectra.[46,51,54] The agreement between experimental and predicted spectra is quantified using maximum numerical values of *Sim*EA, *Sim*ECD, and *Sim*EDF for ECD spectra and *Sim*VA, *Sim*VCD, and *Sim*VDF values for VCD spectra in SSO plots. While *Sim*EA and *Sim*VA values come out to be generally high and not very discriminatory, *Sim*ECD and *Sim*EDF values for ECD, and *Sim*VCD and *Sim*VDF values for VCD, provide good measures of quantitative agreement between experimental observations and QC predictions. Achieving a *Sim*VCD value of +0.4 or higher is recommended for an acceptable agreement between experimental and calculated spectra for assigning the correct molecular structures.[51] Approximately the same value of *Sim*VDF is also recommended. This recommendation arises from comparing the cross correlation values of similarity between diastereomers of 3-Ishwarone and recognizing that diastereomers tend to have quantitatively distinct chiroptical properties.[52] The same criteria are used for *Sim*ECD and *Sim*EDF. This criterion was established by Dr. Polavarapu[23] and Covington et al.[52] while Shen et al. proposed only a magnitude greater than 0.2 for confidence in spectral similarity.[50] My own investigation confirmed that a reliability threshold of magnitudes greater than 0.4 provides better structural discrimination and unique identification of compounds.[55]

1.3.6.3 RMSD of ORD

For similarity analysis of ORD data, an analog of Eq. 1.19 can be formulated to obtain *Sim*ORD, by replacing the integrals with the summation over discrete wavelength data points and replacing $f_{XXX}(\mathrm{x})$ and $g_{XXX}(\mathrm{x})$ with calculated and experimental specific rotations at individual wavelengths. However, *Sim*ORD obtained in this manner did not reveal sufficient discriminating ability for diastereomers due to a limited number of data points. Therefore, for quantitative similarity analysis of ORD, the standard root mean square deviation (RMSD) defined as[56,57]:

$$RMSD = \sqrt{\sum_{i} \frac{(Y_{i,calc} - Y_{i,obs})^2}{N}}. \tag{1.19}$$

is used, where $Y_{i,calc}$ and $Y_{i,obs}$ are, respectively, the calculated and experimental SORs at the *i*th wavelength and N is the number of discrete wavelength data points and the summation runs over discrete wavelength data points.

CHAPTER 2

The AC of (-)-inuloxin C

This work was completed in collaboration with Professor Antonio Evidente of the University of Naples Federico II, Naples, Italy, Professor Nina Berova of Columbia University, New York, NY, and Professor Ana Petrovic of the New York Institute of Technology, New York, NY. This chapter has been reproduced from the *Chirality* article[58] "Absolute configurations of chiral molecules with multiple stereogenic centers without prior knowledge of the relative configurations: A case study of inuloxin C" with permission. My contributions to this work are: conformational analyses, QC optimizations, QC spectral prediction of ECD and ORD, spectral similarity analyses, preparation of figures and tables, drafting the manuscript, and revising the manuscript.

2.1 Introduction

Inuloxin C, a tricyclic eudesmane sesquiterpenoid, was isolated together with inuloxins A, B, and D (belonging to other sequiterpene subgroups) and α-costic acid[59] from the organic extract of *Inula viscosa*. This is an allelopathic Mediterranean wild plant proposed for biological control of broomrapes (*Orobanche* spp.) and dodders (*Cuscuta* spp.) which are parasitic weeds attacking a large number of important crops.[59] These parasites cause severe yield losses[60] leading not only to a complete crop failure but also have a great negative impact over many years, because seeds can survive in soil for a long period of time, preventing the reasonable production of legumes and vegetables in the infested fields.[61] *Inula viscosa* was chosen among other 10 allelopathic Mediterranean plants as its organic extract inhibited both broomrapes and dodders, in particular *Orobanche crenata* and *Cuscuta campestris*.[62] This activity was also preserved by inuloxin C as well as inuloxins A and D.[59] As the relation between AC and the biological activity and mode of action of natural compounds, including bio-pesticides[63,64] was extensively reported,[65,66] the ACs of inuloxins A-C were previously studied by chiroptical methods.[67,68]

When relative configuration of a chiral compound with multiple stereogenic centers is known with a high degree of certainty, chiroptical spectroscopic analysis is considerably simplified because QC calculations of chiroptical properties can then be restricted to only one diastereomer. If the dependence of chiroptical spectroscopic analyses on the knowledge of relative configurations needs to be avoided, or the relative configurations are not known with high degree of certainty, then it is not sufficient to show that QC calculations on one diastereomer match the experimental spectra. Instead, it is necessary to show that QC calculations for no more than one diastereomer provide reasonable match to the experimental data.[4,69,70] If this goal cannot

be realized, then a protocol for the analysis using chiroptical methods needs to be developed to identify one unique diastereomer. Even if the relative configuration of a chiral compound with multiple stereogenic centers is known with high degree of certainty, it will be important to establish that the experimental chiroptical spectra are not reproduced by those calculated for more than one diastereomer.

The structure and relative configuration of inuloxin C was determined through NMR analysis.[59] Using the relative configuration determined from NMR, the AC of inuloxin C was assigned as (5*S*,7*S*,8*S*,10*S*) (see Figure 2.1) based on the comparison of the experimental ECD and ORD data with corresponding QC calculations carried out for only one diastereomer.[67]

In this work, using inuloxin C as a test case, I showed that the AC of inuloxin C can be determined solely from ECD and ORD data, independent of relative configuration determined from NMR. For this purpose, the experimental ECD and ORD for inuloxin C were remeasured, and I undertook QC calculations of ECD and ORD for all diastereomers of inuloxin C. Utilizing these new data, I arrived at two important observations: (1) The individual experimental chiroptical spectroscopic data can match those predicted by QC calculations for more than one diastereomer; (2). Despite this unfavorable observation, a unique diastereomer can still be identified by extending the analyses to similarity in EDF spectra and to a stereochemically unaltered derivative of the parent compound.

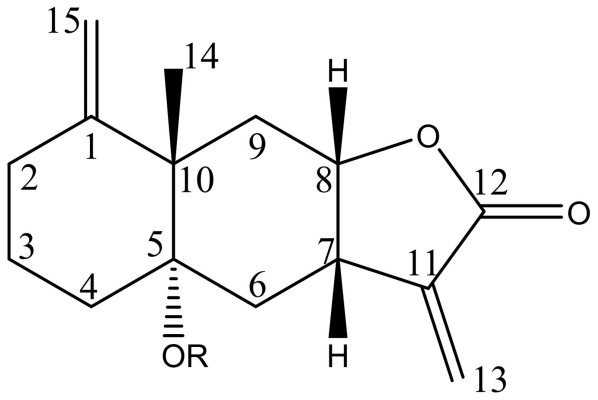

Figure 2.1: (5*S*,7*S*,8*S*,10*S*) diastereomer of inuloxin C (R=H) and acetylated inuloxin C (R=CO(CH_3))

2.2 Methods

The isolation and characterization of inuloxin C was carried out by our Italian collaborators and has been described in the literature.[59] The ORD and ECD spectra used within were measured by Dr. V. Ragahavan while a post doctoral researcher within our lab and details of these measurements are provided below and in the main text.[58]

Optical rotations were measured at room temperature using Autopol IV automatic polarimeter, 0.5 dm

cell and 5.0×10^{-3} g/mL and 2.5×10^{-3} g/mL concentrations in $CHCl_3$. The measured optical rotations were converted to SORs in units of deg cc g^{-1} dm^{-1}. Six discrete wavelengths (633, 589, 546, 436, 405, 365 nm) were utilized for the ORD measurements. For two different separately isolated samples of inuloxin C, the currently measured $[\alpha]_D^{25}$ values are -10° and -13°. In light of these new measurements, previously reported,[59,67] $[\alpha]_D^{25}$ = +13.1° and $[\alpha]_D$ = -0.03° for inuloxin C have to be revised. ECD measurements were performed with JASCO J-810 spectrometer. ECD spectra were measured at 5.6×10^{-6} M concentration for 190 to 230-nm region and at 5.6×10^{-4} M for 230 to 400-nm region, in a 1-cm quartz cell using acetonitrile solvent. The spectra obtained at two concentrations for the separate regions are merged together, after appropriate shifting for continuity, and presented as one combined spectrum. The published experimental ECD and ORD data of acetylated inuloxin C were digitized and used for the analyses presented here.

Inuloxin C was purified from the organic extract of the aerial parts of *I. viscosa* as previously reported.[67]

Number	Diastereomer
1	(5*S*,7*S*,8*S*,10*R*)
2	(5*R*,7*S*,8*S*,10*S*)
3	(5*S*,7*S*,8*S*,10*S*)
4	(5*S*,7*S*,8*R*,10*S*)
5	(5*R*,7*S*,8*S*,10*R*)
6	(5*R*,7*S*,8*R*,10*S*)
7	(5*R*,7*S*,8*R*,10*R*)
8	(5*R*,7*R*,8*S*,10*S*)

Table 2.1: Eight diastereomers of inuloxin C and acetylated inuloxin C studied in this work.

The structure of inuloxin C contains four stereogenic centers (Figure 2.1), resulting in 16 diastereomers. One-half of these are enantiomers of the other half, so ECD and ORD calculations needed to be carried out for only eight diastereomers. For each of the eight selected diastereomers (see Table 2.1), conformers within 20 kcal/mol energy range were determined using CONFLEX program.[71] The geometries of these initial conformers of each diastereomer were further optimized, using Gaussian 09 program (G09),[72] at B3LYP/6-311++G(2d,2p) level. The geometries of optimized conformers within 5 kcal/mol range at B3LYP/6-311++G(2d,2p) level were further optimized using polarizable continuum model (PCM) for $CHCl_3$ solvent, as implemented in G09. The optimized conformers with greater than 1% population were then used for SOR calculation at six different wavelengths (633, 589, 546, 436, 405, and 365 nm) using B3LYP/6-311++G(2d,2p)/ PCM($CHCl_3$). The calculated SORs for the conformers of each diastereomer were multi-

plied with respective populations derived from electronic energies and summed. For the prediction of ECD, the geometries of the conformers of each diastereomer optimized at B3LYP/6-311++G(2d,2p) level within 5 kcal/mol range were further optimized at CAM-B3LYP/6-311+ +G(2d,2p) level using (CH_3CN) solvent represented by the PCM. ECD calculations were then carried out, for the conformers of each diastereomer with population greater than 1%, at CAM-B3LYP/6-311++G(2d,2p)/ PCM(CH_3CN) level.

For calculations on acetylated inuloxin C, the structure of each diastereomer of inuloxin C was acetylated via manual editing. Using these acetylated diastereomer structures, initial conformational search was conducted, as for inuloxin C, with Conflex program. The conformers generated by Conflex were first optimized at B3LYP/6- 31G* level and those optimized conformers that are within 5 kcal/mol range were optimized at B3LYP/6- 311++G(2d,2p) level with PCM($CHCl_3$) implicit solvation for ORD and at CAM-B3LYP/6-311++G(2d,2p)/PCM(CH_3CN) level for ECD calculations, as for inuloxin C.

Additional information on the numbers of conformers used at each of the theoretical levels for inuloxin C and acetylated inuloxin C are summarized in the main manuscript.[58]

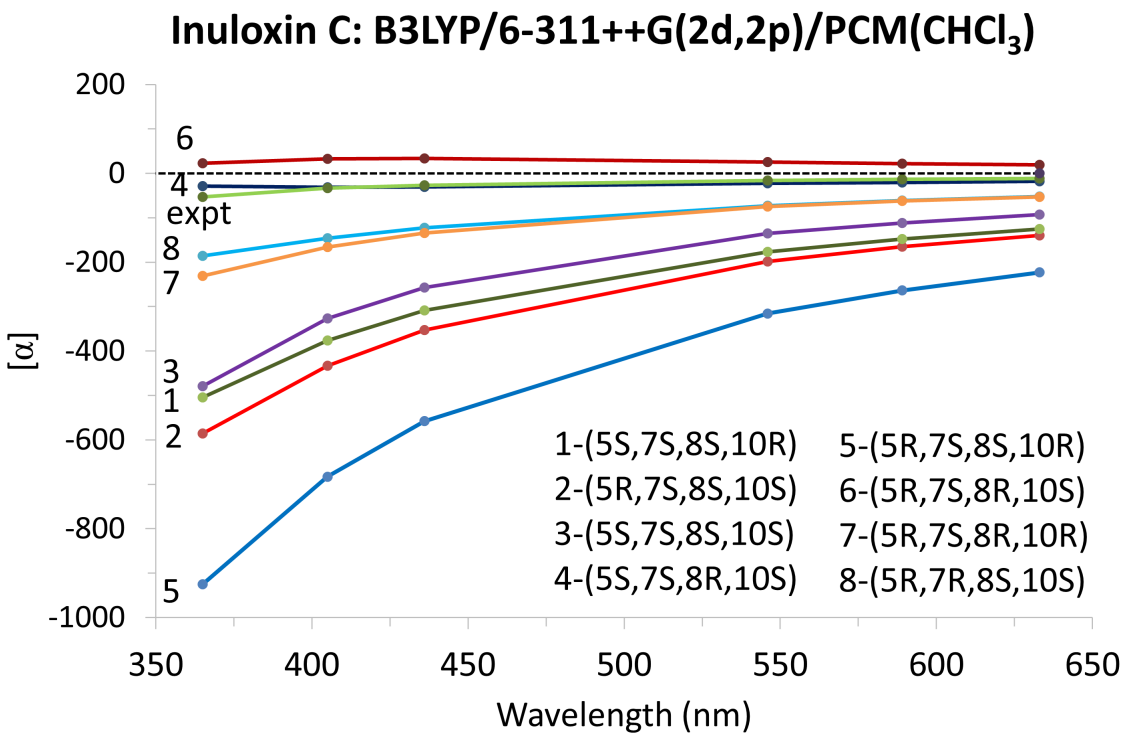

Figure 2.2: Comparison of experimental ORD (green trace labeled "expt") to those calculated for eight diastereomers of inuloxin C

The calculated EA and ECD spectra of the conformers of each diastereomer were multiplied with respective populations derived from electronic energies and summed. EDF spectra were calculated for each of the diastereomers with a molar extinction tolerance of 5 L mol^{-1} cm^{-1} for inuloxin C and 210 L mol^{-1} cm^{-1}

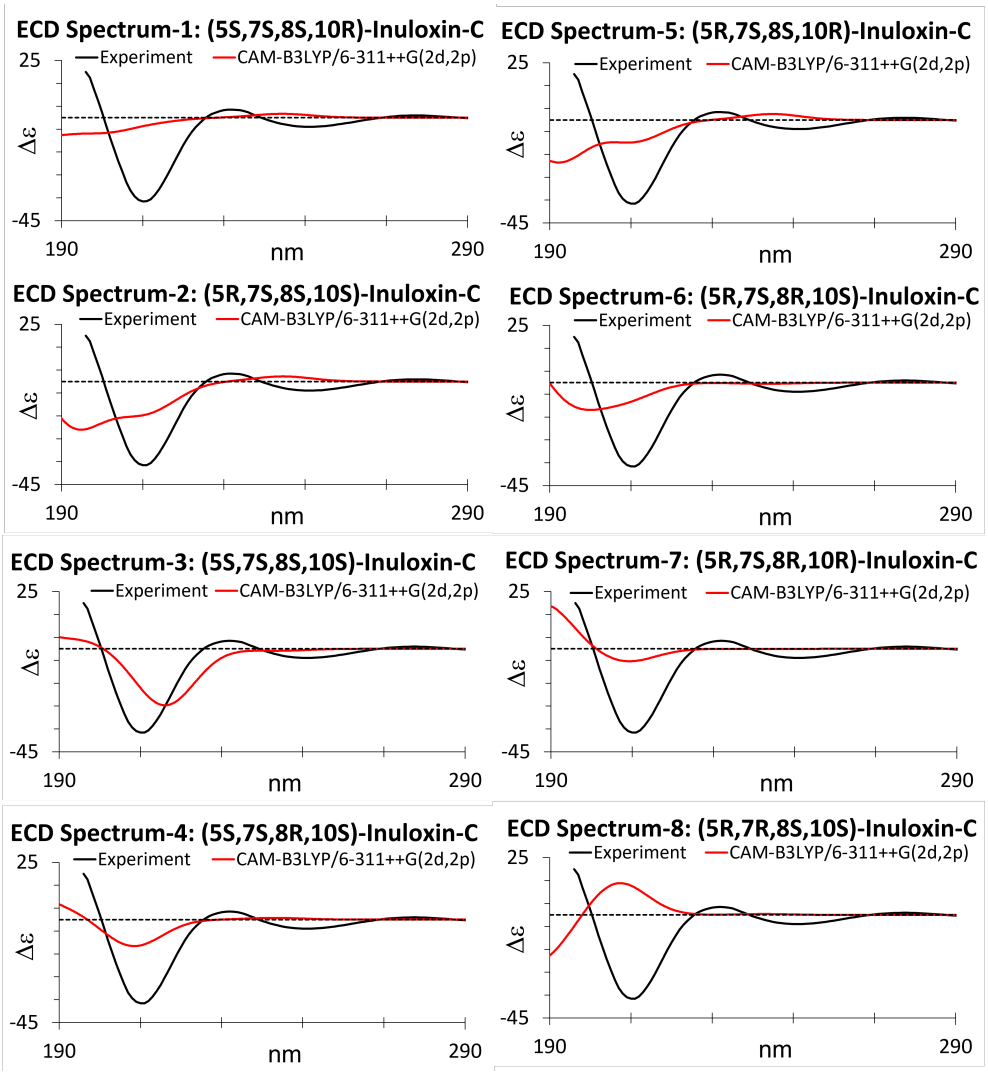

Figure 2.3: Comparison of experimental ECD (black trace) to those calculated for eight diastereomers (red traces) of inuloxin C

for acetylated inuloxin C. This tolerance was applied to both the experimental and calculated spectra. The EA, ECD, and EDF spectral simulations and similarity overlap analyses using Sim functions were undertaken using CDSpecTech program,[52] that is available free of charge.[53] Since the peak intensities for individual EA and ECD bands in the simulated spectra depend on the band widths assumed, gaussian band profiles with three different band widths namely 8, 10, and 12 nm were investigated. The *Sim*EA, *Sim*ECD, and *Sim*EDF values obtained with these different band widths did not alter the final conclusions, and therefore the results reported here are limited to 10-nm band width.

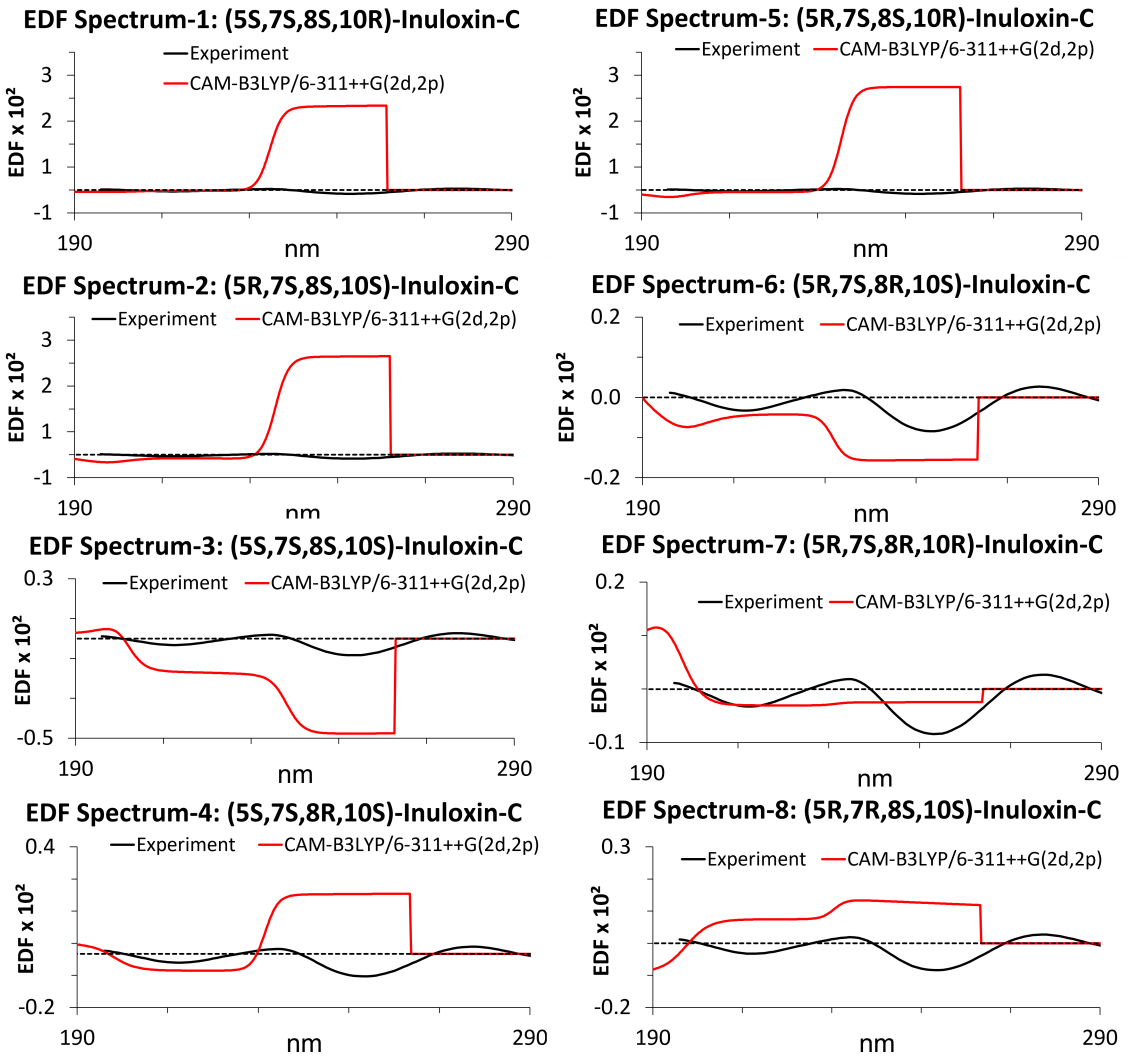

Figure 2.4: Comparison of experimental EDF (black trace) to those calculated for eight diastereomers (red traces) of inuloxin C

2.3 Results and Discussion

The comparison of calculated ORD signs with those in experimental data (Figure 2.2) indicates that seven of the eight diastereomers have the same signed ORD as experimental ORD. The ORD of diastereomer **6**-(5*R*,7*S*,8*R*,10*S*) (trace #6 in Figure 2.2) has opposite sign to that of experimental ORD, indicating that the mirror image diastereomer (5*S*,7*R*,8*S*,10*R*) will be a contender for the AC of inuloxin C, along with the other seven diastereomers listed in Figure 2.2.

The experimental UV spectrum displays absorption around 200 nm followed by a broad tail.[58] The ECD spectrum shows a positive Cotton effect (CE) at 197 nm, a negative CE at 210 nm, a weaker positive CE around 233 nm, a weak negative CE at 252 nm, and a weak positive CE at 278 nm (Figure 2.3). The weak

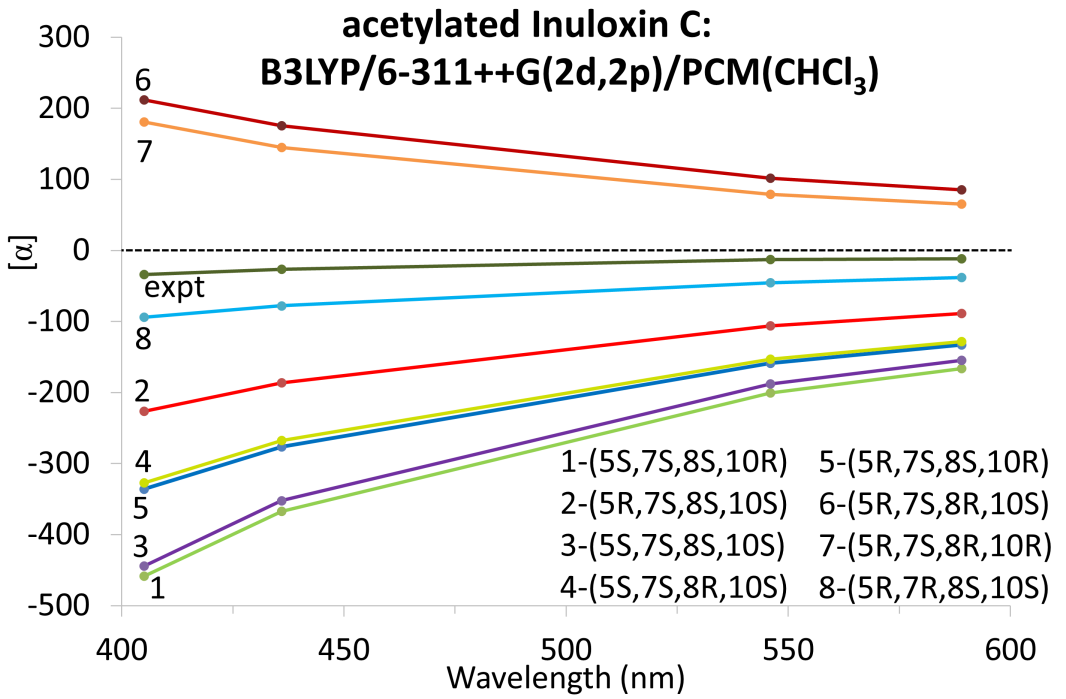

Figure 2.5: Comparison of experimental ORD (trace labeled "expt") to those calculated for eight diastereomers of acetylated inuloxin C

negative CE at 252 nm was not clearly seen in the spectrum reported earlier.[67] In a direct comparison of the most intense experimental CE at 210 nm against those in the calculated spectra, only the diastereomer **8**-(5*R*,7*R*,8*S*,10*S*) (spectrum #8 in Figure 2.3) lacks a corresponding negative CE around 210 nm, and instead has a positive CE. However, since the sign of predicted ORD data for this diastereomer (trace #8 in Figure 2.2) correlates correctly with that of experimental ORD, neither **8**-(5*R*,7*R*,8*S*,10*S*) nor its mirror image (5*S*,7*S*,8*R*,10*R*) can be one of the possible ACs for inuloxin C. Similarly, since the diastereomer **6**-(5*R*,7*S*,8*R*,10*S*) (spectrum #6 in Figure 2.3) has a matching negative CE (shifted slightly to shorter wavelength) but has a reverse ORD sign (trace #6 in Figure 2.2) compared with the experimental ORD, neither this diastereomer nor its mirror image (5*S*,7*R*,8*S*,10*R*) can be considered as one of the possible ACs. These results can be confirmed further using the ECD similarity overlap plots.[58] Note that positive ECD similarity overlap comes from agreement in experimental and predicted ECD signs. Large ECD similarity overlaps with positive magnitude (Table 2.2) are associated with remaining six diastereomers, rendering them possible candidates for ACs.

However, as previously emphasized,[46,73] it is important to analyze EDF, along with ECD, spectra. It was demonstrated previously[54] that a weak ECD band associated with large EDF was critical in establishing the AC of 3-Ishwarone.

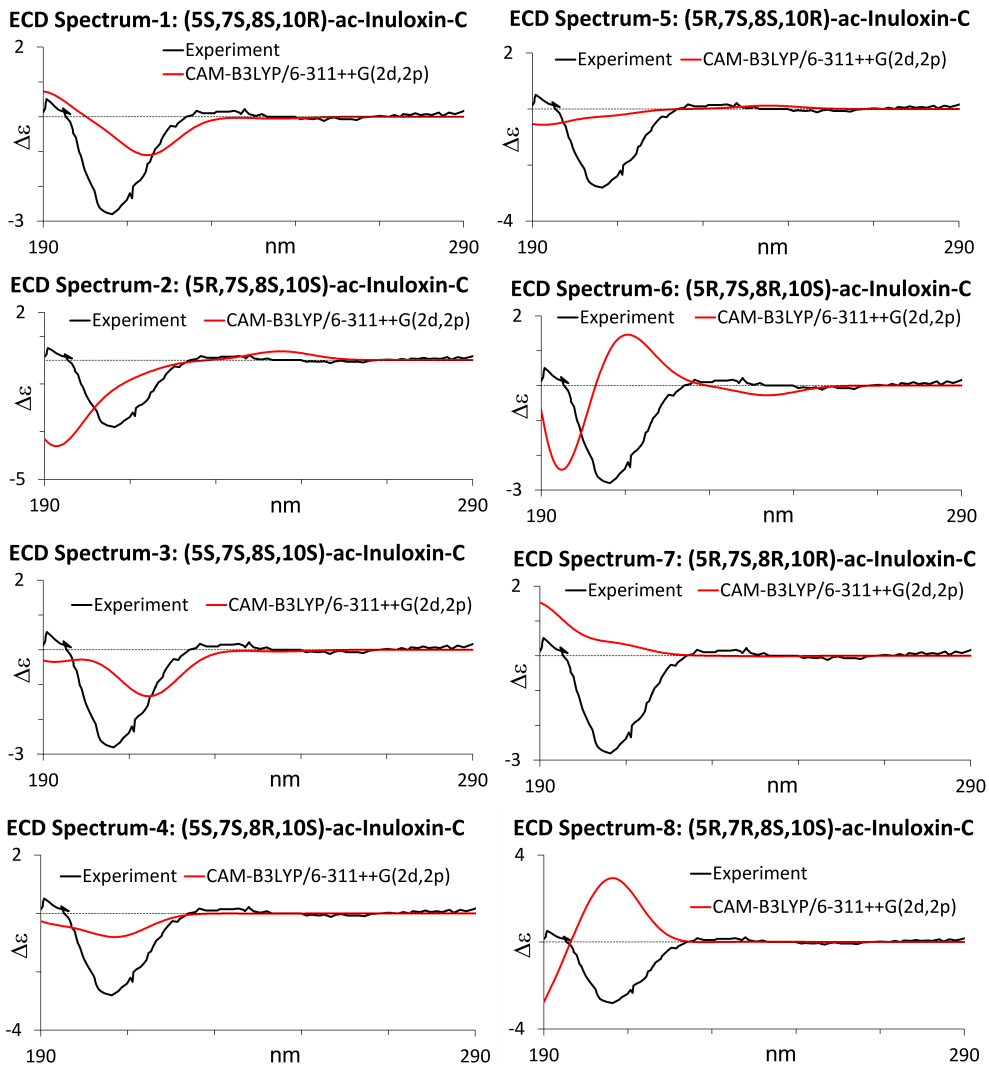

Figure 2.6: Comparison of experimental ECD (black trace) to those calculated for eight diastereomers (red traces) of acetylated inuloxin C

For the analysis of EDFs, I used similarity overlap *Sim*EDF calculated from normalized EDFs. The largest negative CE at 210 nm does not have the largest EDF, but a weak CE at 254 nm does (Figure 2.4). Therefore, the outcome of similarity EDF overlap analysis can be dominated by EDF associated with weak CE at 254 nm. Note that the positive EDF overlap comes from experimental and predicted EDFs of same sign. Further elimination of the diastereomers **1**-(5*S*,7*S*,8*S*,10*R*), **2**-(5*R*,7*S*,8*S*,10*S*), **4**-(5*S*,7*S*,8*R*,10*S*) and **5**-(5*R*,7*S*,8*S*,10*R*) results from the opposite EDF signs at ∼254 nm (Figure 2.4), resulting in negative EDF similarity overlap (Table 2.2). For these diastereomers even though ECD similarity overlap (which is dominated by strong negative CE at 210 nm) is positive, EDF overlap is negative because large EDFs of opposite sign are predicted for the weak CE at 254 nm. In Table 2.2, there are three diastereomers, **3**-(5*S*,7*S*,8*S*,10*S*), **6**-(5*R*,7*S*,8*R*,10*S*),

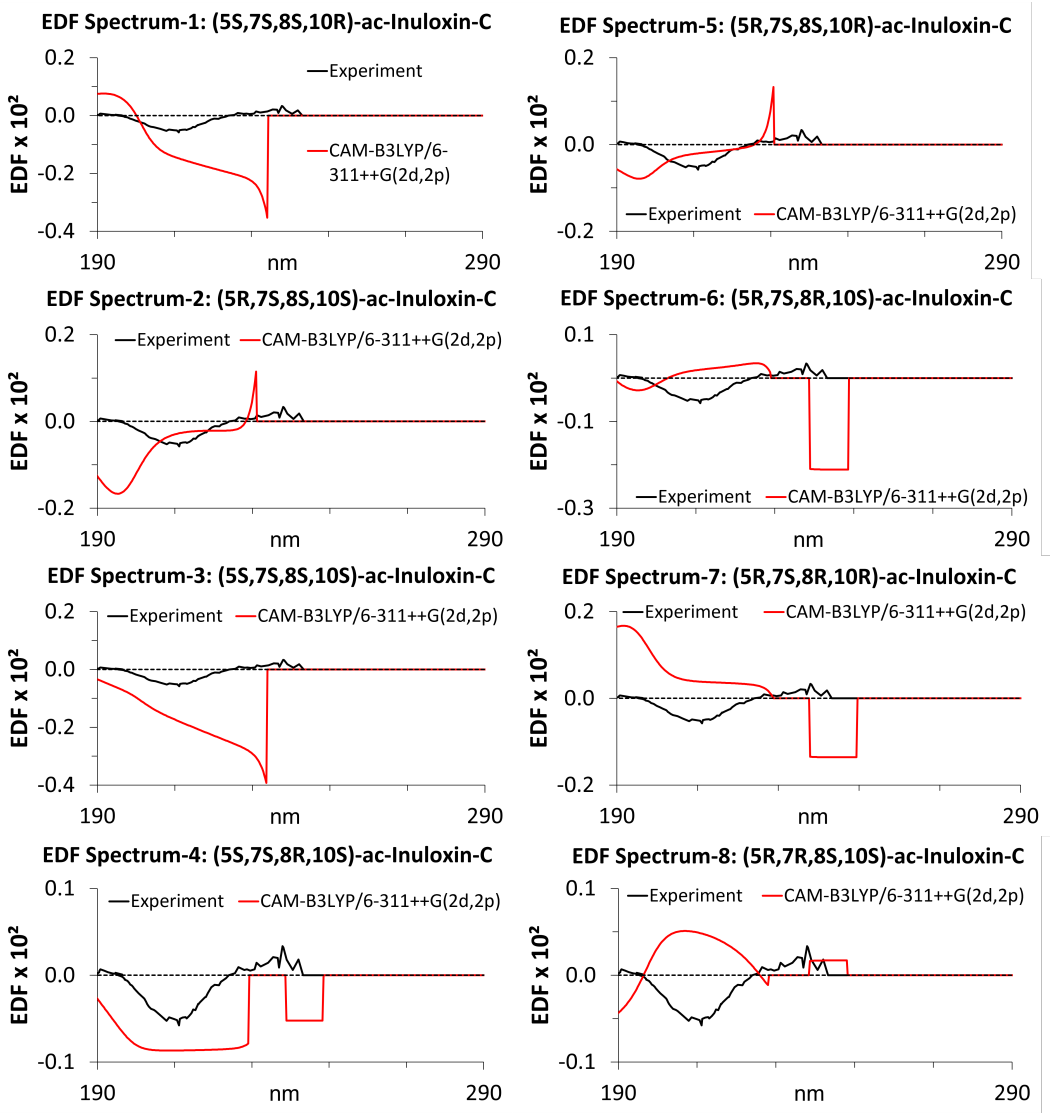

Figure 2.7: Comparison of experimental EDF (black trace) to those calculated for eight diastereomers (red traces) of acetylated inuloxin C

and **7**-(5*R*,7*S*,8*R*,10*R*), with positive EDF overlap, but **6**-(5*R*,7*S*,8*R*,10*S*) has already been eliminated earlier based on ORD analysis. Then, only two diastereomers **3**-(5*S*,7*S*,8*S*,10*S*) and **7**-(5*R*,7*S*,8*R*,10*R*) remain as possible candidates for AC.

Acetylation of inuloxin C was carried out by Evidente et al,[67] to curtail the rotational freedom and hydrogen bonding effects of the O-H group. Acetylation of O-H group does not alter the AC. Comparison of the sign of experimental ORD of acetylated inuloxin C with those calculated for eight diastereomers rules out the diastereomers **6**-(5*R*,7*S*,8*R*,10*S*) and **7**-(5*R*,7*S*,8*R*,10*R*) (traces #6 and 7 in Figure 2.5) as they are predicted to have opposite sign to that of experimental ORD (trace labeled "expt" in Figure 2.5). The ORD, ECD, and

Number	Diastereomer	*Sim*EA	*Sim*ECD	*Sim*EDF
1	(5*S*,7*S*,8*S*,10*R*)	0.94	0.61	-0.53
2	(5*R*,7*S*,8*S*,10*S*)	0.92	0.66	-0.53
3	(5*S*,7*S*,8*S*,10*S*)	0.90	0.92	0.52
4	(5*S*,7*S*,8*R*,10*S*)	0.92	0.94	-0.42
5	(5*R*,7*S*,8*S*,10*R*)	0.90	0.64	-0.52
6	(5*R*,7*S*,8*R*,10*S*)	0.95	0.86	0.48
7	(5*R*,7*S*,8*R*,10*R*)	0.91	0.93	0.33
8	(5*R*,7*R*,8*S*,10*S*)	0.93	-0.97	-0.42

Table 2.2: Maximum SSO values for EA, ECD and EDF spectra of inuloxin C.

EDF analyses of inuloxin C data (Figures 2-4 and Table 2.2) left two possibilities, **3**-(5*S*,7*S*,8*S*,10*S*) and **7**-(5*R*,7*S*,8*R*,10*R*) for the AC assignment. The ORD data of acetylated inuloxin C eliminated the diastereomer **7**-(5*R*,7*S*,8*R*,10*R*) as a possible candidate. That leaves the **3**-(5*S*,7*S*,8*S*,10*S*) as the only remaining possibility for the AC of inuloxin C.

To substantiate this conclusion, I analyzed the ECD and EDF spectra of all diastereomers of acetylated inuloxin C (Figures 6, 7).The experimental ECD spectrum of acetylated inuloxin C showed a strong negative CE (Figure 2.6) around 210 nm and a weak positive CE at 238 nm, with the former associated with stronger EDF (Figure 2.7). The comparisons of experimental and calculated ECD and EDF spectra for acetylated inuloxin C are shown in Figures 6 and 7 and maximum similarity overlaps summarized in Table 2.3.

Table 2.3 highlights five diastereomers of acetylated inuloxin C, with "comment a," for which the ECD and EDF similarity overlaps are large and positive (ie, matching experimental and calculated ECD band signs dominate the overlap); for two diastereomers, **6**-(5*R*,7*S*,8*R*,10*S*) and **7**-(5*R*,7*S*,8*R*,10*R*) highlighted with "comment b," for which the similarity overlaps for ECD or EDF are below the reliability threshold of 0.4 magnitude[23,52,51] and **8**-(5*R*,7*R*,8*S*,10*S*) highlighted with "comment c," the ECD similarity overlaps are large but negative (ie, the dominant experimental and calculated ECD band signs are opposite to each other). Thus, diastereomer **7**-(5*R*,7*S*,8*R*,10*R*) is ruled out from the ECD analyses as well.

All the results discussed so far for inuloxin C and acetylated inuloxin C are summarized in Table 2.4. In this table, each row summarizes the status of a given diastereomer as evaluated from ECD, EDF, and ORD analysis of inuloxin C and acetylated inuloxin C. Under a given chiroptical spectroscopic method, an entry "Y" identifies a given diastereomer as a possible contender for AC, while "M" identifies the mirror image diastereomer as a possible contender. The diastereomer **3**-(5*S*,7*S*,8*S*,10*S*) with all entries as "Y" represents

Number	Diastereomer	*Sim*EA	*Sim*ECD	*Sim*EDF	Comment
1	(5*S*,7*S*,8*S*,10*R*)	0.92	0.95	0.67	a
2	(5*R*,7*S*,8*S*,10*S*)	0.93	0.92	0.80	a
3	(5*S*,7*S*,8*S*,10*S*)	0.91	0.91	0.67	a
4	(5*S*,7*S*,8*R*,10*S*)	0.93	0.84	0.49	a
5	(5*R*,7*S*,8*S*,10*R*)	0.92	0.69	0.66	a
6	(5*R*,7*S*,8*R*,10*S*)	0.94	−0.71	−0.27	b
7	(5*R*,7*S*,8*R*,10*R*)	0.93	−0.90	−0.51	b
8	(5*R*,7*R*,8*S*,10*S*)	0.93	−0.91	−0.73	c

Table 2.3: Maximum SSO values for EA, ECD and EDF spectra of acetylated inuloxin C.

the unique choice for AC of inuloxin C and acetylated inuloxin C. Thus, the combined ECD, EDF, and ORD analyses for all of the diastereomers of inuloxin C and acetylated inuloxin C identify diastereomer **3**-(5*S*,7*S*,8*S*,10*S*) as the only possible choice for AC of inuloxin C and acetylated inuloxin C. This conclusion is independent of the relative configuration determined from NMR.

Number	Diastereomer	Inuloxin C			Acetylated Inuloxin C		
		ECD	EDF	ORD	ECD	EDF	ORD
1	(5*S*,7*S*,8*S*,10*R*)	Y	M	Y	Y	Y	Y
2	(5*R*,7*S*,8*S*,10*S*)	Y	M	Y	Y	Y	Y
3	(5*S*,7*S*,8*S*,10*S*)	Y	Y	Y	Y	Y	Y
4	(5*S*,7*S*,8*R*,10*S*)	Y	M	Y	Y	Y	Y
5	(5*R*,7*S*,8*S*,10*R*)	Y	M	Y	Y	Y	Y
6	(5*R*,7*S*,8*R*,10*S*)	Y	Y	M	M	M	M
7	(5*R*,7*S*,8*R*,10*R*)	Y	Y	Y	M	M	M
8	(5*R*,7*R*,8*S*,10*S*)	M	M	Y	M	M	Y

Table 2.4: Possible contenders for AC of inuloxin C and acetylated inuloxin C as deduced from ECD, EDF, and ORD analyses.

The positive outcome of the current ECD, EDF, and ORD analyses is attributed to the following factors: (1) QC predictions of ECD and ORD for all diastereomers of inuloxin C and acetylated inuloxin C; (2) EDF spectral analysis for inuloxin C has helped eliminate some of the possible diastereomers and accentuated the important role of a weak ECD band at 252 nm. Some limitations in the current analysis for inuloxin C should also be kept in mind. For the diastereomer, **3**-(5*S*,7*S*,8*S*,10*S*), assigned as the correct one: (1) the

predicted ORD magnitudes are significantly larger than those observed in the experiment, both for inuloxin C and acetylated inuloxin C; (2) the predicted magnitude of EA is significantly smaller than that observed in the experiment for inuloxin C; (3) the predicted magnitude of ECD is significantly larger than that observed in the experiment for acetylated inuloxin C. The later two discrepancies would not be revealed in normalized similarity overlap calculations. The sources for these discrepancies have not been determined.

2.4 Conclusion

ORD, ECD, and EDF analyses of all diastereomers of inuloxin C lead to two diastereomers as possible candidates, one of which is **3**-(5*S*,7*S*,8*S*,10*S*), the AC assigned previously.[67] When the ORD, ECD, and EDF data of acetylated inuloxin C are included in the analysis, there is sufficient evidence to favor the **3**-(5*S*,7*S*,8*S*,10*S*) diastereomer for AC assignment. This conclusion is independent of the relative configuration deduced from NMR.[59] Thus, the primary conclusions emerging from this study are as follows: (1) it is important not to ignore weak ECD bands as unimportant because they might be associated with larger EDFs; (2) dissymmetry factor analysis, along with conventional circular dichroism spectral analyses, is helpful to reduce the number of diastereomers that correlate with experimental data; (3) the use of a single parent compound, let alone single chiroptical spectroscopic method, may be insufficient in ascertaining the AC of chiral molecules with multiple stereogenic centers. Experimental and calculated chiroptical spectroscopic data for the stereochemically unaltered derivatives of the parent compounds might help eliminate some of the diastereomers as possible contenders.

CHAPTER 3

The AC of (-)-inuloxin D

This work was completed in collaboration with Professor Antonio Evidente of the University of Naples Federico II, Naples, Italy, Professor Nina Berova of Columbia University, New York, NY, Professor Ana Petrovic of the New York Institute of Technology, New York, NY, and Professor Stefano Superchi of the University of Basilicata, Potenza, Italy. This chapter has been reproduced from the *Chirality* article[56] "Absolute configuration of seco-eudes-manolide inuloxin D from experimental and predicted chiroptical studies of its 4-*O*-acetyl derivative" with permission. My contributions to this work are: QC spectral prediction of VCD, ECD, and ORD, spectral similarity analyses, preparation of figures and tables, drafting the manuscript, and revising the manuscript.

3.1 Introduction

Dittrichia viscosa (syn. *Inula viscosa*), belonging to Asteraceae family tribe Inulae, is a native plant of Mediterranean basin well known for its application in folk medicine and agriculture. A recent review[74] reports the different classes of natural occurring compounds isolated from this plant as well as their biological properties. Besides flavonoids and caffeic acids analogs, the sesquiterpenes are the most represented metabolites isolated.

Four new sesquiterpenes, named inuloxins A–D (**1**,**4**-**6**, see Figure 3.1), belonging to subgroups of germacrane (A and B) eudesmanolide (C) and seco-eudesmanolide (D), were isolated from *D. viscosa* and tested for their potential herbicidal activity on two parasitic plants as *Orobanche crenata* and *Cuscuta campestris*. Inuloxins A, C, and D were the most active on both parasites and caused up to 100% inhibition of the seed germination. Inuloxin B (**5**) was less active on *Cuscuta* and completely inactive against *Orobanche*.[59]

Inuloxins A, C, and D showed strong activity against *Leshmania donovani*, vector of leishmaniosis transmittable to human with inuloxin A (**4**) being the most active.[75] Inuloxin B (**5**) and some semisynthetic derivatives prepared from inuloxins A, C, and D showed a weak or nil activity. These results, in agreement with those previously obtained assaying phytotoxicity, showed that the furanone ring and the exocyclic methylene group are important features to impart this activity. The hydroxyl group of the side chain of inuloxin D (**1**) also seems to play a role in this activity.[76] A potential application of inuloxin A (**4**) as natural herbicide induced further study, as those recently reported on its formulation in β-cyclodextrin which, increasing its solubility in water, could permit its use without any organic solvent.[76]

Considering the noteworthy biological activity of inuloxins and that the absolute configuration (AC) is

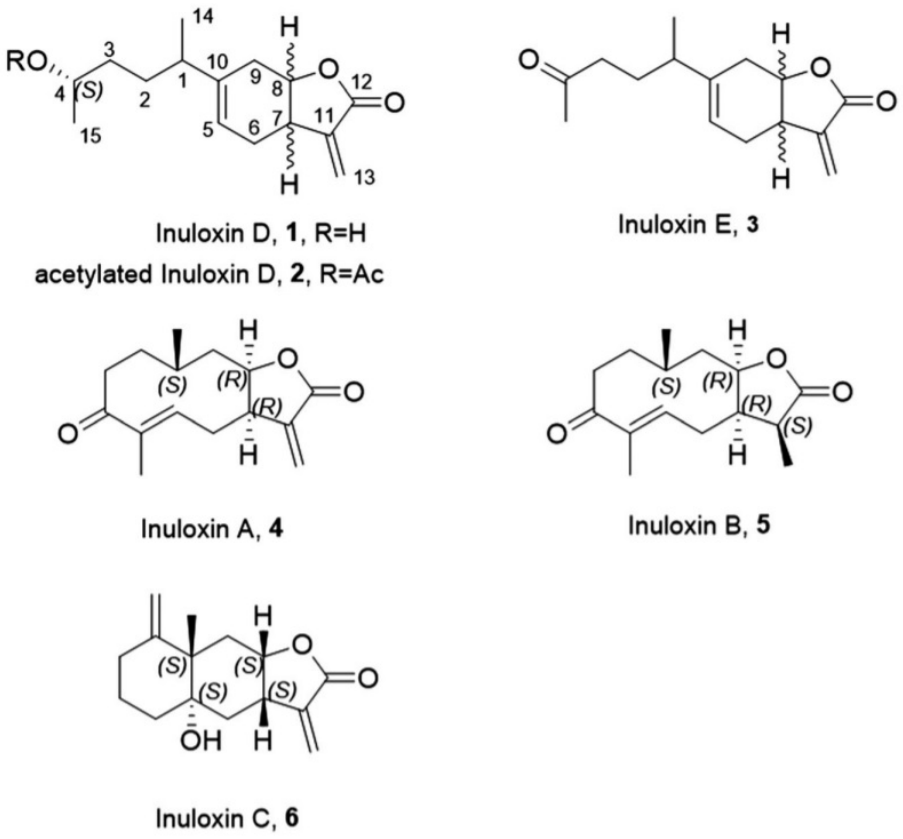

Figure 3.1: Inuloxins **1-6**. The ACs of stereocenters in **1, 2,** and **3** are not shown, except for cis relative configurations at C-7 and C-8

a structural feature strictly connected to biological activity,[66,65] the AC (7*R*,8*R*,10*S*) of (+)-inuloxin A (**4**) was determined by experimental measurements and theoretical simulations of chiroptical properties of three related methods, namely, ORD, ECD, and VCD.[68] Subsequently, the ACs to the naturally occurring (+)-inuloxin B (**5**) and (-)-inuloxin C (**6**) were assigned as (7*R*,8*R*,10*S*,11*S*) and (5*S*,7*S*,8*S*,10*S*), respectively, using the same methods.[67] The AC assigned to **6** was later confirmed also using its acetyl derivative, as a case study to assign the AC to chiral molecules with multiple stereogenic centers without prior knowledge of the relative configurations.[58]

Recently, a new seco-eudesmanolide, named inuloxin E (**3**, Figure 3.1), was isolated from the organic extract of *D. viscosa*. Its structure, determined using spectroscopic data, showed that it differs from inuloxin D (**1**) for having a ketone group in place of the secondary hydroxyl group at C-4 in the hexyl chain attached to C-10 of the α,β-unsaturated γ-lactone moiety.[77] Stereoselective reduction of (+)-inuloxin E (**3**) by $NaBH_4$ afforded naturally occurring (+)-inuloxin D (**1**),[77] showing that these two compounds share the same AC at C-1, C-7 and C-8. Inuloxins D (**1**) and E (**3**) were compared with four different inuloxins (see Figure 3.1) on

seed germination assays of *Orobanche cumana*, *Orobanche minor*, and *Pelipanche ramosa*. **1** and **3** induced germination of the parasitic weed *O. cumana* but were inactive on seeds of *O. minor* and *P. ramosa*.[77] This selective action is important from a practical point view as both **1** and **3** could be used to develop a so-called "suicidal germination" method[63,78,79] for the specific control of the parasitic weed *O. cumana*. This represents an innovative alternative for the control of a very dangerous parasitic weed based on a suitable formulation of plant metabolites that could be obtained conveniently in large amount or by cultivation of *D. viscosa* and/or developing a convenient and environmentally friendly synthesis of both inuloxins.

During the isolation and chemical characterization of inuloxin D carried out by collaborators, only the relative cis configuration, implying either *R,R* or *S,S* AC, was assigned to both the headbridge carbons C-7 and C-8 between the six- and five-membered rings.[59] This assignment was based on the correlations observed in the NOESY spectrum and the $^{3}J_{H,H}$ coupling constants measured in the ^{1}H NMR spectra. The AC at the C-4 of the hexyl side chain was assigned as (*S*) by application of advanced Mosher's methods.[64] This assignment is independently verified by this work. The AC at C-1 remained unassigned. As shown above, inuloxins D and E share the same AC at C-7, C-8, and C-1, and therefore, the determination of AC of inuloxin D will permit the assignment of AC also to the related inuloxin E.

The structures of inuloxin D (**1**), acetylated inuloxin D (**2**), inuloxin E (**3**), inuloxin A (**4**), inuloxin B (**5**), and inuloxin C (**6**) are shown in Figure 3.1. Since the relative configurations at C-7 and C-8 are fixed for **1**, **2**, and **3**, a total of 2^3 stereoisomers are possible for **1** and **2**, of which four are mirror images of the other four. Since enantiomers have oppositely signed ECD, ORD, and VCD, it is sufficient to predict the chiroptical properties for four stereoisomers. The diastereomers used for calculations in this work are as follows: **2A**: (1*S*,4*S*,7*S*,8*S*), **2B**: (1*S*,4*S*,7*R*,8*R*), **2C**: (1*R*,4*S*,7*R*,8*R*), and **2D**: (1*R*,4*S*,7*S*,8*S*). The enantiomers of **2A** and **2D** used for spectra analyses (vide infra) are labeled, respectively, as *ent*-**2A** and *ent*-**2D**. The assignment of ACs to **1**, **2**, and **3** is complicated by the fact that each of these compounds contains multiple stereogenic centers. The high flexibility of the alkyl chain containing C-1 and C-4 renders the AC assignment at C-1 particularly challenging. Intramolecular and intermolecular hydrogen bonding is possible in **1**, but these effects are avoided in **2**, the acetyl derivative of **1**. Since acetylation does not alter the AC, the AC deduced from chiroptical spectroscopic studies on **2** applies to the AC of **1** as well. For these reasons, ECD, VCD, and ORD experiments and calculations were undertaken in the present work for **2**, to determine an unbiased AC for **2**, and, by extension, also for **1** and **3**.

3.2 Methods

The isolation and characterization of inuloxin D and E as well as the acetylation of inuloxin D were carried out by our Italian collaborators and has been described in the literature[56,59,77] and below. The ORD and VCD

spectra used within were measured by Dr. E. Santoro while a post doctoral researcher within our lab and details of these measurements are provided below and in the main text.[56]. ECD measurements utilized for analysis were measured by Dr. A. Petrovic, a former graduate student within our lab, during her tenure at the New York Institute of Technology.

Inuloxin D, **1**, was purified from the organic extract of the aerial parts of *D. Viscosa* as previously reported.[59,77] Its 4-*O*-acetyl derivative **2** was obtained by common acetylation of **1** with pyridine and acetic anhydride as previously reported in detail.[77] ECD, VCD, and ORD were measured for **2** as follows: ECD spectra were measured using Jasco J810 spectrometer in acetonitrile solvent with 1 cm path length at concentration 4.4×10^{-5} M. VCD spectra were measured using ChiralIR instrument in deuterated chloroform solvent with 100 μm path length at concentration 3.07×10^{-1} M. ORD was measured using Autopol IV polarimeter in chloroform solvent with 0.5 dm path length at concentration 3.55×10^{-3} g/ml at six discrete wavelengths (633, 589, 546, 436, 405, and 365 nm).

Label	**Diastereomer**	**No. of conformers**
2A	(5*S*,7*S*,8*S*,10*R*)	35
***ent*-2A**	(5*R*,7*S*,8*S*,10*S*)	35
2B	(5*S*,7*S*,8*S*,10*S*)	37
2C	(5*S*,7*S*,8*R*,10*S*)	40
2D	(5*R*,7*S*,8*S*,10*R*)	42
***ent*-2D**	(5*R*,7*S*,8*R*,10*S*)	42

Table 3.1: Notation for stereoisomers of **2** and the number of investigated conformations.

The initial conformational searches were undertaken within a 20 kcal/mol energy window using Conflex program[71] for four diastereomers of **2**. Subsequent geometry optimizations and spectral calculations were undertaken with Gaussian 16.[80] All conformers of diastereomers of **2** obtained with Conflex were optimized at PM6 level. All unique conformers obtained in PM6 optimization were then optimized at B3LYP/6-31G* level. The unique conformers of diastereomers of **2** were optimized further at B3LYP/6-311++G(2d,2p) level. The lowest energy (within 2 kcal/mol) unique conformers of **2** at this level were further optimized at B3LYP/6-311++G(2d,2p) level with polarizable continuum model (PCM)[44,81] for $CHCl_3$ solvent. ECD, VCD, and ORD calculations for these conformers of **2** were then carried out at B3LYP/6-311++G (2d,2p)/PCM ($CHCl_3$) level, except for ECD calculations which utilized the CAM-B3LYP functional and PCM for CH_3CN. The specific optical rotation (SOR) predicted at 633, 589, 546, 436, 405, and 365 nm was used to generate ORD. The quantitative comparison between experimental and calculated ORD is obtained[57]

by calculating the root mean square deviation (RMSD) as described in section 1.3.6.3:

To assess the differences between ECD predictions performed with PCM ($CHCl_3$) and PCM (CH_3CN), geometry optimizations and ECD predictions were also undertaken at CAM-B3LYP/6-311++G(2d,2p)/PCM (CH_3CN) level for all conformers of diastereomer **2C**. Since no significant differences were found in these calculations, ECD calculations for diastereomers **2A**, **2B**, and **2D** were undertaken with CAM-B3LYP functional and 6-311++G(2d,2p) basis set utilizing PCM(CH_3CN) for implicit solvation and using the optimized geometries and Gibbs energies at B3LYP/6-311++G(2d,2p)/PCM($CHCl_3$) level.

ECD and VCD spectral analyses were undertaken using CDSpecTech program,[52,53] which implements the Tanimoto's similarity function,[82,83] as proposed by Shen et al.[50] and designated here as *Sim*XXX function, where XXX represents the type of spectrum (XXX = VA, VCD, VDF, EA, ECD, EDF etc). Further detail on SSO analysis is described in section 1.3.6.2.

3.3 Results and Discussion

Since there is no possibility for intramolecular hydrogen bonding in **2**, conformational search is less complicated. Nevertheless, there are a large number of conformers due to the flexibility of side chain containing C1-C4. For each diastereomer, the numbers of conformers within 2 kcal/mol electronic energy difference, that have been used for final vibrational frequency and spectral calculations, are listed in Table 3.1. The lowest energy conformer of each diastereomer of **2** has the following populations, based on Gibbs energy: **2A**: 22%, **2B**: 21%, **2C**:15%, and **2D**:14%. It is important to note that since the lowest energy conformer of each diastereomer accounts for only less than 25% of the total population, multiple conformers can be regarded as important for generating the population weighted spectra. Moreover, the population weighting using Gibbs and electronic energies did not significantly alter the final result.

ECD, VCD, and ORD spectral analyses for diastereomers of **2** are discussed below. As mentioned earlier, ECD calculations were undertaken for **2C** at both CAM-B3LYP/6-311++G(2d,2p)/PCM (CH_3CN) and B3LYP/6-311++G(2d,2p)/PCM ($CHCl_3$) levels, using Gibbs energies at the respective optimized geometries, for population weighting. These calculations reveal that there are no significant differences between their results.[56] *Sim*EA, *Sim*ECD, and *Sim*EDF values are 0.94, 0.41, and 0.87 for calculation with PCM ($CHCl_3$) and 0.94, 0.40, and 0.85 for calculation with PCM (CH_3CN). For these reasons, ECD calculations for the other three diastereomers of **2** were carried out using the geometries optimized with PCM ($CHCl_3$), carrying out ECD calculations with PCM (CH_3CN) and using Gibbs energies at the optimized geometries for weighting.

The comparison of calculated EA, ECD, and EDF spectra for *ent*-**2A**, **2B**, **2C**, and *ent*-**2D** with the experimental spectra are given by Johnson et al.[56] The ECD spectra for these four diastereomers, with 7*R*, 8*R* configurations, show relatively strong matches with experimental ECD spectra, having the nega-

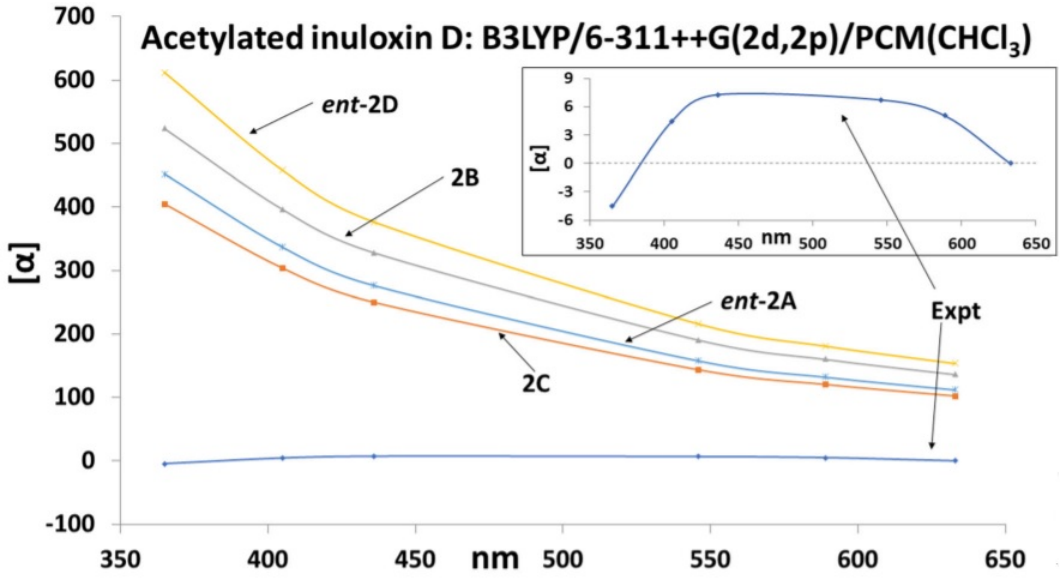

Figure 3.2: Comparison of experimental ORD to those for four diastereomers of acetylated inuloxin D. The inset shows experimental ORD on an expanded y-axis scale

tive–positive–negative pattern of Cotton Effects (CEs). This observation indicates that ECD of **2** is mostly determined by 7*R*, 8*R* configuration of the unsaturated lactone chromophore of the molecule and that ECD of **2** is less sensitive to C-1 and C-4 stereocenters, which are far from the chromophore. Quantitative analysis using the similarity between calculated and experimental spectra, in such cases, especially when there are only few bands for comparison, can be misleading and need not provide additional information about ACs of stereocenters C-1 and C-4. For this reason, numerical values for quantitative similarity measures, *Sim*EA, *Sim*ECD, and *Sim*EDF, are delegated to the article.[56]

Experimental ORD measurements revealed very small positive SOR values, with a sign switch at shorter wavelength (Figure 3.2). ORD predicted for **2A** has been inverted to obtain that for *ent*-**2A**; ORD predicted for **2D** has been inverted to obtain that for *ent*-**2D**. The sign reversal at the shorter wavelength is not reproduced in the calculations for any of the diastereomers, but positive signs observed in the experiment at longer wavelengths are reproduced in the predicted values for *ent*-**2A**, **2B**, **2C**, and *ent*-**2D**, all of them with the 7*R*, 8*R* configuration. Thus, ORD reaffirms the 7*R*, 8*R* configuration that was inferred from ECD. There are some differences among the magnitudes of predicted specific rotations for *ent*-**2A**, **2B**, **2C**, and *ent*-**2D** (Figure 3.2), and the magnitudes predicted for **2C** are closer to the experimental magnitudes. Quantitative analysis[57] indicates that the RMSD values for *ent*-**2A**, **2B**, **2C**, and *ent*-**2D** are, respectively, 272, 320, 244, and 370. Based on these RMSD values, **2C** is the preferred diastereomer for the AC assignment. However, the differences in RMSD magnitudes are not striking and one should not rely on ORD alone for assigning the

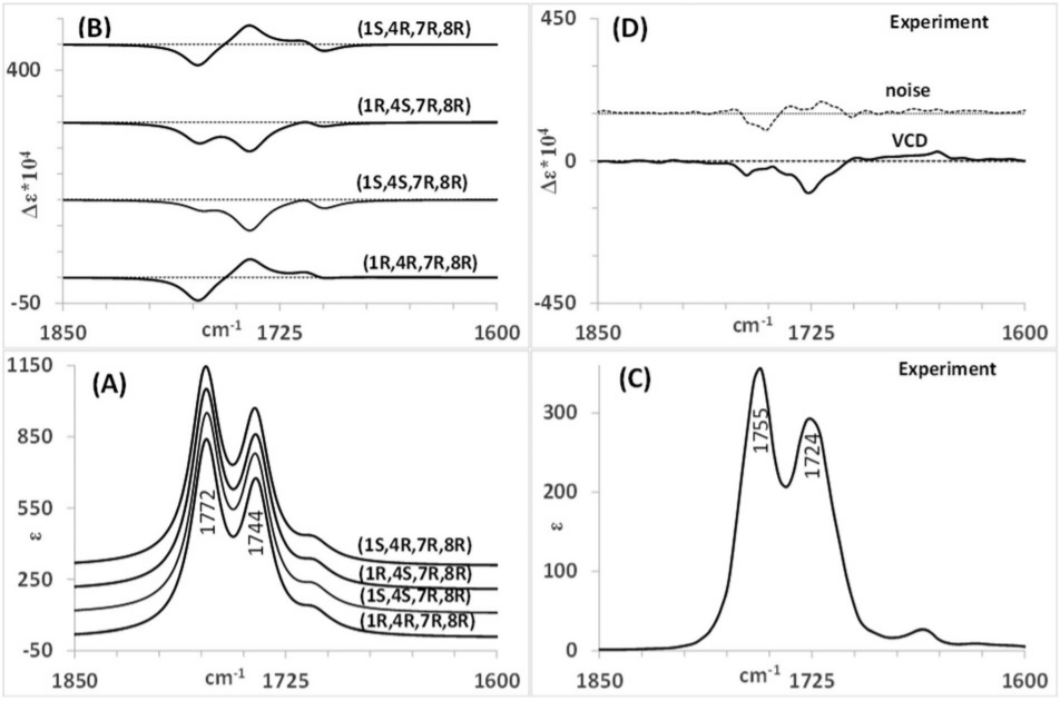

Figure 3.3: Predicted VA and VCD (panels A and B respectively) and experimental VA and VCD (panels C and D respectively) of **2** in the C=O stretching region

unknown AC and this conclusion needs to be verified independently (vide infra).

One might also wonder about the large differences in the magnitudes of predicted and observed specific rotations for **2**, but such large discrepancies were also reported previously for all inuloxins (A, B, C, and acetylated C) investigated to date.[58,67,68] A satisfactory explanation for these large differences remains elusive.

The predicted VA, VCD, and VDF spectra for the diastereomers of **2** are compared with corresponding experimental spectra in the manuscript.[56] Quantitative similarity values are also provided there. The vibrational spectra in the C=O stretching region (1800–1600 cm^{-1} region) and finger print region (1550–950 cm^{-1} region) are discussed separately below.

The C=O group of the α,β-unsaturated γ-lactone moiety has unscaled predicted vibrational stretching frequency of 1772 cm^{-1} and that of acetyl group has 1744 cm^{-1} (see Figure 3.3). The former is associated with predicted negative VCD of smaller intensity for *ent*-**2A**, **2B**, **2C**, and *ent*-**2D**, all with 7*R*,8*R* configuration. Therefore, the predicted negative VCD band at 1772 cm^{-1} appears to be diagnostic of ACs at C-7 and C-8. The predicted acetyl C=O stretching vibrational band at 1744 cm^{-1} has positive VCD for *ent*-**2A** and *ent*-**2D** with 4*R* configuration. This positive VCD changes to negative VCD for the other two diastereomers, **2B** and **2C**, with 4*S* configuration. This observation indicates that for the opposite ACs at C-4, the predicted

VCD sign associated with acetyl carbonyl stretching vibration changes. Therefore, VCD sign associated with acetyl C=O stretching vibration appears to be diagnostic of the AC at C-4. The experimental absorption spectrum indicates two vibrational bands at 1755 and 1724 cm^{-1} , which by correlation with corresponding predicted absorption band positions can be associated with C=O stretching, respectively, of α,β-unsaturated γ-lactone moiety and acetyl group. The former is associated with very weak negative experimental VCD close to the noise level, and the latter has negative experimental VCD above the noise level. The observed weak negative VCD associated with 1755 cm^{-1} band matches with that associated with predicted C=O stretching vibration of the α,β-unsaturated γ-lactone moiety, supporting 7*R*,8*R* configuration, which is affirmed already from ECD and ORD analyses. More importantly, the observed negative experimental VCD at 1724 cm^{-1} matches the negative VCD predicted for 4*S* configuration. Then the negative experimental VCD associated with acetyl group C=O stretching vibration supports (*S*) configuration at C-4, which is in agreement with the earlier conclusion[64] based on Mosher's method.

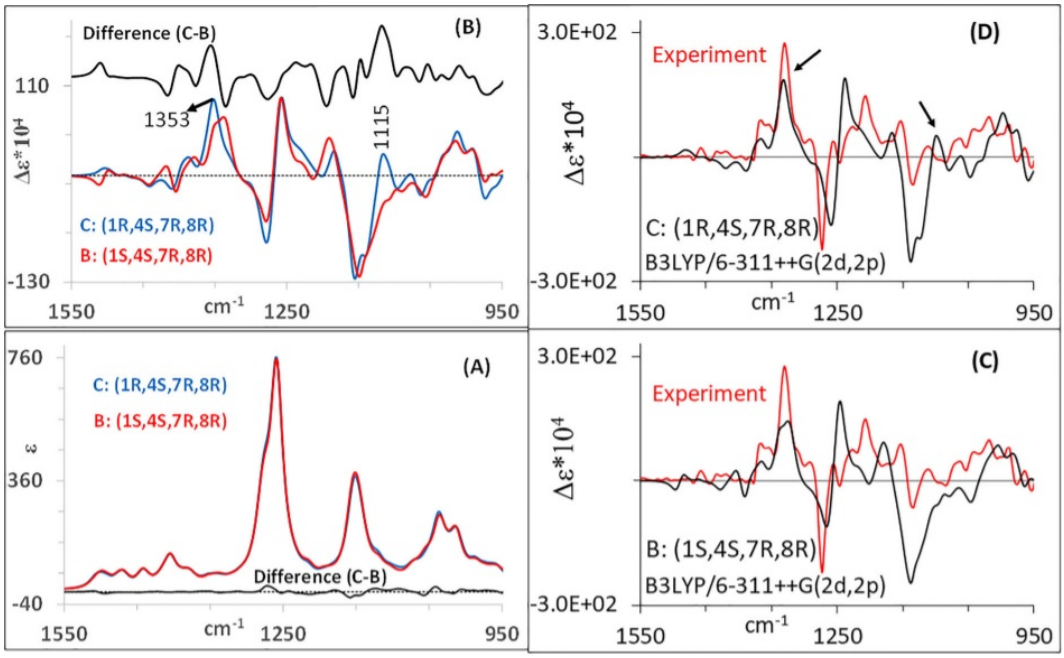

Figure 3.4: Predicted VA (panel A) and VCD (panel B) spectra for **2B** and **2C** and the difference between them. Comparison of experimental VCD spectrum with that predicted for **2B** (panel C) and for **2C** (panel D)

Having established the configurations at C-4, C-7, and C-8, I turned my attention to the AC at C-1. For this purpose, I undertook difference spectral analysis, by enquiring what differences would be predicted if I inverted the configuration at C-1 while keeping the ACs at other three chiral centers the same. The spectra predicted in the 1550–950 cm^{-1} region for **2C** and **2B** and the differences between their spectra are shown in Panels A and B of Figure 3.4. While differences in the absorption spectra are very small, some differences in VCD spectra are noteworthy: (a) the weak positive VCD predicted for unscaled frequency at 1115 cm^{-1}

of **2C** changes sign for **2B** and (b) the positive VCD predicted for unscaled frequency at 1353 cm^{-1} for **2C** is shifted to a slightly lower frequency for **2B** and becomes broad due to a new shoulder band. These differences are reflected as larger magnitudes in the difference spectrum (Panel (B) of Figure 3.4). From the animated vibrations associated with 1115 and 1353 cm^{-1} frequencies, analyzed using Gaussview,[84] these bands are noted to have contributions from the C–H bending at C-1, although several other C–H groups also contribute to these frequencies. In the comparison of the predicted spectra with experimental spectra (Panels C and D of Figure 3.4), the corresponding positions for the two changes noted above are marked with arrows. The predicted frequencies for **2B** were scaled with 0.99 and those for **2C** with 0.985 in Panel C and D and the predicted intensities of these VCD spectra have been multiplied by 2 to appear on the same scale as the corresponding experimental spectra. The sign reversal for predicted weak VCD band at 1115 cm^{-1} when going from **2C** to **2B** makes the comparison with experimental VCD spectrum less favorable for **2B**. This is because predicted VCD for **2B** in the 1050–1150 cm^{-1} region becomes broad negative, which is not seen in the experimental spectrum. Similarly, the shifting, splitting, and broadening of predicted VCD band at 1353 cm^{-1} when going from **2C** to **2B** makes the comparison with experimental positive VCD band at 1330 cm^{-1} less favorable for **2B**. These observations render preference for the (*R*) configuration at C-1. Admittedly, the noted changes are subtle and there is no one obvious band that can be identified as a diagnostic of AC at C-1.

Further support for the (*R*) configuration at C-1 can be obtained from quantitative similarity analysis.[56] The *Sim*VCD and *Sim*VDF values for **2B** are, respectively, 0.30 and 0.42, while the corresponding values for **2C** are 0.37 and 0.52. These quantitative similarity measures also give preference for (*R*) configuration at C-1. Although the magnitude of *Sim*VCD for **2C** is near the confidence threshold,[51] the important point is to note that this magnitude for **2C** is the highest among that of all diastereomers. Thus, VCD analyses, both qualitative and quantitative, for **2** favor assigning *R* configuration at C-1.

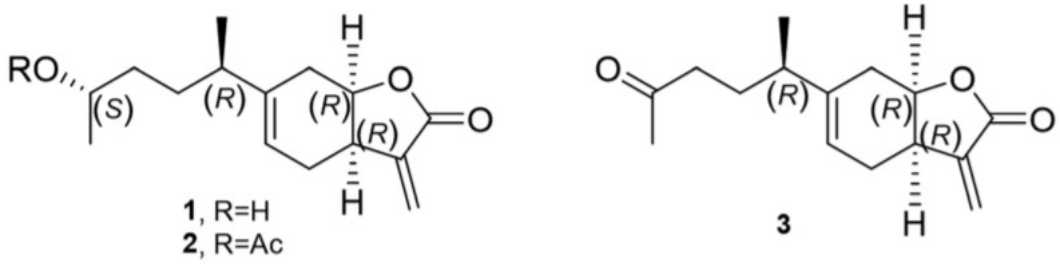

Figure 3.5: Inuloxin D, **1**, acetylated inuloxin D, **2**, and inuloxin E, **3**, with assigned AC

To summarize the above discussed analyses, ECD, ORD, and VCD support 7*R*,8*R* configuration. ORD analyses hinted at 1*R*,4*S* assignment, although not conclusively. VCD associated with the carbonyl stretching vibration of acetyl group supports 4*S* configuration. Difference VCD spectral analysis in the 1500–900 cm^{-1}

region and quantitative VCD similarity analyses suggests preference for 1*R* configuration. Thus, the AC of acetylated inuloxin D (**2**), and by extension of Inuloxin D (**1**), can confidently be assigned as (1*R*,4*S*,7*R*,8*R*), while that of inuloxin E (**3**) can confidently be assigned as (1*R*,7*R*,8*R*) (see Figure 3.5). Even though the current studies offer evidence for *R* configuration at C-1, it would be prudent to confirm this assignment independently using a different technique.

3.4 Conclusion

Based on the analyses of three different chiroptical spectral data, the AC of acetylated inuloxin D, and by extension of the natural product inuloxin D, extracted from aerial parts of *D. viscosa*, can confidently be assigned as (1*R*,4*S*,7*R*,8*R*), while that of inuloxin E (**3**), also extracted from aerial parts of *D. viscosa*, can confidently be assigned as (1*R*,7*R*,8*R*).

CHAPTER 4

The AC of [5]- and [3]-Ladderanoic Acids

This work was completed in collaboration with Professors Thomas Harris of Vanderbilt University, Nashville, TN and the Virginia Institute of Marine Science, Gloucester Point, VA, Professor Nathan Schley of Vanderbilt University, Nashville, TN, and Professor Malgorzata Baranska of Jagiellonian University, Krakow, Poland. This chapter has been reproduced from parts of the *Journal of Natural Products* article[85] "Absolute Configurations of Naturally Occurring [5]- and [3]- Ladderanoic Acids: Isolation, Chiroptical Spectroscopy, and Crystallography" with permission. My contributions to this work are: ROA measurements, conformational analyses, QC optimizations, QC spectral prediction of VCD, ORD, and ROA, spectral similarity analyses, preparation of figures and tables, and revising of the manuscript.

4.1 Introduction

Ammonium ion is essential to all living organisms, being the source of nitrogen in proteins, nucleic acids, and other cellular metabolites. It is produced from atmospheric N_2 by nitrogen-fixing bacteria and in the atmosphere by lightning discharges. In addition, more than 200 million tons of ammonia are produced from N_2 annually using the Haber process, the majority for use as fertilizer. Nature maintains homeostasis via enzymatic pathways of denitrification and anaerobic ammonium oxidation (anammox) to convert ammonia back to N_2.[86,87] Anammox is mediated by bacteria in the Planctomycetes phylum,[88] which can oxidize ammonium while reducing nitrite to produce N_2. Anammox bacteria are ubiquitously dispersed throughout the world, being found in terrestrial and aquatic ecosystems, and are used extensively in wastewater treatment plants to remove ammonium ion from effluent. The anammox pathway, comprising enzymatic reactions for reduction of nitrite to nitric oxide (NO), synthesis of hydroxylamine and hydrazine from ammonium ion and NO, and oxidation of hydrazine to N_2,[87,89,90] occurs in an intracytoplasmic vesicle called the anammoxosome. A phospholipid membrane surrounding the anammoxosome is thought to be unusually dense, thereby retarding escape of the strongly nucleophilic hydroxylamine and hydrazine, which would otherwise react with cellular constituents and thus interfere with essential metabolism.[91,92,93] The membrane has been shown to be less permeable to dyes than typical phospholipid membranes, but reduced permeability to hydroxylamine and hydrazine has not been demonstrated.[94,95]

The phospholipid membrane of the anammoxosome is structurally remarkable, containing fatty acids and analogous alcohols having multiple concatenated cyclobutane rings[91] known as ladderanes. The two major ladderane components of the anammoxosome are [5]-ladderanoic acid **1a** (n = 7) and [3]-ladderanoic

acid **2a** (n = 7) (Figure 4.1). Both **1a** and **2a** are eicosanoids, with [5]-ladderanoic acid 1a containing five linearly fused cyclobutane rings and [3]-ladderanoic acid **2a** having three cyclobutane rings fused to a cyclohexane ring. Lesser quantities of longer and shorter chain analogues (n = 1, 3, 5, 9, and 11) have also been observed.[88,96,97,98,99] Acids **1** and **2** occur in the anammoxosome membrane both as the free acids and as phospholipid esters. In addition, analogous [5]- and [3]-ladderanols **3** and **4** (Figure 4.1) are major constituents and are present as ethers in the phospholipids. Other larger and smaller arrays of concatenated cyclobutanes and combinations of cyclobutanes with cyclohexanes have not, as yet, been found.

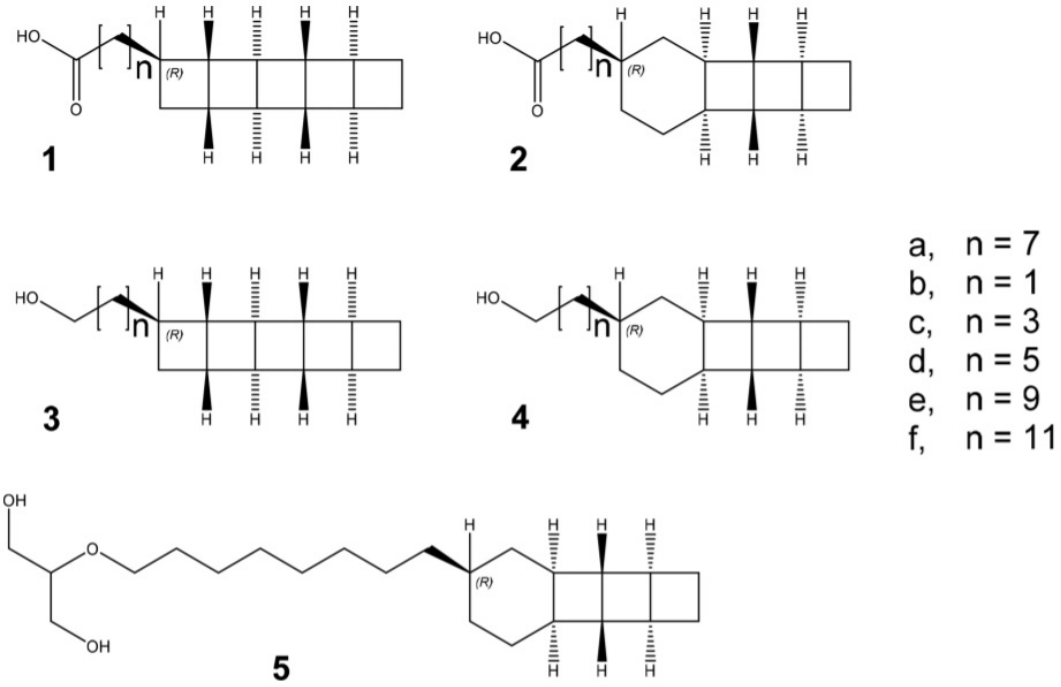

Figure 4.1: Naturally occuring ladderane acids and alcohols **1-5**

The remarkable structure of the ladderane lipids has inspired significant interest in their chemistry and biology, as evident in the review by Nouri and Tantillo[100] and in a patent on use for drug delivery.[101] However, multiple chemical and biological problems have impeded the study of ladderane lipids. Anammox bacteria live in oxygen-minimum environments with slow growth rates leading to doubling times of 2 weeks or more. In no case have the anammox bacteria been obtained in pure culture, and it is likely that they are, in fact, dependent on associated bacteria for, as yet unidentified, essential nutrients. The ladderanoic acids are not readily obtained in pure form. The methyl esters cannot be fully separated from each other by liquid chromatography,[91,94,95,102,103] and their analysis by gas chromatography is made challenging by the thermal instability of acid **1a**, which has a half-life of only 1 h at 140 °C.[104] As a result, the shortage of pure ladderanoic acids has been a major hindrance to their study.

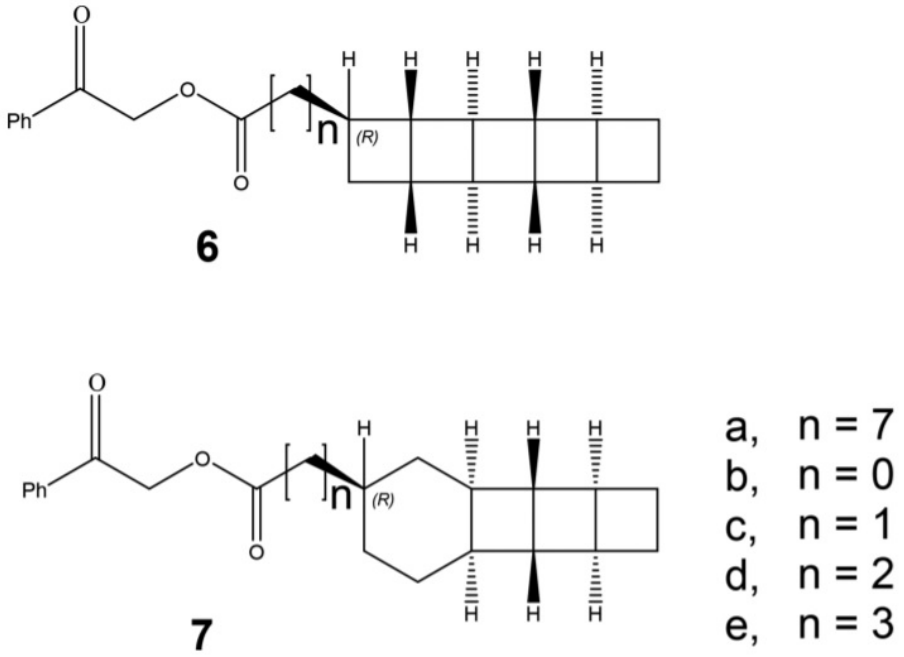

Figure 4.2: Phenacyl esters of ladderanoic acids, **6** and **7**

Throughout this chapter, the stated AC refers to that at the site of attachment of the functionalized side chain: *R* configuration at the site of attachment refers to the structures as depicted in Figures 1 and 2, while *S* configuration at the site of attachment refers to mirror images of the structures depicted in Figures 1 and 2. ACs of natural **4a** and **5a** have been established as *R* by Mercer et al.[102]

An enantioselective synthesis of (*S*)-**1a** was accomplished by Mascitti and Corey via resolution and configurational assignment of a cyclopentanone precursor.[104] They were unable to establish the AC of the natural acid due to inability to obtain an authentic sample with which to compare specific rotation.[105] This inability to gain access to a small reference sample of the acid exemplifies the experimental hurdles associated with isolation and purification of the individual ladderanes.

Recently, enantioselective syntheses of (*R*)-[3]-ladderanol **4a** and its glyceryl ether **5** were reported by Mercer et al. starting from a lipase-resolved cyclohexenediol.[102] The AC of natural **4a** was established to be *R* at the site of chain attachment by comparison of the specific rotations of synthetic and natural glyceryl ether **5**. The synthetic strategy was extended to preparation of (*R*)-[5]-ladderanoic acid **1a**. They were unable to compare their synthetic (*R*)-**1a** with the natural acid due to lack of an authentic sample of the latter.

Configurational assignments are challenging due to the difficulties in separation of the [3]- and [5]-ladderanes and of their derivatives and also due to conformational mobility, which complicate both the calculations and production of crystals suitable for X-ray crystallography. These difficulties are successfully

overcome in this work through a multipronged approach. Initial attempts led to chromatographic separations of the respective phenacyl esters **6a** and **7a** (see Figure 4.2). Later, the acids were able to be separated as **1a** and **2a** themselves. In order to establish the ACs, I employed multiple chiroptical spectroscopic methods[4,23] to overcome the inherent limitations associated with individual chiroptical spectroscopic methods. The ACs of **6a** and **7a** were established by comparison of their experimental specific rotations with those calculated using QC methods and confirmed by the X-ray crystal structure of **6a**. The ACs of **1a** and **2a** were determined by comparison of experimental and QC predictions of specific rotations as well as vibrational Raman optical activity spectra. The ACs deduced in this manner for **1a** and **2a** were further confirmed by respective single-crystal structures.

4.2 Methods

The isolation and NMR characterization of [3]- and [5]- ladderanoic acids were carried out by our American collaborators in Virginia as described in the literature.[85] X-ray crystallography that supports this work was conducted on these compounds by Professor Schley at Vanderbilt University and is described in the full text.[85] The ORD and VCD spectra used within were measured by Dr. V. Ragahavan while a post doctoral researcher within our lab and details of these measurements are provided below and in the main text. Professor Baranska and Dr. Zajac allowed me personal and extended access to their ROA spectrometer in Krakow, Poland to undertake the ROA and Raman measurements and the relevant experimental parameters are presented in the main manuscript[85] and below.

The initial conformational searches were undertaken using the CONFLEX Program[71] with MMFF94S force field. Due to the prohibitively large number of possible conformations in each case, conformational searches for full-size **1a** and **2a** were restricted to a smaller energy window. Starting from Conflex generated 7698 conformations for **1a** and 24,004 conformers for **2a**, within a 5 kcal/mol energy window, and subsequent conformational optimizations at higher theoretical levels, PM6, B3LYP/6-31G* and B3LYP/6-311++G(2d,2p) with PCM for chloroform solvent, were conducted using the Gaussian G09 program as summarized in Tables 4.1 and 4.2.

Out of the 135 conformers optimized at the B3LYP/6-311+ +G(2d,2p)/PCM level for **1a**, two conformers containing imaginary vibrational frequencies were discarded and the remaining 133 conformers of **1a** were used for specific rotation, ROA, and VCD calculations, which were carried out with appropriate Boltzmann population weighting.

Out of the final 185 conformers optimized at the B3LYP/6-311+ +G(2d,2p) level for **2a**, five conformers containing imaginary vibrational frequencies were discarded and the remaining 180 conformers of **2a** were used for specific rotation, ROA, and VCD calculations, with appropriate Boltzmann population weighting.

theoretical level	**starting number of conformers**	**number of unique conformers**	**number of conformers used for the next level**
PM6	7698	5679	5673 within 10.6 kcal/mol
B3LYP/6-31G*	5673	4489	1383 within 5 kcal/mol
B3LYP/6-311++G(2d,2p)	1383	1336	135 within 2 kcal/mol

Table 4.1: Conformers Optimized for (*R*)-5-Ladderanoic Acid, **1a**

theoretical level	**starting number of conformers**	**number of unique conformers**	**number of conformers used for the next level**
PM6	24,004	16,427	10,303 within 10.6 kcal/mol
B3LYP/6-31G*	10,303	6692	307 within 5 kcal/mol
B3LYP/6-311++G(2d,2p)	307	305	185 within 2 kcal/mol

Table 4.2: Conformers Optimized for (*R*)-3-Ladderanoic Acid, **2a**

Specific rotations, VCD, and ROA were predicted using the Gaussian 09 program.[72] Specific rotations for **1a** and **2a** with *R* configuration were calculated at six wavelengths, namely, 365, 405, 436, 546, 589, and 633 nm. Vibrational ROA spectra were predicted using a 532 nm excitation wavelength. ROA spectra were simulated with Lorentzian band shapes with SCP 180° backscattering geometry.

Vibrational Raman and ROA spectra were measured for **1a** in chloroform solvent using ChiralRAMAN spectrometer for 20 hours with using a 532 nm laser at 500 mW of power, although the power density at

the sample is unknown. ChiralRAMAN spectrometer is built for SCP measurements in 180° backscattering geometry. Vibrational Raman and ROA spectra were measured for **2a** in chloroform solvent for 14 hours with 532 nm laser at 500 mW power, although the power density at the sample is unknown. Solvent bands saturated the detector at ∼1200, ∼700 and below ∼400 cm^{-1}. Power was not reduced due to relatively weak Raman intensities originating from the compounds in comparison to the Raman intensities originating from the chloroform solvent.

4.3 Results and Discussion

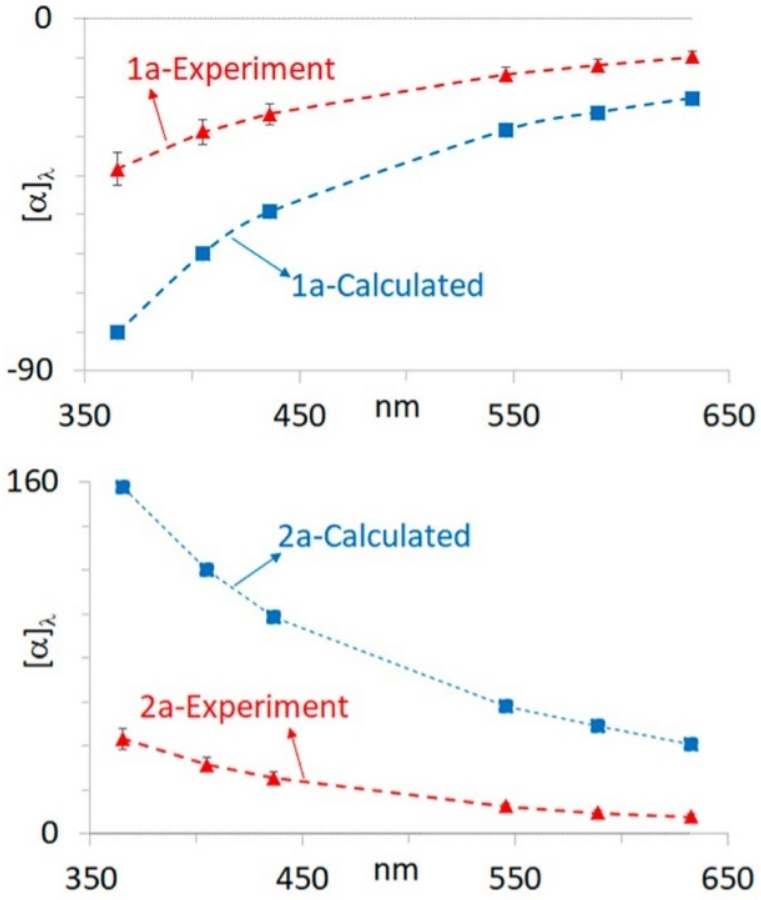

Figure 4.3: Comparison of experimental and QC predicted ORD for **1a** and **2a**

Four methods, namely, ECD,[106] ORD,[107] VCD[23,108,109] and vibrational ROA,[23,109,110] were utilized in this investigation. Each of these methods possesses characteristic, inherent limitations. ECD is not suitable for ladderanoic acids, as there are no visible chromophores attached to chiral centers that can report on the AC. Specific rotations can be used to establish the AC by comparing the signs of measured specific rotations with those obtained using QC predictions. The limitation in these comparisons is that the QC-predicted magnitudes may for various reasons differ from those observed experimentally, restricting the focus to signs

of observed and predicted rotations. VCD and ROA spectroscopies are sensitive to configuration and conformational details, but these spectral measurements require higher concentrations, where dimer formation is possible. Since QC predictions for dimers of molecules as large as **1a** and **2a** cannot be undertaken, it was hoped that QC predictions for monomer structures would capture the essential details needed to interpret the VCD and ROA spectra, along with associated vibrational absorption (VA) and Raman spectra. After the successful isolation of **1a** and **2a** by collaborators, our lab measured the discrete wavelength-resolved specific rotations and VCD spectra and I measured ROA spectra for **1a** and **2a** and analyzed all chiroptical spectra with corresponding QC predictions. Specific rotations are commonly measured and reported only at 589 nm, although, based on Drude's equation[111] for specific rotation, larger specific rotations can be expected at shorter wavelengths. Shorter wavelengths offer advantages for substances that have inherently small rotations or for which the supply is limited. For this reason, specific rotations were also measured at shorter wavelengths: 546, 436, 405, and 365 as well as at 633 nm.

Although a quantum mechanical equation for calculation of specific rotation had been proposed by Rosenfeld[17] in 1929, it could not be put to practical use as a predictive tool until recently. Remarkable advances in quantum chemical methods[11,12,112,113,114,115,116,117] have rendered these predictions reliable enough for a confident assignment of absolute configurations in numerous cases. Beratan, Wipf, and co-workers using the BP-86/aug-ccpVDZ level of theory (with the RI-J approximation) calculated gauge-independent $[\alpha]_D$ values for a series of model [5]-ladderanes bearing linear alkyl chains ranging from methyl to pentyl on a terminal methylene group.[115] The calculations gave positive specific rotations with $[\alpha]_D$ values ranging from 35 to 51 when the site of attachment had an *S* configuration. They acknowledged that this procedure might be expected to give inflated values of specific rotations but judged that the method was sufficiently accurate to be able to predict the correct sign of the specific rotation. They did not carry out calculations on **1a** itself but, on the basis of the results obtained for the alkyl series, predicted that (*S*)-**1a** would be dextrorotatory. However, this prediction needs to be verified. The relative contributions of the multitude of conformers of the long alkyl chain will need to be taken into consideration since they will influence the sign and magnitude of specific rotations.[12,116,117] The presence of functionality at the terminus of the chain is an additional complication to be reckoned with.

The solvent influence in the experimental measurements was represented by the polarizable continuum model (PCM),[44,118] and explicit inclusion of solvent molecules in the calculations is not possible for molecules of the size considered here. Additionally, vibrational contributions to calculated specific rotations are known to be important,[119,120,121] and these contributions can only be incorporated for small molecules. Aggregation at the concentrations employed might occur via stacking of concatenated cyclobutanes, but this cannot be realistically modeled in these calculations. Precedents for aggregation altering the experimental

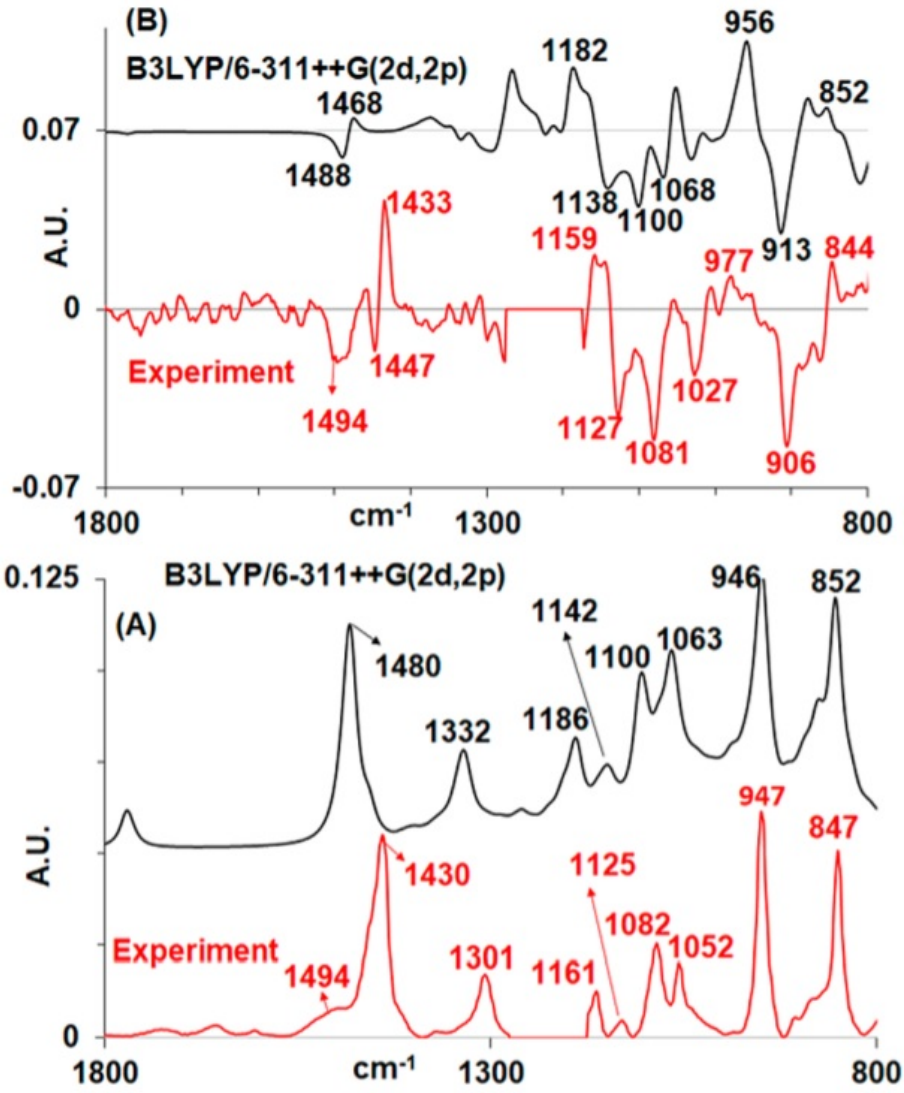

Figure 4.4: Comparison of experimental and QC predicted Raman, ROA, and CID for **1a**

chiroptical properties can be found in studies of chiral surfactants.[122] Despite these limitations, the larger predicted magnitudes of specific rotations are comforting in the sense that they are less prone to have reversed sign when calculations are repeated at a different level of theory.

Conflex[71] initially generated 7698 conformations for **1a** and 24004 conformations for **2a**, which were fully optimized using Gaussian 09[72]. For the lowest energy conformer of **1a** predicted by B3LYP/6-311++G(2d,2p), the four C atoms starting from the second C of the alkyl chain and going along the edge of the ladder make a dihedral angle of $\sim$75°. That is, the alkyl chain is canted with respect to the plane of the ladderane moiety, as also seen in the crystal structures of **1a** and **6a**. For the B3LYP/6-311++G(2d,2p) predicted lowest energy conformer of **2a**, the four C atoms starting from the second C of the alkyl chain and

going along the edge of the ladder make a dihedral angle of ∼180°, as also seen for one of the conformers in the crystal structure of **2a**.[85]

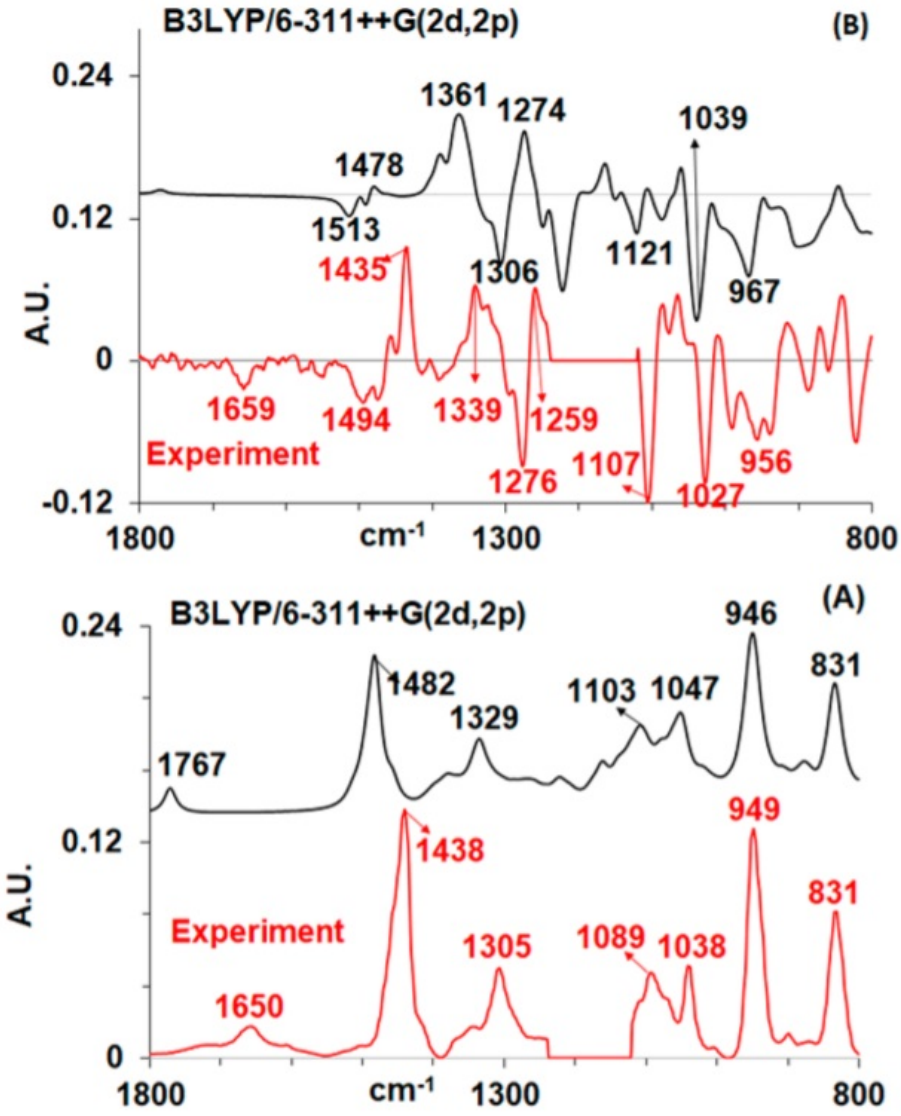

Figure 4.5: Comparison of experimental and QC predicted Raman, ROA, and CID for **2a**

Experimental specific rotations of **1a** and **2a** measured for chloroform solutions at six wavelengths are presented alongside QC-predicted values in Figure 4.3. For naturally occurring **1a** the experimental $[\alpha]_D$ magnitude (-12 ± 2°) corresponds reasonably well to that reported[104] for the synthetic sample by Mascitti and Corey (+16.7°), but the sign of the natural material is negative rather than positive. Their synthetic route led to **1a** having an *S* configuration at the site of attachment of the side chain, which is enantiomeric to natural **1a**. Mercer et al.[102] observed a $[\alpha]_D$ of -15.4° for their synthetic sample of (*R*)-**1a**. The experimental specific rotation values are negative at all six wavelengths measured, in agreement with the corresponding predicted

values for the *R* configuration. The magnitudes of calculated specific rotations are within a factor of 2 of the corresponding experimental specific rotation magnitudes (see Figure 4.3).

The experimental specific rotations of **2a** are positive at all wavelengths measured, and $[\alpha]_D$ is +10 ± 1°. No synthesis of **2a** has been reported, but the sign of the specific rotation is the same as that of the phenacyl ester reported above and that of the glyceryl ether reported by Mercer et al.[102] QC-predicted specific rotations for the *R* configuration calculated at all six wavelengths are positive, indicating that naturally occurring **2a** also has an *R* configuration. The QC-predicted specific rotation magnitudes are 4- to 5-fold larger compared to the corresponding experimental magnitudes. Among the lowest energy conformers, those with the cyclohexane ring in the chair conformation contributed 66% of the total population, with the boat conformation contributing the remaining 34%.

Experimental vibrational Raman and ROA spectra were measured for **1a** and **2a** in chloroform solvent. The vibrational Raman scattering from the solvent interfered in the following regions: ∼1275-1175 cm^{-1}, ∼800-600 cm^{-1}, and <400 cm^{-1}. There were no measurable ROA signals in the 400-600 cm^{-1} region. For these reasons, the presented experimental Raman and ROA spectra are restricted to ∼1800-800 cm^{-1}, without displaying any bands present in the solvent-interfering ∼1275-1175 cm^{-1} region. The QC-predicted Raman and ROA spectra were obtained for the conformers that were used for the specific rotation calculations mentioned earlier.

The Raman and ROA spectra for (*R*)-**1a** are presented in Figure 4.4. The QC-predicted Raman band positions are slightly greater than the corresponding experimental Raman bands. Not withstanding, the calculated spectra visually match the experimental spectra extremely well. To facilitate the comparison between experimental and calculated spectra, the band positions are marked in Figure 4.4.

Eight experimental Raman bands, at 1430, 1301, 1161, 1125, 1082, 1052, 947, and 847 cm^{-1}, are well correlated with the corresponding QC-predicted Raman bands at 1480, 1332, 1186, 1142, 1100, 1063, 946, and 852 cm^{-1}. The experimental Raman band and associated negative ROA band at 1494 cm^{-1} do not have counterparts in the QC-predicted spectra, suggesting they might arise from dimeric structures. The experimental ROA bands also correlate well with QCpredicted ROA bands for **1a**. The experimental bisignate couplet, at 1447(-) and 1433(+) cm^{-1}, is reproduced in the QC-predicted spectrum at 1488(-) and 1468(+) cm^{-1}. The remaining experimental ROA bands at 1159(+), 1127(-), 1081(-), 1027(-), 977(+), 906(-), and 844(+) are faithfully reproduced in the QC-predicted ROA spectrum of **1a** at 1182(+), 1138(-), 1100(-), 1068(-), 956(+), 913(-), and 852(+) cm^{-1}. The overall qualitative agreement between experimental and QC-predicted ROA spectra is excellent, independently confirming the AC assignment for naturally occurring 5-ladderanoic acid as *R*.

The ROA and Raman spectra for (*R*)-**2a** in the 1800-800 cm^{-1} region are presented in Figure 4.5. The

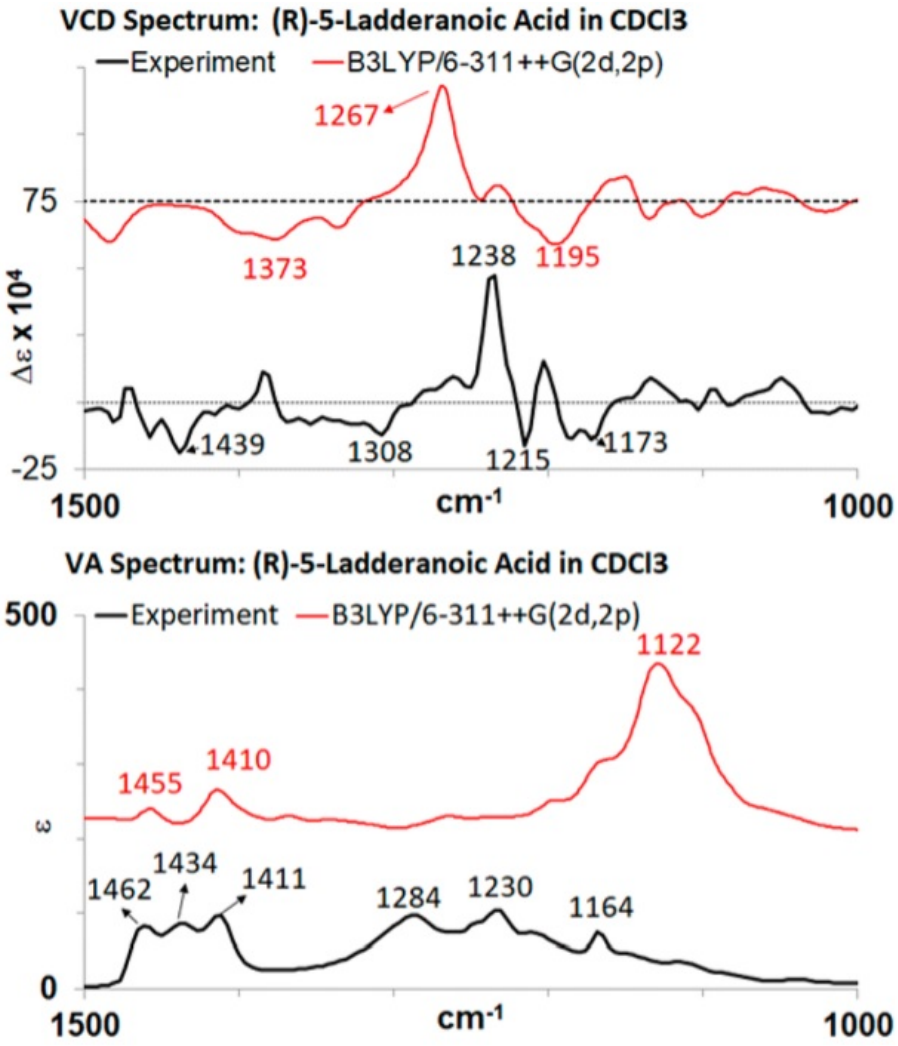

Figure 4.6: Comparison of experimental and QC predicted VA, VCD, and VDF for **1a**

QC-predicted Raman band positions for **2a** are slightly greater than the corresponding experimental Raman bands in the higher wavenumber region (>1000 cm^{-1}) and about the same in the lower wavenumber region (<1000 cm^{-1}). Qualitatively, the calculated spectra are in good agreement with the experimental spectra. To facilitate the comparison between experimental and calculated spectra, the band positions are annotated in Figure 4.5.

Seven experimental Raman bands, at 1650, 1438, 1305, 1089, 1038, 949, and 831 cm^{-1}, can be correlated to the corresponding QC-predicted Raman bands at 1767, 1482, 1329, 1103, 1047, 946, and 831 cm^{-1}. The experimental ROA bands also correlate well with QC-predicted ROA bands for **2a**, except that the weak negative experimental ROA band at 1659 cm^{-1} is not present in the QC-predicted ROA spectrum. The experimental bisignate couplet, at 1494(-) and 1435(+) cm^{-1}, is reproduced in the QC-predicted spectrum

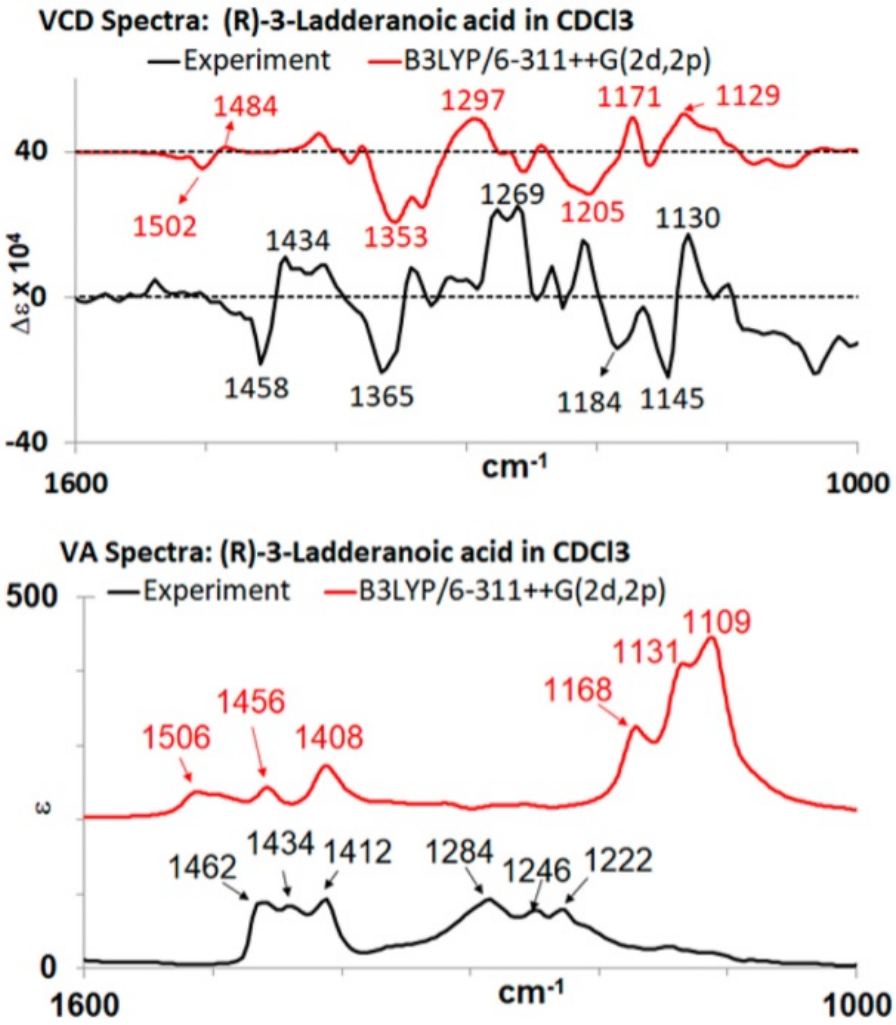

Figure 4.7: Comparison of experimental and QC predicted VA, VCD, and VDF for **2a**

at 1513(-) and 1478(+) cm^{-1}. The three neighboring ROA bands with alternating signs, 1339(+), 1276(-), and 1259(+) cm^{-1}, are faithfully reproduced in the QC-predicted ROA spectrum of **2a** as neighboring bands at 1361(+), 1306(-), and 1274(+) cm^{-1}. The remaining experimental vibrational ROA bands at 1107(-), 1027(-), and 956(-) are also reproduced in the QCpredicted ROA spectrum of **2a** at 1121(-), 1039(-), and 967(-) cm^{-1}. Thus, the overall qualitative agreement between experimental and QC-predicted ROA spectra of **2a** is excellent, independently confirming the AC assignment at the site of sidechain attachement for naturally occurring 3-ladderanoic acid as *R*.

VA and VCD spectra of **1a** and **2a** were measured by a collaborator. After undertaking QC calculations to predict VCD for each ladderanoic acid, visual comparison revealed that VCD cannot be used reliably from these calculations. The experimental and calculated VA spectra for (*R*)-(-)-5-ladderanoic acid (Figure 4.6) in $CDCl_3$ solvent do not appear to have any similarity, indicating that the monomeric structures are not appropri-

ate for reproducing the experimental VA spectrum. The calculated VCD spectrum does show some similarity in that the predicted positive VCD band at 1267 and negative band at 1195 cm^{-1} have corresponding experimental bands at 1238 and 1173 cm^{-1}, respectively. However, these are not trustworthy correlations in the absence of corresponding correlations in the VA bands. In the case of (*R*)-(+)-3-ladderanoic acid also (Figure 4.7), experimental VA spectra in $CDCl_3$ solvent do not appear to have any similarity to the calculated spectra, indicating that the monomeric structures are not appropriate for reproducing the experimental VA spectrum. Any possible correlations between experimental and calculated VCD spectra cannot be trustworthy in the absence of a corresponding correlation in the VA bands. The fact that the QC-predicted monomer Raman spectra matched the experimental Raman spectra extremely well but the QC-predicted monomer VA spectra did not correlate well with the experimental VA spectra may be associated with possible dimer conformers having an outsized impact on the VA spectra relative to the Raman spectra. For these reasons, the assignment of AC of **1a** and **2a** using VA and VCD spectra calculated for monomeric structures appears tenuous. This observation is in line with those derived from previous VA and VCD studies on substituted propanoic acids.[123,124,125]

4.4 Conclusion

The absolute configurations of the naturally occuring [5]-ladderanoic acid and [3]-ladderanoic acids, **1a** and **2a** respectively, are assigned as *R* at the site of side-chain attachment by comparison of experimental ORD with corresponding values predicted using QC calculations. These absolute configurations for **1a** and **2a** were independently verified by comparison of experimental ROA spectra with corresponding QC predicted spectra. QC predicted ROA and Raman for monomeric **1a** and **2a** matched the corresponding experimental data well, but for VA/VCD, the similarity is poor. VA/VCD appear to require QC calculations on the dimeric forms of each compound while ROA and Raman appear to reproduce the experimental spectra without accounting for dimerization. As dimer calculations are far more computationally demanding than that of the monomer, ROA appears to be advantageous for structural prediction when dimerization may occur.

CHAPTER 5

The AC of an Analogue of (-)-Crispine A

This work was completed in collaboration with Professor Ibrahim Ibnusaud of the Institute for Intensive Research in Basic Sciences at Mahatama Gandhi University, Kerala, India. This chapter has been reproduced from the *ACS Omega* article[126] "Dissymmetry Factor Spectral Analysis Can Provide Useful Diastereomer Discrimination: Chiral Molecular Structure of an Analogue of (-)-Crispine A" with permission. My contributions to this work are: VCD measurements, ECD measurements, ORD measurements, QC optimizations, QC spectral prediction of ECD, VCD and ORD, vibrational band assignment, spectral similarity analyses, figure preparation, and drafting and revising of the manuscript.

5.1 Introduction

Isoquinoline alkaloids, such as crispine A, **1**, (Figure 5.1), are important synthetic targets due to their potential biological activity.[127,128,129,130,131,132,133] (+)-**1** isolated from *Carduus crispus* L. was found to exhibit antitumor activity against certain human cancer lines.[134] The chiral solvating agents to determine the enantiomeric purity of crispine A have been investigated.[135] The enantiopure synthesis of an analogue of (-)-**1** has been reported recently using garcinia acid isolated in large amounts from tropical plant sources.[133]

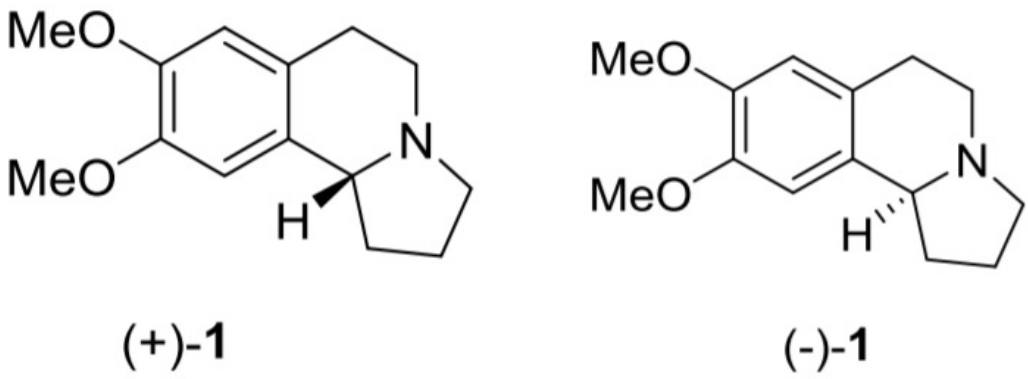

Figure 5.1: Structures of (+) and (-)-crispine A

Determination of the stereostructures of chiral natural products with multiple chiral centers can be challenging. Good-quality crystals are often difficult to grow for use in single-crystal X-ray diffraction methods. In such cases, the use of NMR methods has been the widely adopted approach, although several cases are known where stereostructures were misassigned.[136,137,138] Chiroptical spectroscopy methods have become popular in recent years for the determination of absolute configurations (ACs) of chiral molecules.[23,4] Four

different methods, namely, ORD,[111] ECD,[48] VCD,[139] and vibrational ROA[140] are widely practiced under the banner of chiroptical spectroscopy. Each of these methods has their own specific advantages/disadvantages for interrogating the chiral molecular structures, but simultaneous use of more than one method often leads to more robust structural elucidation.[141] Recent developments in QC methods[142,143,144] and their implementations in computer software[72,145,146] for predicting chiroptical spectra facilitated and popularized the analysis of experimental spectra using QC predicted spectra for deducing the chiral molecular structures.

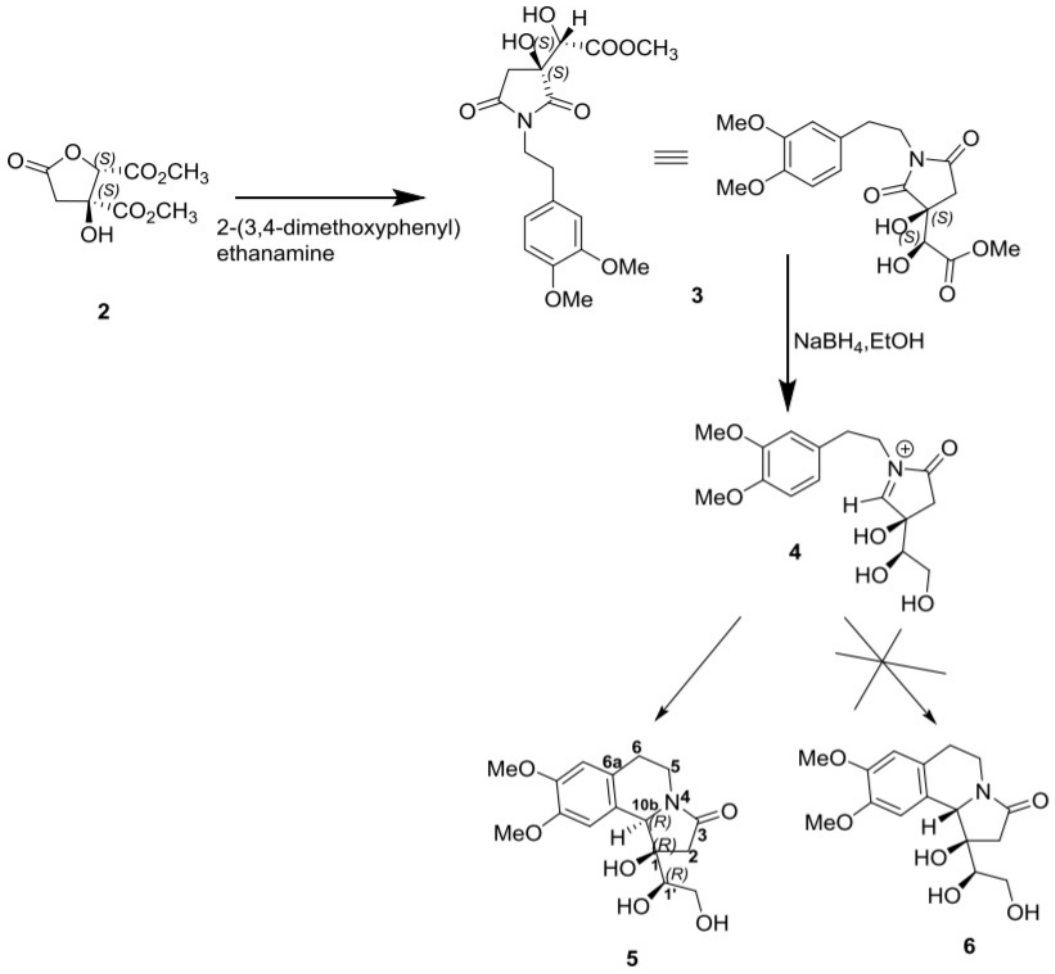

Figure 5.2: Synthetic scheme and atom numbering used for specifying the AC of **5**

In this work, ORD, ECD, and VCD spectroscopies and X-ray crystallography are utilized for investigating the molecular structure of an analogue of (-)-**1** synthesized from enantiopure (2*S*,3*S*)-dimethyl-tetrahydro-3-hydroxy-5-oxo-2,3-furandicarboxylate (garcinia ester, **2**) in two steps involving **3** and **4** (see Figure 5.2). Synthesis and X-ray crystallography were carried out by collaborators. Though there are possibilities for forming two analogues of **1** from the intermediate **4**, only **5**, namely, (1*R*,10*bR*)-1'-((*R*)-1,2-dihydroxyethyl)-1-hydroxy8,9-dimethoxy1,5,6,10*b*-tetrahydropyrrolo [2,1-a]isoquinolin3(2*H*)-one, is obtained. The absolute

configurations (ACs) at the three stereogenic centers of **5** are deduced from the synthetic scheme, and the relative configurations determined from the X-ray structure. To evaluate the capabilities of chiroptical spectroscopic methods for independently establishing the AC, the experimental ORD, ECD, and VCD spectra of **5** are also analyzed using the corresponding QC predicted spectra for all possible diastereomers.

5.2 Methods

The synthesis and X-ray crystallography the analogue of (-)-crispine A used to support this work were carried out by our Indian collaborators as been described in the literature.[126,133] Help with identification of electronic transitions was provided by Dr. E. Santoro while a post doctoral researcher within our lab. Initial calculations were set up by Lee Cantrell, a current PhD student at Vanderbilt University, while he was an undergraduate researcher in our lab. While generally I have access to our own VCD, ECD, and ORD instrumentation, Dr D. Wright allowed me personal and extended access to his lab's ECD spectrometer while ours was out of commission to undertake the ECD measurements that heavily supplemented my VCD and ORD measurements. The relevant experimental parameters are presented in the main manuscript[126] and below.

Conformer	**Energy** $kcal/mol$	**Population**	**N4-C5-C6-C6a**	**C1-C2-C4-N4**
95	0.00	0.86	−49.9	−13.5
90	1.77	0.04	−50.4	−13.9
202	1.85	0.04	−54.1	−21.0
87	2.01	0.03	−50.1	−13.7
60	2.11	0.02	−49.8	−12.7
61	2.39	0.01	−50.1	−12.7

Table 5.1: Energies, Populations, and Ring Puckering Angles of (1*R*,10*bR*,1'*R*) Diastereomer with Deuterated OH Groups at B3Pw91/6-311++G(2d,2p) PCM(Methanol) level.

ORs were measured with an AutoPol IV automatic polarimeter at six discrete wavelengths (633, 589, 546, 436, 405, and 365 nm) at a concentration of 5.87 mg/mL in methanol solvent using a 0.5 dm quartz cell and converted to SORs by dividing the ORs with concentration (g/cm^3) and path length (dm). The experimental EA and ECD spectra were measured at 0.38 mg/mL in methanol solvent with an AVIV 215 ECD spectrometer using a 2 mm quartz cell. EA and ECD spectra were collected from 250 to 400 nm. Since absorbance at wavelengths shorter than 250 nm is too high, a lower concentration of 0.048 mg/ mL in methanol solvent was used to collect EA and ECD spectra from 200 to 250 nm. These two sets of spectra were converted to epsilon units and merged smoothly at 250 nm without adjustments. The experimental

VA and VCD spectra were measured with a ChiralIR VCD spectrometer at a concentration of 45 mg/mL in CD_3OD solvent using the SL3 cell with BaF_2 windows and 100 μm path length. SSO is calculated using the Sim function suggested by Shen et al.[50] Further detail on SSO analysis is described in section 1.3.6.2.

Diastereomer	***Sim*ECD**	***Sim*EDF**	***Sim*VCD**	***Sim*VDF**
(1*R*,10*bR*,1'*R*)	0.84	0.45	0.73	0.65
(1*R*,10*bS*,1'*R*)	-0.97	-0.55	-0.54	-0.37
(1*S*,10*bR*,1'*R*)	0.82	0.47	0.58	0.35
(1*S*,10*bS*,1'*R*)	-0.82	-0.30	-0.58	-0.45

Table 5.2: SSO Values of ECD, EDF, VCD, and VDF Spectra with Maxiumum Magnitudes

The structure of **5** with (*R*)-configuration at each of the three chiral centers was manually built and used as input for conformational search using the CONFLEX program.[71] The geometries of 1730 unique conformers generated by CONFLEX were optimized using the B3LYP functional[34,35,47] and the 6-31G* basis set[49] as implemented in the Gaussian 09 program.[72] This process led to 1236 optimized structures within an energy spread of 19.2 kcal/mol. Of these, the six lowest energy structures account for 97% of the population. The geometries of these six lowest energy conformers were further optimized using the PCM[44] for representing methanol solvent at the CAM-B3LYP/6-311++G(2d,2p) level and the resulting structures used for ECD and ORD calculations at the same level. ECD and ORD calculations were also undertaken for all diastereomers using empirical D3 dispersion corrections.[147] No significant differences could be found between the results obtained with and without dispersion corrections. For VCD calculations, geometry optimizations and QC predictions of VCD were undertaken at the B3PW91/6-311++G(2d,2p)/PCM level. To account for deuterium exchange in the CD_3OD solvent that was used for experimental VCD measurements, the O-H groups of **5** were converted to O-D for VCD calculations. For (1*R*,10*bR*,1'*R*) and (1*R*,10*bS*,1'*R*) diastereomers, full geometry optimizations and VCD calculations were also carried out including empirical D3 dispersion corrections, but the inclusion of dispersion corrections did not provide improvement over those without dispersion corrections. All ECD and VCD calculations were also undertaken using the Gaussian 09 program.[72]

5.3 Results and Discussion

The procedure for chemical synthesis of **5** (see Figure 5.2 for numbered compounds) has been reported before.[133] Specifically, garcinia acid ester ((2*S*,3*S*)-3-hydroxy-5-oxotetrahydrofuran-2,3-dicarboxylate), **2**

(see Figure 5.2), was first reacted with (3,4-dimeth-oxyphenyl)-ethanamine to generate (*S*)-methyl-2-((*S*)-1-(3,4-dimethoxyphenethyl)-3-hydroxy-2,5-dioxopyr-rolidin-3-yl)-2-hydroxyacetate, **3**. During this transformation, the ACs at chiral carbon atoms remain the same. In the subsequent step, there are possibilities for forming **5** and **6** from **3**. The key synthetic step involved here is the cyclization of the *N*-acyliminium ion, **4** (Figure 5.2).[148,149,150] The use of $NaBH_4$ results in the formation of **5** exclusively. Conversion of **3** to **5**, via intermediate **4**, occurs without altering the stereochemistry. Therefore, it is logical to conclude that the ACs of the chiral carbon atoms 1 and 1' are (*R*). However, a new chiral center is formed at position 10*b*, whose AC was suggested[133] using two-dimensional NMR correlations to be *R*. An independent verification of this assignment is needed. For this purpose, single-crystal X-ray diffraction measurements are undertaken to determine the relative configurations at positions 10*b*, 1, and 1'. The crystal structure of **5** is shown in Figure 5.3, which indicates that positions 10*b*, 1, and 1' should all have the same stereodescriptors under Cahn-Ingold-Prelog sequence rules.[151] Therefore, the AC of **5** should be either (1*R*,10*bR*,1'*R*) or (1*S*,10*bS*,1'*S*). Since the ACs at 1 and 1' positions are known from the synthetic scheme to be *R*, the AC of **5** is (1*R*,10*bR*,1'*R*), as depicted in Figure 5.2, and is in agreement with that derived from two-dimensional NMR correlations.[133]

Diastereomer	**ORD**	**ECD**	**EDF**	**VCD**	**VDF**
(1*R*,10*bR*,1'*R*)	Y	Y	Y	Y	Y
(1*R*,10*bS*,1'*R*)	M	M	M	M	U
(1*S*,10*bR*,1'*R*)	Y	Y	Y	Y	U
(1*S*,10*bS*,1'*R*)	M	M	U	M	M

Table 5.3: Possible Contenders for AC of **5** as Deduced from ORD, ECD, EDF, VCD, and VDF Analyses

There are no significant differences in the populations generated from electronic energies (at CAM-B3LYP/6-311++G(2d,2p) level with PCM(methanol)) and Gibbs free energies (B3PW91/6-311++G(2d,2p) level with PCM(methanol)). The six lowest energy conformers of the (1*R*,10*bR*,1'*R*) diastereomer are displayed in Figure 5.4. In these conformers, O-H groups are seen to be oriented for favorable intramolecular O-H···O-H hydrogen-bonding network. The intramolecular H-bonding network among the three O-H groups is depicted for each of the conformers with dashed lines, and H-bond lengths are listed in Angstroms. The relative orientations of O-CH_3 groups differ among these conformers. The lowest energy conformer (#95) with ∼86% population has the methoxy groups turned away from each other, as also seen in the crystal structure; however, unlike in the crystal structure that has cocrystallized water molecules, the hydroxyl groups in

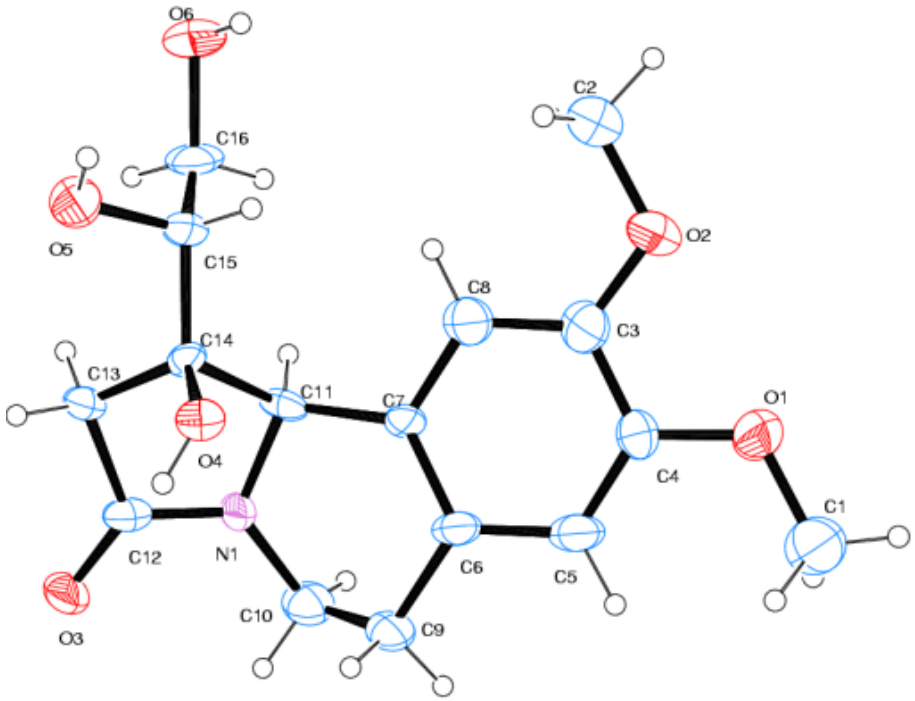

Figure 5.3: Crystal structure of **5**

the isolated molecule adopt an intramolecular hydrogen-bonding network as shown in Figure 5.4.

Table 5.1 reports the Gibbs energies, populations, and ring puckering angles for these six lowest energy conformers. All conformers have nearly the same ring puckering angles. Conformers with ring puckering angles opposite to the ones listed in Table 5.1 have higher energies and are not expected to be populated at room temperature. The lowest energy conformer (#95) is expected to be dominant with ∼86% population.

The predicted SORs for **5** are generated as a sum of those of all six conformers with individual conformer specific rotations weighted by its population. Similarly, predicted VA, VCD, EA, and ECD spectral intensities for **5** are generated as a sum of those of all six conformers with individual conformer spectral intensities weighted by its population.

The experimental SORs measured at six different wavelengths are compared to those predicted at the CAM-B3LYP/6-311++G(2d,2p)/PCM(methanol) level in Figure 5.5. The signs, as well as magnitudes, of the predicted SORs for the (1*R*,10*bR*,1'*R*) diastereomer match those observed in the experiment quite well. The excellent agreement between experimental and QC predicted ORD supports the assigned AC, (1*R*,10*bR*,1'*R*), as well as the predominant conformation listed in Table 5.1. However, it should be noted that since methanol solvent can hydrogen bond to **5**, the predominant conformer of **5** in methanol solution can very well be different from that predicted for an isolated molecule using PCM and that difference may not be transparent

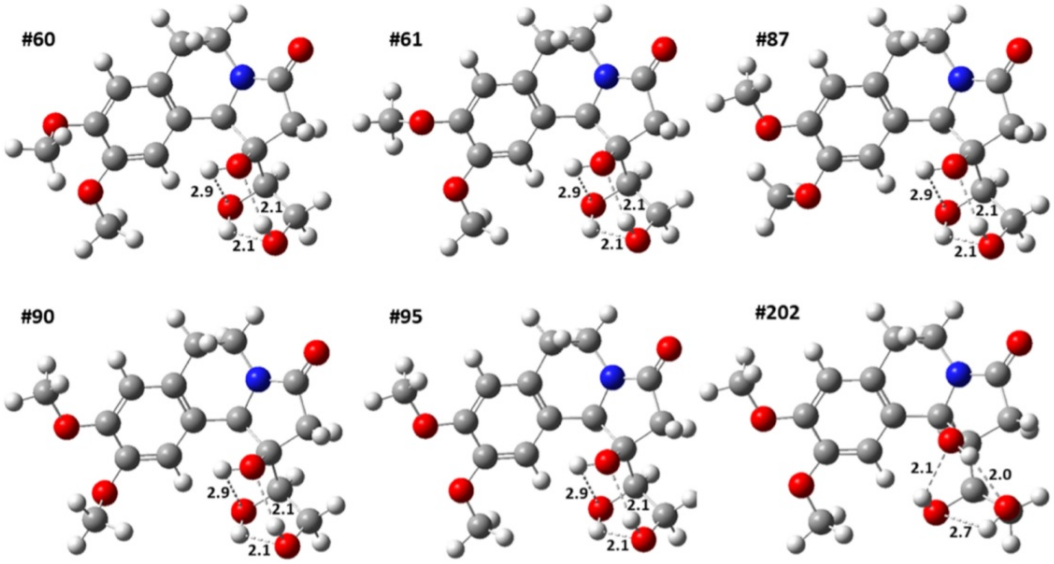

Figure 5.4: Six lowest energy conformers of the (1*R*,10*bR*,1'*R*) diastereomer

from the predicted ORD.

The experimental EA, ECD, and EDF spectra are compared to the corresponding predicted spectra for the (1*R*,10*bR*,1'*R*) diastereomer in Figure 5.6. Since a single conformer (#95) of the (1*R*,10*bR*,1'*R*) diastereomer has the majority of the population (see Table 5.1), the Boltzmann weighted EA and ECD spectra are dominated by those of this conformer at CAM-B3LYP/6-311++G(2d,2p) level with PCM for methanol solvent. In Figure 5.6, the left vertical panel displays the predicted spectra with wavelengths scaled by 1.02 (which corresponds to the maximum *Sim*ECD value) and overlaid on experimental spectra. In the right vertical panel, QC predicted spectra obtained with unscaled band positions are presented. The EA spectrum of **5** in the 180 and 300 nm regions measured in methanol reveals three bands: a stronger intensity band at 207, a weaker intensity band at 233, and a weak-intensity broad band at 286 nm. The QC predicted EA spectrum for the (1*R*,10*bR*,1'*R*) diastereomer at the CAM-B3LYP/6-311++G(2d,2p) level similarly shows three bands at 194, 225, and 253 nm. Inclusion of empirical D3 dispersion corrections did not influence these results. The experimentally measured ECD spectrum for **5** shows four bands at 203, 216, 229, and 242 nm associated with negative, positive, negative, and positive signs, respectively. The negative (229 nm) and positive (242 nm) bisignate couplet appears to be associated with the experimental 233 nm absorption band. The CAM-B3LYP/6-311++G(2d,2p)/PCM(methanol) predicted ECD spectrum reveals only three bands with a sign pattern of negative, positive, and positive at 200, 226, and 252 nm, respectively. The experimental negative cotton effect (CE) at 229 nm is not reproduced in the QC predicted spectrum, resulting in the correlation of the experimental (with predicted) ECD spectrum as follows: 200 (203), 216 (226), and 242(252) nm.

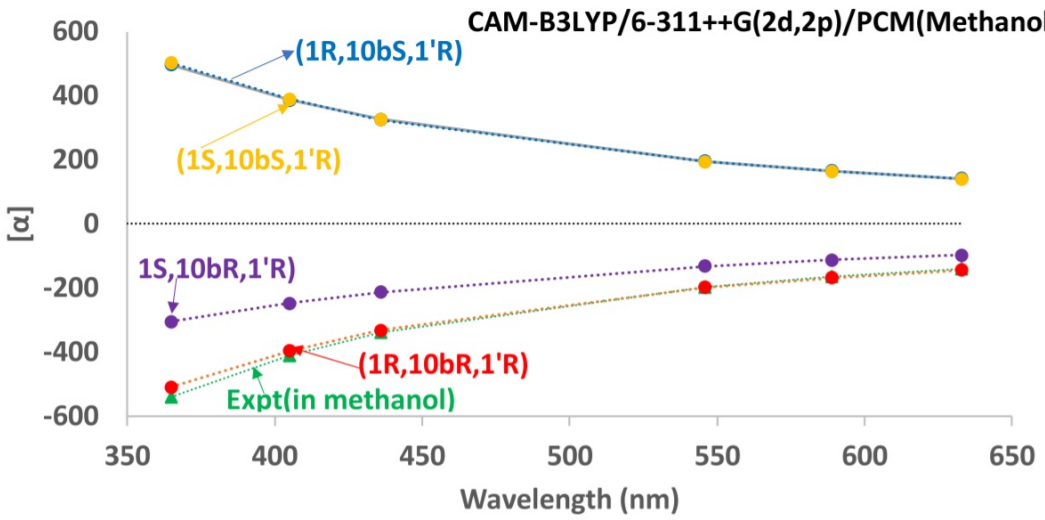

Figure 5.5: Comparison of experimental ORD with ORD predicted for four diastereomers of **5**

Based on the predicted EA and ECD spectra, the electronic transitions observed in the experimental spectrum are considered to arise from those in the benzene ring. The molecular orbitals identifying the electronic transition associated with the predicted CEs are displayed in the corresponding article.[126]

The spectral similarity overlap (SSO) plots provide quantitative assessment of the agreement between experimental and predicted spectra. The maximum magnitudes for *Sim*EA, *Sim*ECD, and *Sim*EDF values seen for (1*R*,10*bR*,1'*R*) are 0.96, 0.84, and 0.45, respectively (Figure 5.7A) at CAM-B3LYP/6-311++G(2d,2p) level with PCM to model methanol solvent. Inclusion of empirical D3 dispersion corrections did not influence these results. Ideally, these values should be close to 1 for perfect agreement, but *Sim*ECD, and *Sim*EDF values greater than $\sim$0.4, are considered necessary to reflect a reliable agreement between experiment and predictions. This criterion was established by Dr. Polavarapu[23] and Covington et al.[52] while Shen et al. proposed only a magnitude greater than 0.2 for confidence in spectral similarity. Therefore, a large *Sim*ECD value, and *Sim*EDF value greater than 0.4, provides further support for the conclusions derived from ORD analyses.

The experimental VA, VCD, and VDF spectra of **5** in the 1800-1160 cm^{-1} region are compared to the corresponding spectra predicted for the (1*R*,10*bR*,1'*R*) diastereomer at B3PW91/6-311++G(2d,2p) level with PCM(methanol), with deuterated OH groups, in Figure 5.8. In the left vertical panel of Figure 5.8, the predicted wavenumbers are scaled with 0.9815 (which corresponds to the maximum *Sim*VCD value) and overlaid on the experimental spectra. In the right vertical panel, predicted spectra with unscaled wavenumbers are stacked over experimental spectra and band positions labeled. The correlations of experimental (and the corresponding unscaled predicted) VA band positions in cm^{-1} noted from the stacked spectra are as follows:

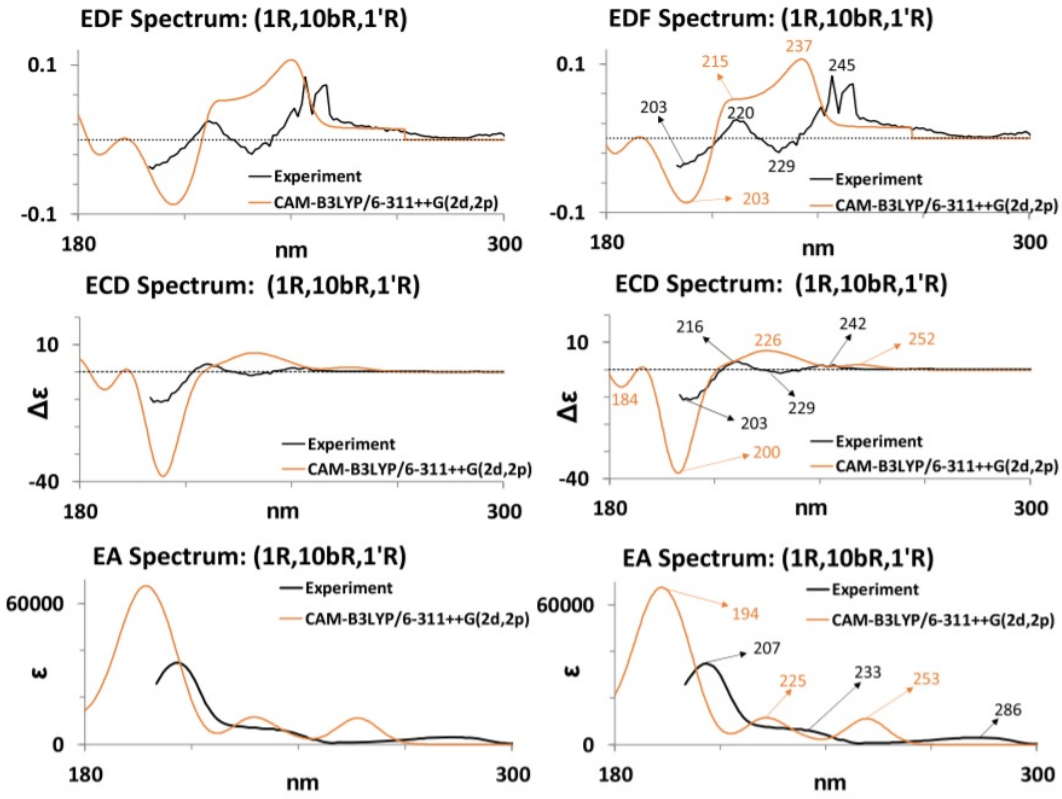

Figure 5.6: Comparison of experimental EA, ECD, and EDF spectra of **5** with those predicted for the (1*R*,10*bR*,1'*R*) diastereomer

1670(1712), 1609(1644), 1516(1551), 1462(1487), 1443(1469), 1408(1434), 1354(1385), 1257(1292), and 1223(1250). Similarly, the correlations of experimental (and the corresponding unscaled predicted) VCD band positions in cm^{-1} are as follows: 1674(1713), 1516(1548), 1470(1491), 1450(1472), 1404(1429), 1296(1323), 1261(1287), 1230(1247), and 1207(1231). The positive experimental VCD band at 1354 cm^{-1} does not appear to have a corresponding positive predicted counterpart. This comparison indicates that most of the experimental bands are satisfactorily reproduced in the predicted spectra. Additionally, the relative intensities for all calculated VCD bands match those in the experimental spectrum.

It should be added that while experimental spectra are measured in the CD_3OD solvent where hydrogen bonding between the solute and solvent can take place, the predicted spectra used PCM, which does not account for solute-solvent hydrogen bonding. As a consequence, one may anticipate poorer agreement for some PCM predicted vibrational bands in hydrogen-bonding solvents, leading to lower SSO values. Contrary to this anticipation, however, the agreement between experimental and predicted VCD spectra can be seen in Figure 5.8 to be excellent. The quantitative agreement between experimental and predicted spectra can be judged using maximum *Sim*VA, *Sim*VCD, and *Sim*VDF values in SSO plots. While *Sim*VA values are

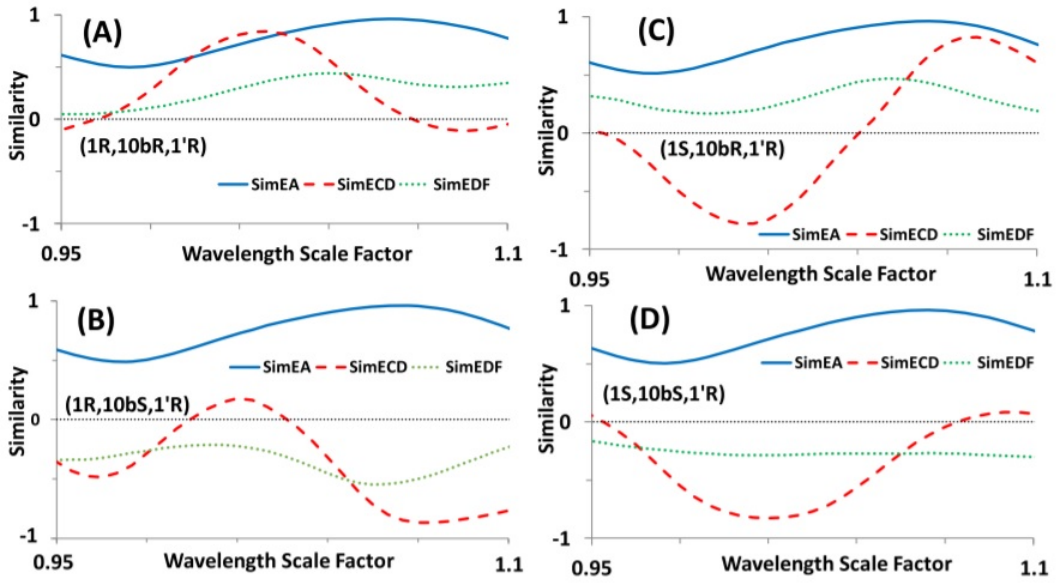

Figure 5.7: SSO plots comparing experimental EA, ECD, and EDF spectra with those predicted for four diastereomers of **5**

not useful for AC assignment, *Sim*VCD values, as well as *Sim*VDF values greater than or equal to ∼0.4,[51] are considered necessary to validate the agreement between experiment and predicted VCD spectra. This criterion was established by Dr. Polavarapu[23] and Covington et al.[52] while Shen et al. proposed only a magnitude greater than 0.2 for confidence in spectral similarity. The similarity analysis for the 1800-1160 cm^{-1} region yields maximum *Sim*VCD and *Sim*VDF values of 0.73 and 0.65, respectively, for (1*R*,10*bR*,1'*R*), at B3PW91/6-311++G(2d,2p)/PCM(methanol) with deuterated OH groups, as seen in Figure 5.9A. Such large values are normally encountered for small rigid molecules in non-hydrogen-bonding solvents and are uncommon for VCD spectra measured in the methanol-d_4 solvent. The quantitative agreement of experimental VCD and VDF spectra of **5** with those predicted for the (1*R*,10*bR*,1'*R*) diastereomer can therefore be considered excellent. Thus, the detailed analysis of the VCD spectra provides further support for the conclusions derived from ORD and ECD analyses.

The successful reproduction of experimental VCD spectra in the methanol-d_4 solvent by B3PW91/6-311++G(2d,2p) predictions, with PCM representing the hydrogen-bonding methanol solvent, can either be fortuitous or suggest one of the following two possibilities: (a) the intramolecular hydrogenbonding network among the three O-H groups (see Figure 5.3) could be providing a stabilizing environment compared to intermolecular hydrogen bonding with solvent molecules or (b) the predicted spectra for **5** with intermolecular hydrogen bonding may not be significantly different from the ones obtained for **5** with intramolecular hydrogen bonding. These later assertions require the verification of the relative stability of, and spectral

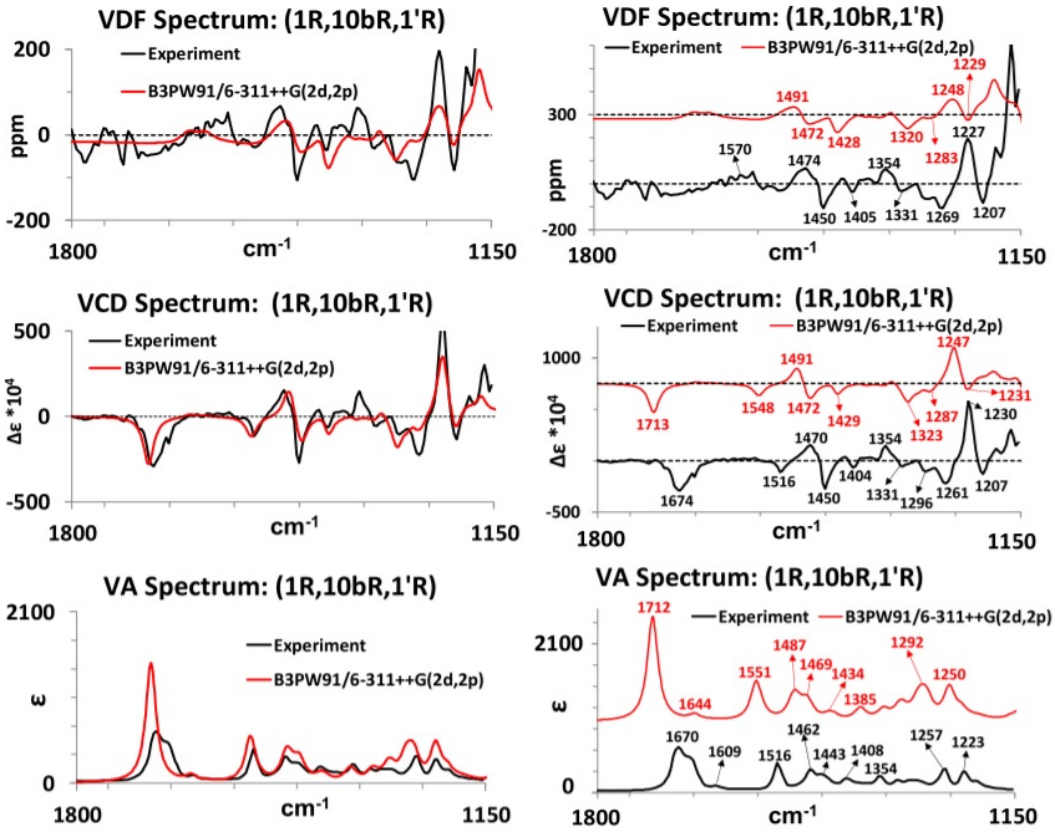

Figure 5.8: Comparison of experimental VA, VCD, and VDF spectra with those predicted for the (1*R*,10*bR*,1'*R*) diastereomer, with deuterated OH groups

predictions for, solute conformers that are hydrogen bonded to methanol solvent molecules. One approach for determining the conformers that are hydrogen bonded to solvent molecules is to incorporate 1-3 explicit solvent molecules hydrogen bonded to one solute molecule and undertake theoretical predictions. However, such approaches lead to arbitrariness in the relative orientations of solvent molecules. Molecular dynamics (MD) simulations of the solute molecule embedded in a methanol solvent bath can provide a more reliable approach for this purpose. The vibrational band assignments for the (1*R*,10*bR*,1'*R*) diastereomer derived from the animated vibrational motions in the GaussView program[84] are provided within the supplementary information of the main manuscript.[126]

For chiral molecules with multiple stereocenters, it has become a common practice in the literature to predict chiroptical spectra for the relative configuration known from some other methods (such as single-crystal X-ray diffraction and/or NMR) and eliminate one of the two possible ACs using chiroptical spectroscopic methods. If the same practice is used here for **5**, and the information derived from X-ray data or the synthetic scheme is imposed as a constraint, then the AC of **5** could be correctly established using any one of the three chiroptical spectroscopic methods (ORD, ECD, or VCD). The relative configuration derived from

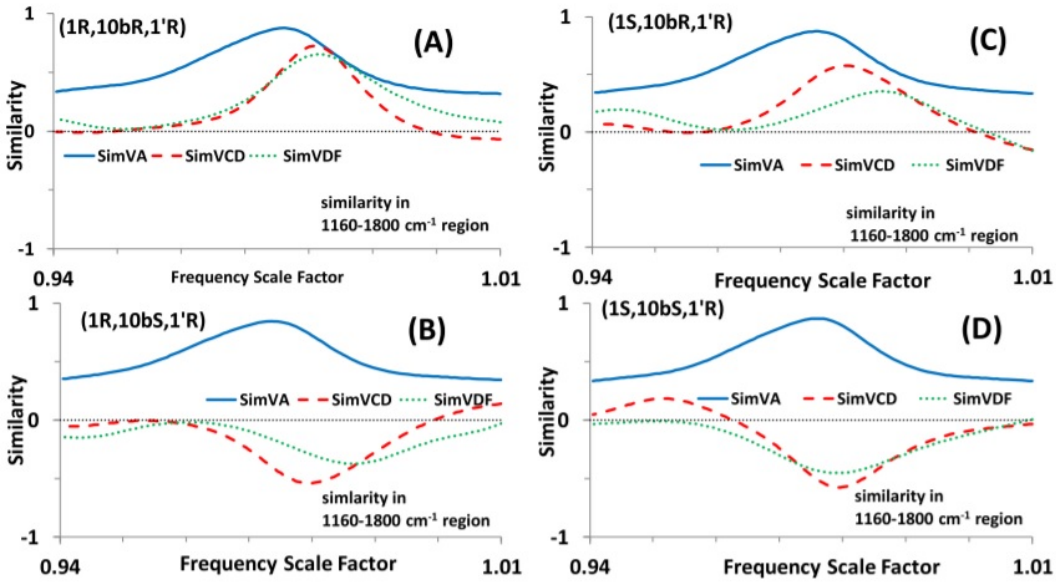

Figure 5.9: SSO plots comparing experimental VA, VCD, and VDF spectra with those predicted for four diastereomers of **5**

X-ray data restricts the consideration of AC to either (1*R*,10*bR*,1'*R*) or (1*S*,10*bS*,1'*S*). ORD, ECD, and VCD analyses individually support the AC assignment as (1*R*,10*bR*,1'*R*) and do not support the AC assignment as (1*S*,10*bS*,1'*S*). The synthetic scheme fixes the ACs at 1 and 1' positions to be (*R*), which restricts the consideration of AC of **5** to either (1*R*,10*bR*,1'*R*) or (1R,10bS,1'R). Again, ORD, ECD, and VCD analyses individually support the AC assignment as (1*R*,10*bR*,1'*R*) and do not support the AC assignment as (1*R*,10*bS*,1'*R*) (vide infra).

The validity of AC derived from chiroptical spectroscopic methods in this manner is subject to the validity of relative configuration that is started with. Therefore, diastereomer discrimination is an important objective for which recently developed microwave chiral detection methods are being pursued.[152,153] To ascertain if the chiroptical spectroscopic methods by themselves can provide enough diastereomer discrimination and determine the AC without depending on prior knowledge of relative configuration information, or of ACs at some of the stereogenic centers, one has to undertake QC predictions of chiroptical spectra for all possible diastereomers and evaluate if the correct diastereomer can be ascertained.[58,69,70,154] For this purpose, I undertook conformational analysis and QC predictions for all possible diastereomers. Due to the presence of three chiral centers, eight diastereomers are possible for **5**. Among these, four diastereomers are mirror images of the other four, so investigations are undertaken for four diastereomers: (1*R*,10*bR*,1'*R*), (1*R*,10*bR*,1'*R*), (1*S*,10*bR*,1'*R*), and (1*S*,10*bS*,1'*R*). The previous sections dealt with the (1*R*,10*bR*,1'*R*) diastereomer that is known to represent the correct AC for **5** (vide supra). Here I enquired if, in the absence of prior knowledge on

the AC of **5**, the analyses of experimental and predicted ORD, ECD, and VCD data can determine the correct diastereomer.

The conformational analysis for the remaining diastereomers follows that used for (1*R*,10*bR*,1'*R*). While (1*S*,10*bR*,1'*R*) and (1*S*,10*bS*,1'*R*) diastereomers have a single conformation with significantly higher populations than others, the (1*R*,10*bS*,1'*R*) diastereomer has three conformations with significantly higher populations than others.

The population-weighted ORD values predicted for these diastereomers are compared to the experimental ORD in Figure 5.5. It is seen that wavelength-dependent SORs of the (1*S*,10*bR*,1'*R*) diastereomer have the same signs and relatively closer magnitudes to those of experimental values, and therefore (1*S*,10*bR*,1'*R*) remains a viable candidate for assigning the AC of **5** based on ORD alone. For the (1*R*,10*bS*,1'*R*) and (1*S*,10*bS*,1'*R*) diastereomers, the signs of wavelength resolved SORs are opposite to those of experimental values, and the absolute magnitudes are comparable to the experimental magnitudes. Then, the antipodes of these diastereomers, namely, (1*S*,10*bR*,1'*S*) and (1*R*,10*bR*,1'*S*), remain viable candidates for assigning the AC of **5**. Thus, the criterion of matching the predicted ORD with the experimental ORD alone cannot be used as the sole criterion for assigning the unknown ACs of chiral molecules.[58]

The SSO plots of EA, ECD, and EDF spectra for (1*R*,10*bS*,1'*R*), (1*S*,10*bR*,1'*R*), and (1*S*,10*bS*,1'*R*) diastereomers are shown in Figure 5.7B-D, respectively, whereas individual comparisons of experimental and predicted EA and ECD spectra are provided in the main manuscript.[126] In the case of (1*S*,10*bR*,1'*R*), the SSO plot shows both positive and negative extremes of larger magnitudes for *Sim*ECD. Therefore, in principle, it becomes uncertain to decide whether this diastereomer or its antipode matches the experimental spectra. However, the negative extreme occurs for a scale factor <1. Since predicted electronic transitions, at the theoretical level employed, generally occur at shorter wavelengths, a scale factor >1 is preferred, and therefore the positive extreme is considered for *Sim*ECD. The maximum SSO magnitudes for four diastereomers are summarized in Table 5.2. From the maximum SSO magnitudes in Table 5.2, it is seen that the (1*S*,10*bR*,1'*R*) diastereomer yields positive *Sim*ECD and *Sim*EDF values, whereas (1*R*,10*bS*,1'*R*) and (1*S*,10*bS*,1'*R*) diastereomers yield negative *Sim*ECD and *Sim*EDF values. The *Sim*EDF value for the (1*S*,10*bS*,1'*R*) diastereomer is not of sufficient magnitude to place reliance on the agreement. Thus, based on *Sim*ECD and *Sim*EDF values, (1*S*,10*bR*,1'*R*) and antipode of (1*R*,10*bS*,1'*R*) remain viable candidates for assigning the AC of **5**.

The SSO plots of VA, VCD, and VDF spectra for (1*R*,10*bS*,1'*R*), (1*S*,10*bR*,1'*R*), and (1*S*,10*bS*,1'*R*) diastereomers are shown in Figure 5.9B-D, respectively, whereas individual comparisons of experimental and predicted VA and VCD spectra are provided in the main manuscript.[126] The maximum SSO magnitudes are summarized in Table 5.2. From Table 5.2, it is seen that the diastereomer (1*S*,10*bR*,1'*R*)) yields maximum

positive *Sim*VCD and *Sim*VDF values. The (1*R*,10*bS*,1'*R*) and (1*S*,10*bS*,1'*R*) diastereomers yield maximum negative *Sim*VCD and *Sim*VDF values. Based only on the magnitudes of *Sim*VCD, (1*R*,10*bS*,1'*R*), (1*S*,10*bR*,1'*R*), and (1*S*,10*bS*,1'*R*) diastereomers have large enough magnitudes to consider the (1*S*,10*bR*,1'*R*) diastereomer and the antipodes of (1*R*,10*bS*,1'*R*) and (1*S*,10*bS*,1'*R*) diastereomers as viable candidates for AC assignment. The magnitude of *Sim*VDF is also large enough to consider the antipode of (1*S*,10*bS*,1'*R*) as a viable candidate, but those for (1*R*,10*bS*,1'*R*) and (1*S*,10*bR*,1'*R*) diastereomers are close to ∼0.4. Therefore, the latter two would require further scrutiny for AC assignment.

It should be noted that the *Sim*VDF value depends on the experimental reliability criterion as discussed in section 1.3.6.2. If I considered only the experimental VCD signals with ΔA/A greater than 10 ppm as reliable, then *Sim*VDF values for (1*R*,10*bS*,1'*R*) and (1*S*,10*bR*,1'*R*) diastereomers come out as -0.35 and 0.36, respectively, and their magnitudes are still close to ∼0.4. But if I considered that only the experimental VCD signals with ΔA/A greater than 40 ppm as reliable (as has been recommended previously[46]), then *Sim*VDF values for (1*R*,10*bS*,1'*R*) and (1*S*,10*bR*,1'*R*) diastereomers drop to -0.23 and 0.16, respectively, ruling out these two diastereomers as viable candidates. For the (1*S*,10*bS*,1'*R*) diastereomer, *Sim*VDF values are -0.46 and -0.44, respectively, for 10 ppm and 40 ppm criteria, leaving its antipode in the consideration for assigning AC. Based on *Sim*VCD and *Sim*VDF values alone, the antipode of the (1*S*,10*bS*,1'*R*) diastereomer remains a viable candidate for assigning the AC of **5**.

Thus, ORD, ECD, and VCD analyses, independently, cannot provide a unique solution to the AC of **5**. A different situation emerges when I combined the results obtained from all three methods. Table 5.3 summarizes the conclusions derived from ORD comparisons and maximum SSO values derived from ECD and VCD spectral analyses. In this table, a "Y" entry indicates that the diastereomer belonging to that entry is a reliable candidate for the correct AC assignment; an "M" entry indicates that the antipode of the diastereomer belonging to that entry is a reliable candidate for correct AC; entries with "U" cast doubt on the reliability of AC assignment. The row belonging to (1*R*,10*bR*,1'*R*) has all Y entries, so the assignment of AC to this diastereomer is supported independently by ORD, ECD, EDF, VCD, and VDF. This is not the case for the other three rows. (1*S*,10*bR*,1'*R*) and the antipode of (1*R*,10*bS*,1'*R*) are not supported by the low magnitudes of VDF, whereas the antipode of the (1*S*,10*bS*,1'*R*) diastereomer is not supported by the low magnitude of EDF. However, when the conclusions from EDF and VDF analyses are combined, the (1*R*,10*bR*,1'*R*) diastereomer emerges as the single reliable candidate, which is also the AC determined from combined X-ray and synthetic analyses (vide supra).

It is important to note that if EDF and VDF are not included in Table 5.3 for assigning the AC of **5**, then four different diastereomers remain viable candidates, without offering diastereomer discrimination. Therefore, in the case of (-)-crispine A analogue investigated here, combined similarity analysis of EDF and

VDF spectra turned out to be essential for determining the correct diastereomer.

For chiral molecules with multiple chiral centers, the situation encountered in the current work can occur more often than desired. To recognize new opportunities, and any limitations, chiroptical spectroscopic predictions should be undertaken for all possible diastereomers (regardless of the availability of prior knowledge on relative configurations), as recommended previously.[58,69,70,154]

5.4 Conclusion

An analogue of (-)-crispine A with three stereogenic centers has been synthesized and its AC established using the combined information derived from the synthetic scheme and single-crystal X-ray diffraction data. Either by constraining the relative configuration to be that derived from X-ray data or by constraining the ACs at two of the chiral centers to be those derived from the synthetic scheme, the AC of (-)-crispine A analogue can be correctly established using any one of the three chiroptical spectroscopic methods (ORD, ECD, or VCD). In the absence of such outside information, separate independent analysis of ORD, ECD, and VCD does not entail a unique solution to the AC, and multiple diastereomers present themselves as viable candidates for the AC of (-)-crispine A analogue. Combined EDF and VDF spectral analysis, however, rules out the incorrect diastereomers. Thus, the combined EDF and VDF spectral analysis is seen as a useful diastereomer discrimination tool.

CHAPTER 6

The AC of an Analogue of (+)-Crispine A

This work was completed in collaboration with Professor Ibrahim Ibnusaud of the Institute for Intensive Research in Basic Sciences at Mahatama Gandhi University, Kerala, India, Dr. Cody Covington of Austin Peay State University, Clarksville, TN, and Professor Nathan Schley of Vanderbilt University, Nashville, TN. This chapter has been reproduced from parts of the *Spectrochimica Acta Part A* article[57] "How important are the intermolecular hydrogen bonding interactions in methanol solvent for interpreting the chiroptical properties?" with permission. My contributions to this work are: VCD measurements, ECD measurements, ORD measurements, crystallization of compound, QC optimizations, QC spectral prediction of ECD, VCD and ORD, vibrational band assignment, spectral similarity analyses, MD simulations, RDF calculations, figure preparation, and revising of the manuscript.

6.1 Introduction

Two crispine A analogues and tetrahydrofuro[2,3-*b*]furan-3,3a(6a*H*)-diol, endowed with hydroxyl groups that can participate in intramolecular hydrogen bonding, have been synthesized and experimental VCD and ORD data have been measured in CD_3OD/CH_3OH solvents. The absolute configurations (ACs) of these compounds have been determined using their synthetic schemes, supplemented wherever possible with X-ray diffraction data. The ACs are also analyzed with quantum chemical (QC) calculations of VCD and ORD utilizing implicit solvation as well as explicit solvation models, with the later employing classical molecular dynamics (MD) simulations. This chapter focuses on my work for this project, which was in the chiroptical measurements and calculations of spectra for an analogue of (+)-crispine A, **3**.

When chiroptical spectroscopic measurements are undertaken in non-polar and non-hydrogen bonding solvents (such as CCl_4, CS_2 etc.), the experimental observations can often be explained satisfactorily with QC predictions in vacuum (i.e. for isolated molecules). In the case of measurements for non-aggregating molecules undertaken in polar and non-hydrogen bonding solvents (such as $CHCl_3$, CH_2Cl_2 etc.), QC calculations utilizing dielectric continuum models, such as polarizable continuum model (PCM) using the integral equation formalism variant (IEFPCM)[44,81] or conductor-like screening model (COSMO),[155] for representing the solvent, provide satisfactory predictions. The results obtained in such calculations are also referred to as implicit solvation model predictions. However, neither vacuum nor implicit solvation model predictions may provide satisfactory explanations of experimental observations, when hydrogen bonding solvents (such as H_2O/D_2O, $(CH_3)_2SO$, CH_3CN, or CH_3OH/CD_3OD) are used for experimental measurements.[156] In the case

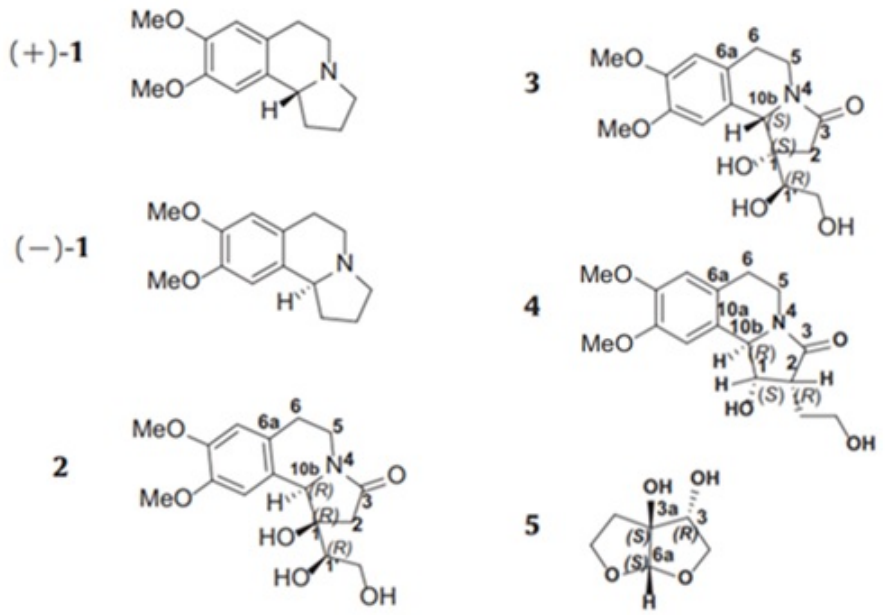

Figure 6.1: Structures of compounds relative to this work

of H_2O/D_2O solvents, explicit placement of solvent molecules interacting with solute molecule is considered to be essential.[157,158,159,160,161] The same situation applies to $(CH_3)_2SO$ solvent.[162,163] The solvent influence to a large extent depends on the functional groups in the solute molecule. For example, complications arising from inter and intramolecular hydrogen bonding in carboxylic acids is inevitable, but such complications can be avoided by converting the acids to corresponding to sodium salts or anhydrides, where possible,[164] or to methyl esters.[124,125]

The explicit placement of solvent molecules could be handled in two different ways: (a). molecular dynamics (MD) simulations of solute molecule in a solvent bath and utilizing the MD snap shots for QC calculations;[157,158,165] (b). placing selected number of solvent molecules strategically around the solute molecules and performing QC calculations on solute-solvent clusters. The influence of explicit solvation of $CDCl_3$, CD_3CN and $(CD_3)_2SO$ solvents on the VCD spectra has been explored by Merten and coworkers for chiral carboxylic acids,[166] alcohols,[167,168] and diols.[169] However, a realistic level of solvation may be difficult to achieve for nucleic acids using either method.[170]

Although there are now several VCD studies investigating the influence of explicit solvation in H_2O/D_2O, $CDCl_3$, CD_3CN and $(CD_3)_2SO$ solvents, measurements in methanol solvent are limited. VCD investigations of carboxylic acids,[166] agathisflavone[171] and cascarosides,[172] explored the influence of explicit solvation in the methanol. While carboxylic acids have characteristic interactions with solvent, agathisflavone and cascarosides are relatively large systems that are not amenable to generalizations. Therefore, the influence

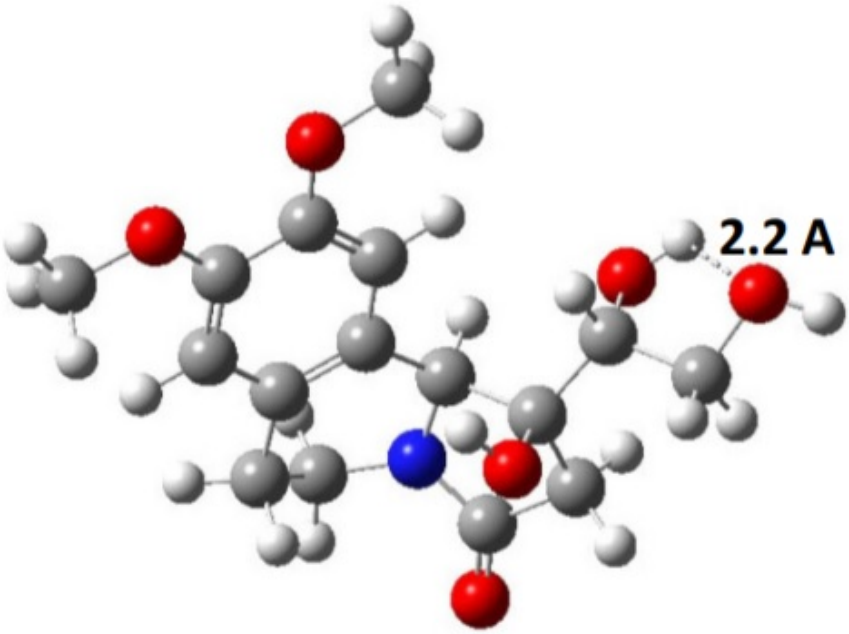

Figure 6.2: Predominant conformation of (1*S*,10*bS*,1'*R*)-**3** (74%)

of CH_3OH/CD_3OD solvent on VCD spectra is not as clear as that for H_2O/ D_2O/CD_3CN and $(CD_3)_2SO$ solvents, especially when the solute molecule has two or more –OH groups that are in close proximity for facilitating intramolecular hydrogen bonding. In such situations, intermolecular hydrogen bonding interactions between solute and solvent molecules may be energetically favored over intramolecular hydrogen bonding interactions within the solute molecule. However, for determining the ACs it was not clear if the intermolecular hydrogen bonding interactions would alter the VCD spectra in the mid infrared region significantly enough to warrant modeling these interactions. The purpose of this work was to gain further understanding on this topic.

Pyrrolo[2,1-a]isoquinoline alkaloids, such as Crispine A (Figure 6.1) and its analogues, are important synthetic targets due to their potential antitumor activity against certain human cancer lines.[127,130,131] Isolation[134] of, and different synthetic strategies[126,133,173,174,175,176] for, crispine A analogues have been reported. Semi-synthetic strategies for the synthesis of analogues of (-)- and (+)-crispine A, including **2**, **3**, and **4**, have recently been reported employing chiral 2-hydroxycitric acid lactones isolated in large amounts from tropical plant sources.[126,177] In this work, recently synthesized[177] compounds **3** and **4** are used.

In a previous investigation[126] on **2**, an analogue of (-)-**1**, it was found that the experimental chiroptical spectroscopic properties can be satisfactorily explained using QC predictions utilizing PCM for representing the CD_3OD solvent. Significantly, **2** has three hydroxyl groups close to each other facilitating intramolecular

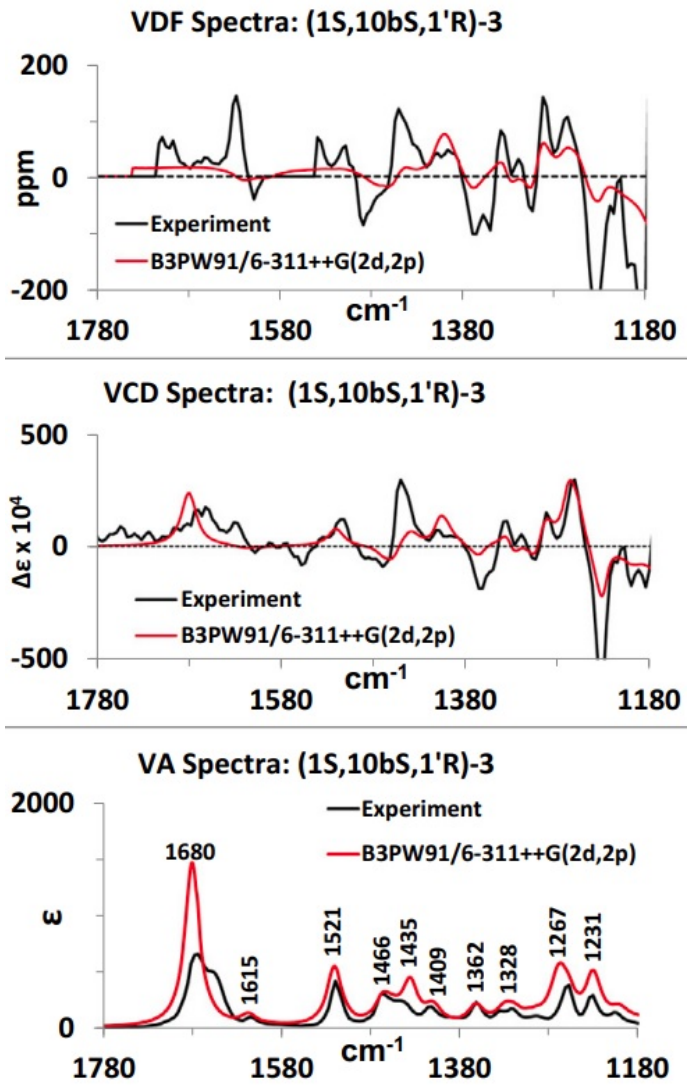

Figure 6.3: Experimental and predicted VA, VCD and VDF spectra for (1*S*,10*bS*,1'*R*)-**3**

hydrogen bonding. The observations on **2** raise important considerations because, if QC calculations with PCM are adequate for satisfactorily reproducing the chiroptical properties measured in methanol solvent, then expensive and time demanding QC calculations utilizing either MD snap shots, or solute-solvent clusters, can be avoided. Therefore, I was led to enquire if the observations made for **2** are fortuitous, or general enough to apply for other molecules with intramolecular hydrogen bonding. For this purpose, I have undertaken new experimental and QC investigations on **3**, while collaborators have investigated two additional compounds, each containing two or three OH groups in close proximity for facilitating intramolecular hydrogen bonding. Compound **3** is an analogue of (+)-1 with three hydroxyl groups. The experimental and QC predictions of VCD and ORD for these three compounds, in association with those previously reported for **2**, are utilized to

achieve practical and useful general conclusions.

6.2 Methods

The synthesis of the analogue of (+)-crispine A used to support this work were carried out by our Indian collaborators as been described in the literature.[57,126,133] X-ray crystallography that supports this work was conducted on this compound by Professor Schley at Vanderbilt University and is described in the full text[57] after I carefully crystallized the compound. The investigation of two additional compounds by Dr. E. Santoro while a post doctoral researcher within our lab, Andrew Puente a current PhD student in our lab, and Dr C. Covington of Austin Peay State University helped solidify the conclusions within and are detailed in the main manuscript.[57] While I took on VCD, ECD, and ORD measurements and corresponding calculations on the analogue of (+)-crispine A, these three collaborators undertook similar experiments. Professor McCabe of Vanderbilt University provided me references for undertaking Radial Distribution Function (RDF) calculations. Initial optimization calculations were set up by Lee Cantrell, a current PhD student at Vanderbilt University, while he was an undergraduate researcher in our lab. The relevant experimental parameters are presented in the main manuscript[57] and below.

SORs were measured in CH_3OH solvent using a 0.5 dm quartz cell and AutoPol IV automatic polarimeter at six discrete wavelengths (633, 589, 546, 436, 405, and 365 nm). The concentrations used for **3** was 2.9 mg/mL. The measured ORs were converted to SORs by dividing with concentration (g cc^{-1}) and path length (dm). The experimental VA and VCD spectra were measured in CD_3OD solvent using SL3 cell with BaF_2 windows and 100 μm path length and ChiralIR VCD spectrometer. The amount of sample dissolved in 100 μL solvent for VCD measurements was 1.73 mg for **3**.

For each diastereomer of a given chiral compound, a conformational search with MMFF94S force field was performed using CONFLEX program,[71] within an energy window of 20 kcal/mol. All of the conformers generated by CONFLEX were subjected to geometry optimization using B3LYP functional and 6-31G* basis set, using Gaussian 16 program,[80] to determine the most stable conformers. In the case of 3 and 4, an additional intermediate optimization at PM6 level was carried out and the unique conformations obtained therein were then optimized at B3LYP/6-31G* level. The unique conformers obtained at B3LYP/6-31G* level in 5 kcal/mol range were further optimized at B3LYP/6-311++G(2d,2p) level. The unique conformers obtained in this step in 2 kcal/mol range were further optimized at B3LYP/6-311++G(2d,2p)/PCM(CH_3OH) level, and used for VCD and ORD calculations at the same level. Gibbs free energies have been used to derive the populations of conformers. The simulated spectra were obtained as the sum of population weighted conformer spectra. For VCD calculations, the O-H groups of the solute molecules were replaced with O-D. Predicted spectra were simulated using Lorentzian band profiles and 10 cm^{-1} bandwidth.

Classical MD simulations for (1*S*,10*bS*,1'*R*)-**3** were performed using Amber.[178] The parameters for the GAFF force field were assigned using Antechamber, which also assigns atomic charges using the AM1-BCC method.[178,179,180,181,182] A single solute molecule was placed in a rectangular box of 194 methanol molecules for (1*S*,10*bS*,1'*R*)-**3** using tleap.[178] To prepare the system for a sampling run, 250 steps steepest descent and then 750 steps conjugate gradient minimizations were performed. Next an equilibration run was performed for 20 ns for (1*S*,10*bS*,1'*R*)-**3** using a Langevin thermostat set to 300 K and constant pressure periodic boundary conditions of 1 atm. For (1*S*,10*bS*,1'*R*)-**3**, a rectangular box was used with box vectors of 26.2, 24.3, and 20.4 Å and a non-bonding cutoff of 8 Å. Simulated annealing was carried out due to the conformational flexibility of the molecule. The system was heated to 500 K and cooled back to 300 K over a 100 ps time frame and an additional equilibration was performed. Lastly a 20-ns run was performed under NTP conditions and was used for analysis of hydrogen bonding. Radial distribution functions (RDFs) were used to analyze the proximity of solvent molecules from hydroxyl and C=O groups of **3**. RDF analyses were processed using CPPTRAJ.[183]

6.3 Results and Discussion

The ACs of chiral compounds can be deduced from the employed synthetic scheme, X-ray crystallography and/or chiroptical spectroscopy. The synthetic scheme[177] employed for **3** suggested the AC to be (1*S*,10*bS*,1'*R*). An independent verification of this AC is obtained here from X-ray diffraction.

For assigning the AC using chiroptical spectroscopy, the experimental VCD and ORD measurements are analyzed with QC predictions of corresponding properties. Since the ACs of studied compounds have been established using synthetic scheme/X-ray crystallographic methods, one may be tempted to limit the QC calculations for the known ACs and verifying that predicted properties match the corresponding experimentally observed properties. However, to independently determine the AC from, and to evaluate the discriminating ability of, chiroptical spectroscopic data it was emphasized previously[184] to undertake QC calculations on all possible diastereomers of chiral compounds. Chiral molecules with multiple chiral elements can exist in diastereomeric forms. For a chiral molecule with "n" chiral elements, 2^n diastereomers are possible. Of these, one-half of them are mirror images of the other-half and mirror image molecules yield oppositely signed chiroptical spectroscopic properties. For this reason, it is sufficient to predict chiroptical properties for 2^{n-1} non-enantiomeric diastereomers.

The first step in undertaking the theoretical chiroptical spectroscopic predictions is to investigate the complete conformational space of a given diastereomer. For this purpose, I undertook a conformational search for the four non-enantiomeric diastereomers of **3**, using different levels of QC optimizations. These are summarized in the main text.[57] For (1*S*,10*bS*,1'*R*)-**3**, the diastereomer of **3** established from synthetic/X-

ray methods, 1509 conformers were generated from CONFLEX. 1055 unique conformers remained following optimization at B3LYP/6-31G* level. Of these 1055, only 14 were within 5 kcal/mol of the lowest energy conformer. 14 unique conformers remained following optimization at B3LYP/6-311++G(2d,2p) level. These same 14 unique conformers remained following the final optimization at the same level but with PCM for methanol solvent included. All 14 were unique and within 2 kcal/mol of the lowest energy conformer.

Analysis of the final set of conformers helped ascertain the intramolecular hydrogen bonding in a given diastereomer. The predominant conformer of (1*S*,10*bS*,1'*R*)-**3**, the diastereomer of **3** established from synthetic/Xray methods, is displayed in Figure 6.2. Among the 14 conformers of (1*S*,10*bS*,1'*R*)-**3**, one has a dominant population (74%) with intramolecular hydrogen bonding between two of the hydroxyl groups at a distance of 2.2 Å. From this analysis, it is apparent that intramolecular hydrogen bonding is dominant for **3**, in the implicit solvation environment i.e. in the absence of an external environment providing intermolecular hydrogen bonding.

The experimental VCD spectrum measured in CD_3OD solvent and Boltzmann population-weighted predicted VCD spectrum at B3LYP/6- 311++G(2d,2p)/PCM(CH_3OH) level for (1*S*,10*bS*,1'*R*)-**3** are shown in Figure 6.3. The data for the remaining diastereomers are presented in the full text.[57] The visual comparison in Figure 6.3 reveals a very good one-to-one correspondence between experimental and predicted VCD bands. In the VA spectra, however, some differences do exist in absorption band intensities between observed and calculated spectra.

Diastereomer	***Sim*VA**	***Sim*VCD**	***Sim*VDF**	**SI**	$\Sigma_{fg,VCD}$	$\Sigma_{fg,VDF}$
(1*R*,10*bR*,1'*R*)	0.88	−0.57	−0.49	0.93	0.02	0.04
(1*R*,10*bS*,1'*R*)	0.85	0.51	0.24	0.92	0.74	0.55
(1*S*,10*bR*,1'*R*)	0.88	−0.55	−0.28	0.93	0.03	0.11
(1*S*,10*bS*,1'*R*)	**0.88**	**0.58**	**0.55**	**0.93**	**0.80**	**0.77**

Table 6.1: Maximum similarity overlap values for **3** over the 1710-1180 cm^{-1} region

The quantitative comparison between experimental and predicted spectra can be assessed from the SSO plot. Details on SSO analysis are provided in Section 1.3.6.2. The SSO plot for (1*S*,10*bS*,1'*R*)-**3** is shown in Figure 6.4. The maximum similarity overlap values, extracted from SSO plots are given in Table 6.1, where the entries for the diastereomer determined from synthetic scheme/X-ray methods are bolded.

There are two separate questions that need to be addressed from the SSO plot and the data in Table 6.1. The first question pertains to whether or not a satisfactory agreement is present between experimental

observations and those predicted for the AC known from synthetic scheme/X-ray methods. To answer this question, I analyzed the maximum similarity overlap values for the known AC, (1*S*,10*bS*,1'*R*)-**3**.

The VA spectra reveal some differences among absorption band intensities between experimental and predicted spectra, as intermolecular hydrogen bonding effects appear to be more evident in experimental VA spectra than in corresponding VCD spectra. Predicted band intensities are larger for some bands compared to those of corresponding bands in the experimental spectra. These differences in absorption band intensities are less transparent in *Sim*VA values, which in general are large and less sensitive to ACs of diastereomers.

For the known AC of **3**, (1*S*,10*bS*,1'*R*), *Sim*VCD value is 0.58; *Sim*VDF value is 0.55; $\Sigma_{fg,VCD}$ value is 0.80; $\Sigma_{fg,VDF}$ value is 0.77. *Sim*VDF and $\Sigma_{fg,VDF}$ values are more sensitive to differences, in VCD as well as VA intensities. As a result, *Sim*VDF and $\Sigma_{fg,VDF}$ values are usually smaller than the corresponding *Sim*VCD values. Nevertheless, *Sim*VDF and $\Sigma_{fg,VDF}$ values are also significant. This level of quantitative agreement, as reflected by *Sim*VCD, $\Sigma_{fg,VCD}$, *Sim*VDF and $\Sigma_{fg,VDF}$ values, is quite satisfactory for confirming the AC determined from synthetic scheme/X-ray methods for **3**. Thus, both from visual comparisons and quantitative similarity comparisons, it can be concluded that a satisfactory reproduction of experimental VCD spectra of **3** in CD_3OD solvent can be obtained with implicit solvation using PCM, and without having to use explicitly solvated molecules, for determining the AC.

Here it is necessary to address why QC/PCM predictions, with PCM representing CH_3OH solvent, satisfactorily match the experimental VCD spectra in the $\sim$1700–1180 cm^{-1} region, even though intermolecular hydrogen bonding is expected to be dominant in CD_3OD solvent. The vibrations most effected by intermolecular hydrogen bonding are those associated with C=O and O-D groups. For the C=O stretching vibration at $\sim$1680 cm^{-1}, the experimental VCD band position and sign are reproduced in the QC/PCM predictions. Then it is likely that intermolecular hydrogen bonding between C=O and CD_3OD solvent may not be influencing the VCD associated with C=O stretching. The O-D stretching vibrations, appearing in the $\sim$2700–2500 cm^{-1} region are outside the accessible region for our instrument. The predicted O-D bending vibrations are at frequencies lower than $\sim$1000 cm^{-1}. Specifically, for the lowest energy conformer of different compounds, unscaled vibrational frequencies with major contribution from O-D bending motion, as viewed from vibrational displacements with GaussView 6 program,[84] are as follows: 905, 910, 999 cm^{-1} for (1*R*,10*bR*,1'*R*)-**2**; 858, 911, and 951 cm^{-1} for (1*S*,10*bS*,1'*R*)-**3**; 944 and 855 cm^{-1} for (1*S*,2*R*,10*bR*)-**4**; 863 and 889 cm^{-1} for (3*R*,3*aS*,6*aS*)-**5**. Experimental VCD associated with the O-D bending vibrations, appearing in the 1150–900 cm^{-1} region, could not be measured because that region is overwhelmed by the solvent absorption interference. The 1500–1150 cm^{-1} region, normally investigated for VCD measurements in CD_3OD solvent, is mostly determined by C-H bending vibrations and C-C and C-O stretching vibrations, which appear in the present cases to be less sensitive to intermolecular hydrogen bonding.

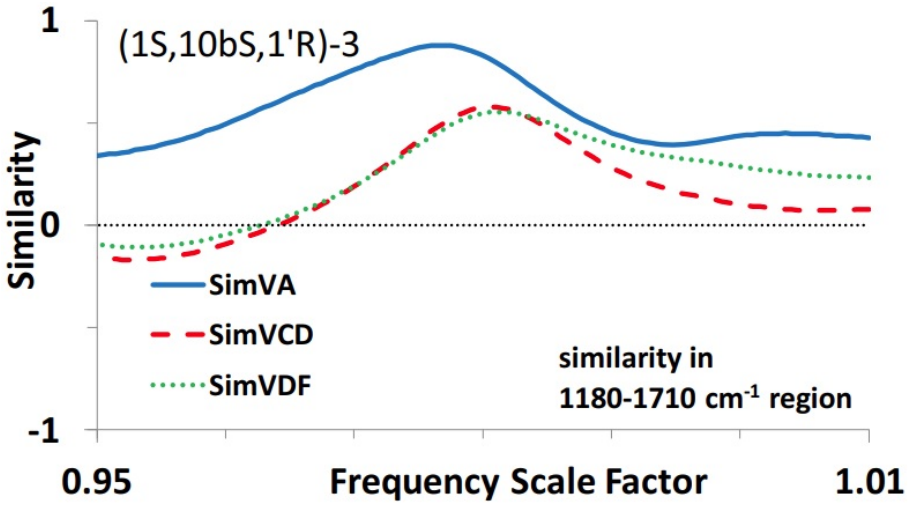

Figure 6.4: SSO plot for (1*S*,10*bS*,1'*R*)-**3**

The second question, facing chiroptical spectroscopy community is whether or not different diastereomers yield sufficiently different VCD properties for unambiguously determining the ACs of chiral compounds. To answer this question, I compared the maximum *Sim*VCD, *Sim*VDF, $\Sigma_{fg,VCD}$ and $\Sigma_{fg,VDF}$ values obtained for all diastereomers of a given compound. These four similarity values are clearly larger for (1*S*,10*bS*,1'*R*)-**3** among the diastereomers of **3**. Thus, VCD in combination with VDF and spectral similarity analysis may be stated to provide diastereomer discrimination for this compound. This observation is encouraging for the potential of VCD spectroscopy to independently establish the ACs.

A word of caution, however, is in order. The similarity measures for some diastereomers are not strikingly different. If calculations were done only for one selected diastereomer, then (1*S*,10*bS*,1'*S*)-**3**, stands good chance of being chosen for the AC assignments, because similarity measures for these diastereomers are satisfactory. However, this diastereomer are not supported by the synthetic scheme/X-ray studies. This observation drives home the point that one needs to be extra careful in AC assignments through a comparison of the experimental data with those predicted for a single diastereomer. Moreover, a general assessment of whether or not the predicted VCD spectral properties have sufficient diastereomer discrimination can only be obtained when VCD calculations for all possible diastereomers [184] are undertaken and analyzed with experimental observations.

Combining all these observations, and similar observations for two additional compounds investigated by collaborators, a reasonably confident statement can be stated as follows: For chiral compounds containing two or more hydroxyl groups with favorable intramolecular hydrogen bonding, VCD spectra obtained in CD_3OD solvent in the 1700–1150 cm^{-1} region may in some cases be satisfactory reproduced using PCM for representing the solvent. In this context, the possibility of predicted spectra for more than one diastereomer

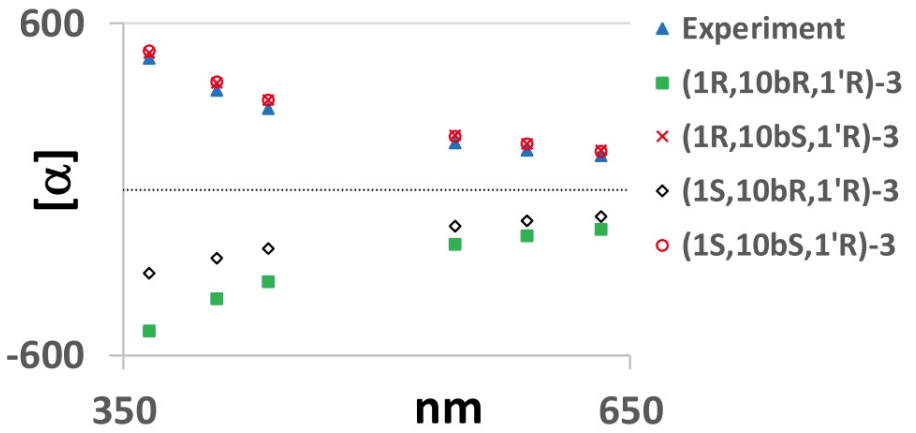

Figure 6.5: Comparison of experimental and QC predicted ORD for diastereomers of **3**

satisfactorily matching the experimental VCD spectra should always be kept in mind.

As with the analysis of VCD data, there are two separate questions that need to be addressed with ORD data. The first question pertains to whether or not a satisfactory agreement is present between experimental ORD observations and those predicted for the ACs known from synthetic scheme/X-ray methods. The experimental ORD traces measured in CH_3OH solvent and predicted traces at B3LYP/6-311++G (2d,2p)/PCM(CH_3OH) level for (1*S*,10*bS*,1'*R*)-**3** are shown in Figure 6.5. The observed SOR values vary from +121 at 633 nm to +475 at 365 nm and these signs and large magnitudes are reproduced extremely well by the calculations for (1*S*,10*bS*,1'*R*)-3.

To address why QC/PCM predictions of ORD, with PCM representing CH_3OH solvent, for the known AC of 3 satisfactorily reflect the experimental ORD data, it should be remembered that ORD and VCD are governed by different phenomenological processes. ORD, being predominantly an electronic property, the difference between intra and intermolecular interactions might be less important for ORD.

It is important to note that ORD by itself should not be relied upon to assign the unknown ACs because predicted ORD traces for more than one diastereomer can reasonably reflect the experimental ORD traces. This is evident from the comparison of experimental ORD with those predicted for all diastereomers, also in Figure 6.5. From visual analysis, calculated ORD curves for four different diastereomers, (1*S*,10*bS*,1'*R*)-**3**, (1*R*,10*bS*,1'*R*)-**3**, (1*R*,10*bS*,1'*S*)-**3** and (1*S*,10*bS*,1'*S*)-**3** (the latter two are the mirror images of (1*S*,10*bR*,1'*R*) and (1*R*,10*bR*,1'*R*) shown in Figure 6.5), have the same sign and trend as seen in the experimental ORD for **3**.

For quantitative comparison of experimental and calculated ORD curves I compared the root mean square deviations (RMSDs) obtained for measurements at six wavelengths. Details on RMSD analysis can be found in Section 1.3.6.3. RMSDs for (1*S*,10*bS*,1'*R*)-**3**, (1*R*,10*bS*,1'*R*)-**3**, (1*R*,10*bS*,1'*S*)-**3** and (1*S*,10*bS*,1'*S*)-**3** are,

respectively, 27, 26, 91, and 33, which indicates better agreement not only for (1*S*,10*bS*,1'*R*)-**3** (the correct diastereomer), but also for (1*R*,10*bS*,1'*R*)-**3** (an incorrect diastereomer). Thus, quantitative comparisons of ORD lead to ambiguity in determining the AC of **3**.

From these observations, a reasonably confident statement regarding ORD in methanol solvent can be stated as follows: For chiral compounds containing two or more hydroxyl groups with favorable intramolecular hydrogen bonding, larger magnitude ORD values obtained in CH_3OH solvent may in some cases be satisfactorily reproduced using PCM for representing the solvent, but smaller magnitude ORD values can be difficult to reproduce. When used in isolation, the diastereomer discriminating ability of ORD can be quite low.

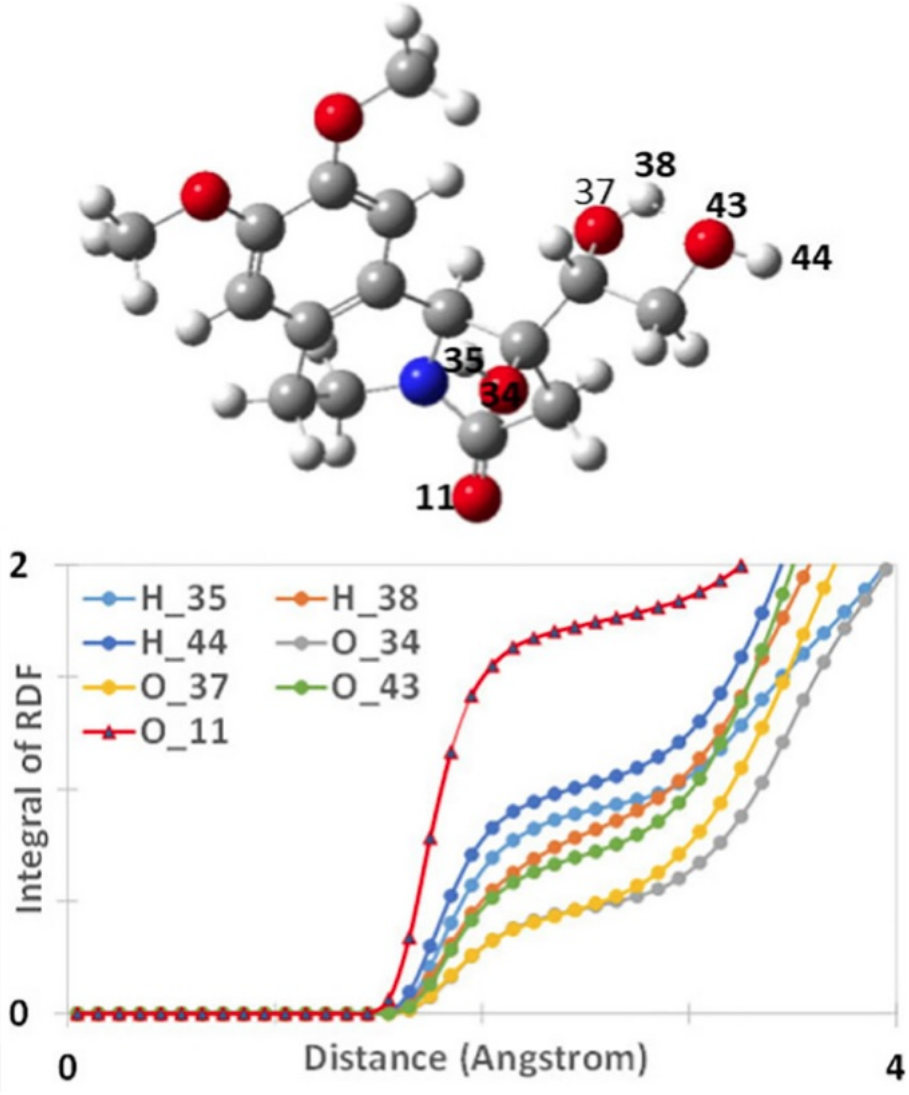

Figure 6.6: Integral RDF plot of **3** with relevant atoms labeled

Some insight is gained on the intermolecular hydrogen bonding for **3** in methanol solvent using RDFs

obtained from MD simulations. The integral RDF values as a function of distance (Figure 6.6) indicate that there is about one methanol molecule near each hydroxyl hydrogen of **3** at a distance of 2.5 Å. A distance of 3 Å is needed for one full methanol molecule to be near O atom of hydroxyl group. However, C=O group in **3** has higher propensity for hydrogen bonding to methanol, because one methanol molecule is predicted at a distance of 1.85 Å from the O atom of C=O.

From these results, it is clear that MD simulations indicate the presence of intermolecular hydrogen bonding for **3** in methanol solvent. However, as mentioned already, the influence of intermolecular hydrogen bonding in CD_3OD solvent is less evident in VCD spectra in the 1800–1150 cm^{-1} region because the O-D stretching and bending vibrations that are more directly influenced by intermolecular hydrogen bonding occur outside of this region.

6.4 Conclusion

Combining all these observations, and similar observations for two additional compounds(an analogue of crispine A, **4**, and (3*R*,3*aS*,6*aS*)-tetrahydrofuro[2,3-b]furan-3,3a(6aH)-diol, **5**) investigated by collaborators (*supra vida*), it appears that for molecules containing two or more hydroxyl groups with favorable intramolecular hydrogen bonding, implicit solvation model calculations utilizing PCM reproduce satisfactorily the experimental VCD properties in CD_3OD solvent and ORD properties in CH_3OH solvent. Explicit solvation was not necessary to model these chiroptical properties. Nevertheless, intermolecular hydrogen bonding is evident in methanol solvent, as reflected by the integral RDFs. The intermolecular hydrogen bonding, however, does not seem to significantly influence the VCD properties in the $\sim$1800–1150 cm^{-1} region mainly because the most sensitive O-D stretching and bending vibrations appear outside of this region. In the case of ORD, which is mostly governed by electronic properties, the inter- vs intra-molecular hydrogen bonding does not appear to have a major influence in the compounds studied here.

CHAPTER 7

Chiral Molecular Structures of Substituted Indans

This chapter has been reproduced from the *ACS Omega* article [162] "Chiral Molecular Structures of Substituted Indans: Ring Puckering, Rotatable Substituents, and Vibrational Circular Dichroism" with permission. My contributions to this work are: conformational analyses, explicit systematic solvation of 1-indanol in DMSO, QC optimizations, QC spectral prediction of VCD, spectral similarity analyses, preparation of figures and tables, drafting the manuscript, and revising the manuscript.

7.1 Introduction

Determining the molecular structures of chiral molecules is often a challenging task. X-ray crystallography offers the most direct approach to establish molecular structures, but its utility is restricted due to the difficulty in obtaining good quality single crystals for the samples of interest. The second popular approach is the use of NMR, where chiral shift reagents are needed for studying chiral molecules. Furthermore, the NMR time scale limits the achievable information to average molecular structures. Chiroptical spectroscopy [4,23] represents an important alternative to X-ray and NMR methods. Since optical transitions occur on a much faster time scale, chiroptical spectroscopy is more appropriate for investigating the molecular conformers and is increasingly becoming popular for probing the structures of chiral molecules. As molecular vibrational transitions are expected to be sensitive to minor structural changes, the vibrational transition-based VCD [108,109] and vibrational ROA [109,110] methods offer greater promise for chiral molecular structure determination. The electronic transition-based ECD [111] and ORD [185] methods are also sensitive for certain structural changes.

Substituted indans have attracted interest for their chiroptical properties since early 1970s, mainly in terms of ECD associated with n-σ*, σ-π*, π-σ*, and π-π* transitions. [186,187] More recently, various jet-cooled spectroscopies were undertaken for substituted indans, where low-frequency ring puckering modes were investigated. [188,189,190,191,192,193] In the early stages of VCD research, emphasis was placed on spectra-structure correlations and qualitative interpretations. In this connection, the methine-hydrogen bending mode has been suggested to generate a VCD band with its sign correlating with the AC in a series of related compounds. This observation was further investigated [194] for (*S*)-1-methylindan and (*R*)-1-methylindan-1-d and (*R*)-1-aminoindan, by measuring the experimental VA and VCD spectra for neat liquid samples and identifying the VCD associated with methine-hydrogen (or deuterium) bending mode. The sign of VCD band associated with methine-hydrogen bending mode in these three indans was also suggested [194] to correlate with the AC.

In recent years, the advances in VCD theory [142,144,195] and QC computational methods [4,196] have shifted the focus of VCD research to interpreting the experimental VCD spectra using QC predictions of corresponding spectra and deducing the chiral molecular structures therefrom. This approach has facilitated probing not only the AC but also the molecular conformational space, in a variety of systems, containing both small and large molecules. The availability of faster computer processors with larger memory and storage space has led to the use of molecular dynamics (MD) for generating the geometries of solute-solvent clusters and simulating the QC-predicted spectra for solvated species. [156,157,197] Attempts to analyze the experimental chiroptical spectra using corresponding QC-predicted spectra have now become a routine practice. [198] Recent works have also emphasized on quantitative comparisons between experimental and calculated spectra using the similarity index [199] and SSO plots. [46,52,53] The puckering of a five-membered ring common to substituted indans can shed light on the role of ring puckering conformation on VCD spectra. The contributions to VA and VCD spectra from the ring puckering conformations have not been identified before, and there is a need to identify conformer-specific VA and VCD bands. Second, (*R*)-1-aminoindan has a freely rotatable amino group whose influence on VCD spectra needs to be explored. A similar situation, with a freely rotatable O-H group, exists for (*S*)-1-indanol whose experimental VA and VCD spectra in CCl_4 and dimethyl sulfoxide (DMSO) solvents were reported recently. [163] In this chapter, I identified the puckering conformer-specific VA/VCD bands in 1-methylindans and the conformational sensitivity of rotatable NH_2/OH groups to VA/VCD spectra. In addition, I carried out comparisons of the experimental and QC-predicted VA and VCD spectra of (*S*)-1-methylindan, (*R*)-1-methylindan-1-d, (*R*)-1-aminoindan, and (*S*)-1-indanol, using quantitative SSO measures. Investigations presented here on four 1-substituted indans lead to consistent conclusions as follows: (a) the positive ring puckering of a five-membered ring, leading to a pseudoequatorial substituent at the 1-position, dominates the contributions to VCD spectra in all molecules. Nevertheless, evidence for the signatures of negative ring puckering, leading to a pseudoaxial substituent, is also present in the experimental spectra. (b) Even though the rotatable NH_2 and OH substituents at 1-position add additional conformational complexity, the observed VA and VCD spectra are reproduced remarkably well by QC predictions at the optimized geometries. (c) Hydrogen-bonding DMSO solvent facilitates the formation of the 1-indanol/DMSO complex, requiring exploration of numerous conformers. Even then, with proper care in probing the conformational space, the experimental spectra of 1-indanol/DMSO-d_6 complex are remarkably well reproduced by QC predictions at the optimized geometries. (d) The excellent correlation of experimental spectra with corresponding QC-predicted spectra, obtained for all four substituted indans, including for a complex with a hydrogen bonding solvent, identifies successful pathways for future chiral molecular structure determination.

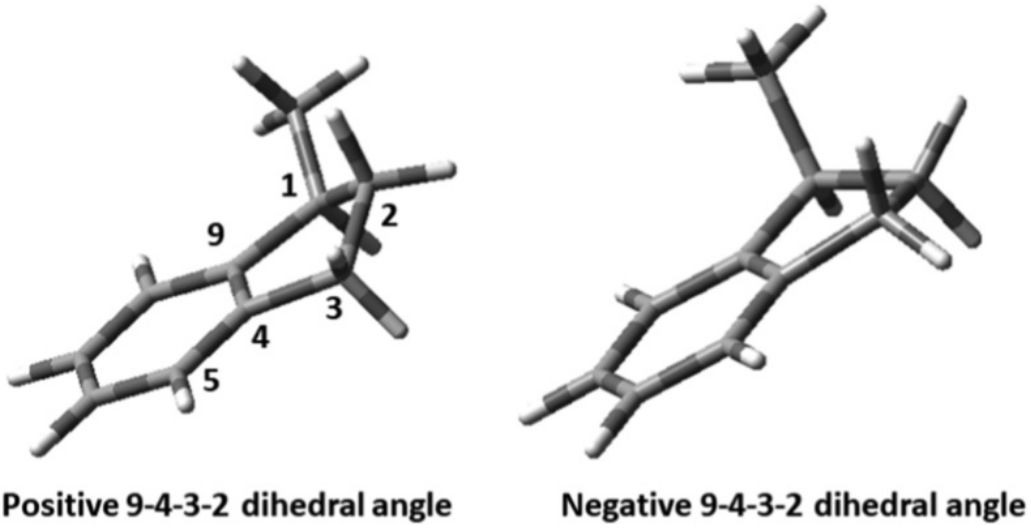

Figure 7.1: Two conformers of (*S*)-1-methylindan with opposite puckering of the 5-membered ring

7.2 Methods

For all molecules studied here, an initial conformational search was conducted using CONFLEX[71], unless otherwise stated. Final optimizations were carried out using B3LYP functional[34,35,47] and 6-311++G(2d,2p) basis set,[49] as implemented in the Gaussian program,[72] followed by QC prediction of VA and VCD at the same level. VCD spectra were obtained using magnetic field perturbation theory,[142,144,195] as implemented in the Gaussian 09 program.[72] The QC-predicted spectral intensities were Boltzmann population-weighted using Gibbs free energies and spectra simulated using Lorentzian band profiles with 10 cm^{-1} band width. For all four indan molecules considered, calculations were done for the (*S*) configuration. Since experimental spectra were reported for the (*R*)-enantiomers of 1-methylindan-1-d and 1-aminoindan, the corresponding calculated VCD spectra were multiplied by -1 for comparison to the experimental spectra. The comparisons between experimental and QC-predicted spectra and similarity analysis were carried out using CDSpecTech.[52,53] SSO plots were also generated using CDSpecTech.

The experimental spectra of (*S*)-1-methylindan, (*R*)-1-methylindan-1-d, and (*R*)-1-aminoindan were reported previously[194] where all spectra were recorded for neat liquids in absorbance units, but y-axis values cannot be discerned with sufficient quantitative accuracy. The spectra for (*S*)-1-indanol in CCl_4 and in DMSO-d_6 solvents were reported by Zehnacker et al.,[163] where y-axis values were not available. These literature experimental VA and VCD spectra were digitized using WebPlotDigitizer[200] and normalized for carrying out similarity analyses. The excessive level of absorbance in some spectral regions restricted the analysis of experimental VCD spectra to the following regions: 925-1320 cm^{-1} for 1-methylindan; 920-1345 cm^{-1} for methylindan-1-d; 950-1450 cm^{-1} for 1-aminoindan; 900-1350 cm^{-1} for 1-indanol in CCl_4;

and 1100-1700 cm^{-1} in DMSO-d_6. SSO values are for these ranges are given in Table 7.3. The y-axes values in the displayed VA and VCD spectra are those of predicted spectra in molar extinction units; VDF spectral intensities are presented as ppm. The experimental spectral intensities were scaled to be on the same scale as those for predicted spectra.

7.3 Results and Discussion

7.3.1 (*S*)-1-Methylindan

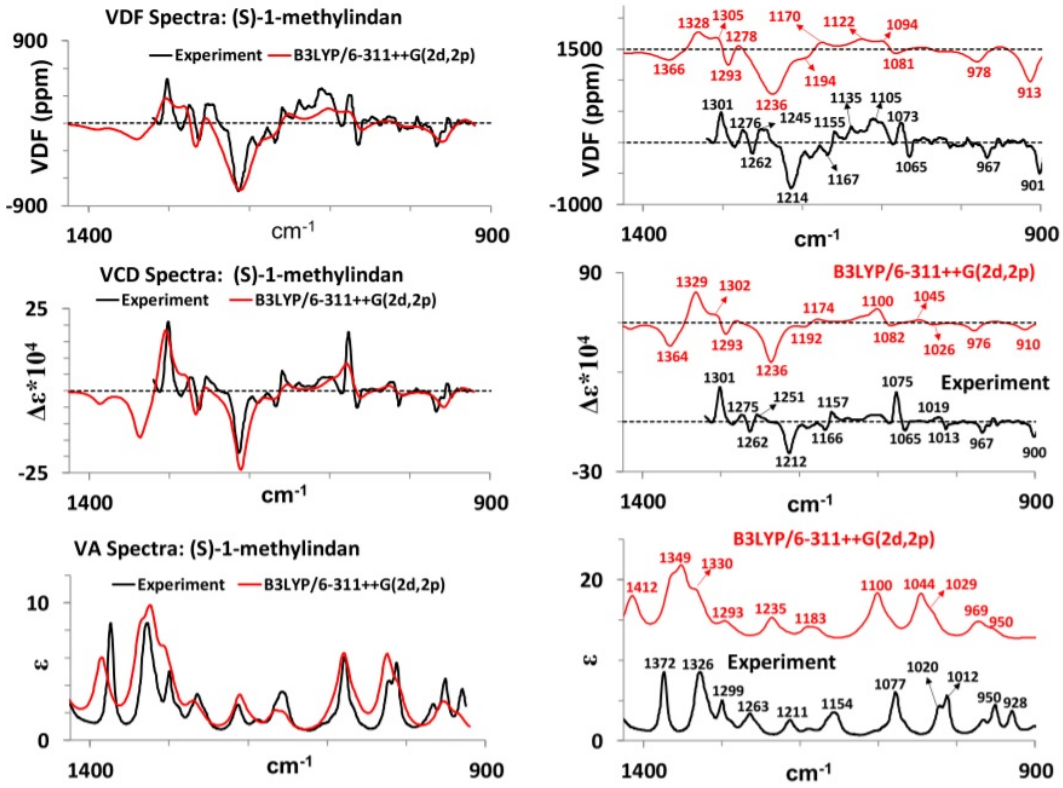

Figure 7.2: Experimental and predicted VA, VCD, and VDF spectra for (*S*)-1-methylindan

Relaxed energy scan[162] for 1-methylindan as a function of 9-4-3-2 dihedral angle (see Figure 7.1) using the B3LYP functional[34,35,47] and 6-311++G(2d,2p) basis set,[49] indicates two minima at +17 and -15°. Thus, two conformers with opposite ring puckering angles are possible for this molecule. Positive C9-C4-C3-C2 dihedral angle leads to the methyl group to be in the pseudoequatorial position, and negative C9-C4-C3-C2 dihedral angle leads to pseudoaxial position. The energies and populations of these conformers for the isolated molecule are summarized in Table 7.1. These conformers, along with atom numbering for identifying the ring puckering angle, are shown in Figure 7.1. The conformation with a positive C9-C4-C3-C2 dihedral angle has the lowest Gibbs energy and is expected to be present with 75% population in vacuum. The experimental VA and VCD spectra obtained for the neat liquid sample are compared to the population-weighted

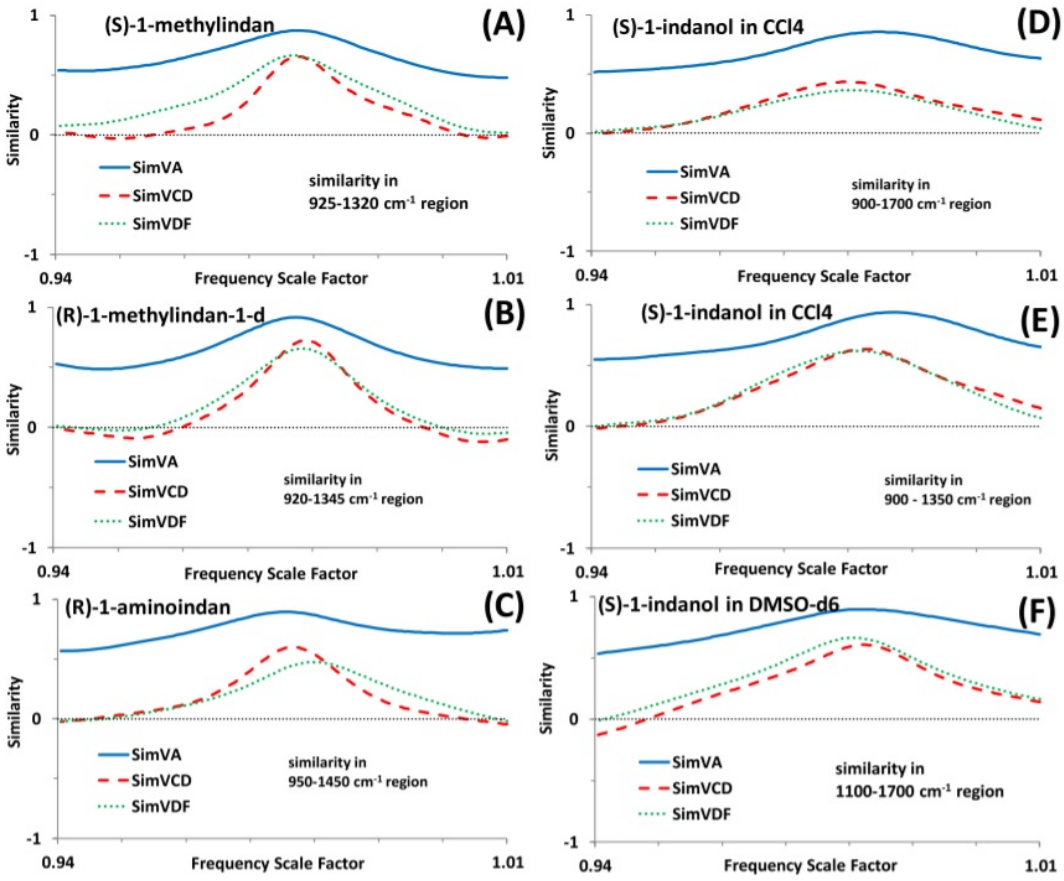

Figure 7.3: SSO plots (A-F) for investigated substituted indans

QC-predicted spectra for the isolated molecule in Figure 7.2.

conf #	**Gibbs Energy**	**Energy** $kcal/mol$	**population**	**C9-C4-C3-C2**
1	−388.25842	0.0	0.75	17.0
2	−388.25741	0.63	0.25	−15.2

Table 7.1: Energies, Populations, and dihedral angle of Conformers of 1-methylindan

In Figure 7.2, the left vertical panels display predicted spectra with vibrational frequencies scaled with 0.9785 (which corresponds to maximum SSO of VCD spectra) as overlaid with experimental spectra whereas the right vertical panels provide stacked display of predicted spectra with unscaled frequencies. Although the experimental VA spectra are shown until 1425 cm^{-1}, the experimental VCD was not displayed above 1320

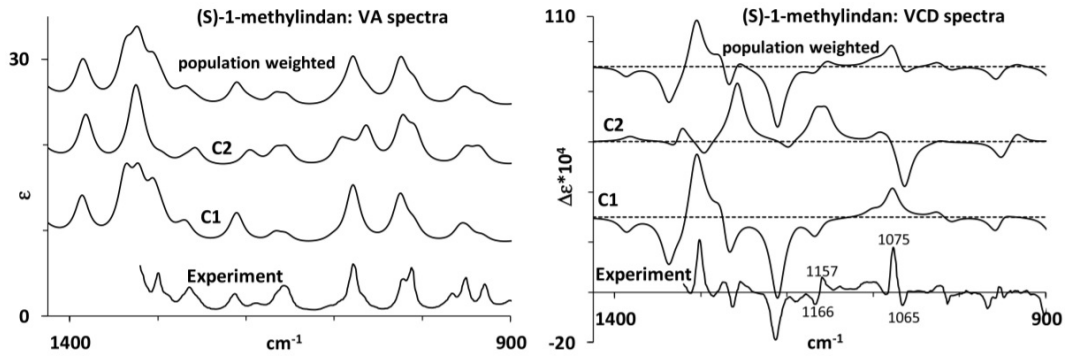

Figure 7.4: Comparison of individual conformer and population-weighted predicted VA (left panel) and VCD (right panel) with corresponding experimental spectra for (*S*)-1-methylindan

cm^{-1} due to excessive absorbance present in the 1325-1425 cm^{-1} region. Experimental VA and VCD are taken from Fontana et al.[194] From the overlaid experimental and predicted spectra, it can be seen that the predicted spectra faithfully reproduce the bands seen in the experimental spectra and band-by-band correlation in experimental and predicted spectra can be seen. The correlation of experimental (and corresponding unscaled predicted) VA band positions (in cm^{-1} units) noted from the stacked spectra is as follows: 1372 (1412), 1326 (1349), 1299 (1330), 1263 (1293), 1211 (1235), 1154 (1183), 1077 (1100), 1020 (1044), 1012 (1029), 950 (969), and 928 (950). Similarly, the correlation of experimental (and corresponding unscaled predicted) VCD band positions is as follows: 1301 (1329), 1275 (1302), 1262 (1293), 1212 (1236), 1166 (1194), 1157 (1174), 1075 (1100), 1065 (1082), 1019 (1045), 1013 (1026), 967 (976), and 900 (910). This comparison indicates that almost all experimental bands are satisfactorily reproduced in the predicted spectra. The same type of correlation can also be seen in VDF spectra, as is clearly apparent from the overlaid spectra.

The similarity or agreement between experimental and predicted spectra is quantified using maximum *Sim*VA, *Sim*VCD, and *Sim*VDF values (see SSO plots in Figure 7.3A), which are respectively, 0.87, 0.65, and 0.66 for (*S*)-1-methylindan. The *Sim*VCD and *Sim*VDF values obtained here are excellent, because these higher magnitudes are not that routine, and for typical cases, they can be as low as 0.4.[51] Achieving a *Sim*VCD value of +0.4 or higher is recommended for an acceptable agreement between experimental and calculated spectra for assigning the correct molecular structures.[51] Approximately the same value of *Sim*VDF is also recommended. This criterion was established by Dr. Polavarapu[23] and Covington et al.[52] while Shen et al. proposed only a magnitude greater than 0.2 for confidence in spectral similarity.

Since there are two ring puckering conformers, it is useful to understand how the experimental spectra are influenced by the individual conformers. The spectra predicted for individual conformers and their Boltzmann population-weighted spectra are compared to the experimental spectra in Figure 7.4. Experimental VA and VCD are taken from Fontana et al.[194] The predicted frequencies are scaled with 0.9785 (which corresponds

to maximum SSO of VCD spectra).

From this comparison, it can be seen that the major VA and VCD spectral features seen in the experimental spectra have counterparts in those predicted for the conformation with a positive puckering angle (#1 in Table 7.1). However, it is important to note that the presence of both conformers is required for reproducing the experimental positive-negative bisignate VCD couplet (positive at 1075 and negative at 1065 cm^{-1}), with positive VCD mostly coming from conformer 1 and negative VCD mostly coming from conformer 2. Similarly, the presence of both conformers is required for reproducing the experimental weak negative-positive bisignate VCD couplet (negative at 1166 and positive at 1157 cm^{-1}), with negative VCD mostly coming from conformer 1 and positive VCD coming from conformer 2. This information on conformer-specific VCD bands is useful for future investigations on temperature-dependent conformer equilibrium and associated thermodynamic parameters.

7.3.2 (*R*)-1-Methylindan-1-d

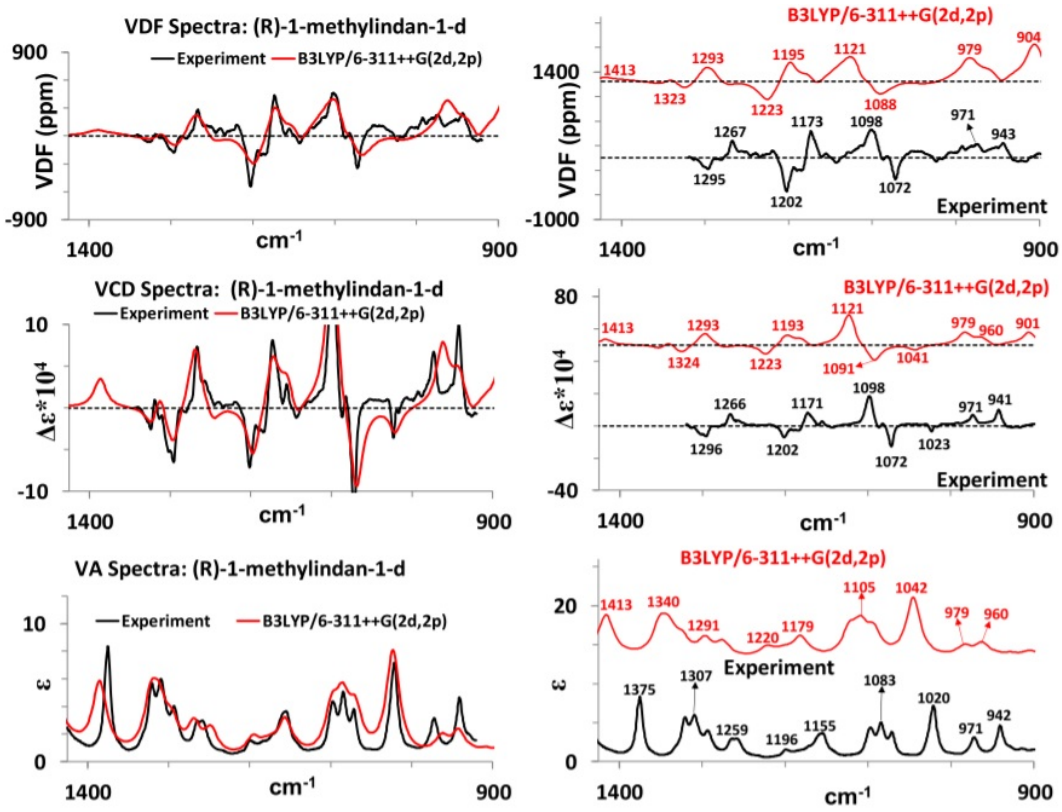

Figure 7.5: Experimental and predicted VA, VCD, and VDF spectra for (*R*)-1-methylindan-d

Note that the substitution of hydrogen atom with deuterium does not alter the conformational space or the optimized geometries. However, Gibbs energies and hence populations of conformers can be different

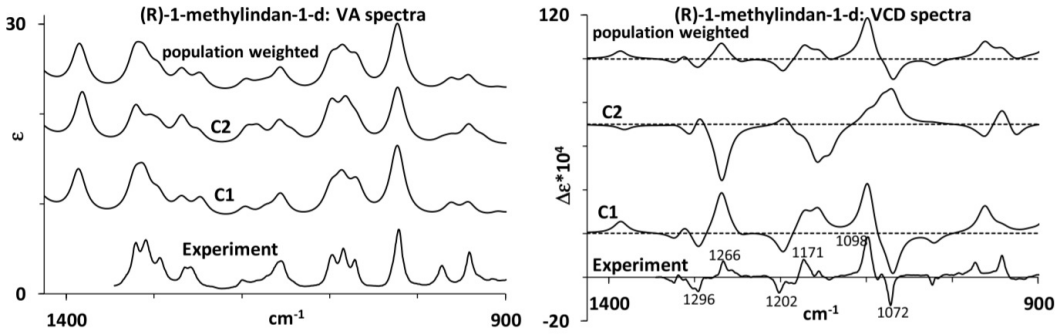

Figure 7.6: Comparison of individual conformer and population-weighted predicted VA (left panel) and VCD (right panel) with corresponding experimental spectra for (*R*)-1-methylindan-1-d

upon deuterium substitution. The energies, populations, and ring puckering angles for the two conformers are summarized in Table 7.2.

The deuterium substitution has changed the populations of the two conformers only slightly. The conformer with the positive 9-4-3-2 dihedral angle is expected to be present with 74% population (as opposed to 75% in 1-methylindan) in vacuum. The experimental VA and VCD spectra obtained for the neat liquid sample are compared to the population-weighted QC-predicted spectra for isolated molecules in Figure 7.5. In the left vertical panels of Figure 7.5, predicted spectra with vibrational frequencies scaled with 0.978 (which corresponds to the maximum SSO of VCD spectra) are overlaid on experimental spectra. In the right vertical panels of Figure 7.5, predicted spectra with unscaled vibrational frequencies are stacked above experimental spectra and individual band peak positions are labeled. Experimental VA and VCD are taken from Fontana et al.[194] Although the experimental VA spectra are shown until 1425 cm^{-1}, the experimental VCD was not displayed above 1345 cm^{-1} due to excessive absorbance present for the 1375 cm^{-1} band. As with (*S*)-1-methylindan, it can be seen that the predicted spectra faithfully reproduce the bands seen in the experimental spectra for (*R*)-1-methylindan-1-d, and band-by-band correlation in experimental and predicted spectra can be seen. The correlation of experimental (and corresponding unscaled predicted) VA band positions noted from the stacked spectra is as follows: 1375 (1413), 1307 (1340), 1259 (1291), 1196 (1220), 1155 (1179), 1083 (1105), 1020 (1042), 971 (979), and 942 (960). Similarly, the experimental (and the corresponding unscaled predicted) VCD band positions are as follows: 1296 (1324), 1266 (1293), 1202 (1223), 1171 (1193), 1098 (1121), 1072 (1091), 1023 (1041), 971 (979), and 941 (960). This comparison indicates that almost all of the experimental bands are satisfactorily reproduced in the predicted spectra. The same type of correlation is apparent in the overlaid VDF spectra.

The agreement between experimental and predicted spectra is quantified using maximum *Sim*VA, *Sim*VCD, and *Sim*VDF values (see SSO plots in Figure 7.3B), which are, respectively, 0.92, 0.72, and 0.66. These val-

conf #	Gibbs Energy	Energy $kcal/mol$	population	C9-C4-C3-C2
1	−388.26188	0.0	0.74	17.0
2	−388.2609	0.61	0.26	−15.2

Table 7.2: Energies, Populations, and dihedral angle of Conformers of 1-methylindan-1-d

ues are even better than those for (*S*)-1-methylindan (see the comparison in Table 7.3) and reflect the excellent agreement between experimental and predicted spectra.

As with (*S*)-1-methylindan, the role of two individual conformer contributions to the experimental spectra of (*R*)-1-methylindan-1-d can be analyzed. For this purpose, the predicted spectra for two individual conformers and their population-weighted average are compared to the experimental spectra in Figure 7.6. The predicted frequencies are scaled with 0.978 (which corresponds to the maximum SSO of VCD spectra). The dominant VCD features seen in the experimental spectrum are clearly present in the predicted spectrum of conformer with a positive puckering angle (#1 in Table 7.2). For example, the dominant positive-negative bisignate VCD couplet with positive VCD at 1098 cm^{-1} and negative VCD at 1072 cm^{-1} originates predominantly from conformer #1. It is of interest to note that the signs of intense VCD bands of conformer #2 are opposite to those of conformer #1. As a result, the experimental intensities of bisignate VCD couplets (1296(-)/1266(+) and 1202(-)/1171(+)) of (*R*)-1-methylindan-1-d can be seen to result from a significant reduction of the VCD band intensities of conformer #1 by the opposing contributions from conformer #2. As with (*S*)-1-methylindan, this information on conformer specific VCD bands is useful for future investigations on temperature-dependent conformer equilibrium and associated thermodynamic parameters.

7.3.3 (*R*)-1-aminoindan

The energies, populations, and ring puckering angles for the six conformers of 1-aminoindan are summarized in Table 7.4. The structures are displayed in Figure 7.7. Conformers 5, 3, and 1, with pseudoequatorial orientation (positive ring puckering angle) contribute to 80% of the population. Conformers 6, 4, and 2, with negative C9-C-4-C3-C2 dihedral angle with pseudoaxial orientation contribute to 20% of the population. Thus, as with 1-methylindans, the pseudoequatorial orientation is predicted to be the dominant contributor to vibrational spectra.

The experimental VA and VCD spectra obtained for the neat liquid sample are compared to the population-

Molecule	*Sim*VA	*Sim*VCD	*Sim*VDF
(*S*)-1-methylindan	0.87	0.65	0.66
(*R*)-1-methylindan-1-d	0.92	0.72	0.66
(*R*)-1-aminoindan	0.89	0.60	0.47
(*S*)-1-indanol in CCl_4	0.94	0.63	0.62
(*S*)-1-indanol in DMSO-d_6	0.89	0.61	0.66

Table 7.3: Maximum SSO Values in Similarity Analysis for Substituted Indans

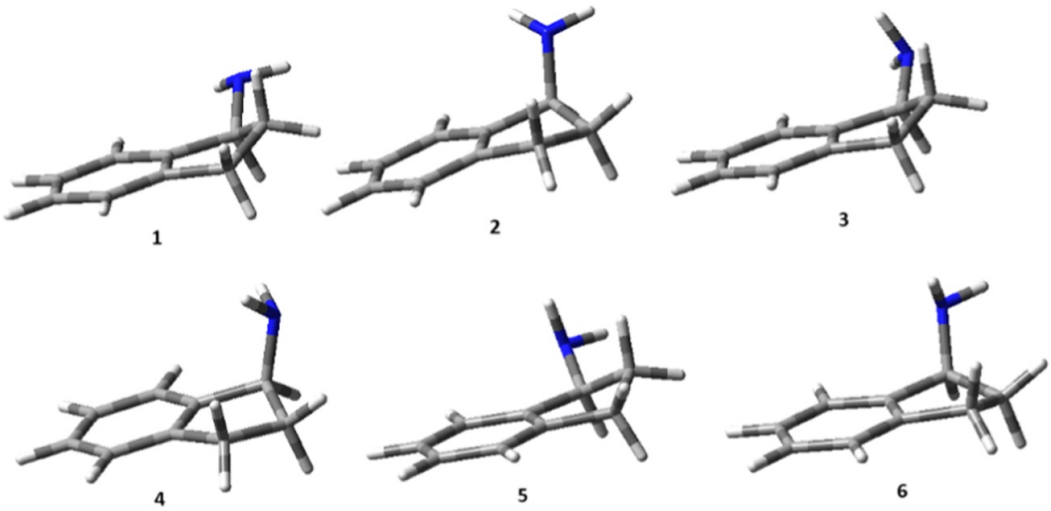

Figure 7.7: Six conformers of (*S*)-1-aminoindan

weighted QC-predicted spectra for isolated molecules in Figure 7.8. The predicted VA and VCD spectra of individual conformers are presented by Johnson et al.[162] In the left vertical panels of Figure 7.8, predicted spectra with vibrational frequencies scaled with 0.977 (which corresponds to maximum SSO of VCD spectra) are overlaid on experimental spectra. In the right vertical panels of Figure 7.8, predicted spectra with unscaled vibrational frequencies are stacked above experimental spectra and individual band peak positions are labeled. Experimental VA and VCD are taken from Fontana et al.[194] The experimental region above 1450 cm^{-1} is associated with excessive absorbance and not displayed. As with (*S*)-1-methylindan and (*R*)-methylindan-1-d, it can be seen that the predicted spectra faithfully reproduce the vibrational bands seen in the experimental spectra.

The correlation of experimental (and corresponding unscaled predicted) VA band positions noted from the stacked spectra is as follows: 1377 (1413), 1317 (1349), 1296 (1327), 1259 (1293), 1215 (1240), 1184 (1204),

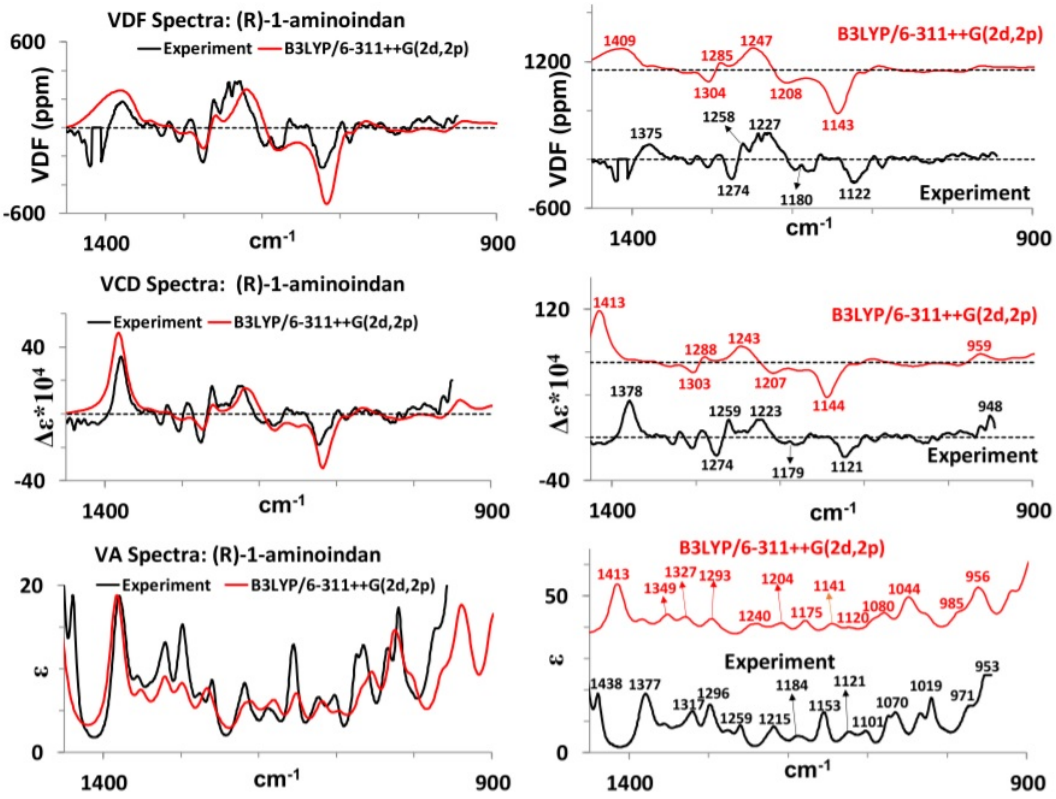

Figure 7.8: Experimental and predicted VA, VCD, and VDF spectra for (*R*)-1-aminoindan

1153 (1175), 1121 (1141), 1101 (1120), 1070 (1080), 1019 (1044), 971 (985), and 953 (956). Similarly, the correlation of experimental (and corresponding unscaled predicted) VCD band positions is as follows: 1378 (1413), 1274 (1303), 1259 (1288), 1223 (1243), 1179 (1207), 1121 (1144), and 948 (959). Quantification of the agreement between experimental and predicted spectra can be gleaned from SSO plot (see Figure 7.3C and Table 7.3), where the maximum *Sim*VA, *Sim*VCD, and *Sim*VDF values are, respectively, 0.89, 0.60, and 0.47. The *Sim*VCD and *Sim*VDF magnitudes for 1-aminoindan are slightly smaller than those for 1-methylindan and 1-methylindan-1-d, which is probably because of the additional conformational degrees of freedom arising from the NH_2 group (each of the two puckered ring conformers is associated with three different orientations of the NH_2 group). Nevertheless, the agreement between experimental and predicted VA and VCD spectra of 1-aminoindan is remarkably good.

It is useful to note that the liquid-phase experimental spectra are satisfactorily reproduced by gas-phase calculations for 1-methylindan, 1-methylindan-1-d, and 1-aminoindan. This observation raises the relevance of solute-solute interactions in condensed media. Although polarizable continuum model (PCM)[44] calculations might not be useful for shining light on this issue, liquid-phase MD simulations might be able to.[201]

conf #	Energy *kcal/mol*	population	C9-C4-C3-C2	C9-C1-N-H1	C9-C1-N-H2
5	0.00	0.43	17.1	65.9	−174.8
3	0.18	0.32	16.8	−70.3	48.8
4	0.98	0.08	−13.7	−64.8	51.7
6	1.10	0.07	−14.4	55.9	175.2
2	1.18	0.06	−16.0	−172.7	−54.2
1	1.28	0.05	16.6	179.8	−63.0

Table 7.4: Energies, Populations, and dihedral angle of Conformers of 1-aminoindan

7.3.4 (*S*)-1-Indanol in Nonhydrogen-bonding Solvent CCl_4

Whereas six stable conformations were found for 1-aminoindan, only four initial conformations of different energies were generated by CONFLEX.[71] To avoid missing the remaining two conformations, I manually recreated the six conformations and optimized their geometries at B3LYP/6- 311++G(2d,2p) level, followed by VA and VCD calculations at the same level of theory. These six conformers are identical to those reported recently by Zehnacker et al.[163] The geometries of conformers for (*S*)-1-indanol were also optimized, with PCM representing CCl_4 solvent at B3LYP/6-311++G- (2d,2p) level, followed by VA and VCD calculations at the same level of theory. The use of PCM has only a minor influence on conformational populations and predicted spectra, as already pointed out by Zehnacker et al. For this reason, the comparison of experimental spectra is restricted to the QC predicted spectra obtained with PCM (see Table 7.5, Figures 9 and 3).

conf #	Energy *kcal/mol*	population	C9-C4-C3-C2	C9-C1-O-H
5	0.00	0.33	16.5	54.6
4	0.27	0.21	−12.9	38.3
6	0.45	0.16	−14.5	−47.2
3	0.59	0.12	15.7	−76.8
1	0.62	0.12	16.1	−167.3
2	1.01	0.06	−14.9	−176.1

Table 7.5: Energies, Populations, and dihedral angle of Conformers of 1-indanol in CCl_4

The experimental VA, VCD, and VDF spectra obtained for (*S*)-1-indanol in CCl_4 are compared to the

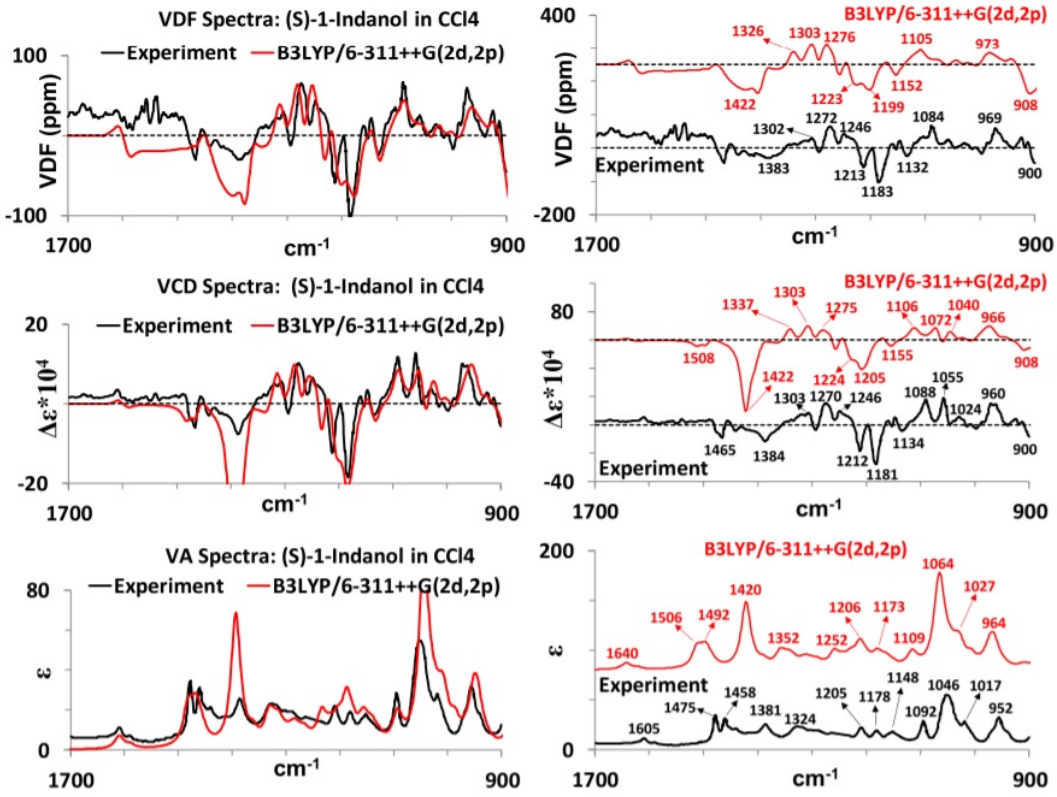

Figure 7.9: Experimental and predicted VA, VCD, and VDF spectra for (*S*)-1-indanol

populationweighted QC-predicted spectra, with PCM representing the solvent environment, in Figure 7.9. In the left vertical panels of Figure 7.9, predicted spectra with vibrational frequencies scaled with 0.98 (which corresponds to the maximum SSO of VCD spectra) are overlaid on experimental spectra. In the right vertical panels of Figure 7.9, predicted spectra with unscaled vibrational frequencies are stacked above experimental spectra and individual band peak positions are labeled. Experimental VA and VCD are taken from Le Barbu-Debus et al[163] with permission from the Royal Society of Chemistry. There is a good correlation of experimental (and predicted) bands, which in VA spectra is as follows: 1605 (1640), 1475 (1506), 1458 (1492), 1381 (1420), 1324 (1352), 1205 (1252), 1178 (1206), 1148 (1173), 1092 (1109), 1046 (1064), 1017 (1027), and 952 (964) and that in VCD is as follows: 1465 (1508), 1384 (1422), 1303 (1337), 1270 (1303), 1246 (1275), 1212 (1224), 1181 (1205), 1134 (1155), 1088 (1106), 1055 (1072), 1024 (1040), and 960 (966).

The overlaid VA and VCD spectra show an excellent match in the 1350-900 cm^{-1} region of the experimental spectrum. But neither VA nor VCD has a good match in the 1350-1450 cm^{-1} region of the experimental spectrum, mainly because the predicted VA spectrum has very large intensity at 1420 cm^{-1} compared to that of the corresponding experimental VA band at 1381 cm^{-1}. The same discrepancy appears between the corresponding predicted VCD at 1422 cm^{-1} and experimental VCD at 1384 cm^{-1}. The maximum *Sim*VA,

*Sim*VCD, and *Sim*VDF values obtained, for the 900-1700 cm^{-1} region of the experimental spectrum (Figure 7.3D and Table 7.3), are 0.86, 0.43, and 0.36, respectively. Of these, *Sim*VCD and *Sim*VDF values are not as large as those seen for (*S*)-1-methylindan, (*R*)-1-methylindan-1-d, or (*R*)-1-aminoindan. The reason for lower SSO values for (*S*)-1-indanol in CCl_4 is due to the mismatch of the large intensities associated with the predicted band at 1420 cm^{-1}. Both absorption and VCD intensities predicted at 1420 cm^{-1} are significantly higher than the corresponding intensities observed in the experiment. If the region of similarity analysis is restricted to 900-1350 cm^{-1} region of the experimental spectra, then the *Sim*VA, *Sim*VCD, and *Sim*VDF values increase to 0.94, 0.63, and 0.62, respectively (see Figure 7.3E), which are better than those for 1-aminoindan and closer to those of 1-methylindans (see Table 7.3). The higher VA and VCD intensities predicted at 1420 cm^{-1} originate from coupled O-H and C-H bending motions at 1-position.

VA and VCD spectra of (*S*)-1-indanol in CCl_4 have been analyzed previously by Zehnacker et al.[163] They reported predicted spectra using the B3LYP functional and 6-31++G(d,p) basis set and adopting two different approaches: (1) conventional static optimized geometries, nuclear velocity perturbation (NVP) theory[202,203,204] for calculating VCD, and PCM for representing the CCl_4 solvent. This approach was referred to as static optimized geometry method. (2) Ab initio molecular dynamics for capturing the dynamical behavior of solute configurations, optimizing the geometries of snapshots extracted from the MD trajectories and NVP theory for predicting VCD. This approach was referred to as first principle molecular dynamics (FPMD) method. The analysis of Zehnacker et al. was limited to qualitative visual comparison of selected VCD bands. Their focus was on large intensities predicted for the 1420 cm^{-1} band and incorrect relative intensities predicted for the negative doublet at 1200 cm^{-1}. Although these discrepancies were considered to have been corrected in their FPMD calculations,[163] the following negative points in the FPMD predictions can be noticed: the predicted spectrum in the FPMD approach did not show bands corresponding to the experimental positive VCD doublet at 1088 and 1055 cm^{-1} ; the FPMD-predicted spectrum also did not reproduce the experimental negative VCD band at 1465 cm^{-1} (instead the FPMD-predicted spectrum showed a negative VCD couplet; positive on the higher frequency side and negative on the lower frequency side) in the 1425-1475 cm^{-1} region. Thus, although FPMD calculations may have appeared to have corrected the deficiencies associated with the predicted band intensities at 1420 and 1200 cm^{-1} , two other regions (1000-1100 and 1400-1500 cm^{-1}) of the experimental spectra were negatively influenced. As a result, the overall spectral similarity between FPMD-predicted and experimental VCD spectra of (*S*)-1-indanol in CCl_4 appears less than desired.

The current simulated spectra at the B3LYP/6-311+ +G(2d,2p) level, using conventional static optimized geometries, match the experimental spectra in the 1350-900 cm^{-1} region (see Table 7.3 and Figure 7.3E, and overlaid spectra in Figure 7.9) and also the 1450-1480 cm^{-1} region very well. The discrepancy associ-

ated with higher intensities predicted for the 1420 cm^{-1} band of (*S*)-1-indanol in CCl_4 (compared with the corresponding experimental intensities), however, remains to be resolved.

It appears then that, identifying a method to correct for the overestimated intensities predicted for 1420 cm^{-1} band for 1-indanol in CCl_4, without negatively influencing other regions of the spectra, will be useful. Different directions were pursued for this purpose.

First, I followed the recent findings of Nicu et al.,[205] where the overestimated intensities in the 1160-1380 cm^{-1} region for vibrations involving COH bending in 3-methyl-1-(methyldiphenylsilyl)-1-phenylbutan-1-ol could be satisfactorily corrected by averaging over the spectra obtained for thermal fluctuations in C-O-H angle. Taking cues from this study, I changed the C-O-H angle in the fully optimized geometries of 1-indanol conformers in three steps of +2° increment each and three steps of -2° increment each and performed constrained geometry optimizations. Some of the resulting geometries with positive increments were found to have imaginary frequencies. The Gibbs energy-derived Boltzmann population-weighted spectra, including those of additional C-O-H angle variations that did not have imaginary frequencies, are then compared to the experimental spectra and similarity analyses performed. Unfortunately, these results (*Sim*VA = 0.80, *Sim*VCD = 0.40, and *Sim*VDF = 0.36 for the 900-1700 cm^{-1} region;[162]) also did not provide improvement over the spectra shown in Figure 7.9 and Table 7.3. Second, individual conformer spectra indicate that out of the four lowest energy conformers, two of the conformers (C5 and C4) make dominant contributions to VA intensity and three conformers (C5, C4, and C3) make dominant contributions to VCD intensity at 1420 cm^{-1}. The C6 conformer does not contribute to this band. Assuming that Boltzmann populations determined from Gibbs energies may not be reflecting the actual populations in the experiment, I considered the following: (a) the populations derived from electronic energies; (b) equal populations for all six conformers, and (c) the influence of increasing the population of C6 and decreasing C4. With these altered populations, spectra were simulated for each case and similarity analysis carried out. The spectra derived from populations with electronic energies yielded slightly inferior similarity values (*Sim*VA = 0.84, *Sim*VCD = 0.41, and *Sim*VDF = 0.36) for the 900-1700 cm^{-1} region. The simulated spectra with equal populations yielded lower intensities for the 1420 cm^{-1} band and improved overall quantitative similarity for the 900-1700 cm^{-1} region in VA and VCD spectra slightly (*Sim*VA = 0.87 and *Sim*VCD = 0.48), but *Sim*VDF has decreased from 0.36 to 0.28. When the population of C6 was increased and that of C4 decreased, via the combination C5(0.29), C6(0.27), C3(0.14), C4(0.1), C2(0.1), and C1(0.1), the predicted intensities of the 1420 cm^{-1} band were lowered and some improvement was found for the similarity values for the 900-1700 cm^{-1} region (*Sim*VA = 0.88, *Sim*VCD = 0.46, and *Sim*VDF = 0.40) compared to those in Table 7.3. It turned out to be difficult to improve the similarity values further without negatively influencing the VCD doublet at 1088 and 1055 cm^{-1}. Thus, a satisfactory procedure for fully correcting the overestimated intensities of the 1420 cm^{-1} band of 1-

indanol in CCl_4, without negatively influencing other regions of spectra, remained elusive. Even if a different combination of populations happens to improve the similarity values, the task of finding experimental proof for such hypothesized populations remains to be undertaken.

7.3.5 (*S*)-1-Indanol in Hydrogen-bonding Solvent, DMSO-d_6

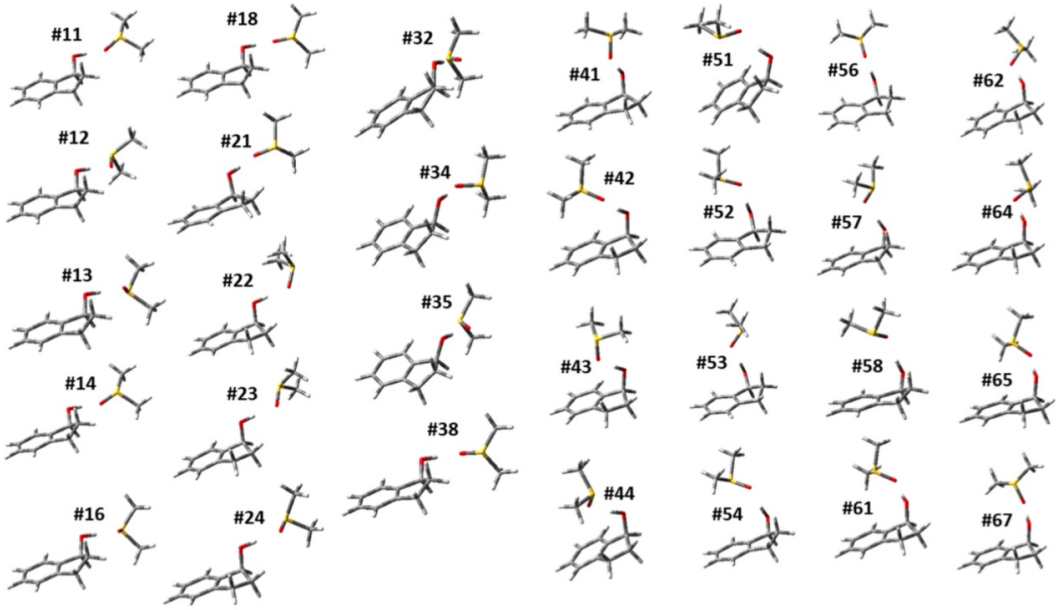

Figure 7.10: Different conformers of (*S*)-1-indanol-DMSO complexes

To analyze the experimental spectra of (*S*)-1-indanol obtained in DMSO-d_6 solvent, two different calculations are undertaken: (a) one set of calculations is carried out with PCM representing the DMSO solvent. However, PCM cannot realistically represent the solute-solvent interactions, especially when hydrogen bonding is involved.[69] (b) DMSO solvent has very high propensity for forming hydrogen bonds with solute molecules containing O-H and N-H groups, and in such cases, successful reproduction of experimental spectra by the predicted spectra without incorporating explicit solvent molecules becomes a challenge. For this reason, an explicit DMSO molecule is added in two different orientations to each of the six conformations of (*S*)-1-indanol. These complexes are optimized at B3LYP/6-311++G(2d,2p) level with PCM representing DMSO solvent environment. The complexed DMSO molecule in each of the optimized structures is then rotated to 90, 180, and 270° around the hydrogen bond to explore the conformational space of the 1-indanol/DMSO complex. This process led to investigating a total of 48 conformers. The optimized 30 lowest energy structures at B3LYP/6-311++G(2d,2p) level, with PCM representing DMSO solvent environment, are used for VA and VCD calculations on the 1-indianol/DMSO-d_6 complex at the same level of theory, and none of these structures had imaginary vibrational frequencies. The energies, populations, and dihedral angles as-

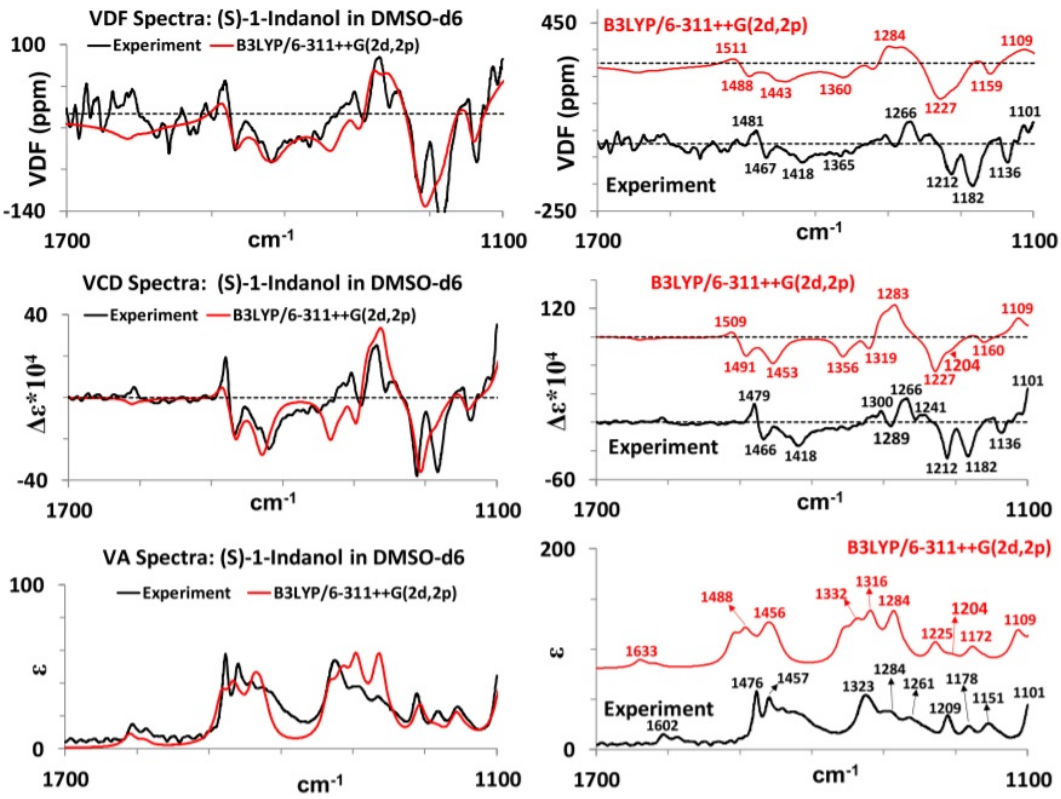

Figure 7.11: Comparison of experimental and predicted VA, VCD, and VDF spectra for (*S*)-1-indanol-DMSO-d_6

sociated with these 30 conformers are presented in Table 7.6 and conformer structures displayed in Figure 7.10.

Table 7.6 lists some conformers with nearly the same C9-C4-C3-C2 and C9-C1-O-H dihedral angles. For example, the five lowest energy conformers (#18, #14, #38, #16, and #12) have nearly the same C9-C4-C3-C2 and C9-C1-O-H dihedral angles, making one wonder if they are in fact distinctly separate conformers. In this connection, it is useful to note that the orientation of DMSO molecules can be different even for the same C9-C4-C3-C2 and C9-C1-O-H dihedral angles. For this reason, an additional dihedral angle C1-O···O-S, relating the C1-O bond of the parent molecule to the S-O bond of the solvent molecule is also included in Table 7.5. With the three dihedral angles listed, one can identify the differences in conformers that have the C9- C4-C3-C2 and C9-C1-O-H dihedral angles. However, even with the use of these three dihedral angles, the two higher energy structures #41 and #43 do not appear to be significantly different. But they are in fact different in terms of the inversion at S atom of the solvent molecule (see Figure 7.10). It is important to note that (*S*)-1-indanol/DMSO complexes adopt both positive and negative ring puckering angles for the five-membered ring, with 77% population contribution coming from the conformers with positive ring puckering angle. Also, since O-H group can have free rotation around C-O bond, the hydrogen-bonded DMSO

conf #	Energy $kcal/mol$	population	C9-C4-C3-C2	C9-C1-O-H	C1-O···O=S
18	0.00	0.11	16.46	−156.29	−78.0
14	0.19	0.08	16.46	−157.57	−91.8
38	0.20	0.08	16.44	−157.22	−63.5
16	0.32	0.07	16.47	−156.33	111.8
12	0.40	0.06	16.44	−156.31	144.5
35	0.42	0.06	16.40	−98.20	85.7
13	0.46	0.05	16.48	−158.42	97.8
54	0.57	0.04	16.31	59.84	−110.4
61	0.70	0.03	−14.25	−74.15	105
56	0.70	0.03	16.40	60.81	−173.6
32	0.71	0.03	16.38	−97.05	99.1
11	0.71	0.03	16.47	−157.76	−107.8
34	0.78	0.03	16.21	−83.62	−97.3
62	0.86	0.03	−14.11	−73.60	−129
21	0.87	0.03	−14.52	−169.88	−89.3
22	0.89	0.03	−14.55	−170.44	157.8
24	1.01	0.02	−14.47	−170.72	105.2
67	1.04	0.02	−14.32	−73.53	147.8
51	1.06	0.02	16.32	58.73	5.3
52	1.07	0.02	16.27	59.56	−102.2
57	1.17	0.02	16.43	58.51	95.5
64	1.18	0.02	−14.12	−74.14	−115.7
58	1.18	0.02	16.37	59.00	−5.1
23	1.19	0.02	−14.59	−170.75	140.2
53	1.21	0.01	16.41	58.85	112.2
42	1.33	0.01	−13.02	54.90	−101.1
65	1.35	0.01	−14.28	−73.85	111
41	1.48	0.01	−13.42	55.67	175.3
44	1.57	0.01	−13.21	55.06	126.2
43	1.76	0.01	−13.57	57.49	178.1

Table 7.6: Energies, Populations, and dihedral angle of Conformers of 1-indanol in DMSO-d_6

molecule is in a different relative orientation with respect to the parent molecule. This should not be viewed as a hydrogen-bonded DMSO molecule rotating with O-H group; instead, this could be viewed as a different DMSO molecule from the solvent bath hydrogen bonding to O-H in a different orientation. The experimental VA and VCD spectra of 1-indanol in DMSO-d_6 solvent are compared to the population-weighted QC-predicted spectra for 1-indanol/DMSO-d_6 complex in Figure 7.11, and SSO plots are displayed in Figure 7.3F. The left vertical panels display overlaid spectra after scaling the calculated vibrational frequencies with 0.9825, whereas the right vertical panels display stacked spectra with unscaled frequencies and labeled band

positions. Experimental VA and VCD are taken from Le Barbu-Debus et al[163] with permission from the Royal Society of Chemistry. The excellent correlation between experimental and predicted spectra features is surprising, given that modeling hydrogen bonding interactions with the solvent is often challenging.[69] The correlation of experimental (and predicted) spectra is as follows. In VA spectra: 1602 (1633), 1476 (1488), 1457 (1456), 1323 (1332), 1284 (1316), 1261 (1284), 1209 (1225), 1178 (1204), 1151 (1172), and 1101 (1109); in VCD spectra: 1479 (1509), 1466 (1491), 1418 (1453), 1300 (1356), 1289 (1319), 1266 (1283), 1212 (1227), 1182 (1204), 1136 (1160), and 1101 (1109). Of these, there are three places where some differences are apparent: The weak positive VCD band at 1300 cm^{-1} seen in the experimental spectrum does not match in sign with the predicted VCD at 1356 cm^{-1} . The weak negative experimental VCD band at 1289 cm^{-1} is not as well developed as the corresponding predicted negative VCD at 1319 cm^{-1} . The weak positive VCD band seen in the experimental spectrum at 1241 cm^{-1} is not resolved in the predicted spectrum. Except for the differences associated with weak bands mentioned above, the overlaid VA and VCD spectra show good match in the entire 1700-1100 cm^{-1} region of the experimental spectrum. The quantification of this match is reflected in a similarity analysis (see Figure 7.3H and Table 7.3), which yields 0.89, 0.61, and 0.66 for *Sim*VA, *Sim*VCD, and *Sim*VDF, respectively. The *Sim*VCD and *Sim*VDF values obtained here are better than those for 1-aminoindan and 1-indanol in CCl_4. Also, the *Sim*VDF value obtained for 1-indanol/DMSO-d_6 complex is larger than that for 1-aminoindan and as high as that for 1-methylindan and 1-methylindan-1-d. VA and VCD spectra of (*S*)-1-indanol in DMSO-d_6 have been analyzed previously by Zehnacker et al.[163] They reported predicted spectra using NVP theory, B3LYP functional, and 6- 31++G(d,p) basis set adopting different approaches: (1) conventional static optimized geometries and PCM representing the DMSO solvent continuum; (2) conventional static optimized geometries for 1:1 indanol/DMSO complex and PCM representing the DMSO solvent continuum; (3) conventional static optimized geometries for 1:2 indanol/ DMSO complex and PCM representing the DMSO solvent continuum; and (4) FPMD for capturing the dynamical behavior of 1-indanol in a box of DMSO solvent molecules, optimizing the geometries of snapshots extracted from the MD trajectories. Optimization of snapshots from FPMD trajectories resulted in both pseudoequatorial and pseudoaxial conformers and provided improved comparison between predicted and experimental VCD spectra, although disagreement (resulting from a large negative doublet in the predicted spectrum) in the region around 1300 cm^{-1} still persisted. The main point leading to the differences in the previous and present results can be associated with the following aspect: in the previous conventional static optimized geometry calculations,[163] only four major geometries of the 1:1 indanol/DMSO complex, all with pseudoequatorial orientation of O-H group, were considered for VCD predictions on 1-indanol/DMSO-d_6 complex. This is different from the current finding that nearly 30 different static geometries (with 77% population coming from pseudoequatorial conformers and 23% from pseudoaxial) are needed to describe the

1-indanol/ DMSO conformational space. The critical point to recognize here is that rotating the O-H group around C-O bond and complexing a DMSO molecule in each of the new O-H orientations is important for exploring the conformational space of the 1-indanol/DMSO complex. With these precautions, the present conventional static optimized geometry calculations can be seen to be a successful (see Table 7.3, Figure 7.3 and overlaid spectra in Figure 7.11) alternative to FPMD simulations in reproducing the experimental VCD spectra in the 1700-1100 cm^{-1} region of the 1-indanol/ DMSO-d_6 complex. This observation should not be interpreted as discouraging the general use of MD simulations, because the analyses of solution-phase spectra using MD geometries have certainly helped in many situations.[156,157,197]

7.4 Conclusion

(*S*)-1-Methylindan, and (*R*)-1-methylindan-1-d have just two conformations resulting from opposite ring puckering angles. Ring puckering conformer populations that result in pseudoequatorial vs pseudoaxial substituent are in the ratio of 80:20. Ring puckering conformer-specific VCD bands could be clearly identified in the experimental spectra of 1-methylindans, and pseudoaxial contribution is important for reproducing the experimental spectra. (*R*)-1-Aminoindan and (*S*)-1-indanol in CCl_4 have six conformations, resulting from the rotation of the NH_2/OH substituent for each of the two ring puckering conformations. On the contrary, (*S*)-1-indanol in DMSO has numerous hydrogen-bonded conformations with 80% arising from ring puckering that results in the pseudoequatorial substituent and 20% from the pseudoaxial substituent. The QC-predicted spectra for the minimum energy geometry optimized conformers of (*S*)-1-methylindan, (*R*)-1-methylindan-1-d, (*R*)-1-aminoindan, and (*S*)-1-indanol in CCl_4 and (*S*)-1-indanol/DMSO-d_6 complex are found to be in excellent agreement with corresponding experimental VA and VCD spectra. The methods to correct overestimated VA and VCD intensities for the 1420 cm^{-1} band of 1-indanol in CCl_4, without negatively influencing other regions of spectra, however, remain to be identified.

CHAPTER 8

The Vibrational ROA of (1S)-camphor

This work was completed in collaboration with Professor Malgorzata Baranska of Jagiellonian University, Krakow, Poland. This chapter has been reproduced from the *Journal of Raman Spectroscopy* article[206] "Vibrational Raman optical activity of camphor: The importance of electric-dipole—electric-quadrupole polarizability contribution" with permission. My contributions to this work are: measurement of ROA, QC optimizations, QC spectral prediction of ROA, vibrational band assignment, spectral similarity analyses, preparation of figures and tables, drafting the manuscript, revising the manuscript, and recognizing the bands with predominant δ_k^2 contributions.

8.1 Introduction

Raman optical activity (ROA) is the measurement of differential intensity of Raman scattering between right and left circularly polarized light.[23,24,109,110] The experimental vibrational Raman and ROA spectral intensities depend on various experimental parameters, including the incident laser power density, and therefore cannot be quantitatively compared with the intensities calculated using QC methods. However, dimensionless CID, the ratio of ROA spectrum to corresponding Raman spectrum, eliminates the dependence on experimental parameters when both ROA and Raman spectra are measured simultaneously and permits quantitative comparisons in the analysis of QC predictions. In this work, the importance of the electric-dipole-electric-quadrupole polarizability contribution to ROA and CID of (1*S*)-camphor is examined. Additionally, the importance of these polarizability contributions for (1*S*,4*R*)-(+)-fenchone and dimethyloxirane are also investigated. ROA originates from the products of the normal coordinate derivatives of three polarizability tensors. For detailed information on the equations governing the Raman, ROA, and CID see section 1.3.4.

For camphor (Figure 8.1) and related molecules, the depolarized ICP VROA spectra in CS_2 solvent in the 2,000–80 cm^{-1} region were reported by Barron very early on, where the emphasis was on explaining the observed VROA couplets.[207] The predicted ROA spectra of camphor were reported by Bour et al.,[208] with an emphasis on the use of transition polarizability model for predicting induced resonance vibrational ROA.

The importance of the electric-dipole—magnetic-dipole polarizability and the electric-dipole—electric-quadrupole polarizability tensor contributions was assessed experimentally from a combination of different vibrational ROA measurements: (a) polarized and depolarized ICP ROA measurements and (b) depolarized ICP and backscattering DCP_I or 180° backscattering SCP vibrational ROA measurements.

In the right angle scattering geometry, the coefficient a_3 in Equation 1.14 is zero for depolarized ICP

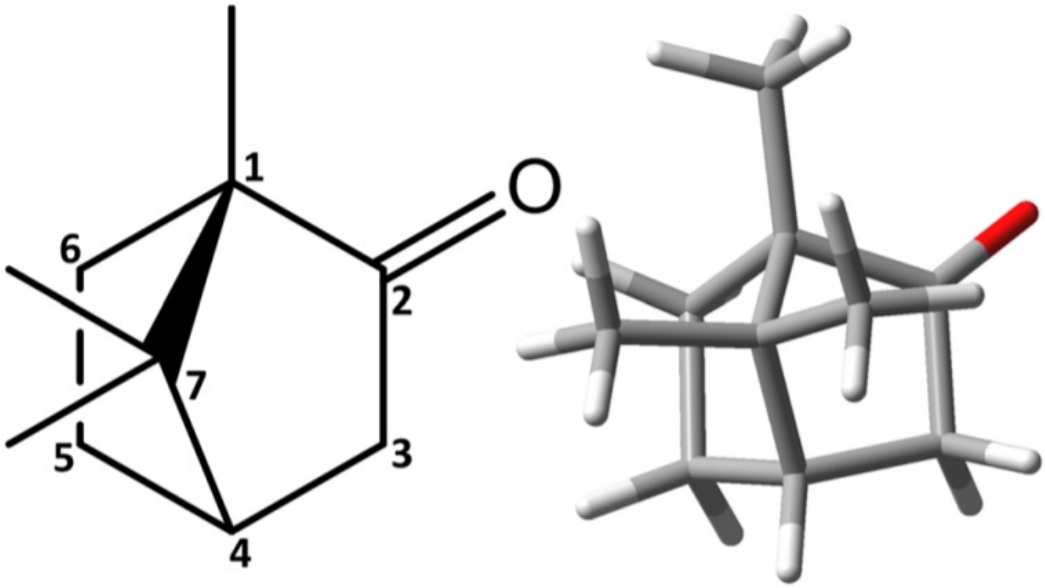

Figure 8.1: (1*S*)-camphor with carbon atom-labeled skeletal structure and its 3D depiction

ROA; also, the ratio a_4/a_5 becomes -3 in depolarized ICP ROA and 7 in polarized ICP ROA. In the approximation that $\omega^{-1}\bar{\alpha}_k\bar{G}'_k$ contribution to polarized ICP ROA is negligible, the ratio of polarized to depolarized ICP ROA will reduce to 2 when $\delta_k^2 = \gamma_k^2$ and 6/8 when $\delta_k^2 = -\gamma_k^2$.[209,210,211,212,213,214] As the ratio of polarized to depolarized ICP ROA for most of the vibrational bands in the experimental spectra for trans-pinane, (+)-α-phenylethylamine, (+)-3-methylcyclohexanone, (-)-α-phellandrene, and α-pinene was found to be between 2 and 1, Barron et al.[209] emphasized that "electric-quadrupole contributions, although smaller than magnetic dipole contributions, are still usually significant." For menthyl chloride, this experimental ratio was found to be $\sim$2 for nearly all bands,[214] indicating that $\delta_k^2 \sim \gamma_k^2$. In the experimental spectra of menthol and neomenthol, some bands were to found to exhibit a ratio of 2, whereas others exhibited a ratio of $\sim$1.[214] In the experimental spectra of thioneomenthol, some bands were found to have larger intensities in depolarized ICP ROA than in polarized ICP ROA.[214] Based on these observations, Barron et al. have suggested that "depolarized vibrational ROA as large as, or larger than, polarized indicates a dominant contribution from the electric-quadrupole mechanism of opposite sign to that from the magnetic dipole mechanism." The overlap between heteroatom Rydberg p orbitals and orbitals on other parts of the molecule was considered to be a source for larger electric-quadrupole contributions in these molecules. In the experimental spectra of dimethyloxirane,[210] depolarized ICP VROA associated with the antisymmetric ring deformation band at 722 cm^{-1} was found to have the opposite sign and approximately half the magnitude as that in the polarized ICP VROA, a situation representing $\delta_k^2 = 6\gamma_k^2$. The depolarized ICP VROA associated with the symmetric ring deformation band of dimethyloxirane at 813 cm^{-1} was found to have the same sign and approximate

magnitude as that in polarized ICP ROA. This observation was considered to represent roughly equal contributions from $\omega^{-1}\gamma_k^2$ and $\omega^{-1}\delta_k^2$ with opposite signs, or predominantly the $\omega^{-1}\gamma_k^2$ contribution. Here again, Rydberg orbitals were thought to be responsible for the dominance of $\omega^{-1}\delta_k^2$ contribution. Even though QC vibrational ROA predictions for dimethyloxirane were undertaken,[211] individual contributions from $\omega^{-1}\gamma_k^2$ and $\omega^{-1}\delta_k^2$ were not analyzed. Although the depolarized ICP VROA sign associated with antisymmetric ring deformation mode at 722 cm^{-1} of dimethyloxirane was reproduced in separate measurements,[210,211,212] that reported a few years later[213] showed opposite sign with weaker intensity compared with those reported earlier.[210,211,212] This latter observation[213] leaves uncertainty regarding the experimental depolarized VROA sign associated with the antisymmetric ring deformation band at 722 cm^{-1} of dimethyloxirane.

The coefficient a_3 in Equation 1.14 is zero for SCP and DCP_I VROA in 180° backscattering geometry (as it is in right angle depolarized ICP ROA), but the ratio a_4/a_5 becomes +3 for both (as opposed to -3 in right angle depolarized ICP ROA). As a result, the comparison of right angle depolarized ICP ROA with backscattering DCP_I ROA or 180° backscattering SCP ROA spectra provided an opportunity to extract two ROA invariants, $\omega^{-1}\gamma_k^2$ k and $\omega^{-1}\delta_k^2$ k . Che and Nafie[215] isolated these two VROA invariants from the experimental ICP and DCP_I VROA spectra for (-)-α-pinene and (+)-trans-pinane. They concluded that "the electric-quadrupole invariant contributes at a level at least three times smaller than the magnetic dipole anisotropic invariant, the influence of the electric-quadrupole scattering mechanism for the molecules considered in this paper is relatively minor and nearly unimportant." In the bond polarizability model, using axially symmetric bonds,[209,216,217] ROA equations simplify due to the emerging relations therein as $\gamma_k^2 = \delta_k^2$ and $\bar{\alpha}_k\bar{G}'_k = 0$. These predictions were also analyzed[215] by Che and Nafie for (-)-α-pinene and (+)-trans-pinane. Because different scattering geometries adopt different scattering cones, and ROA measurement efficiencies differ among different scattering cones, Hecht indicated[218] that isolation of invariant contributions from experimental data should be limited to single vibrational bands and not for the entire spectrum.

From ICP and DCP_I ROA measurements carried out for some terpenes,[219] including camphor, it was concluded that "the magnetic dipole scattering contribution is, on average, approximately an order of magnitude more important than the electric-quadrupole contribution, and that the anisotropic magnetic-dipole ROA invariant is two to three times larger than the corresponding electric-quadrupole ROA invariant."

The relative importance of $\omega^{-1}\gamma_k^2$ and $\omega^{-1}\delta_k^2$ terms to vibrational ROA spectra has been discussed in other places as well.[220,221,222,223,224,225,226] The contributions from the ROA invariants in various conformers of heptasilane were theoretically investigated by Liégeois,[222] where they found that $\omega^{-1}\gamma_k^2$ and $\omega^{-1}\delta_k^2$ terms are of similar order of magnitude for some conformers; for the TGTG' conformer, the intensity of the mode at 481 cm^{-1} was dominated by the $\omega^{-1}\delta_k^2$ term and $\omega^{-1}\gamma_k^2$ term dominated that at 585 cm^{-1}. In an article advancing a simplified method for vibrational ROA calculations,[223] Zuber et al. stated that "Indeed, previous

VROA calculations[224] on a large set of small chiral molecules show that the quadrupole term contributes to less than 20% of the total vibrational ROA intensity in the back and depolarized right angle scattering measurements and less than 8% in the polarized right angle setup" and chose not to include the $\omega^{-1}\delta_k^2$ term in their simplified method. These views were also expressed in an article that focused on efficient calculation of ROA tensors with analytical gradients and fragmentation.[225] Luber et al. have reviewed[226] some of the existing vibrational ROA papers and noted the varying conclusions on the relative importance of $\omega^{-1}\delta_k^2$ term. In an attempt to derive a conclusion on the relative importance of $\omega^{-1}\gamma_k^2$ and $\omega^{-1}\delta_k^2$ terms based on statistical analysis, Luber et al. have undertaken QC VROA calculations for a large number of molecules. Their statistical analysis led to the conclusion that "the vanishing electric-dipole—electric-quadrupole contribution is found to be a general feature for all vibrations of the organic molecules studied, which do not involve C–H stretching modes, and especially when the backscattering and polarized experimental setups are employed."

The factor of 3 larger coefficient associated with the $\omega^{-1}\gamma_k^2$ term in SCP, DCP_I, and depolarized ICP ROA renders the $\omega^{-1}\delta_k^2$ term relatively smaller. Therefore, in order for $\omega^{-1}\delta_k^2$ to be dominant, the magnitude of $\omega^{-1}\delta_k^2$ itself should be at least three times larger than that of $\omega^{-1}\gamma_k^2$, especially for situations where $a_4\omega^{-1}\gamma_k^2$ and $a_5\omega^{-1}\delta_k^2$ terms have opposite signs. Polavarapu et al. have investigated[227] the SCP backscattering vibrational ROA spectra of five terpenes and displayed the simulated spectra using the following equations for ROA activities, P_k, and Raman activities, S_k:

$$P_k = \frac{2\pi}{\lambda} \times 16 \times (3\omega^{-1}\gamma_k^2 + \omega^{-1}\delta_k^2) = P_k(\gamma_k^2) + P_k(\delta_k^2), \tag{8.1}$$

$$S_k = 2 \times (45\bar{\alpha}_k^2 + 7\beta_k^2). \tag{8.2}$$

In Equation 9.1, $P_k(\gamma_k^2)$ and $P_k(\delta_k^2)$ were referred[227] to, respectively, as "γ_k^2 and δ_k^2 contributions" to vibrational ROA and were displayed separately. The γ_k^2 and δ_k^2 contributions from individual vibrations were not analyzed,[227] but the overall spectral displays indicated that latter are generally smaller than the former.

Even though the individual experimental vibrational ROA bands were analyzed in the past using the corresponding theoretical predictions, the specific cases of experimental VROA bands originating from the dominating δ_k^2 contributions have not yet been identified using corresponding QC predictions. In this work, I reported the first observation that δ_k^2 contribution to certain vibrational bands below 2,000 cm^{-1} have unique importance, using the experimental VROA measurements and corresponding QC predictions for (1*S*)-(-)-camphor. The experimental VROA intensity associated with 1,125 cm^{-1} band of camphor cannot be reproduced in the predicted QC VROA spectra without the inclusion of the δ_k^2 contribution. This observation prompted us to also analyze the predicted QC ROA predictions for individual vibrations of dimethyloxirane,

α-pinene, and fenchone and compare them with corresponding previously reported experimental vibrational ROA spectra for assessing the importance of δ_k^2 contributions.

8.2 Methods

Professor Baranska and Dr. Zajac allowed me personal and extended access to their ROA spectrometer in Krakow, Poland to undertake the ROA and Raman measurements with a 532 nm wavelength laser source and the relevant experimental parameters are presented in the main manuscript[206] and below. A special thanks is warranted here for Dr Covington and his production of CDSpecTech[52,53] while a graduate student in our lab. Without this software and the ease in which it allowed viewing δ_k^2 contributions to ROA, I would likely never have noticed the ROA bands that lead to this work.

The experimental Raman and VROA spectra of (-)-camphor were measured in CCl_4, at a concentration of 181 mg/ml, using commercially available ChiralRaman Spectrometer (Biotools, Jupiter, FL). Vibrational ROA and Raman spectra were measured simultaneously in SCP backscattering with a 532-nm laser source using 1,000 mW power, although the power density at the sample is unknown. The instrument was calibrated with a one minute ROA measurement for a neat liquid sample of alpha-pinene. Signal averaging was done for 72 hours. The experimental spectra shown here are solvent spectra subtracted.

The single conformer of (1*R*)-camphor was optimized at various levels of theory (Table 8.1). Vibrational Raman and ROA spectral intensities were calculated using the same level of theory as that employed for geometry optimization, using Gaussian16 program,[80] with PCM representing the CCl_4 solvent. CDSpecTech[52] was used to simulate all theoretical vibrational Raman and ROA spectra, including the separate γ_k^2 and δ_k^2 contributions 9.1 using Lorentzian band shapes. The same program was used for the similarity overlap analysis. Predicted VROA spectra for (*R*)-configuration of camphor have been multiplied by -1 to match the absolute configuration of the sample used for experimental measurements.

8.3 Results and Discussion

The vibrational Raman, ROA, and CID spectra of (1*S*)-(-)-camphor (Figure 8.2) were analyzed in two regions, from ~1,800–900 cm^{-1} and from ~730–505 cm^{-1}. Due to the interference from solvent vibrational bands in the region between 900 and 730 cm^{-1} and in the region below ~505 cm^{-1} , these regions are excluded from the analysis in the current work. QC calculations were carried out using three different density functionals and a variety of basis sets, as summarized in Table 8.1. Dispersion corrections were also included in some of the calculations. However, the displayed spectra are limited to those predicted at B3PW91/aug-cc-PVDZ/PCM (CCl_4) level, for illustration purposes.

Considering the 1,800–900 cm^{-1} region, the QC predicted spectrum in this region is seen to match the

Level Denoted as	Functional	Basis Set	Dispersion
A	B3LYP	6-311++G(2d,2p)	No
B	B3LYP	6-311++G(2d,2p)	Yes
C	B3LYP	aug-cc-pVDZ	No
D	B3LYP	aug-cc-pVDZ	Yes
E	B3LYP	aug-cc-pVTZ	No
F	B3LYP	aug-cc-pVTZ	Yes
G	B3PW91	6-311++G(2d,2p)	No
H	B3PW91	aug-cc-pVDZ	No
I	B3PW91	aug-cc-pVDZ	Yes
J	B3PW91	aug-cc-pVTZ	No
K	B3PW91	aug-cc-pVTZ	No
L	wB97x	aug-cc-pVDZ	No
M	wB97x	aug-cc-pVDZ	Yes
N	wB97x	aug-cc-pVTZ	Yes
P	wB97x	6-311++G(2d,2p)	Yes

Table 8.1: Various levels of theory (and their designations) used for optimization and QC predicted spectra of (1*S*)-(-)-camphor

experimental spectrum qualitatively (Figures 3 and 4). Predicted band positions were scaled by 0.989, which corresponds to the x-axis scale factor that yields the maximum *Sim*ROA value. Although calculated positions for most of the vibrational bands closely match the experimental band positions, there is some disparity in the C=O stretching band position among the experimental and predicted spectra. Despite this mismatch for the C=O band position, the similarity overlap between experimental and predicted spectra is still quite high.[206] Because C=O band position is far removed from the next closest vibrational band, the similarity analysis was be limited to the 1,550–900 cm^{-1} region. The omission of C=O stretching band from the similarity overlap calculation does not influence the final conclusions but does increase the similarity overlap magnitudes. The SSO plot is shown in Figure 8.5, and the similarity overlap values are summarized in Table 8.3.

In the 730–505 cm^{-1} range also, the QC predicted spectra matched the experimental Raman, ROA, and CID both qualitatively and quantitatively. This is seen from the similarity overlap values shown in Table 8.2 and the overlaid and stacked spectra, as well as the SSO plot in Johnson et al.[206] This region provides further support for the satisfactory predictions of the QC methods used.

In the 1,550–900 cm^{-1} region, there are three negative VROA bands that appear in the experimental vibrational ROA spectrum at 1,125, 1,013, and 955 cm^{-1} that are identified to be dominated by the δ_k^2

Level	*Sim*Raman	*Sim*ROA	*Sim*CID
A	0.82	0.84	0.92
B	0.83	0.87	0.93
C	0.84	0.88	0.93
D	0.84	0.88	0.92
E	0.83	0.84	0.92
F	0.84	0.87	0.93
G	0.82	0.88	0.92
H	0.84	0.89	0.91
I	0.84	0.89	0.90
J	0.84	0.88	0.92
K	0.84	0.90	0.92
L	0.85	0.88	0.87
M	0.85	0.91	0.92
N	0.84	0.89	0.93
P	0.84	0.87	0.92

Table 8.2: Maximum similarity overlap between experimental and QC predicted vibrational Raman, ROA, and CID spectra for (1*S*)-(-)-camphor at various levels of theory in the 730-505 cm^{-1}

Level	*Sim*Raman	*Sim*ROA	*Sim*CID
A	0.84	0.49	0.66
B	0.85	0.49	0.67
C	0.91	0.65	0.76
D	0.92	0.65	0.75
E	0.85	0.52	0.69
F	0.86	0.52	0.70
G	0.92	0.70	0.79
H	0.94	0.77	0.77
I	0.94	0.74	0.75
J	0.94	0.76	0.81
K	0.95	0.77	0.80
L	0.89	0.59	0.61
M	0.92	0.75	0.69
N	0.94	0.73	0.71
P	0.93	0.70	0.69

Table 8.3: Maximum similarity overlap between experimental and QC predicted vibrational Raman, ROA, and CID spectra for (1*S*)-(-)-camphor at various levels of theory in the 1550-900 cm^{-1}

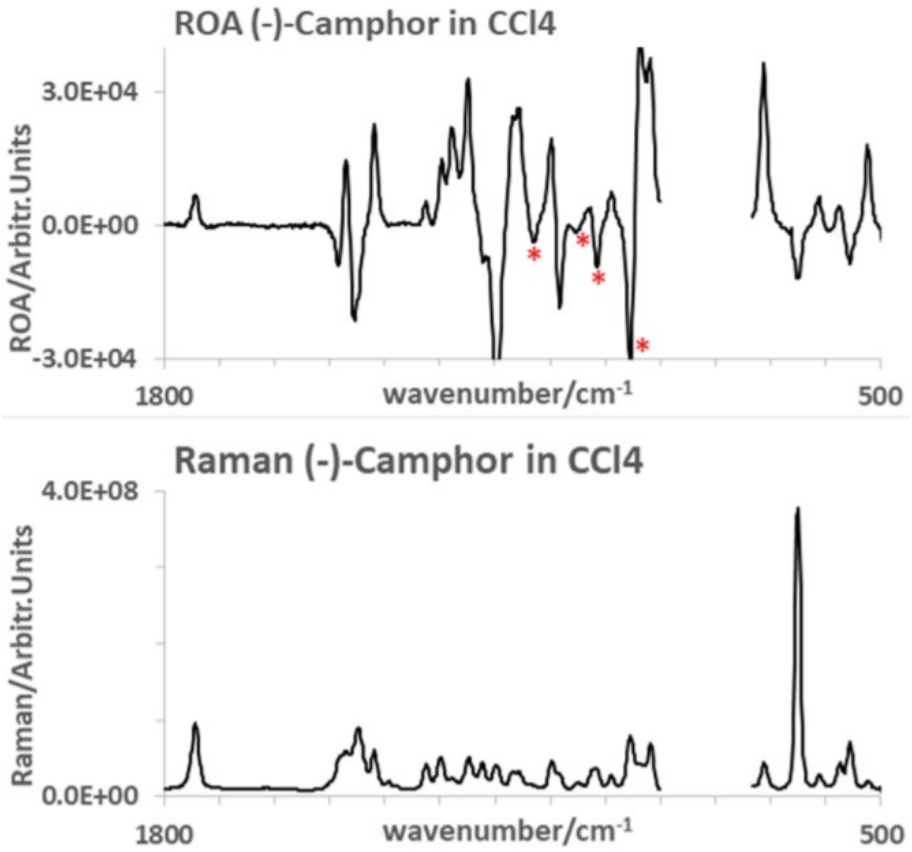

Figure 8.2: Experimental vibrational Raman and ROA of (-)-camphor in CCl_4. The ROA bands marked with * are identified as having dominating δ_k^2 contributions

contribution. These bands are marked with a * in Figures 2 and 4. There is also a positive vibrational ROA couplet in the experimental spectrum between 1,030 and 1,050 cm^{-1}, where the positive signal at 1,030 cm^{-1} correlates to a positive band that is only apparent in the predicted δ_k^2 contribution. The bands referred to above are highlighted with vertical dashed lines in Figure 8.6, where the spectra are displayed in the 1,150–900 cm^{-1} region on an expanded x-axis scale. The importance of δ_k^2 contributions can be more easily recognized in the CID spectra shown in Figure 8.7, where the simulated γ_k^2 and δ_k^2 contributions in the 1,150–900 cm^{-1} region are displayed separately. Predicted band positions for Figures 6 and 7 were scaled by 0.989, which corresponds to the x-axis scale factor that yields the maximum *Sim*ROA value. When the δ_k^2 contributions were excluded in simulating the spectra, the negative VROA band seen in the experimental spectrum at 1,125 cm^{-1} does not appear at all in the simulated spectra. This is considered to be the first case of a band in the vibrational ROA experimental region below 2,000 cm^{-1} to be clearly identified as arising from the electric-dipole—electric-quadrupole polarizability tensor contribution using QC calculations. The dominance of δ_k^2 contribution for above-mentioned experimental VROA bands of (1*S*)-camphor can also be

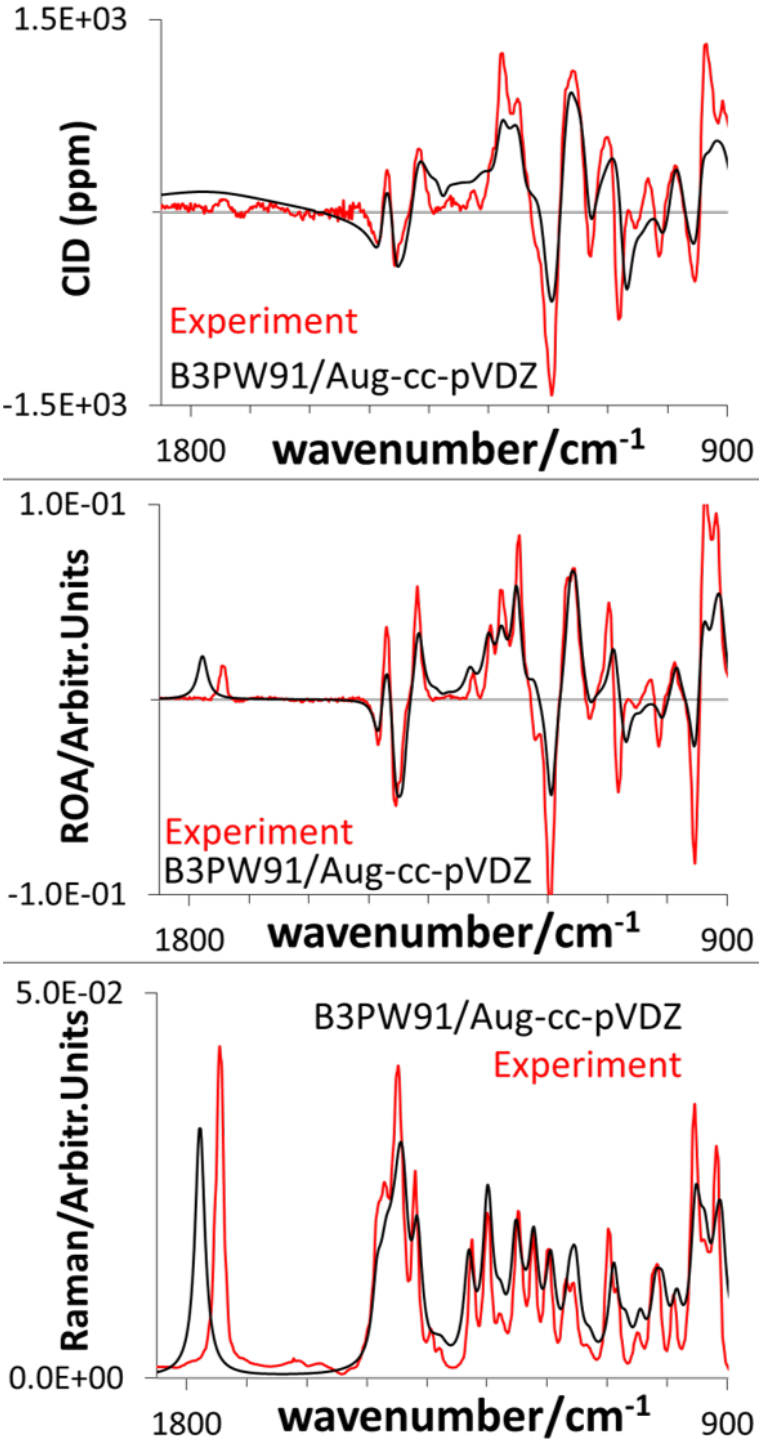

Figure 8.3: Experimental vibrational Raman, ROA, and CID spectra of (-)-camphor overlaid with simulated spectra for (1*S*)-camphor

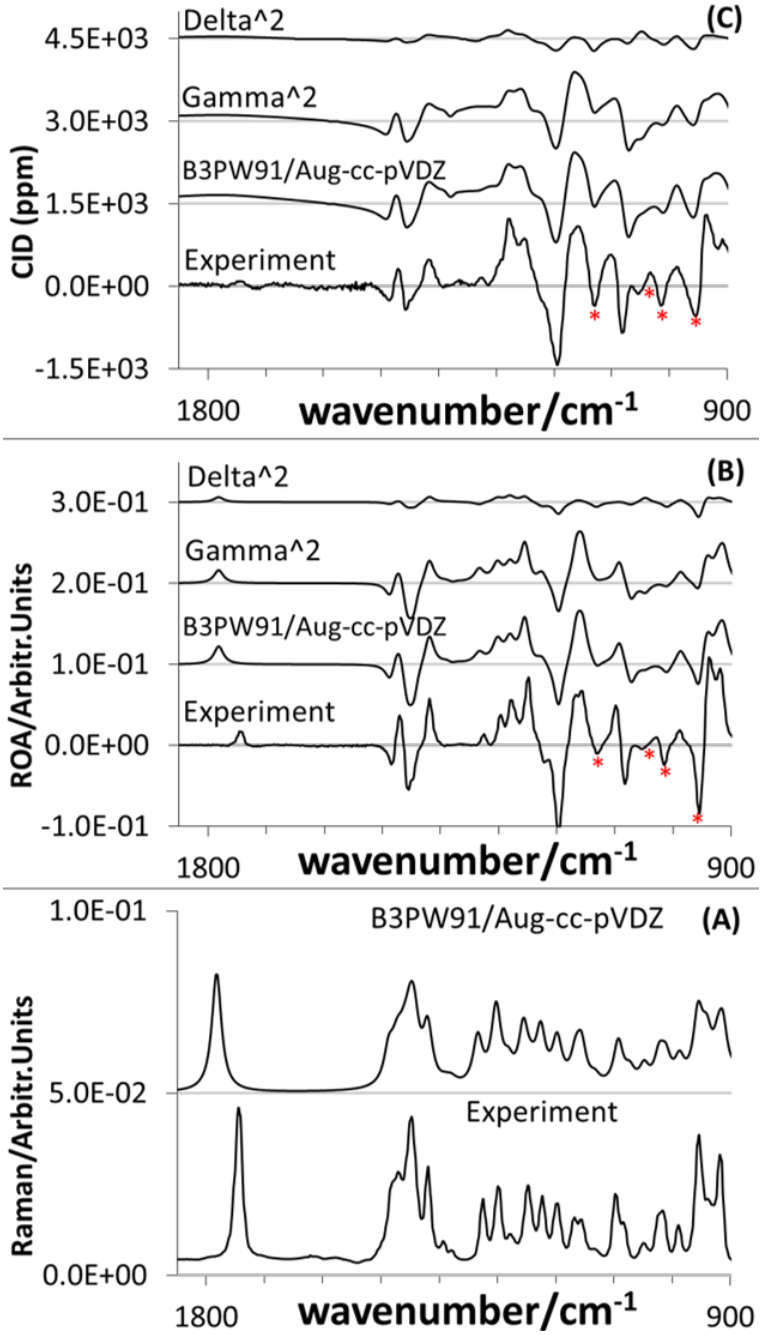

Figure 8.4: Experimental vibrational Raman, ROA, and CID spectra of (-)-camphor stacked with simulated spectra for (1*S*)-camphor. The ROA bands marked with * are identified as having dominating δ_k^2 contributions

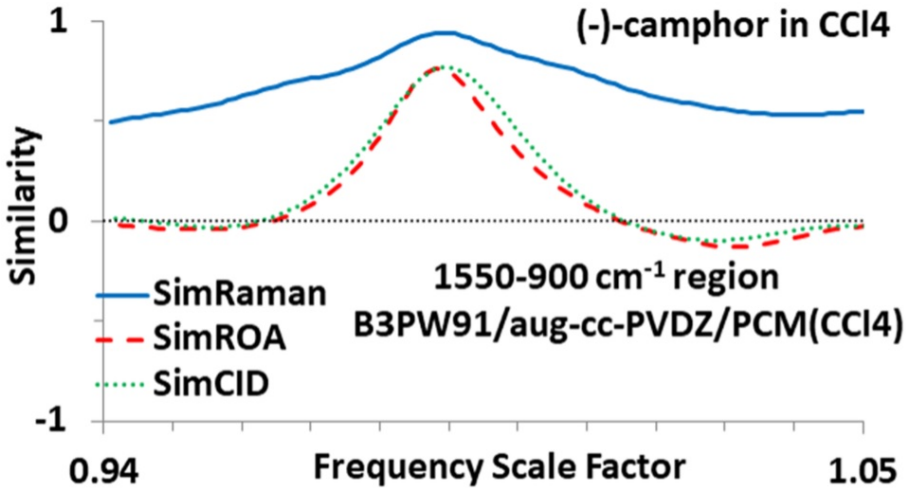

Figure 8.5: Spectral similarity overlap plot of (1*S*)-(-)-camphor

Level	**- at 1125**	**+ at 1030**	**- at 1013**	**- at 955**
A	Yes	Yes	No	Yes
B	Yes	Yes	No	Yes
C	Yes	Yes	No*	Yes
D	Yes	Yes	No*	Yes
E	Yes	Yes	No*	Yes
F	Yes	Yes	No	Yes
G	Yes	Yes	Yes	Yes
H	Yes	Yes	Yes	Yes
I	Yes	Yes	Yes	Yes
J	Yes	Yes	Yes	Yes
K	Yes	Yes	Yes	Yes
L	Yes	Yes	No*	Yes
M	Yes	Yes	No*	Yes
N	Yes	Yes	No*	Yes
P	Yes	Yes	No*	Yes

Table 8.4: Validation of calculated δ_k^2 as the predominant contribution towards experimental ROA bands for (1*S*)-(-)-camphor at various levels of theory

seen from the QC predicted values of the invariants, $\omega^{-1}\bar{\alpha}_k\bar{G}'_k$, $\omega^{-1}\gamma_k^2$, and $\omega^{-1}\delta_k^2$, summarized by Johnson et al.[206]

The VROA signals that are identified to be originating predominantly from the δ_k^2 contributions are confirmed by the predictions at most levels of theory investigated in this work. Table 8.4 enumerates the levels of theory that showed significant δ_k^2 contributions corresponding to each of the above-mentioned VROA

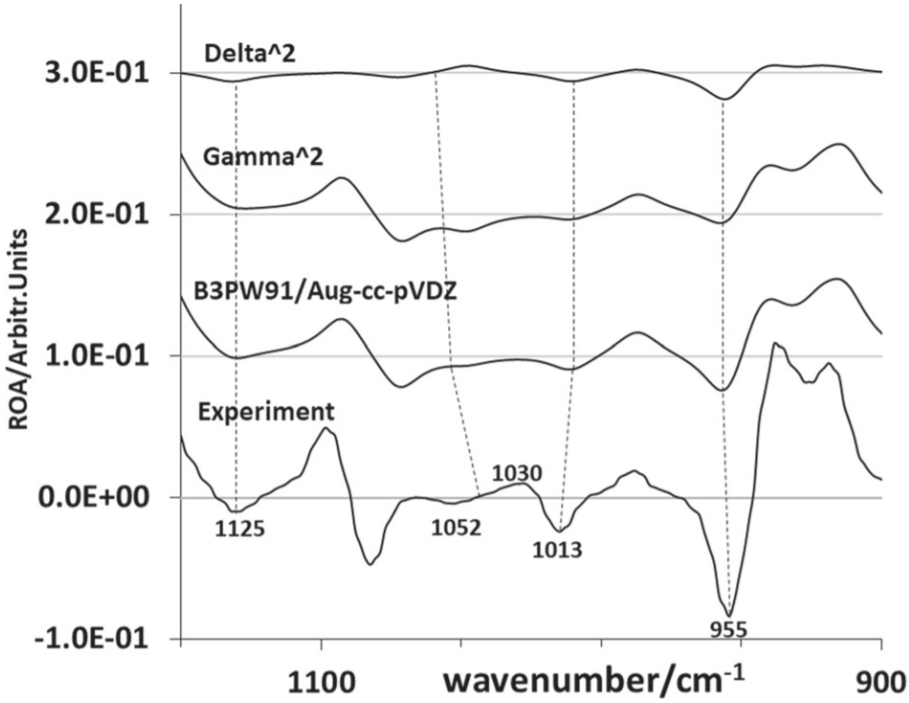

Figure 8.6: Experimental vibrational ROA spectrum (bottom trace) of (-)-camphor stacked with simulated spectra for (1*S*)-camphor displaying separate contributions from δ_k^2 (top trace), γ_k^2 (second trace from the top) and total (third trace from the top)

bands. It can be seen from Table 8.4 that at all levels of theory used to investigate the VROA spectrum of camphor, the VROA bands at 1,125, 1,030, and 955 cm^{-1} are always predominantly of δ_k^2 origin. Although the negative band at 1,013 cm^{-1} is generally reproduced at each level of theory, its dominance from the γ_k^2 or δ_k^2 contributions is somewhat ambiguous. The calculations using B3PW91 functional predict this band to be coming predominantly from δ_k^2 contributions, whereas those using wB97x functional indicate the predominance from the γ_k^2 contributions (despite the presence of significant δ_k^2 contributions to this band). The calculations using B3LYP functional generally attribute this band to the γ_k^2 contribution, with significant δ_k^2 contribution depending on the basis set.

GaussView[84] was utilized to visualize the vibrational mode associated with a particular wavenumber. The analysis of vibrations responsible for VROA bands that had predominant δ_k^2 contribution is as follows: The negative band at 1,125 cm^{-1} arises from CH_3 bending of the methyl group attached to C1 and CH_2 bending of C3. This band is particularly interesting as it does not appear at all in the simulated γ_k^2 contribution spectra as can be seen in Figures 6 and 7. The vibrational mode associated with the positive couplet, seen in the experimental region of 1,030–1,050 cm^{-1}, is due to a CH bend at C4 and CH_2 bending at C5 and C6. Although the couplet does not appear in the overall predicted vibrational ROA, the positive part of it is only apparent in the δ_k^2 contribution. The negative band at 1,013 cm^{-1} comes from CH_3 bending at

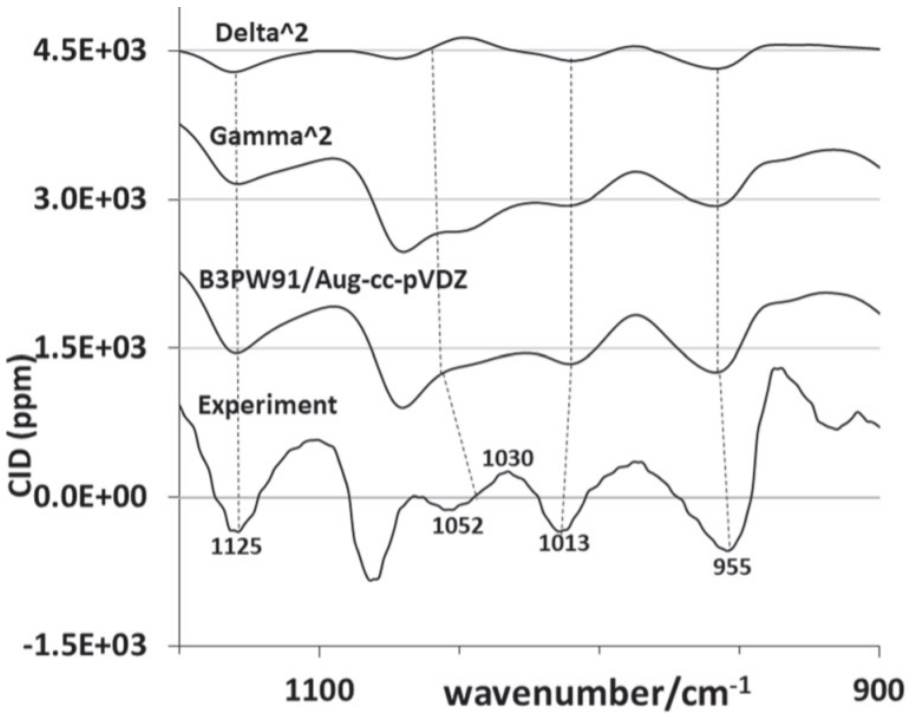

Figure 8.7: Experimental CID spectrum (bottom trace) of (-)-camphor stacked with simulated spectra for (1*S*)-camphor displaying separate contributions from δ_k^2 (top trace), γ_k^2 (second trace from the top) and total (third trace from the top)

the methyl groups attached at C1 and C7, CH_2 bending at C3, C5, and C6, and CH bending at C4. The complexity of this vibrational mode may have something to do with the discrepancy among different levels of theory as to whether γ_k^2 or δ_k^2 contribution is dominant for this vibration. The negative signal at 955 cm^{-1} in the experimental VROA originates from CH_3 bending of the methyl groups attached at C1 and C7 and CH bending at C4.

In Figure 8.8, the similarity between experimental spectra of camphor and predicted γ_k^2 contributions (excluding δ_k^2 contribution) is shown as SSO plot. Even though γ_k^2 contribution alone provides sufficient similarity with the experimental spectra, the *Sim*ROA and *Sim*CID values are lower at every level of theory investigated when the δ_k^2 contributions were not included (Table 8.5). When δ_k^2 contributions are included, the *Sim*ROA and *Sim*CID values increased in individual calculations by up to $\sim$12%, whereas the average increases in *Sim*ROA and *Sim*CID values are $\sim$9.0% (Table 8.5).

The current observations identifying the importance of δ_k^2 contributions in camphor prompted us to verify whether or not this situation occurs also in other molecules. For this purpose, I have selected a few additional molecules, (2*R*,3*R*)-dimethyloxirane, (1*S*,5*S*)-(+)-α-pinene, and (1*S*,4*R*)-(+)-fenchone, all with a single conformation, and analyzed their QC predictions. The invariants, $\omega^{-1}\bar{\alpha}_k\bar{G}'_k$, $\omega^{-1}\gamma_k^2$, and $\omega^{-1}\delta_k^2$, for

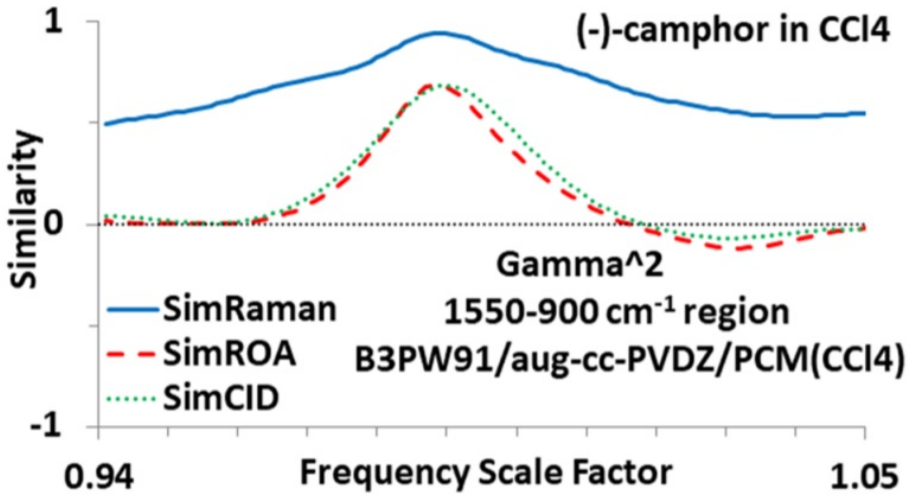

Figure 8.8: Spectral similarity overlap between experimental vibrational ROA spectrum and predicted γ_k^2 contribution spectrum for (1*S*)-(-)-camphor

Level	***Sim*ROA**	***Sim*ROA from γ_k^2 alone**	**% Increase of *Sim*ROA**	***Sim*CID**	***Sim*CID from γ_k^2 alone**	**% Increase of *Sim*CID**
A	0.49	0.47	5.7	0.66	0.62	6.5
B	0.49	0.46	6.3	0.67	0.64	6.2
C	0.65	0.60	8.5	0.76	0.69	10.2
D	0.65	0.59	9.3	0.75	0.68	9.7
E	0.52	0.49	6.5	0.69	0.64	7.3
F	0.52	0.48	6.9	0.70	0.65	6.9
G	0.70	0.64	8.8	0.79	0.71	10.4
H	0.77	0.69	10.9	0.77	0.69	12.4
I	0.74	0.67	11.4	0.75	0.67	11.9
J	0.76	0.69	9.7	0.81	0.73	11.0
K	0.77	0.70	10.5	0.80	0.72	11.3
L	0.59	0.54	8.2	0.61	0.56	7.9
M	0.75	0.68	9.8	0.69	0.63	9.0
N	0.73	0.67	9.5	0.71	0.66	7.6
P	0.70	0.64	8.7	0.69	0.64	7.2

Table 8.5: Maximum similarity between experimental and QC predicted ROA and CID spectra with and without the δ_k^2 contributions for (1*S*)-(-)-camphor at various levels of theory in the 1550-900 cm^{-1} range

all vibrations of these molecules are summarized by Johnson et al.[206] and the vibrational modes for which $|\delta_k^2| > |3\gamma_k^2|$ are highlighted in bold text. For dimethyloxirane, α-pinene, and fenchone, I presented the predictions using two different density functionals, namely, B3LYP and B3PW91. Some of the predictions vary

among these two levels of theory, so only the modes for which the predictions are consistent at both levels of theory have been focus upon.

For dimethyloxirane, the antisymmetric ring deformation mode (at 741 cm^{-1} in B3LYP/6-311++G(2d,2p) calculation and 762 cm^{-1} in B3PW91/aug-cc-pVDZ calculation) is predicted to have $|\delta_k^2|>|3\gamma_k^2|$. The symmetric ring deformation mode (at 814 cm^{-1} in B3LYP/6-311++G(2d,2p) calculation and 824 cm^{-1} in B3PW91/aug-cc-pVDZ calculation) is predicted to have $|\delta_k^2|\sim|2\gamma_k^2|$. The simulated polarized ICP ROA bands for these two modes have opposite signs, whereas those for depolarized ICP ROA bands have the same signs. These predictions match the early prognostications[210] based on the analysis of the experimental ratio for right angle polarized and depolarized ICP ROA (vide supra). However, some uncertainty remains with the experimental VROA sign associated with the antisymmetric ring deformation mode at $\sim$722 cm^{-1} , because, unlike in earlier measurements,[210,211,212] it was not reproduced in a later measurement.[213] Regardless of this sign uncertainty, the symmetric and antisymmetric ring deformation modes of dimethyloxiane do have dominating electric-dipole—electric-quadrupole polarizability contribution, as originally suggested.[210] It should also be noted that there are some vibrational modes of dimethyloxirane for which $|\delta_k^2|\sim|\gamma_k^2|$, and therefore, δ_k^2 contributions cannot be generally ignored for dimethyloxirane.

For α-pinene, predictions at B3PW91/aug-cc-pVDZ level indicate a few vibrational modes that exhibit $|\delta_k^2|>|3\gamma_k^2|$. However, the VROA bands associated with each of these modes are overlapped by those associated with one or more modes in the immediate vicinity. As a consequence, it becomes difficult to experimentally resolve or identify the bands that have dominant δ_k^2 contribution in α-pinene. Nevertheless, there are some vibrational modes of α-pinene for which $|\delta_k^2|\sim|\gamma_k^2|$, and therefore, δ_k^2 contributions cannot be generally ignored for α-pinene as well.

For (1*S*,4*R*)-(+)-fenchone, the C=O stretching mode (at 1,793 cm^{-1} in B3LYP/aug-cc-pVDZ and 1,813 cm^{-1} in B3PW91/aug-cc-pVDZ) is predicted to have $|\delta_k^2|>|3\gamma_k^2|$. The positive experimental vibrational ROA observed[227] for this band at $\sim$1,740 cm^{-1} is weak, but the observed sign is reproduced in the simulated spectrum from δ_k^2 contribution but not in simulated spectrum from γ_k^2 contribution. A lower wavenumber mode (232 cm^{-1} in B3LYP/aug-cc-pVDZ calculation and 228 cm^{-1} in B3PW91/aug-cc-pVDZ calculation) also has $|\delta_k^2|>|3\gamma_k^2|$. A close inspection of the experimental VROA spectrum indicates that the negative VROA at $\sim$229 cm^{-1} can only be reproduced in simulated spectrum from δ_k^2 contribution and not in the simulated spectrum from γ_k^2 contribution. Therefore, the C=O stretching vibrational band and a low wavenumber band at $\sim$230 cm^{-1} of fenchone can also be counted as the confirmed examples of dominant δ_k^2 contribution.

8.4 Conclusion

Analysis of the experimental VROA and CID spectra of (-)-camphor with corresponding predicted spectra at multiple levels of theory indicated that a few VROA bands of (1*S*)-(-)-camphor in the 1,150–900 cm^{-1} region have dominant electric-dipole—electricquadrupole polarizability contributions. The similarity overlap between experimental and predicted spectra increased by up to $\sim$12%, when electric-dipole—electric-quadrupole polarizability contributions are included in the simulation of VROA spectra, and these increases are present at all levels of theory examined. When the electric-dipole—electric-quadrupole polarizability contributions were excluded in simulating the VROA spectra, the negative band seen in the experimental spectrum of (-)-camphor at 1,125 cm^{-1} does not appear at all in the simulated spectra. Two vibrational bands of fenchone at $\sim$1,740 and 220 cm^{-1} are also confirmed to be originating from dominating electric-dipole—electric-quadrupole polarizability contribution. The symmetric and antisymmetric ring deformation modes of dimethyloxirane also have dominating electric-dipole—electric-quadrupole polarizability contribution. These are considered to be the first cases of bands in the typical VROA experimental region to be clearly identified as arising from the dominant electric-dipole—electricquadrupole polarizability contribution. These observations establish the importance of the electric-dipole—electric-quadrupole polarizability contribution to vibrational ROA and CID spectra. For experimentally identifying the dominant electric-dipole—electric-quadrupole polarizability contributions, it is important to analyze the experimental VROA spectra with corresponding predicted spectra from electric-dipole—magnetic-dipole and electric-dipole—electric-quadrupole polarizability contributions separately.

CHAPTER 9

Chiral Molecular Structure Determination from VOA and Molecular Formula

This chapter has been reproduced from the *Journal of Physical Chemistry A* article[55] "Chiral Molecular Structure Determination for a Desired Compound Just from Its Molecular Formula and Vibrational Optical Activity Spectra" with permission. My contributions to this work are: conformational analyses, QC optimizations, QC spectral prediction of VCD and ROA, spectral similarity analyses, cross correlation analyses, preparation of figures and tables, drafting the manuscript, and revising the manuscript.

9.1 Introduction

For any new synthesized compound, the first step in its characterization constitutes the determination of its molecular formula, which is commonly achieved through elemental analysis. In the next step, the functional groups present in that compound can be determined using a variety of experimental methods. There are several commonly used organic reactions that can establish the presence of different functional groups. Spectroscopically, the functional groups in a given compound can be identified through characteristic vibrational bands present in its infrared spectrum and characteristic chemical shifts in its NMR spectrum. Recently, it was reported that molecular formula, as well as functional groups that are likely to be present in a given compound, can be identified from the data obtained using rotational spectroscopy.[228] A computer assisted approach to detect the presence/absence of 16 functional groups from the experimental infrared spectra was also reported.[229]

Molecular structure determination, another important step, is relevant and often vital across all fields of chemistry. Crystal structure determination using X-ray diffraction remains the first choice, but most compounds are not amenable for obtaining good quality crystals. For determining the unknown molecular structures of non-crystalline achiral compounds, NMR is often the preferred choice. A more recent entry for molecular structure determination is the cryo-EM method.[230] When it comes to chiral compounds, the determination of molecular structure is more challenging than that for achiral molecules because one has to determine the absolute configurations (ACs) associated with individual chiral elements. In this context, NMR methods are being used for determining the relative ACs, and different chiroptical spectroscopic methods[4] are being utilized for determining the ACs.

In the context of determining the molecular structure of a desired chiral compound, I addressed the following question: is it possible to uniquely determine the chiral molecular structure when only its molecular formula and experimental vibrational optical activity (VOA) spectra are provided? VOA[23,108,109,110] encom-

passes vibrational circular dichroism (VCD) and vibrational Raman optical activity (ROA), both of which probe vibrational transitions of chiral molecules. VCD is the differential absorption of left and right circularly polarized infrared light. ROA is the differential vibrational Raman scattering of left and right circularly polarized visible light.

The approach to addressing the above mentioned question is as follows. (1) Measure the experimental VOA spectra for a desired compound with known molecular formula. (2) Generate all possible chiral structural isomers associated with that molecular formula. (3) Predict VOA spectra for all possible chiral structural isomers and analyze their similarity with each other (referred to as cross-correlations). (4) Analyze the similarity of predicted VOA spectra for all possible chiral structural isomers with experimental spectra of a desired compound. (5) Assess if a unique chiral molecular structure can be identified for the desired compound from this analysis.

Using different test molecules, I arrived at the proof-of-concept conclusion that it is feasible to uniquely determine the chiral molecular structure of a desired compound from only the knowledge of its molecular formula and VOA spectra. In complex cases, the information on functional groups contained in the desired compound will be helpful. These observations pave the way for a new paradigm in chiral molecular structure determination and identify its limitations.

9.2 Methods

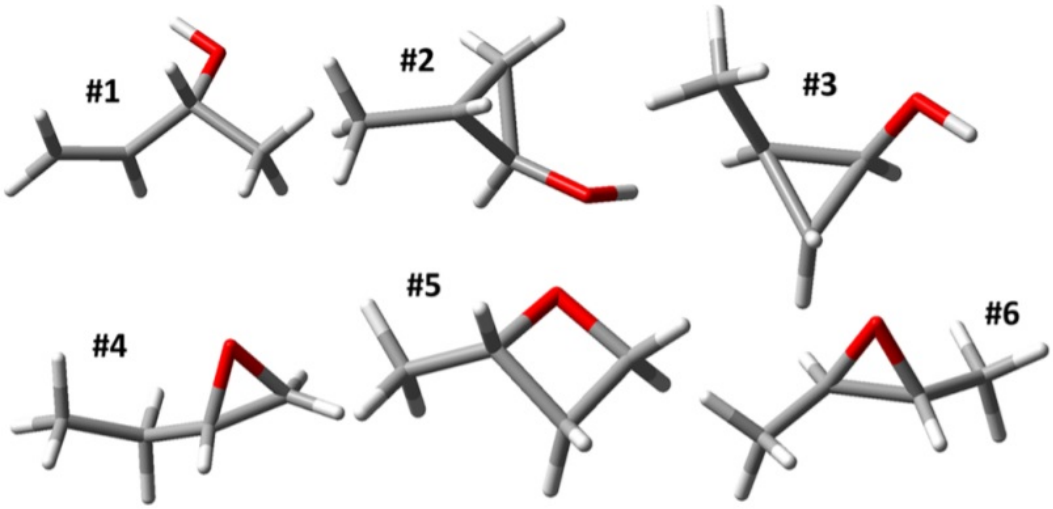

Figure 9.1: Chiral structural isomers with molecular formula C_4H_8O

Three compounds, (2*R*,3*R*)-(+)-dimethyloxirane[211,213] (C_4H_8O), (*S*)-(+)-epichlorohydrin,[231,232,233] also known as (+)-2-(chloromethyl)oxirane, (C_3H_5ClO), and (*R*)-(+)-3-methylcyclopentanone[234,235] ($C_6H_{10}O$), whose VCD and ROA spectra have been investigated before were selected as the desired compounds for

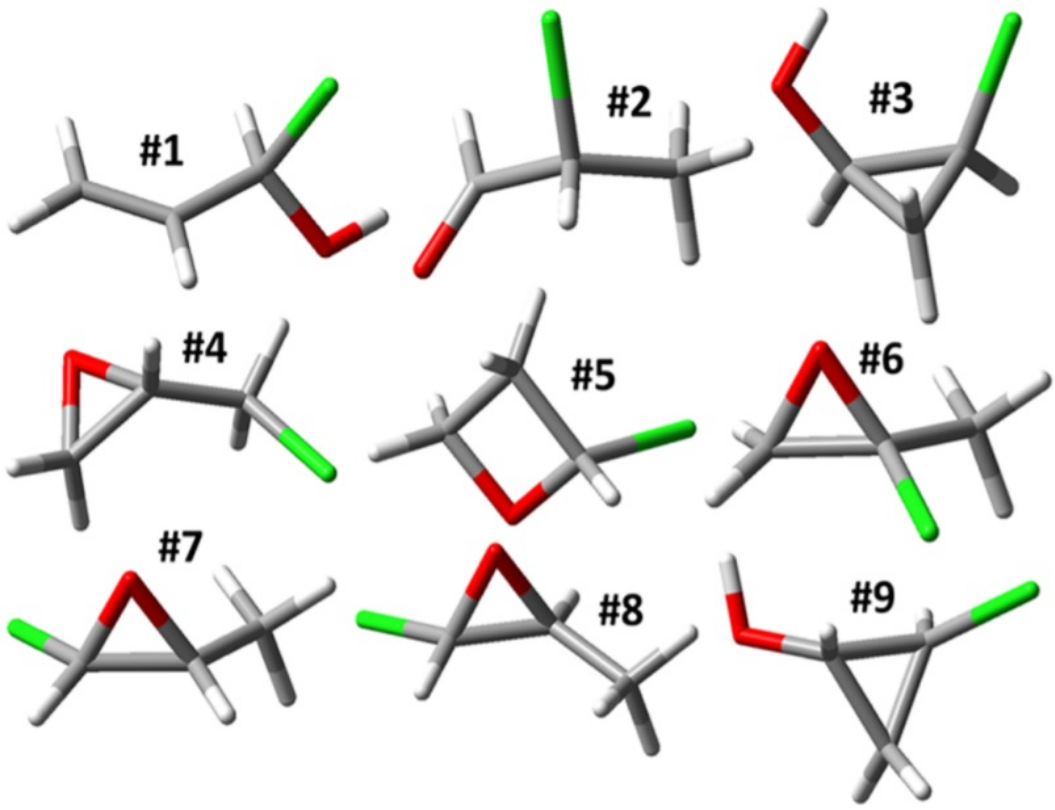

Figure 9.2: Chiral structural isomers with molecular formula C_3H_5ClO

the current studies. Recently detected *N*-hydroxyoxaziridine[236,237] (also known as 2-hydroxyoxaziridine, or 1,2-oxaziridin-2-ol), a chiral structural isomer of nitromethane (CH_3NO_2), is also examined.

Among the available computer algorithms for generating non-isomorphic chemical structures[238] associated with a given molecular formula, MOLGEN[239] is widely used. It may be noted that MOLGEN does not recognize chirality and *E*/*Z* isomerism. As a result, manual intervention is needed for (a) discarding the achiral structural isomers, (b) incorporation of diastereomers in the presence of multiple stereocenters, and (c) taking *E*/*Z* isomerism, wherever appropriate, into account.

I used the MOLGEN program to generate possible structural isomers associated with the molecular formulae C_4H_8O, C_3H_5ClO, $C_6H_{10}O$, and CH_3NO_2. For the molecular formulae C_4H_8O, C_3H_5ClO, and CH_3NO_2, the above mentioned process resulted in, respectively, 6, 9, and 4 unique chiral structural isomers (Figures 1-3). For $C_6H_{10}O$, however, greater than 800 unique chiral structural isomers were found. This large number necessitated devising means to reduce the number of structural isomers to a reasonable subset for enabling practical investigations. To this end, chiral structural isomers of $C_6H_{10}O$ containing a C=O bond and/or a five-membered ring were selected, which resulted in 75 chiral structural isomers (Figure 9.4). The chemical names of all structural isomers were derived using the ChemDraw program and summarized in

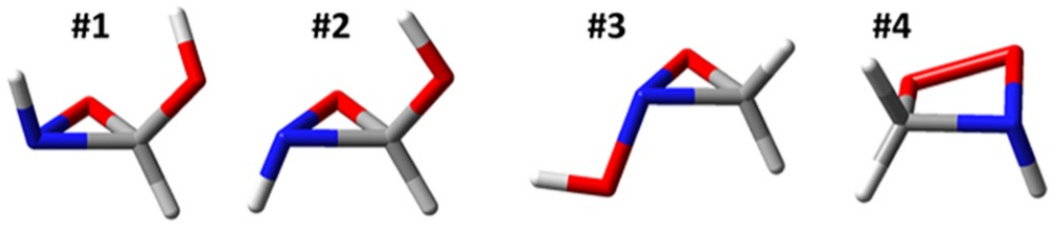

Figure 9.3: Chiral structural isomers with molecular formula CH_3NO_2

Tables in the main text.[55]

All of the unique chiral structural isomers associated with each molecular formula were subjected to full conformational search using the CONFLEX program[71] with a built-in MMFF94s[240] force field.

For the molecular formula C_4H_8O, the geometries of all conformers associated with each of the structural isomers 1-6 were optimized using the B3LYP functional[35,34] and 6-311++G(2d,2p) basis set[241,242,243] for isolated molecules (labeled as B3LYP/6-311++G(2d,2p)) or with polarizable continuum model (PCM)[44,81] representing the CCl_4 solvent, labeled as B3LYP/6-311++G(2d,2p)/PCM(CCl_4). The numbers of predominant conformers associated with each of these structural isomers are, respectively, 9, 2, 2, 3, 1, and 1. The fully optimized geometries were used to predict VOA spectra at the same level as that used for geometry optimization. Experimental ROA spectra for the desired compound, (2*R*,3*R*)-(+)-dimethyloxirane, were measured by Polavarapu et al.[213] for a neat liquid sample using incident circular polarization (ICP) and depolarized 90° scattering geometry with 488 nm incident laser wavelength.[211,213] The predicted ROA spectra for all chiral structural isomers were obtained at the B3LYP/6-311++G(2d,2p) level using the same scattering geometry and incident laser wavelength. Even though theoretical methods to predict VOA spectra for neat liquids are available now,[200,244] for the present ROA analysis, I considered in vacuo prediction adequate. Experimental VCD spectra were measured by Black et al.[211] for (2*R*,3*R*)-(+)-dimethyloxirane in CCl_4 and CS_2 solvents. Predicted VCD spectra were obtained with PCM representing the CCl_4 solvent at the B3LYP/6-311++G(2d,2p)/PCM(CCl_4) level.

For the molecular formula C_3H_5ClO, the geometries of all conformers associated with each of the chiral structural isomers 1-9 were optimized at the B3LYP/6-311++G(2d,2p)/PCM(CCl_4) level followed by ROA and VCD predictions at the same level. The numbers of predominant conformers of structures 1-9 are, respectively, 5, 3, 2, 3, 1, 1, 1, 1, and 2. The experimental ROA spectra of the desired compound, (*S*)-(+)-epichlorohydrin (also known as 2-(chloromethyl)oxirane), were obtained using scattered circular polarization (SCP) and unpolarized 180° backscattering geometry with 532 nm incident laser wavelength by Tommasini et al.[233] The predicted ROA spectra for all structural isomers 1-9 were obtained with PCM representing the CCl_4 solvent using the same scattering geometry and incident laser wavelength. Experimental VCD of

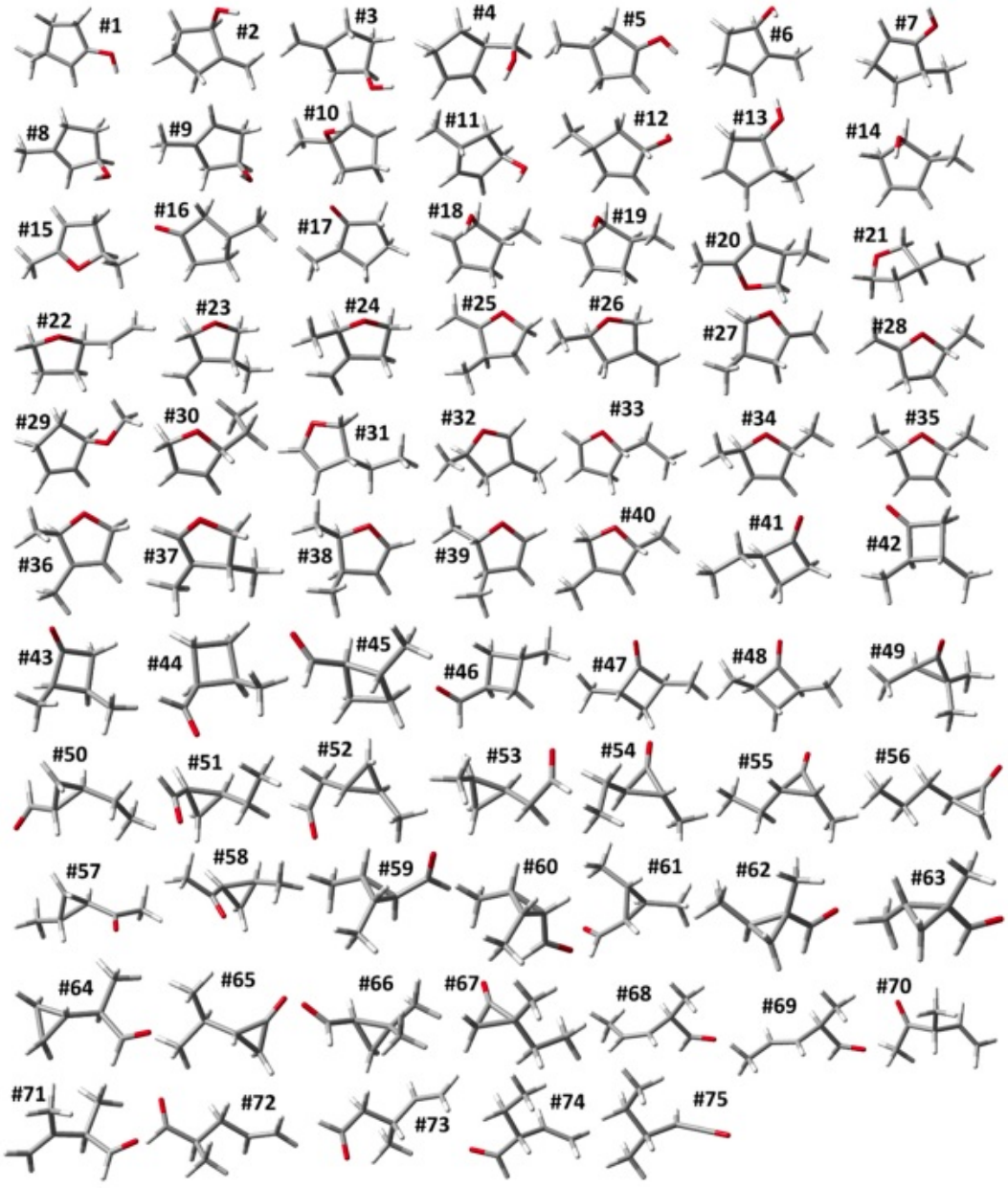

Figure 9.4: Chiral structural isomers with molecular formula $C_6H_{10}O$ containing a C=O group and/or a five-membered ring

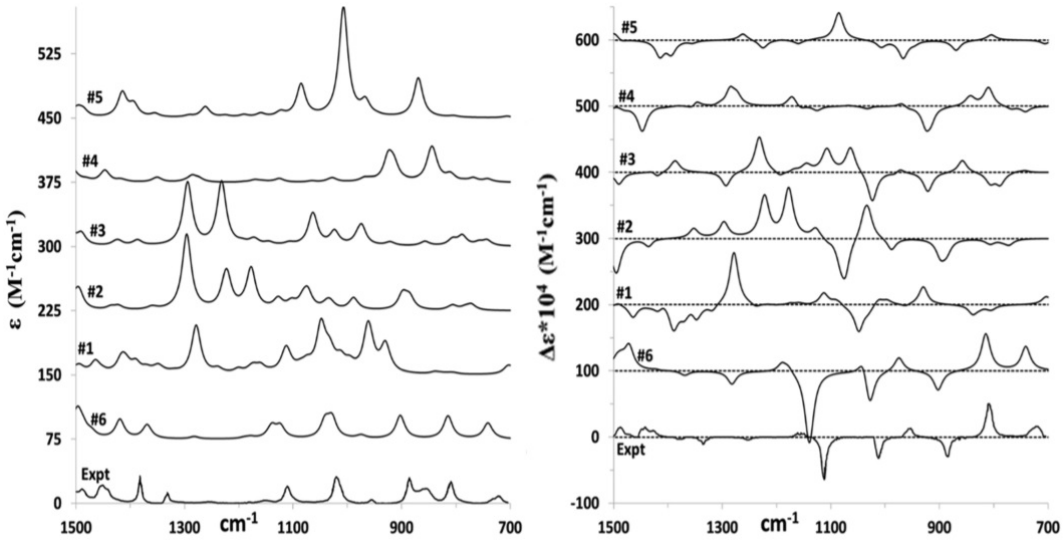

Figure 9.5: Predicted VA (left) and VCD (right) spectra of the six chiral structural isomers with molecular formula C_4H_8O and the experimental spectra of (2*R*,3*R*)-(+)-dimethyloxirane

(*S*)-(+)-epichlorohydrin was measured by Wang et al.[231]

For the molecular formula $C_6H_{10}O$, the geometries of all conformers associated with 75 chiral structural isomers were optimized at B3LYP/6-311++G(2d,2p) and B3LYP/6-311++G(2d,2p)/PCM(CCl_4) levels. Individual structural isomers exhibited different numbers of low-energy conformers, ranging from 1 to 27.[55] Experimental ROA spectra for the desired compound, (*R*)-(+)-3-methylcyclopentanone, were measured by Polavarapu et al.[235] for the neat liquid using ICP and depolarized 90° scattering geometry with 488 nm incident laser wavelength. The predicted ROA spectra were obtained for all 75 chiral structural isomers at the B3LYP/6-311++G(2d,2p) level using the same scattering geometry and incident laser wavelength. Experimental VCD spectra were measured14 for (*R*)-(+)-3-methylcyclopentanone in CCl_4. Experimental VCD of (*R*)-(+)-3-methylcyclopentanone was measured by He et al.[234] Predicted VCD spectra for all 75 chiral structural isomers were obtained with PCM representing the CCl_4 solvent at the B3LYP/6-311++G(2d,2p)/PCM(CCl_4) level.

For the molecular formula CH_3NO_2, the geometries of all conformers associated with each of the four chiral structural isomers were optimized at the B3LYP/6-311++G(2d,2p) level, followed by ROA and VCD predictions at the same level. The numbers of predominant conformers associated with structural isomers 1-4 are, respectively, 2, 2, 2, and 1. The predicted ROA spectra for the four chiral structural isomers were obtained with 532 nm incident laser wavelength for SCP and unpolarized 180° backscattering geometry. All geometry optimizations and spectral predictions were undertaken using the Gaussian 16 program.[80] Recent developments[203] allowed the inclusion of anharmonic contributions and explicit solvent influence for spectral

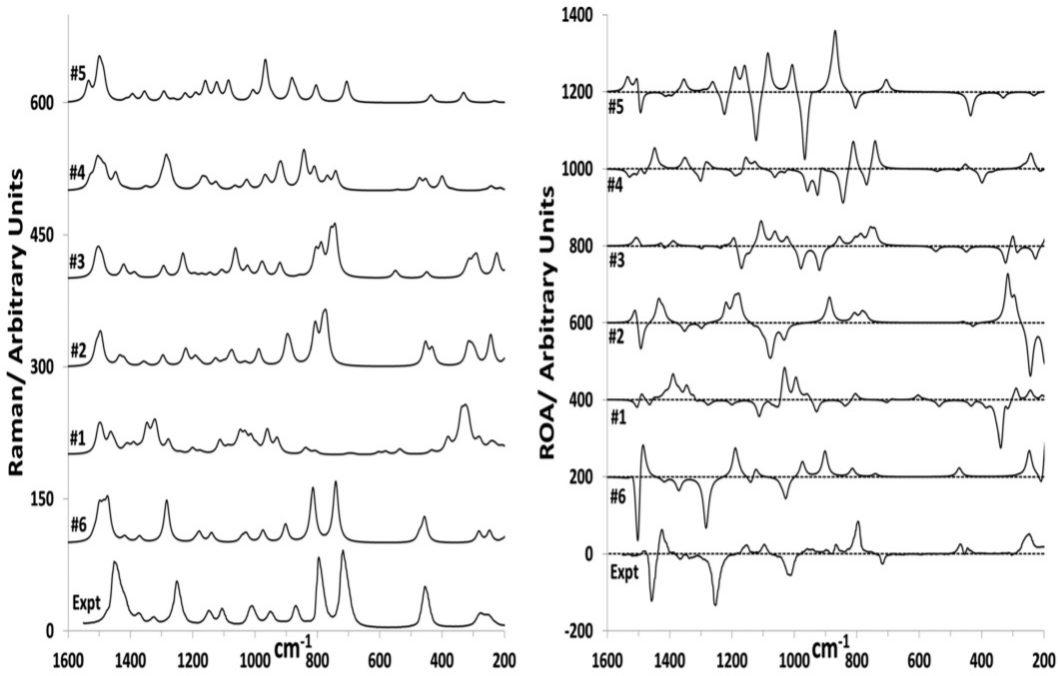

Figure 9.6: Predicted Raman (left) and ROA (right) spectra of the six chiral structural isomers with molecular formula C_4H_8O and the experimental spectra of (2*R*,3*R*)-(+)-dimethyloxirane

predictions. These contributions were not included in the present analyses. The predicted spectrum for each of the structural isomers was obtained as the sum of population-weighted spectra of its conformers using Gibbs energies. Experimental VCD and ROA spectra for (2*R*,3*R*)-(+)-dimethyloxirane, (*S*)-(+)-epichlorohydrin, and (*R*)-(+)-3-methylcyclopentanone were digitized using WebPlotDigitizer.[200] The predicted spectra for all investigated chiral structural isomers associated with a given molecular formula are compared quantitatively with each other (referred to as cross-correlations) and with the experimental spectra of the desired compounds. Quantitative similarities of vibrational absorption (VA), VCD, vibrational dissymmetry factor (VDF), Raman, ROA, and circular intensity difference (CID) are determined by calculating spectral similarity overlap (SSO) as a function of frequency scale factor, using the CDSpecTech program.[52,53] The Tanimoto's similarity function,[82] as suggested by Shen et al.[50] and labeled as the Sim function,[57] was used for the SSO calculations. The range for SSO values of VA and Raman spectra is 0-1, where 1 indicates perfect spectral similarity and 0 indicates no spectral similarity. The range for SSO values of VCD, VDF, ROA, and CID spectra is -1 to +1. A value of +1 indicates perfect agreement with the spectra of the structure used for calculations, and -1 indicates perfect agreement with the spectra of enantiomeric structure used for calculations.[57,126] The SSO is generally maximized at a frequency scale factor that aligns the most intense bands in the predicted and experimental spectra. A vibrational mode with strong VA intensity may not be associated with strong Raman intensity or vice versa. For this reason, a frequency scale factor that maximizes SSO in VA spectra can

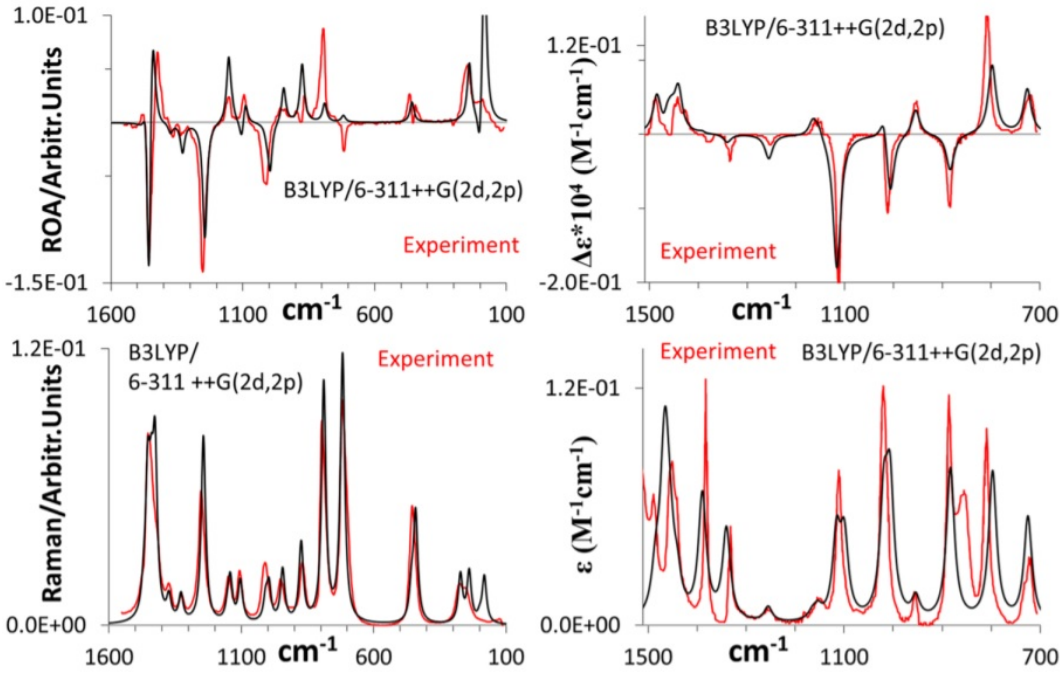

Figure 9.7: Predicted Raman and ROA (left), VA and VCD (right) spectra for structural isomer #6 and the experimental spectra of (2*R*,3*R*)-(+)-dimethyloxirane

be slightly different from that maximizing SSO in Raman spectra. The same explanation holds for slightly different scale factors normally seen for maximizing SSO in VA and VCD spectra and in Raman and ROA spectra.

Previously it was suggested,[51] based on investigations on a variety of molecules, that in order for two different spectra to be considered reliably similar, a Sim value of at least 0.4 is needed. This criterion was established by Dr. Polavarapu[23] and Covington et al.[52] Some laboratories have suggested a lower threshold (∼0.2),[50] but these studies confirmed that a reliability threshold of 0.4 provides better structural discrimination. It is important to have a reliable *Sim*VA value to accompany the corresponding reliable *Sim*VCD value. Similarly, a reliable *Sim*VDF value should accompany the corresponding reliable *Sim*VCD value. In other words, reliable *Sim*VA, *Sim*VCD, and *Sim*VDF values should all be greater than the reliability threshold for stringent discrimination among structural isomers. Likewise, reliable *Sim*Raman, *Sim*ROA, and *Sim*CID values should all be greater than the reliability threshold.

9.3 Results and Discussion

9.3.1 C_4H_8O

The main work[55] summarizes the chemical names and the number of conformations of each of the six chiral structural isomers of C_4H_8O. Predicted VCD and ROA spectra of these structural isomers are displayed in

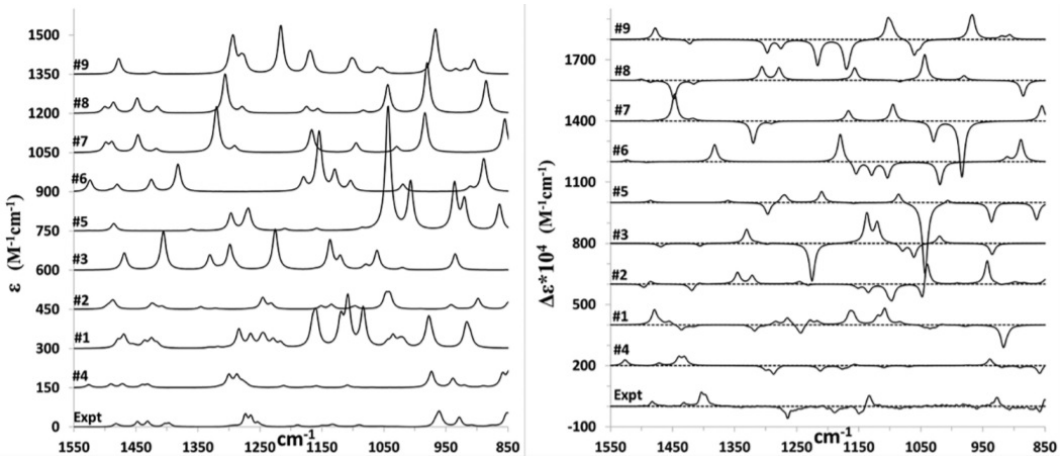

Figure 9.8: Predicted VA (left) and VCD (right) spectra of the nine chiral structural isomers with molecular formula C_3H_5ClO and the experimental spectra of (*S*)-(+)-epichlorohydrin

Figures 5 and 6, along with the experimental spectra of the desired compound (2*R*,3*R*)-(+)-dimethyloxirane. The experimental VA and VCD spectra obtained in CCl_4 and CS_2 solvents were combined to present the spectra in the 1500-700 cm^{-1} region. Experimental VA and VCD spectra were adapted with permission from *J. Am. Chem. Soc.* 1990, 112, 1479-1489.[211] Experimental Raman and ROA spectra were measured as a neat liquid with ICP 90° depolarized scattering geometry and 488 nm laser source and adapted with permission from *J. Phys. Chem. A.* 1993, 97, 1793-1799.[213]

Visual comparison of VA and VCD spectra in Figure 9.5 clearly indicates a close resemblance of the experimental spectra to the predicted spectra for structure #6. The same is true for Raman and ROA spectra in Figure 9.6. However, visual comparison and analysis of many different spectra can be time consuming, stressful, and biased. For this reason, a simpler and quantifiable analysis is preferred. Sim values for the predicted spectra of structural isomers of C_4H_8O with each other are presented by Johnson et al.[55] Each structure is a perfect match with itself for each type of spectra, leading to the diagonal Sim values of 1. The off-diagonal Sim values (representing cross-correlations) represent the similarity between the spectra of different structural isomers. The Sim values at or above the reliability threshold magnitude of 0.4 are highlighted.[55] It can be seen that no two structural isomers of C_4H_8O have large enough values for *Sim*VCD, *Sim*VDF, *Sim*ROA, or *Sim*CID for significant similarity. This observation bodes well for identifying the correct chiral molecular structure from VOA spectra of compounds with molecular formula C_4H_8O. VA or Raman spectra by themselves, however, would not be adequate for discrimination among the structural isomers of C_4H_8O because in several instances, SimVA or SimRaman values are greater than 0.4.

The Sim values for the predicted spectra of structural isomers of C_4H_8O, obtained using the experimental spectra of the desired compound (2*R*,3*R*)-(+)-dimethyloxirane, are listed in Table 9.1. A comparison of these

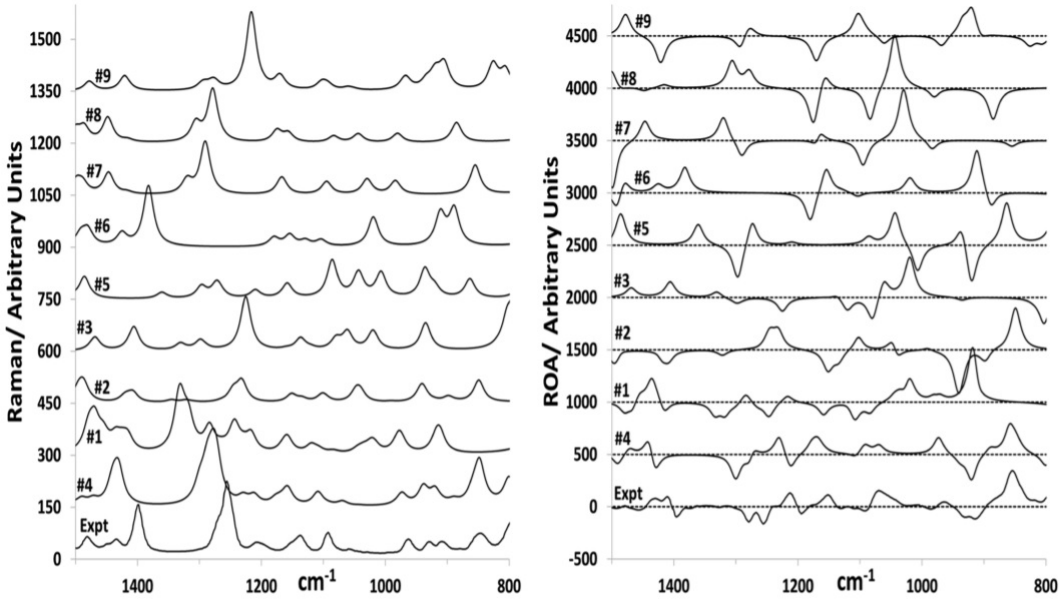

Figure 9.9: Predicted Raman (left) and ROA (right) spectra of the nine chiral structural isomers with molecular formula C_3H_5ClO and the experimental spectra of (*S*)-(+)-epichlorohydrin

Structure	***Sim*VA**	***Sim*VCD**	***Sim*VDF**	***Sim*Raman**	***Sim*ROA**	***Sim*CID**
1	0.34	−0.08	−0.19	0.33	0.05	0.11
2	0.30	−0.10	−0.12	0.50	−0.24	−0.44
3	0.23	−0.19	−0.23	0.63	0.24	0.22
4	0.59	0.19	0.11	0.69	0.14	0.18
5	0.55	−0.12	−0.13	0.50	−0.06	−0.10
6	**0.67**	**0.67**	**0.58**	**0.91**	**0.45**	**0.48**

Table 9.1: Similarity compared to the corresponding experimental measurements of (2*R*,3*R*)-(+)-dimethyloxirane

values indicates the following. (1) *Sim*VA, *Sim*VCD, *Sim*VDF, *Sim*Raman, *Sim*ROA, and *Sim*CID values are all above the reliable threshold only for one structural isomer, #6, which in fact is the correct structure for (2*R*,3*R*)-(+)-dimethyloxirane. Thus, VCD and/or ROA spectra can be considered to be uniquely associated with chiral structural isomers of C_4H_8O. (2) *Sim*CID value for structural isomer #2 is greater than the reliability threshold but the corresponding *Sim*ROA value is not. Such instances are not considered reliable. (3) VA and Raman spectra by themselves are not suitable for discrimination among different structural isomers of C_4H_8O. This is because five structural isomers (#2, 3, 4, 5, and 6) have *Sim*Raman magnitudes greater than the threshold value, and three structural isomers (#4, 5, and 6) have *Sim*VA magnitudes greater than the

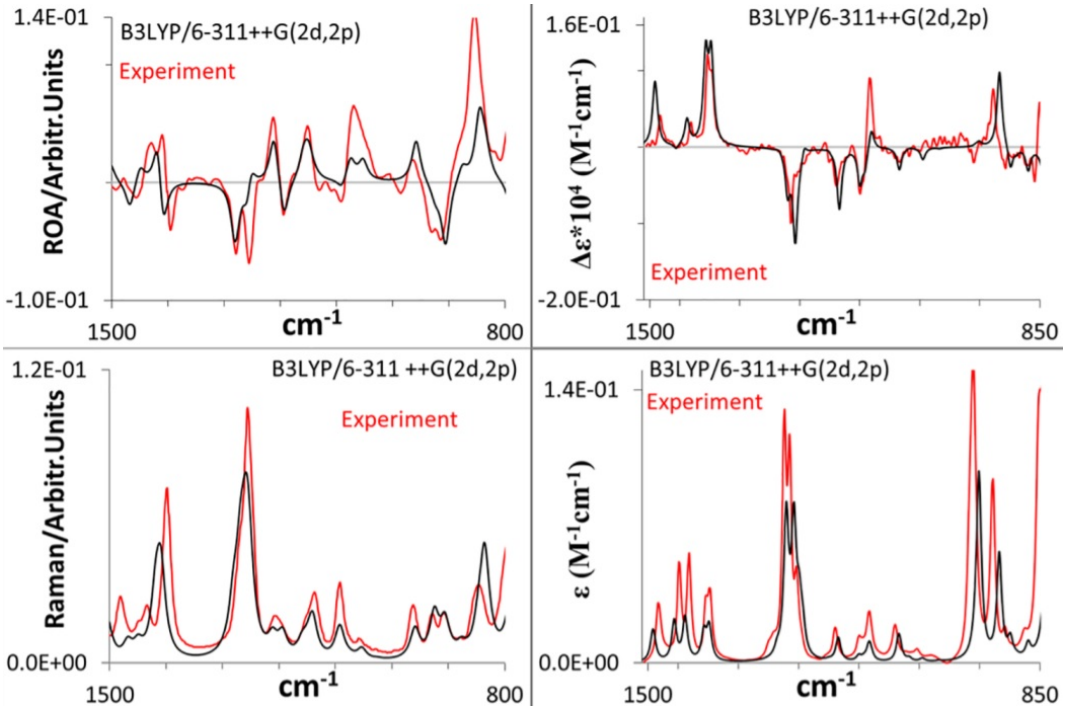

Figure 9.10: Predicted Raman and ROA (left), VA, and VCD (right) spectra for structural isomer #4 and the experimental spectra of (*S*)-(+)-epichlorohydrin

threshold value.

The strong similarity between the predicted spectra for structural isomer #6 and the experimental spectra of the desired compound, (2*R*,3*R*)-(+)-dimethyloxirane, can be clearly seen in the overlaid spectra in Figure 9.7. The QC predicted Raman and ROA spectra were simulated for structure #6 with molecular formula C_4H_8O at the B3LYP/6-311++G(2d,2p) level in vacuo with 488 nm incident light and band positions scaled by 0.969, which corresponds to the x-axis scale factor that yields the maximum *Sim*ROA value. A few features in ROA spectra at ∼1100, ∼800-900, and ∼100 cm^{-1} have mismatching intensities, which may be corrected at higher theoretical level calculations. The QC predicted VA and VCD spectra were simulated for structure #6 with molecular formula C_4H_8O at the B3LYP/6-311++G(2d,2p)/PCM(CCl_4) level and band positions scaled by 0.979, which corresponds to the x-axis scale factor that yields the maximum *Sim*VCD value.

9.3.2 C_3H_5ClO

The main work[55] summarizes the chemical names and the number of conformations of each of the nine chiral structural isomers with molecular formula C_3H_5ClO. The predicted VCD and ROA spectra for these structural isomers are displayed in Figures 8 and 9, along with the experimental spectra of the desired compound (*S*)-

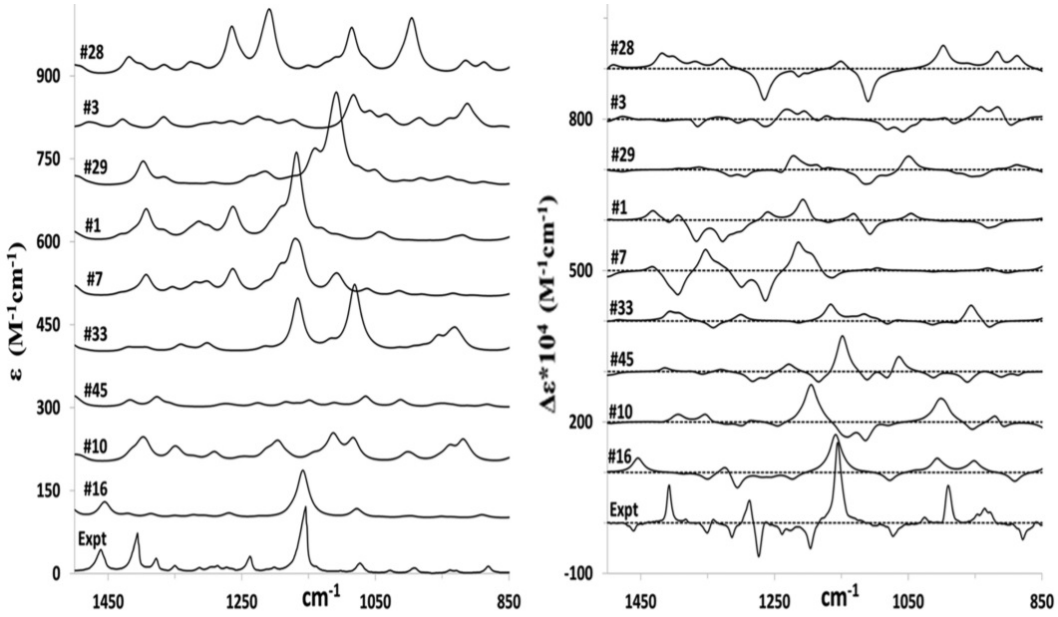

Figure 9.11: Predicted VA and VCD spectra of nine selected chiral structural isomers with molecular formula $C_6H_{10}O$ and the experimental spectra of (*R*)-(+)-3-methylcyclopentanone

(+)-epichlorohydrin. Experimental VA and VCD spectra were adapted with permission from *J. Phys. Chem. A.* 2000, 104, 6189-6196.[231] Experimental Raman and ROA spectra were measured in CCl_4 solvent with SCP 180° unpolarized scattering geometry and 532 nm laser light. The experimental Raman and ROA spectra below 800 cm^{-1} are not used due to interference from solvent Raman bands. Experimental Raman and ROA spectra were adapted with permission from *J. Chem. Theory Comput.* 2014, 10, 5520-5527.[233]

The *Sim*VA, *Sim*VCD, and *Sim*VDF values for the predicted spectra of nine structural isomers, comparing their spectra with each other, and the corresponding *Sim*Raman, *Sim*ROA, and *Sim*CID values are given in Johnson et al.[55] The off-diagonal Sim values above the reliability threshold magnitude of 0.4 are once again highlighted. The following observations can be noted from the cross-correlation Sim values. (1) As can be seen, there is no significant similarity between ROA or CID spectra of the structural isomers, indicating the uniqueness of ROA spectra for chiral structure discrimination of structural isomers with molecular formula C_3H_5ClO. (2) Structural isomers 5 and 8 have *Sim*VCD magnitude of >0.4, but the corresponding *Sim*VDF is <0.4. Similarly, structural isomers 8 and 9 have *Sim*VDF magnitudes of >0.4, but the corresponding *Sim*VCD values are <0.4. It should be noted that when both *Sim*VCD and *Sim*VDF are not simultaneously greater than the reliability threshold, the individual magnitudes of *Sim*VCD and *Sim*VDF should not be used with confidence. This criterion is needed to identify the uniqueness of the VCD spectra for chiral structure discrimination of structural isomers with molecular formula C_3H_5ClO. (3) There are several structural isomers of C_3H_5ClO for which the cross-correlation *Sim*VA or *Sim*Raman magnitudes are greater than the re-

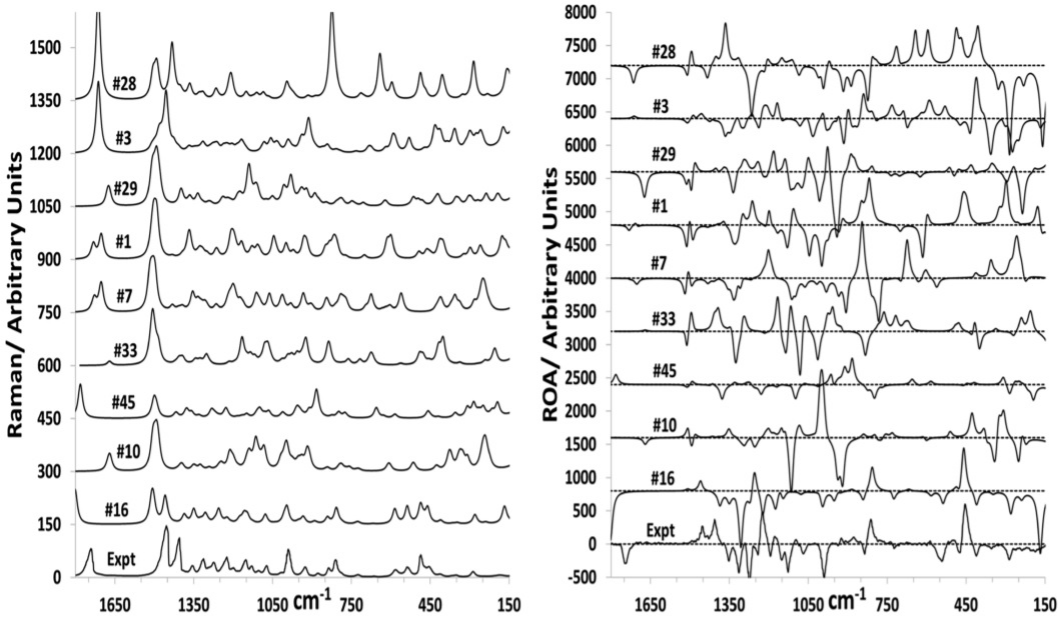

Figure 9.12: Predicted Raman and ROA spectra of nine selected chiral structural isomers with molecular formula $C_6H_{10}O$ and the experimental spectra of (*R*)-(+)-3-methylcyclopentanone

liability threshold. Therefore, VA and Raman spectra by themselves would not be adequate for discrimination among structural isomers of C_3H_5ClO.

Structure	***Sim*VA**	***Sim*VCD**	***Sim*VDF**	***Sim*Raman**	***Sim*ROA**	***Sim*CID**
1	0.30	0.16	0.21	0.80	0.25	0.22
2	0.26	−0.09	−0.13	0.67	0.27	0.25
3	0.21	0.07	0.11	0.72	−0.16	−0.16
4	**0.79**	**0.59**	**0.63**	**0.90**	**0.61**	**0.53**
5	0.18	−0.07	−0.31	0.50	0.32	0.32
6	0.15	−0.13	−0.14	0.43	0.23	0.16
7	0.45	0.16	0.05	0.81	−0.14	−0.08
8	0.54	−0.19	−0.24	0.75	−0.34	−0.24
9	0.43	0.13	0.20	0.71	−0.22	−0.16

Table 9.2: Similarity of each chiral structural isomer compared to the corresponding experimental measurements of (*S*)-(+)-epichlorohydrin

The Sim values for the predicted spectra of chiral structural isomers of C_3H_5ClO, obtained using the experimental spectra of the desired compound (*S*)-(+)-epichlorohydrin, are given in Table 9.2, where the Sim values larger than the reliability threshold are bolded. From a comparison of these values, the following

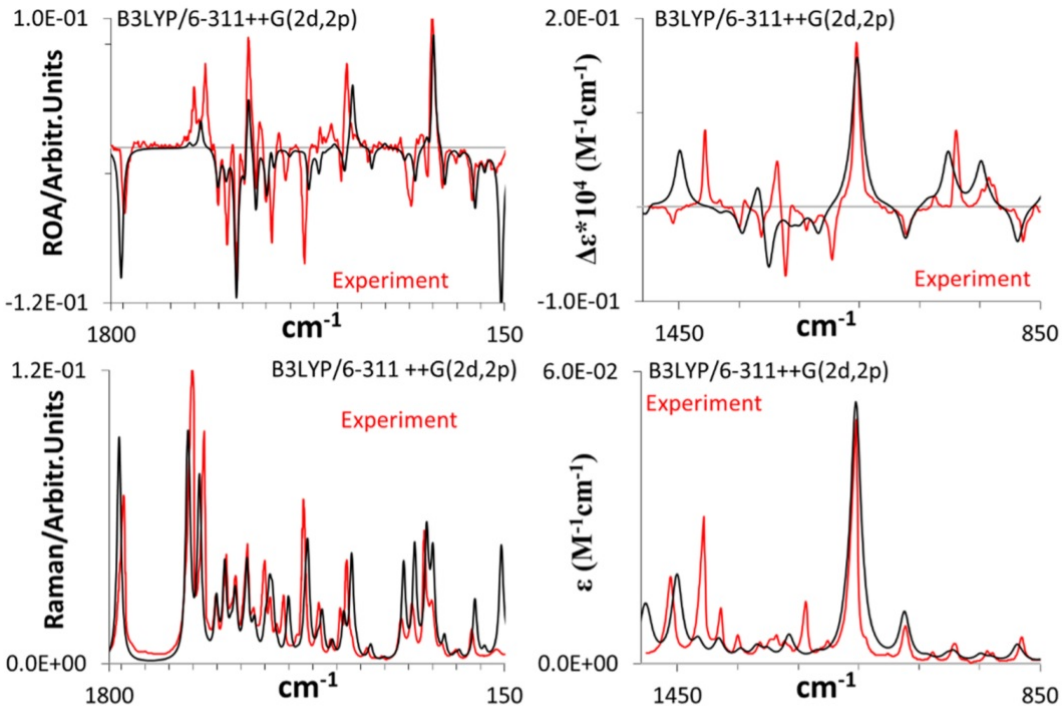

Figure 9.13: Predicted Raman and ROA (left), VA, and VCD (right) spectra for structural isomer #16 and the experimental spectra of (*R*)-(+)-3-methylcyclopentanone

points can be noted. (1) The magnitudes of *Sim*VA, *Sim*VCD, *Sim*VDF, *Sim*Raman, *Sim*ROA, and *Sim*CID values are all above the reliability threshold only for structure #4, which in fact is the correct structure for (*S*)-(+)-epichlorohydrin. Thus, VCD and/or ROA spectra can be considered to be uniquely associated with chiral structural isomers of C_3H_5ClO, as also seen for C_4H_8O. (2) VA and Raman spectra by themselves do not provide enough discrimination for structural isomers of C_3H_5ClO. This is because all structural isomers of C_3H_5ClO have *Sim*Raman magnitudes greater than the threshold value, and four structural isomers (#4, #7, #8, and #9) have *Sim*VA magnitudes greater than the threshold value.

The strong similarity between the predicted spectra for structural isomer #4 and the experimental spectra of (*S*)-(+)-epichlorohydrin can also be seen visually in the overlaid spectra (Figure 9.10). The QC predicted Raman and ROA spectra are simulated from the Boltzmann weighted spectra of all 3 conformers of structural isomer #4 with molecular formula C_3H_5ClO at the B3LYP/6-311++G(2d,2p)/PCM(CCl_4) level with 532 nm incident light and band positions scaled by 0.985, which corresponds to the x-axis scale factor that yields the maximum *Sim*ROA value. QC predicted VA and VCD spectra of #4 were simulated from the same level of theory and band positions scaled by 0.977, which corresponds to the x-axis scale factor that yields the maximum *Sim*VCD value.

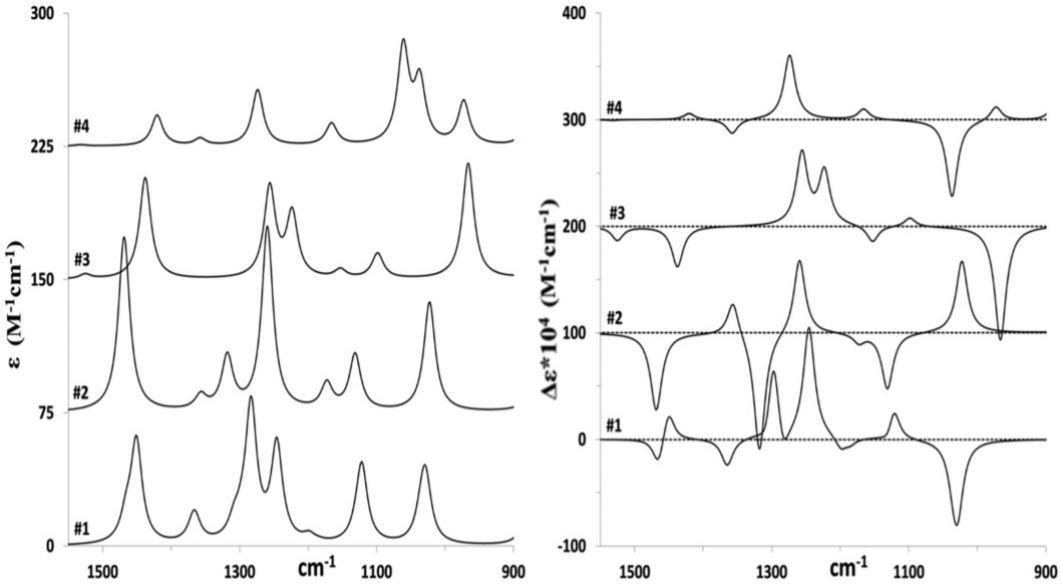

Figure 9.14: Predicted VA (left) and VCD (right) spectra of the four chiral structural isomers (#1-4) with molecular formula CH_3NO_2

9.3.3 $C_6H_{10}O$

The main text[55] summarizes the chemical names and the number of conformations of each of the 75 chiral structural isomers with molecular formula $C_6H_{10}O$. The predicted VCD and ROA spectra for nine selected structural isomers with higher similarity are displayed in Figures 11 and 12, along with the experimental spectra of the desired compound (*R*)-(+)-3-methylcyclopentanone. Experimental VA and VCD spectra were measured in CCl_4 and adapted with permission from *J. Phys. Chem. B.* 2004, 108, 20451-20457.[234] Experimental Raman and ROA spectra were measured for the neat liquid sample with ICP 90° depolarized scattering geometry and 488 nm laser wavelength and adapted with permission from *J. Phys. Chem. A.* 1993, 97, 11211-11215.[235]

It is not practical to present the cross-correlation table for chiral structural isomers of $C_6H_{10}O$ due to its large size (76 rows), nor the table comparing similarity of structural isomers to the experiment for the same reason. Therefore, I have restricted this discussion to the Sim values in Table 4 of Johnson et al.[55] obtained for 75 chiral structural isomers of $C_6H_{10}O$, using the experimental spectra of the desired compound, (*R*)-(+)-3-methylcyclopentanone. From a comparison of these values, the following observations can be noted. (1) Only one structural isomer #16, which happens to be the correct structure of (*R*)-(+)-3-methylcyclopentanone, has similarity overlap magnitudes satisfying the reliability threshold for all Sim values (*Sim*VA, *Sim*VCD, *Sim*VDF, *Sim*Raman, *Sim*ROA, and *Sim*CID). (2) Structural isomer #10 has Sim values satisfying the threshold for *Sim*VCD, and *Sim*VDF but not for *Sim*VA. Here, it should be noted that a reasonable similarity in

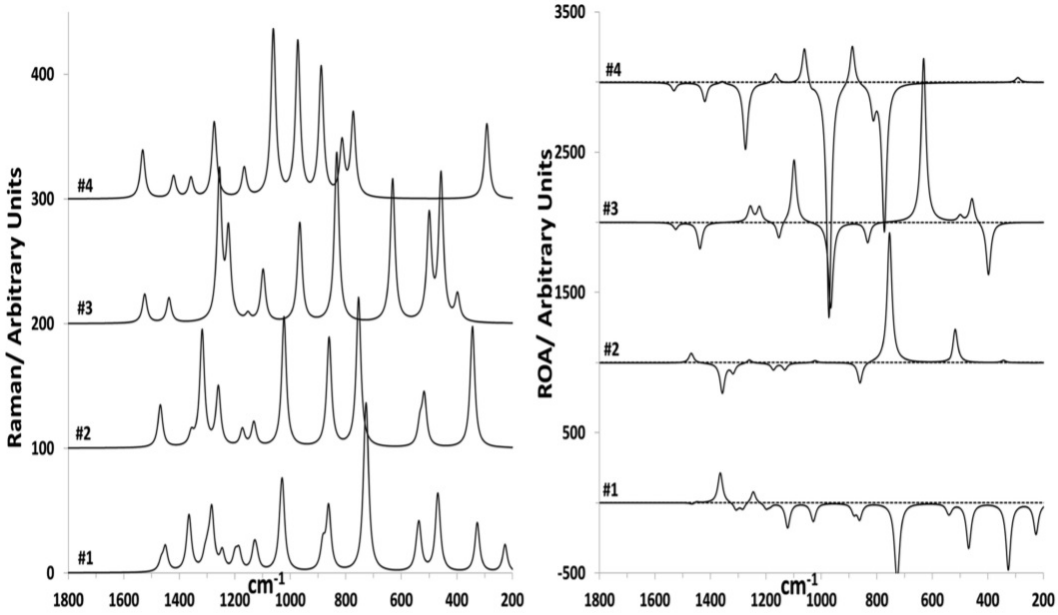

Figure 9.15: Predicted Raman (left) and ROA (right) spectra of the four chiral structural isomers (#1-4) with molecular formula CH_3NO_2

VCD spectra should not be viewed with confidence unless the corresponding VA spectrum has a strong similarity. Moreover, *Sim*ROA and *Sim*CID values for this structural isomer are quite low, ruling out #10 as a possible structure. This observation also indicates that ROA appears to provide better discrimination for structural isomers of $C_6H_{10}O$ and maybe more generally so because ROA spectra are measured and analyzed in a wider vibrational spectral region than VCD. Importantly, it is recommended to use both ROA and VCD spectra simultaneously for unknown structural determination. (3) Multiple structural isomers of $C_6H_{10}O$ have significant *Sim*VA and *Sim*Raman values with the experimental spectra of (*R*)-(+)-3-methylcyclopentanone, emphasizing the inadequacy of VA/Raman spectra for discrimination of structural isomers of $C_6H_{10}O$.

The markedly low *Sim*VCD, *Sim*VDF, *Sim*ROA, and *Sim*CID values for all examined structural isomers, except for #16, provide further confidence that VOA spectra can be considered to be uniquely associated with chiral structural isomers of $C_6H_{10}O$. The strong similarity between the predicted spectra for #16 and the experimental spectra for the desired compound, (*R*)-(+)-3-methylcyclopentanone, can be visually seen in the overlaid spectra in Figure 9.13. QC predicted Raman and ROA spectra were simulated from the Boltzmann weighted spectra of both conformers of structural isomer #16 with molecular formula $C_6H_{10}O$ at the B3LYP/6-311++G(2d,2p) level in vacuo with 488 nm incident light and vibrational frequencies scaled by 0.976, which corresponds to the x-axis scale factor that yields the maximum *Sim*ROA value. QC predicted VA and VCD were simulated at the same level, with PCM for CCl_4 solvent, and scaled by 0.9965 which corresponds to the x-axis scale factor that yields the maximum *Sim*VCD value.

9.3.4 CH_3NO_2

N-Hydroxyoxaziridine (also known as 1,2-oxaziridin-2-ol) was only recently detected[231] in the gas phase of subliming binary ices of methane and nitrogen dioxide through tunable vacuum ultraviolet photoionization reflectron time-of-flight mass spectrometer.[236,237] Johnson et al.[55] summarizes the chemical names and the number of conformations of each of the four chiral structural isomers with molecular formula CH_3NO_2. The predicted VCD and ROA spectra corresponding to these structural isomers are displayed in Figures 14 and 15. Without experimental VCD or ROA spectra to compare to the simulated spectra, the validity and accuracy of the predicted VCD and ROA spectra cannot be known. However, cross-correlation of the similarities among the predicted spectra for these structural isomers can show the level of similarity among the predicted spectra for structural isomers of CH_3NO_2.

Johnson et al.[55] summarizes the similarities among the predicted spectra of four chiral structural isomers of CH_3NO_2 when compared to one another. Although structural isomers #1 and #4 have cross-correlation *Sim*VCD and *Sim*VDF magnitudes larger than the reliability threshold, the corresponding *Sim*VA value is below the threshold. Also, the cross-correlation *Sim*ROA and *Sim*CID magnitudes for structural isomers #1 and #4 are well below the reliability threshold. It is important to note that no single pair of structural isomers gives significant spectral similarity values for all *Sim*VA, *Sim*VCD, *Sim*VDF, *Sim*Raman, *Sim*ROA, and *Sim*CID parameters. Thus, VOA spectra of structural isomers of CH_3NO_2 can also be considered to be uniquely characteristic of the respective structures.

Visual comparison in Figures 14 and 15 indicates that ROA spectra provide better chiral structure discrimination than VCD among the structural isomers of CH_3NO_2, which is also reflected in Sim values. Then, it is reasonable to suggest that when ROA spectral measurements become available for the elusive enantiomer of *N*-hydroxyoxaziridine, the predicted ROA spectra for all chiral structural isomers with molecular formula CH_3NO_2 are sufficiently different to allow for the assignment of its AC.

9.4 Conclusion

Examination of the predicted VOA spectra for all chiral structural isomers of C_4H_8O indicates that the correct structure of the desired compound, (2*R*,3*R*)-(+)-dimethyloxirane, can be readily determined just from its molecular formula and experimental VOA spectra. The same conclusion is obtained for determining the structure of another desired compound, (*S*)-(+)-epichlorohydrin, just from its molecular formula and the experimental VOA spectra. For the desired compound, (*R*)-(+)-3-methylcyclopentanone, with molecular formula $C_6H_{10}O$, only the structural isomers containing C=O double bond and/or five-membered ring structure were considered to overcome the complexity in analysis. A comparison of the VOA spectra for the resulting 75 structural isomers is still cumbersome. It may be noted that the analysis for $C_6H_{10}O$ would have been

much simpler if I had excluded the heterocyclic five-membered rings, but I did not do that in the interest of generality. Yet, the quantitative comparison of the predicted VOA spectra with experimental spectra of (*R*)-(+)-3-methylcyclopentanone led to the correct structure determination. One might consider restricting the structural isomers to those containing the C=O group or five-membered ring as a limitation of the current approach when molecular formulae with a large number of atoms are involved. However, several methods are now able to adequately assign, besides the molecular formula, the correct functional groups. Therefore, with currently available methods, restricting the analysis of structural isomers to those with known functional groups need not be viewed as a limitation.

As evidenced by the cross-correlation of spectral similarities among all chiral structures of C_4H_8O, C_3H_5ClO, and CH_3NO_2, an unknown compound with known molecular formula should have uniquely identifiable VOA spectra. It is noteworthy that either similarities of VA, VCD, and VDF combination or of Raman, ROA and CID combination alone may correctly predict the structural isomer corresponding to the desired compound. However, one might encounter some ambiguities when only one of these two combinations is utilized. Therefore, it is strongly recommended that both ROA and VCD spectra be used simultaneously and similarities for VA, VCD, VDF, Raman, ROA, and CID combination be used for unknown chiral molecular structure determination. This recommendation falls in line with the suggestion of using multiple methods[51,58,141] to add further confidence to the final assessments.

The current proof-of-concept study shows that if an unknown compound with known molecular formula can be isolated and its VOA spectra measured, the chiral molecular structure of that unknown compound could readily be discerned. The steps involved in this approach include the measurement of experimental VOA spectra for a desired chiral compound with known molecular formula and then predicting the VOA spectra for all chiral structural isomers associated with that molecular formula. Similarity analysis of the predicted VOA spectra for all possible chiral structural isomers among each other and with experimental spectra of a desired compound can lead to the identification of a unique chiral molecular structure for the desired compound.

A bottleneck in the current approach is the generation of all possible chiral structural isomers associated with a given molecular formula. The available structure generator programs can be used to generate structural isomers, but the identification of *E*/*Z* isomerism, chirality, and diastereomers need to be handled manually. For the molecular formulae where the number of associated chiral structural isomers is less than $\sim$100, as may be the case for most medium size molecules, the procedure demonstrated in this article can be applied readily. However, for the molecular formulae representing larger size molecules, the number of chiral structural isomers associated with such molecular formulae can become impractically large. In such cases, if the number of chiral structural isomers can be pared down to a practical level by making use of the known information

on functional groups, ring types and their sizes, etc., then I anticipate that the approach presented here can be readily applied to larger size molecules. Keeping this limitation in mind, the observations reported in this article should pave the way for a new paradigm in chiral molecular structure determination.

CHAPTER 10

Conclusions

I have been fortunate in my years at Vanderbilt University to encounter many unique natural products and small synthetic molecules. I have been able to undertake measurements on the chiroptical properties of many of these compounds, including two trips to Poland for ROA measurements. I have also spent years performing calculations on these various natural products and small synthetic molecules. The veracity of those results depend as much upon the model and methods one utilizes as it does the complexity of the system one is investigating. However, the computational cost should not be the deciding factor or justification in not pursuing a particular problem. There are a wide variety of computational methods available, and even more clever ways to tease out desired details with thoughtful design of the investigation.

The determination of the AC of inuloxin C was made possible by the concerted investigation of inuloxin C with its acetylated derivative. Electronic circular dichroism, EDF, and ORD could not sufficiently discriminate amongst the eight diastereomers of inuloxin C, or its derivative, but the combined results clearly establish the AC of both. This result identified a pathway for determining the ACs of chiral compounds with multiple stereogenic centers whether relative configurations of the molecule are known or not. For inuloxin D and E, the assignments of AC were made possible through careful analysis of chiroptical properties of acetylated inuloxin D. ECD and ORD were used to assign two of the 4 stereocenters, while carbonyl stretching vibrations were used to ascertain the AC of a third stereocenter. Utilizing the differences of VCD spectra between epimers (see Figure 3.4) allows for the determination of the final stereocenter and the AC assignment of acetylated inuloxin D, and, by extension, that of inuloxins D and E. This firmly established the usefulness at examining stereochemically unaltered derivatives of compounds for investigation of their ACs. It also highlights the failure to directly determine the AC of two of these compounds through chiroptical methods. The incorporation of larger basis sets, grid sizes, or anharmonicity will likely serve to improve similar calculations in the future.

In the study on the ladderanoic acids, the AC was already established by X-ray crystallography and ORD. My investigation of these showed not only could the AC of these ladderanes be established independently with ROA, but that ROA (but not VCD) calculations on monomers can adequately reproduce experimental spectra even when dimer formation is likely. This itself leads to an immense reduction in the computational cost necessary for ROA calculations on dimer forming molecules. It will be beneficial to use VCD in a similar manner, although in this case VCD prediction on the monomer conformations were not sufficient for reproducing the experimental spectra of either ladderanoic acid. Converting the acids into methyl esters or sodium

salts shows some promise in simplifying not only the calculations, but also the experiments by preventing dimerization. However, having a method of accurately predicting VCD spectra of dimer forming compounds on monomer conformations would be useful as well, as then the compound examined needs no experimental modifications such as esterification. To this end, I speculate that in some cases of dimer forming compounds, the monomer conformations can be hydrogen bonded to formic acid, the simplest carboxylic acid, for QC prediction of chiroptical properties. This would not increase the conformational space considered for the calculations, and would better model the interactions at the carboxylic acid group of dimer forming compounds. This method would likely work well with compounds such as the ladderanoic acids where the carboxylic acid group is not near a stereogenic center or other source of chirality, but may be less useful when carboxylic acid groups are in close proximity to a stereocenter.

Determination of the AC of an analogue of (-)-crispine A utilized ORD, ECD, and VCD. However, it is only when the dissymmetry factor spectra (EDF and VDF) are included that one can reliably assign the AC, establishing dissymmetry factor spectral analysis as a useful tool in diastereomer discrimination. Having this additional tool in our toolbox provides additional spectroscopic data, and thus confidence in QC predictions of chiroptical properties, especially in cases where we are limited to certain chiroptical methods. For example, when ROA is not available due to the sample fluorescing, the addition of EDF and VDF data may provide enough information to determine the AC of the compound.

Generally it is expected that when hydrogen bonding solvents are used, explicit solvent molecules or MD simulations are necessary. However, I measured the VCD of this analogue of (-)-crispine A in deuterated methanol and obtained useful results from QC prediction of the VCD without including explicit methanol solvent molecules. This lead to a multi-pronged investigation of compounds in methanol solvent. For my part in this investigation, I measured and analyzed ORD, ECD, and VCD of an analogue of (+)-crispine A in methanol solvent, while collaborators investigated two other compounds. These compounds had known ACs through their synthetic schemes, however I was able to show that for molecules with two or more hydroxyl groups with favorable intramolecular hydrogen bonding, explicit solvation is not necessary to satisfactorily reproduce VCD and ORD properties. This can lead to large reduction in computation cost through avoiding cumbersome MD simulations, processing, and the additional solvent molecules that would typically be included in the QC optimizations and spectral predictions. However, it is important to note that the low energy optimized intramolecular hydrogen bonding conformations may not be the low energy conformers that exist in hydrogen bonding solvent. In these cases, we can expect that the general skeleton of the low energy conformers are accurate, but the atoms participating in H bonding, and adjacent atoms, may not accurately reflect the true low energy conformers in solution. Thus, this is a useful method for quickly establishing the AC of a hydrogen bonding compound, but accurately gauging the conformational space of the compound would

likely require inclusion of explicit solvent molecules. When explicit solvation is necessary for reproducing the experimental VCD, such as with 1-indanol in DMSO, I showed that systematically adding explicit solvent molecules to each conformation of 1-indanol allowed for good reproduction of experimental VCD. The prediction of VCD appeared even better than similar predictions made utilizing MD for explicit solvation. This is not to say that MD does not have a place in the prediction of chiroptical properties, but that in some cases systematic placement of solvent molecules can reproduce chiroptical properties as well as or better than MD approaches, at a significantly reduced computational cost. Ultimately this provides a fast and efficient method for predicting chiroptical properties when explicit solvent molecules are necessary for accurate predictions. This method also should accurately model the conformational space of the solute-solvent complex, where additional accuracy can be obtained by including more intervals of rotation in the process of systematically solvating the compound.

In seeking to reduce computational costs, sometimes one can overreach. The QC prediction of ROA is outlined in section 1.3.4 and chapter 8. Some suggestions of lesser importance of the δ_k^2 contribution to ROA from QC predictions, as predominant δ_k^2 had not been seen in QC predicted ROA. I identified several ROA bands of (1*S*)-camphor that arose predominantly from the δ_k^2 contribution. While not including the δ_k^2 contribution will generally lead to a satisfactory reproduction, in every level of theory examined for the ROA of (1*S*)-camphor, the similarity between experimental and predicted ROA was larger when including the δ_k^2 term, on average by 9%. The inclusion of δ_k^2 term does not significantly increase computational cost, while it may significantly improve the similarity between experimental and QC predicted ROA spectra. Ultimately, the results from my study on the ROA of (1*S*)-camphor established the δ_k^2 contribution to be unavoidable for the QC interpretation of ROA and that its inclusion can increase the similarity between predicted ROA and corresponding experimental ROA measurements.

If feasible and timely, all QC calculations would proceed with the highest accuracy possible. However, there are many concessions worth making at times. The most obvious is that generally quantum chemists utilize implicit solvation formalisms to model solvent effects of non-hydrogen bonding solvents. This works extremely well, and the inclusion of explicit solvation would likely only increase the cost of doing the calculation, not its accuracy. In several of my investigations, there is a choice to be made between time and accuracy. When I established the importance of the δ_k^2 contribution to ROA with many different QC levels of theory, I was examining a molecule with only a single conformation. While I did see significant increases with the similarity of my predicted ROA and CID to the corresponding experimental spectra at every level of theory examined, I can understand why someone investigating a more flexible molecule or more complicated system may opt to reduce their computational time by ignoring the δ_k^2 contribution to ROA. Their results will be less accurate but likely still accurate enough for confident assignments of AC for a compound. In a

similar fashion, this is why I accepted the loss of accuracy for the analogues of (-)- and (+)-crispine A by allowing their intramolecular hydrogen bonding to account H-bonding interactions, as these compounds were large, flexible, and had several hydrogen bonding groups. However, with 1-indanol, a small molecule with only a few conformations and one H-bonding group, it was no problem to explicitly solvate the system to obtain very accurate, and yet still quick, results. Similarly, for studying the ladderanoic acids, modeling the dimer formation of compounds with such immense conformational spaces would increase the accuracy of my results. However, considering how long it took to even perform a reduced conformational search on the monomer units, I do not believe I would have been able to complete my study on the ladderanoic acids within my time at Vanderbilt if I had included dimeric structures. Of course, this loss of accuracy limited my study to ORD and ROA analyses, despite having nice experimental VCD spectra for each ladderane. I suspect that utilizing a formic acid cap on dimer forming structures such as the ladderanoic acids would better model the system than just examining monomer conformations, and allow for VCD analysis, although likely still with some loss of accuracy or discrepancies. Luckily, some loss of accuracy in QC predictions is acceptable. Assuming one can correctly identify the chemical formula of a compound, which can readily be accomplished with routine methods such as mass spectrometry, then the vibrational optical activity spectra can be expected to uniquely identify the compound amongst all other compounds with the same chemical formula as I established in chapter 9 with my work on chiral molecular structure determination for a desired compound just from its molecular formula and VOA spectra.[55] The inclusion of ratio spectra (EDF, VDF, CID) allows for better discrimination amongst diastereomers while increasing the confidence in the assessments made. All of these findings come together to allow for quick and accurate determination of a compound, its absolute configuration, and often additional structural information.

While these methods are widely applicable to many areas, the pharmaceutical industry would benefit greatly from applying these practices in the process of drug screening and synthesis. For example, if we imagine that a drug company is interested in a hypothetical oily compound that they recently isolated and exhibited anti-tumor properties. This hypothetical compound contains four hydroxyl groups and is soluble in only methanol. After rudimentary analysis, such as mass spectrometry or NMR, the scientists working for this drug company might be able to determine the structure of the compound, except the configuration at each of the stereocenters. Suppose there are five stereocenters, then that will result in 16 non-enantiomeric diastereomers to investigate. ECD measurements can be carried out and ECD spectral predictions made for each diastereomer. In this hypothetical situation, even though there are four hydroxyl groups in this compound, they are in close proximity to each other and explicit solvation for calculations was not modeled.[57,162] Suppose the ECD data alone rules out only two of the diastereomers from low similarity. By extending ECD analysis to include EDF, which does not require additional computational time, suppose eight diastereomers

and their enantiomers are eliminated as possible contenders for the correct compound.[126] At this point, the diastereomers with the largest similarity to the experimental ECD and EDF are the most likely configurations for the AC. If there is only a couple with large similarity with ECD and EDF, it may be worthwhile for the synthetic chemist to attempt synthesizing these diastereomers for direct experimental comparisons of ECD and EDF. Suppose, however, that eight remaining diastereomers share similar levels of spectral similarity overlap with the experimental ECD and EDF, then further analysis may be a more efficient route at determining the AC of the compound. There are several options which can be considered. The compound can be altered, such as through acetylation or esterification; if the alteration leaves the compound unaltered stereochemically, then ECD and EDF analysis of this derivative and the original compound can be combined to reduce the number of diastereomers that are possible contenders for the AC, as in the cases of acetylated inuloxin C and acetylated inuloxin D.[56,58] Besides utilizing stereochemically unaltered derivatives for this investigation, the scientists could instead use the difference spectra of epimers to gain insight to the AC or configuration at a particular stereocenter, as in the case of acetylated inuloxin D.[56] While ECD spectroscopy is more commonly used in general, by using vibrational optical activity instead of or in conjunction with ECD, the diastereomer with the correct AC might be uniquely identified.[55] Seemingly, the scientists could likely use only VCD or ROA along with their corresponding ratio spectrum to determine the AC of the compound,[55] although in academic research the use of multiple chiroptical methods is recommended for confidence.[55,56,85] Suppose the synthetic chemist is able to successfully synthesize the expected diastereomer (or few most likely diastereomers), then direct experimental comparison of whichever chiroptical properties measured can also provide confidence in the assignment of AC. Utilizing some, all, or even just one of these techniques will reduce the work the synthetic chemist must do to unambiguously determine the configuration of this or any compound of interest.

Computational investigation on even small chiral molecules is not always as straightforward as predicting chiroptical spectra and comparing them to the corresponding experimental spectra. There are limits to computational methods. These limits will depend on the properties or information that is desired as well as the system of interest. Orbital free DFT was used in 2009 to calculate the energy of over one million atoms in only 5 hours[245] and linear scaling DFT was used to calculate the energy of over two million atoms in 2010 in just under 2 hours.[246] However in each of these cases, the energy calculations were done on bulk matter of a single element, which is not particularly interesting to the chemist. The latter calculation also required over 4000 cores to compute so quickly. While these computations demonstrate that modified DFT can be scaled up significantly for bulk calculations, I imagine that parallel communication between the cores becomes a bottleneck for these types of calculations as they scale up further. A more intriguing calculation is the QC energy optimization of a large carbohydrate formed from 26 D-glucose residues and containing 546 atoms which showed great reproduction of the X-ray crystal structure.[247] Schnupf et al. were able to

perform such a large optimization by utilizing a smaller basis set for less-essential atom interactions while maintaining a larger basis set for critical aspects of the calculation. However, if chemically accurate calculations of chiroptical properties are desired, using this reduced basis set method will likely be insufficient for reproducing experimental data, especially for a compound with over one hundred stereocenters. The largest compound I have investigated is [3]-ladderanoic acid with 54 atoms. While larger compounds can be used for prediction of chiroptical properties, the limit comes down to two predominant factors, the number of stereogenic elements and flexibility of the compound. Since a compound with n stereocenters will have up to $\frac{1}{2} * 2^n$ diastereomers worth investigating, compounds with more stereocenters become exponentially more challenging. Flexible compounds lead to larger conformational spaces. The computational time needed for predicting chiroptical properties increases roughly proportionally with the amount of conformers examined. To this end, I expect that compounds with approximately 250 atoms are the extent to which meaningful and accurate QC predictions of chiroptical properties can be carried out currently. Particularly contrived compounds or systems could have twice as many atoms incorporated in the successful prediction of chiroptical properties. There is no doubt that computational power will continue to grow. The models used in the prediction of chiroptical properties will continue to be improved. Clever methods for reducing the complexity of the system will continue to emerge. The prediction of chiroptical properties will ultimately be limited by the conformational space of a system, the number of sources of chirality a system contains, the accuracy desired, and the time one can or is willing to spend. It is my dream to one day see accurate QC predictions of chiroptical properties on systems with thousands of atoms like whole proteins.

References

[1] E. Merzbacher. *Quantum Mechanics*, Wiley: New York, 1998.

[2] J. Knobloch, J. D. Shaughnessy, and U. Ruther. Thalidomide induces limb deformities by perturbing the bmp/dkk1/wnt signaling pathway. *FASEB*, 21(7):1410–1421, 2007.

[3] L. Friedman and J. G. Miller. Odor incongruity and chirality. *Science*, 172(3987):1044–1046, 1971.

[4] *Comprehensive Chiroptical Spectroscopy*, Wiley: Hoboken, New Jersey:Vol. 1–2, 2012.

[5] T. D. Crawford, M. C. Tam, and M. L. Abrams. The current state of ab initio calculations of optical rotation and electronic circular dichroism spectra. *J. Phys. Chem. A*, 111(48):12057–12068, 2007.

[6] J. Kypr, I. Kejnovska, K. Bednarova, and M. Vorlickova. *Circular Dichroism Spectroscopy of Nucleic Acids*, Wiley: New York, 2012.

[7] J. Kypr, I. Kejnovska, D. Renciuk, and M. Vorlickova. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res*, 37(6):1713–1725, 2009.

[8] S. Ma, X. Cao, M. Mak, A. Sadik, C. Walkner, T. B. Freedman, I. K. Lednev, R. K. Dukor, and L. A. Nafie. Vibrational circular dichroism shows unusual sensitivity to protein fibril formation and development in solution. *J. Am. Chem. Soc.*, 129(41):12364–12365, 2007.

[9] J.B. Biot. Optical activity. *Mem. Institut.*, 13(1), 1812.

[10] P.L. Polavarapu, A. Petrovic, and F. Wang. Intrinsic rotation and molecular structure. *Chirality*, 15:143–149, 2003.

[11] P.L. Polavarapu. Ab initio molecular optical rotations and absolute configurations. *Mol. Phys.*, 91(3):551–554, 1997.

[12] P.J. Stephens, F.J. Devlin, J.R. Cheeseman, and M.J. Frisch. Calculation of optical rotation using density functional theory. *J. Phys. Chem. A*, 105(22):5356–5371, 2001.

[13] J. Autschbach, S. Patchkovskii, and T. Ziegler. Chiroptical properties from time-dependent density functional theory. II. optical rotations of small to medium sized organic molecules. *J. Chem. Phys.*, 117(2):581–592, 2002.

[14] R. Ditchfield. Self-consistent perturbation theory of diamagnetism. *Mol. Phys.*, 27(4):789–807, 1974.

[15] J. Autschbach. Computing chiroptical properties with first-principles theoretical methods: Background and illustrative examples. *Chirality*, 21(1E):E116–E152, 20009.

[16] A. Cotton. Absorption inegale des rayons circulaires droit et gauche dans certains corps actifs. *Compt. Rend*, 120:989–991, 1895.

[17] L. Rosenfeld. Quantenmechanische theorie der natürlichen optischen aktivität von flüssigkeiten und gasen. *Zeitschrift für Physik A Hadrons and Nuclei*, 52(3-4):161–174, 1929.

[18] G. Holzwarth, E. C. Hsu, H. S. Mosher, T. R. Faulkner, and A. Moscowitz. Infrared circular dichroism of carbon-hydrogen and carbon-deuterium stretching modes. *J. Am. Chem. Soc.*, 96(1):251–252, 1974.

[19] L. A. Nafie. Dual polarization modulation: A real-time, spectral-multiplex separation of circular dichroism from linear birefringence spectral intensities. *Appl. Spectrosc.*, 54(11):1634–1645, 2000.

[20] V. P. Nicu and E. J. Baerends. *Complexation, Solvation, and Chirality Transfer in Vibrational Circular Dichroism*, John Wiley & sons: New York:Vol. 1, 2012.

[21] L. A. Nafie. *Infrared vibrational optical activity: Measurement and Instrumentation*, John Wiley sons: Hoboken:Vol. 1, 2012.

[22] P. J. Stephens and N. Harada. ECD cotton effect approximated by the gaussian curve and other methods. *Chirality*, 22(2):229–233, 2010.

[23] P.L. Polavarapu. *Chiroptical spectroscopy: Fundamentals and applications*, Taylor & Francis: Boca Raton, Florida, 2017.

[24] P.L. Polavarapu. *Vibrational spectra: Principles and applications with emphasis on optical activity*, Elsevier: Amsterdam, Netherlands, 1998.

[25] L.D. Barron and C.J. Johnston. Rotational Raman optical activity in chiral symmetric tops. *J. Raman Spectrosc*, 16(3):208–218, 1985.

[26] P.L. Polavarapu. Rotational optical activity. *J. Chem. Phys*, 86:1136–1139, 1987.

[27] L.D. Barron, M.P. Bogaard, and A.D. Buckingham. Raman scattering of circularly polarized light by optically active molecules. *J. Am. Chem. Soc*, 95(2):603–605, 1973.

[28] W. Hug and H. Surbeck. Vibrational Raman optical activity spectra recorded in perpendicular polarization. *Chem. Phys. Lett*, 60(2):186–192, 1979.

[29] L. Hecht, L.D. Barron, and W. Hug. Vibrational Raman optical activity in backscattering. *Chem. Phys. Lett*, 158(5):341–344, 1989.

[30] K.M. Spencer, T.B. Freedman, and L.A. Nafie. Scattered circular polarization Raman optical activity. *Chem. Phys. Lett*, 149(4):367–374, 1988.

[31] D. Che, L. Hecht, and L.A. Nafie. Dual and incident circular polarization Raman optical activity backscattering of (-)-trans-pinane. *Chem. Phys. Lett*, 180(3):182–190, 1991.

[32] G.S. Yu and L.A. Nafie. Isolation of preresonance and out-of-phase dual circular polarization Raman optical activity. *Chem. Phys. Lett*, 222(4):403–410, 1994.

[33] P. W. Atkins and R. S. Friedman. *Molecular quantum mechanics*, Oxford University Press, 2010.

[34] A. D. Becke. Density-functional thermochemistry. III. the role of exact exchange. *J. Chem. Phys.*, 98(7):5648–5652, 1993.

[35] C. Lee, W. Yang, and R. G. Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Physical Review B*, 37(2):785–789, 1988.

[36] P. J. Stephens, D. M. McCann, J. R. Cheeseman, and M. J. Frisch. Determination of absolute configurations of chiral molecules using ab initio time-dependent density functional theory calculations of optical rotation: How reliable are absolute configurations obtained for molecules with small rotations? *Chirality*, 17(S1):S52–S64, 2005.

[37] T. D. Crawford. Ab initio calculation of molecular chiroptical properties. *Thor. Chem. Acc.*, 115, 2005.

[38] T. D. Crawford. *High-Accuracy Qunatum Chemistry and Chiroptical Properties*, Wiley: New York:Vol. 1, 2012.

[39] T. D. Crawford and W. D. Allen. Optical activity in conformationally flexible molecules: a theoretical study of large-amplitude vibrational averaging in (R)-3-chloro-1-butene. *Mol. Phys.*, 107(8-12):1041–1057, 2009.

[40] T. D. Crawford, L. S. Owens, M. C. Tam, P. R. Schreiner, and H. Koch. Ab initio calculation of optical rotation in (p)-(+)-[4]triangulane. *J. Am. Chem. Soc.*, 127(5):1368–1369, 2005.

[41] T. D. Crawford and K. Ruud. Coupled-cluster calculations of vibrational Raman optical activity spectra. *ChemPhysChem*, 12(17):3442–3448, 2011.

[42] T. D. Crawford and P. J. Stephens. Comparison of time-dependent density-functional theory and coupled cluster theory for the calculation of the optical rotations of chiral molecules. *J. Phys. Chem*, 112(6):1339–1345, 2008.

[43] S. Tsuzuki, T. Uchimaru, and K. Tanabe. Conformational analysis of n-alkanes using density functional theory. comparison with ab initio calculations. *Chem. Phys. Lett.*, 246(1-2):9–12, 1995.

[44] G. Scalmani and M. J. Frisch. Continuous surface charge polarizable continuum models of solvation. I. general formalism. *J. Chem. Phys.*, 132(11):114110, 2010.

[45] B. Mennucci, J. Tomasi, R. Cammi, J. R. Cheeseman, M. J. Frisch, F. J. Devlin, S. Gabriel, and P. J. Stephens. Calculations of solvent effects on optical rotations of chiral molecules. *J. Phys. Chem. A*, 106:6102–6113, 2002.

[46] C.L. Covington and P.L. Polavarapu. Similarity in dissymmetry factor spectra: a quantitative measure of comparison between experimental and predicted vibrational circular dichroism. *J. Phys. Chem. A*, 117(16):3377–3386, 2013.

[47] S.H. Vosko, L. Wilk, and M. Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys*, 58(8):1200–1211, 1980.

[48] G. Pescitelli, L.D. Bari, and N. Berova. Application of electronic circular dichroism in the study of supramolecular systems. *Chem. Soc. Rev*, 43:5211–5233, 2014.

[49] W.J. Hehre, L. Radom, P.v.R. Schleyer, and J.A. Pople. *Ab Initio Molecular Orbital Theory*, Wiley: New York, 1986.

[50] J. Shen, C. Zhu, S. Reiling, and R. Vaz. A novel computational method for comparing vibrational circular dichroism spectra. *Spectrochim. Acta, Part A*, 76:418–422, 2010.

[51] P.L. Polavarapu, C.L. Covington, and V. Raghavan. To avoid chasing incorrect chemical structures of chiral compounds: Raman optical activity and vibrational circular dichroism spectroscopies. *ChemPhysChem*, 18:2459–2465, 2017.

[52] C.L. Covington and P.L. Polavarapu. A single software suite for multiple chiroptical spectroscopic analyses. *Chirality*, 29:178–192, 2017.

[53] C.L. Covington and P.L. Polavarapu. Cdspectech: Computer programs for calculating similarity measures of comparison between experimental and calculated dissymmetry factors and circular intensity differentials. *https://sites.google.com/site/cdspectech1/*, 22, 2016.

[54] F.M. Junior, C.L. Covington, M.B.d. Amorim, L.S. Velozo, M.A. Kaplan, and P.L. Polavarapu. Absolute configuration of a rare sesquiterpene: (+)-3-ishwarone. *Chirality*, 77:1881–1886, 2014.

[55] J.L. Johnson and P.L. Polavarapu. Chiral molecular structure determination for a desired compound just from its molecular formula and vibrational optical activity spectra. *J. Phys. Chem. A*, 125:8000–8013, 2021.

[56] J.L. Johnson, E Santoro, R. Zatout, A.G. Petrovic, A. Cimmino, S. Superchi, A. Evidente, N.D. Berova, and P.L. Polavarapu. Absolute configuration of seco-eudesmanolide inuloxin D from experimental and predicted chiroptical studies of its 4-O-acetyl derivative. *Chirality*, 33:233–241, 2021.

[57] P.L. Polavarapu, E. Santoro, C.L. Covington, J.L. Johnson, A.R. Puente, N.D. Schley, Z. Kallingathodi, P.C. Prakasan, S. Haleema, A.A. Thomas, and I. Ibnusaud. How important are the intermolecular hydrogen bonding interactions in methanol solvent for interpreting the chiroptical properties? *Spectrochim. Acta A Mol. Biomol. Spectrosc*, 247:119094, 2021.

[58] J.L. Johnson, V. Raghavan, A. Cimmino, A. Moeini, A.G. Petrovic, S. Superchi, N. Berova, A. Evidente, and P.L. Polavarapu. Absolute configurations of chiral molecules with multiple stereogenic centers without prior knowledge of the relative configurations: A case study of inuloxin C. *Chirality*, 30:1206–1214, 2018.

[59] Andolfi A, Zermane N, Cimmino A, F. Avolio, A. Boari, M. Vurro, and A. Evidente. Inuloxins A-D, phytotoxic bi-and tri-cyclic sesquiterpene lactones produced by inula viscosa: potential for broomrapes and field dodder management. *Phytochemistry*, 86:112-120, 2013.

[60] Parker C. Observations on the current status of orobanche and striga problems worldwide. *Pest Manag Sci*, 65(5):453-459, 2009.

[61] Joel DM, Hershenhorn J, Eizenberg H, Aly R, Ejeta G, Rich PJ, Ransom JK, Sauerborn J, and Rubiales D. Biology and management of weedy root parasites. in: Janick J, ed. *Horticultural Reviews*, 2007.

[62] Zermane N, Vurro M, Boari A, Avolio F, Andolfi A, and Evidente A. Towards broomrape and field dodder management using natural metabolites from plants. *Abstracts of 11th World Congress on Parasitic Plants. Martina Franca, Ital*, page 82, 2011.

[63] Cimmino A, Masi M, Evidente M, Superchi S, and Evidente A. Fungal phytotoxins with potential herbicidal activity: chemical and biological characterization. *Nat Prod Rep.*, 32(12):1629-1653, 2015.

[64] Cimmino A, Masi M, Evidente M, Superchi S, and Evidente A. Application of mosher's method for absolute configuration assignment to bioactive plants and fungi metabolites. *J Pharm Biomed Anal.*, 144:59-89, 2017.

[65] Evidente A, Cimmino A, and Andolfi A. The effect of stereochemistry on the biological activity of natural phytotoxins, fungicides, insecticides and herbicides. *Chirality*, 25(2):59-78, 2013.

[66] Evidente A, Andolfi A, and Cimmino A. Relationships between the stereochemistry and biological activity of fungal phytotoxins. *Chirality*, 23(9):674-693, 2011.

[67] Evidente M, Santoro E, Petrovic AG, Cimmino A, Joshoubu J, Evidente A, Berova N, and Superchi S. Absolute configurations of phytotoxic inuloxins B and C based on experimental and computational analysis of chiroptical properties. *Phytochemistry*, 130:328-334, 2016.

[68] Santoro E, Mazzeo G, Petrovic AG, Cimmino A, Koshoubu J, Eviente A, Berova N, and Superchi S. Absolute configurations of phytotoxins seiricardine A and inuloxin A obtained by chiroptical studies. *Phytochemistry*, 116:359-366, 2015.

[69] Polavarapu P.L. Molecular structure determination using chiroptical spectroscopy: where we may go wrong? *Chirality*, 24(11):909-920, 2012.

[70] Polavarapu P.L., Donahue E.A., Shanmugam G., Scalmani G., Hawkins E.K., Rizzo C., Ibnusaud I., Thomas G., Habel D., and Sebastian D. A single chiroptical spectroscopic method may not be able to establish the absolute configurations of diastereomers: dimethylesters of hibiscus and garcinia acids. *J Phys Chem A.*, 115(22):5665-5673, 2011.

[71] CONFLEX. *High Performace Conformation Analysis*, Tokyo(CONFLEX Corporation), 2000.

[72] Frisch MJ, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, Williams-Young D. Sonnenberg, J. L., F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, Jr. J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox. Gaussian 09. Wallingford, Connecticut, 2016.

[73] Covington CL and Polavarapu PL. Solvation dependence observed in the electronic dissymmetry factor spectra: how much information are we missing by analyzing the circular dichroism spectra alone? *PCCP*, 18(20):13912-13917, 2016.

[74] Grauso L, Cesarano G, Zotti M, Ranesi M, Sun W, Bonanomi G, and Lanzotti V. Exploring dittrichia viscosa (L.) greuter phytochemical diversity to explain its anti- microbial, nematicidal and insecticidal activity. *Phytochem Rev.*, 19(3):659–689, 2020.

[75] Avolio F, Rimando AM, Cimmino A, Andolfi A, Jain S, Tekwani B, and Evidente A. Inuloxins A-D and derivatives as antileishmanial agents: structure-activity relationship study. *J Antibiot*, 67(8):597–601, 2014.

[76] Moeini A, Masi M, Zonno MC, Boari A, Cimmino A, Tarallo O, Vurro M, and Evidente A. Encapsulation of inuloxin A, a plant germacrane sesquiterpene with potential herbicidal activity, in β-cyclodextrins. *Org Biomol Chem.*, 17(9):2508–2515, 2019.

[77] Masi M, Fernández-Aparicio M, Zatout R, Boari A, Cimmino A, and Evidente A. Inuloxin E, a new seco-eudesmanolide isolated from dittrichia viscosa, stimulating orobanche cumana seed germination. *Molecules*, 24(19):3479–3489, 2019.

[78] Fernández-Aparicio M, Westwood JH, and Rubiales D. Agronomic, breeding, and biotechnological approaches to parasitic plant management through manipulation of germination stimulant levels in agricultural soils. *Botany*, 89(12):813–826, 2011.

[79] Zwanenburg B, Mwakaboko AS, and Kannan C. Suicidal germination for parasitic weed control. *Pest Manag Sci.*, 72(11):2016–2025, 2016.

[80] Frisch MJ, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, Williams-Young D. Sonnenberg, J. L., F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, Jr. J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox. Gaussian 16 rev. Wallingford, Connecticut, 2016.

[81] Tomasi J, Mennucci B, and Cammi R. Quantum mechanical continuum solvation models. *Chem Rev.*, 105(8):2999–3094, 2005.

[82] Tanimoto TT. An elementary mathematical theory of classification and prediction. *New York: International Buisiness Machine Corporation*, 1958.

[83] Cheng C, Maggiora G, Lajiness M, and Johnson M. Four association coefficients for relating molecular similarity measures. *J Chem Inf Comput Sci.*, 36(4):909–915, 1996.

[84] R. Dennington, T.A. Keith, and J.M. Millam. *GaussView, Version 6.1*, Shawnee Mission, Kansas(Semichem Inc.), 2016.

[85] V. Raghavan, J.L. Johnson, D.F. Stec, B. Song, G. Zajac, M. Baranska, C.M. Harris, N.D. Schley, P.L. Polavarapu, and T.M. Harris. Absolute configurations of naturally occurring [5]- and [3]- ladderanoic acids: Isolation, chiroptical spectroscopy, and crystallography. *J. Nat. Prod*, 81:2654–2666, 2018.

[86] E. Z Broda. Two kinds of lithotrophs missing in nature. *Allg. Mikrobiol.*, 17:491–493, 1977.

[87] M. Strous, J. J. Heijnen, J. G. Kuenen, and M. S. M. Jetten. The sequencing batch reactor as a powerful tool for the study of slowly growing anaerobic ammonium-oxidizing microorganisms. *Appl. Microbiol. Biotechnol*, 50:589–596, 1998.

[88] J. E. Rattray, J. van de Vossenberg, E. C. Hopmans, B. Kartal, L. van Niftrik, W. I. C. Rijpstra, M. Strous, M. S. M. Jetten, S. Schouten, and J. S. S. Damste. Ladderane lipid distribution in four genera of anammox bacteria. *Arch. Microbiol.*, 190:51–66, 2008.

[89] M. S. M. Jetten, L. van Niftrik, M. Strous, B. Kartal, J. T. Keltjens, and H. J. M. Op den Camp. Biochemistry and molecular biology of anammox bacteria. *Crit. Rev. Biochem. Mol. Biol.*, 44:65–84., 2009.

[90] B. Kartal and J. T. Keltjens. Anammox biochemistry: a tale of heme c proteins. *Trends Biochem. Sci.*, 41:998– 1011., 2016.

[91] J. S. S. Damste, M Strous, W. I. C. Rijpstra, E. C. Hopmans, J. A. J. Geenevasen, A. C. T. van Duin, L. A. van Niftrik, and M. S. M. Jetten. Linearly concatenated cyclobutane lipids form a dense bacterial membrane. *Nature*, 419:708–712, 2002.

[92] L. A. Niftrik, J. A. Fuerst, J. S. S. Damste, J. G. Kuenen, M. S. M. Jetten, and M. FEMS Strous. The anammoxosome: an intracytoplasmic compartment in anammox bacteria. *Microbiol. Lett*, 233:7–13., 2004.

[93] A. Dietl, C. Ferousi, W. J. Maalcke, A. Menzel, S. de Vries, J. T. Keltjens, M. S. M. Jetten, B. Kartal, and T. R. M. Barends. The inner workings of the hydrazine synthase multiprotein complex. *Nature*, 527:394–397., 2015.

[94] H. A. Boumann, M. L. Longo, P. Stroeve, B. Poolman, E. C. Hopmans, M. C. A. Stuart, J. S. S. Damste, and S. Schouten. Biophysical properties of membrane lipids of anammox bacteria: I. ladderane phospholipids form highly organized fluid membranes. *Biochim. Biophys. Acta, Biomembr.*, 1788:1444–1451., 2009.

[95] H. A. Boumann, P. Stroeve, M. L. Longo, B. Poolman, J. M. Kuiper, E. C. Hopmans, M. S. M. Jetten, J. S. S. Damste, and S. Schouten. Biophysical properties of membrane lipids of anammox bacteria: II. impact of temperature and bacteriohopanoids. *Biochim. Biophys. Acta, Biomembr.*, 1788:1452–1457., 2009.

[96] H. A. Boumann, E. C. Hopmans, I. van de Leemput, H. J. M. Op den Camp, J. Van De Vossenberg, M. Strous, M. S. M. Jetten, J. S. S. Damste, and S. Schouten. Ladderane phospholipids in anammox bacteria comprise phosphocholine and phosphoethanolamine headgroups. *FEMS Microbiol. Lett*, 258:297– 304., 2006.

[97] J. E. Rattray, J. van de Vossenberg, A. Jaeschke, E. C. Hopmans, S. G. Wakeham, G. Lavik, M. M. M. Kuypers, M. Strous, M. S. M. Jetten, S. Schouten, and J. S. S. Damste. Impact of temperature on ladderane lipid distribution in anammox bacteria. *Appl. Environ. Microbiol.*, 76:1596–1603, 2010.

[98] D. Rush, A. Jaeschke, E. C. Hopmans, J. A. J. Geenevasen, S. Schouten, and J. S. S. Damste. Short chain ladderanes: Oxic biodegradation products of anammox lipids. *Geochim. Cosmochim. Acta*, 75:1662–1671, 2011.

[99] D. Rush, E. C. Hopmans, S. G. Wakeham, S. Schouten, and J. S. S. Damste. Occurrence and distribution of ladderane oxidation products in different oceanic regimes. *Biogeosciences*, 9:2407–2418, 2012.

[100] D. H. Nouri and D. J. Tantillo. They came from the deep: Syntheses, applications, and biology of ladderanes. *Curr. Org. Chem.*, 10:2055–2074, 2006.

[101] N. Burns, S. R. Shuken, J. A. M. Mercer, and C. M. Cohen. Ladderane lipid compounds and liposomes and methods of preparing and using the same. *WO 2018045094*, 2018.

[102] J. A. M. Mercer, C. M. Cohen, S. R. Shuken, A. M. Wagner, M. W. Smith, III Moss, F. R., M. D. Smith, R. Vahala, A. Gonzalez-Martinez, S. G. Boxer, and N. Z. Burns. Chemical synthesis and self-assembly of a ladderane phospholipid. *J. Am. Chem. Soc.*, 138:15845–15848, 2016.

[103] J. S. S. Damste, W. I. C. Rijpstra, G. J. A. Geenevasen, M. Strous, and Jetten M. S. M. Structural identification of ladderane and other membrane lipids of planctomycetes capable of anaerobic ammonium oxidation (anammox). *FEBS J*, 272:4270–4283., 2005.

[104] V. Mascitti and E. J. Corey. Enantioselective synthesis of pentacycloanammoxic acid. *J. Am. Chem. Soc.*, 128:3118–3119, 2006.

[105] J. S. S. Damste to E. J. Corey cited by Mascitti and Corey. *In a personal communication*, (19), 2006.

[106] D. A. Lightner and J. E. Gurst. Organic conformational analysis and stereochemistry from circular dichroism spectroscopy. *Wiley-VCH*, 2000.

[107] C. Djerassi. Optical rotatory dispersion:applications to organic chemistry. *McGraw-Hill: New York*, 1960.

[108] P. J. Stephens, F. Devlin, and J. R. Cheeseman. VCD spectroscopy for organic chemists. *CRC Press*, 2012.

[109] L. A. Nafie. Vibrational optical activity: Principles and applications. *John Wiley and Sons: New York*, 2011.

[110] L.D. Barron. Molecular light scattering and optical activity; 2nd ed. *Cambridge University Press: Cambridge, UK*, 2004.

[111] T. M. Lowry. Optical rotatory power. *Dover Publications: New York*, 1964.

[112] R. K. Kondru, P. Wipf, and D. N. Beratan. Theory-assisted determination of absolute stereochemistry for complex natural products via computation of molar rotation angles. *J. Am. Chem. Soc.*, 120:2204–2205, 1998.

[113] R. K. Kondru, P. Wipf, and D. N. Beratan. Structural and conformational dependence of optical rotation angles. *J. Phys. Chem. A*, 103:6603–6611, 1999.

[114] P. L. Polavarapu. Optical rotation: Recent advances in determining the absolute configuration. *Chirality*, 14:768–781, 2002.

[115] G. Zuber, M. R. Goldsmith, D. N. Beratan, and P. Wipf. Assignment of the absolute configuration of [n]-ladderanes by TD-DFT optical rotation calculations. *Chirality*, 17:507–510, 2005.

[116] K. M. Specht, J. Nam, D. M. Ho, N. Berova, R. K. Kondru, D. N. Beratan, P. Wipf, R. A. Pascal, and D. Kahne. Determining absolute configuration in flexible molecules: A case study. *J. Am. Chem. Soc*, 123:8961–8966, 2001.

[117] P. J. Stephens, D. M. McCann, J. R. Cheeseman, and M. J. Frisch. Determination of absolute configurations of chiral molecules using ab initio time-dependent density functional theory calculations of optical rotation: how reliable are absolute configurations obtained for molecules with small rotations? *Chirality*, 17:S52–S64, 2005.

[118] B. Mennucci, J. Tomasi, R. Cammi, J. R. Cheeseman, M. J. Frisch, F. J. Devlin, S. Gabriel, and P. J. Stephens. Polarizable continuum model (PCM) calculations of solvent effects on optical rotations of chiral molecules. *J. Phys. Chem. A*, 106:6102–6113, 2002.

[119] K. Ruud and R. Zanasi. The importance of molecular vibrations: The sign change of the optical rotation of methyloxirane. *Angew. Chem. Int. Ed*, 44:3594–3596, 2005.

[120] B. C. Mort and J. Autschbach. Magnitude of zero-point vibrational corrections to optical rotation in rigid organic molecules: A time-dependent density functional study. *J. Phys. Chem. A*, 109:8617–8623, 2005.

[121] P. Lahiri, K. B. Wiberg, and P. H. Vaccaro. Intrinsic optical activity and large-amplitude displacement: Conformational flexibility in (R)-glycidyl methyl ether. *J. Phys. Chem. A*, 119:8311–8327, 2015.

[122] V. Raghavan and P. L Polavarapu. Chiroptical spectroscopic studies on soft aggregates and their interactions. in chiral analysis: Advances in spectroscopy, chromatography and emerging methods, 2nd ed. *Elsevier: Amsterdam, The Netherlands*, 2018.

[123] J. He and P. L. Polavarapu. Determination of intermolecular hydrogen bonded conformers of α-aryloxypropanoic acids using density functional theory predictions of vibrational absorption and vibrational circular dichroism spectra. *J. Chem. Theory Comput*, 1:506–514, 2005.

[124] J. He and P. L. Polavarapu. Determination of the absolute configuration of chiral α-aryloxypropanoic acids using vibrational circular dichroism studies: 2-(2-chlorophenoxy) propanoic acid and 2-(3-chlorophenoxy) propanoic acid. *Spectrochim. Acta, Part A*, 61:1327 – 1334, 2005.

[125] J. He, F. Wang, and P. L. Polavarapu. Absolute configurations of chiral herbicides determined from vibrational circular dichroism. *Chirality*, 17:S1–S8, 2005.

[126] J.L. Johnson, D.S. Nair, S.M. Pillai, D. Johnson, Z. Kallingathodi, I. Ibnusaud, and P.L. Polavarapu. Dissymmetry factor spectral analysis can provide useful diastereomer discrimination: Chiral molecular structure of an analogue of (-)-crispine A. *ACS Omega*, 4(4):6154–6164, 2019.

[127] A. Kapat, P. S. Kumar, and S. Baskaran. Synthesis of crispine a analogues via an intramolecular schmidt reaction. *Beilstein J. Org. Chem.*, 3:49, 2007.

[128] F. D. King. A facile three-step synthesis of ($\pm$)-crispine a via an acyliminium ion cyclisation. *Tetrahedron*, 63:2053–2056, 2007.

[129] K. R. Bailey, A. J. Ellis, R. Reiss, T. J. Snape, and N. J. Turner. A template-based mnemonic for monoamine oxidase (mao-n) catalyzed reactions and its application to the chemo-enzymatic deracemisation of the alkaloid ($\pm$)-crispine a. *Chem. Commun.*, 35:3640–3642, 2007.

[130] J. Szawkało, A. Zawadzka, K. Wojtasiewicz, A. Leniewski, J. Drabowicz, and Z. Czarnocki. First enantioselective synthesis of the antitumour alkaloid (+)-crispine a and determination of its enantiomeric purity by 1H NMR. *Tetrahedron: Asymmetry*, 16:3619–3621, 2005.

[131] H.-J. Knölker and S. Agarwal. Total synthesis of the antitumor active pyrrolo[2,1-a]isoquinoline alkaloid ($\pm$)-crispine A. *Tetrahedron Lett.*, 46:1173–1175, 2005.

[132] P. S. Kumar, A. Kapat, and S. Baskaran. An intramolecular schmidt reaction strategy for the synthesis of a methyl analogue of crispine A. *Tetrahedron Lett.*, 49:1241–1243, 2008.

[133] C. P. Prasanth, E. Joseph, A. Abhijith, D. S. Nair, I. Ibnusaud, J. Raskatov, and B. Singaram. Stabilization of $NaBH_4$ in methanol using a catalytic amount of naome. reduction of esters and lactones at room temperature without solvent-induced loss of hydride. *J. Org. Chem.*, 83:1431–1440, 2018.

[134] Q. Zhang, G. Tu, Y. Zhao, and T. Cheng. Novel bioactive isoquinoline alkaloids from carduus crispus. *Tetrahedron*, 58:6795–6798, 2002.

[135] F. Yuste, R. Sańchez-Obregoń, E. Díaz, and M.A. García-Carrillo. Enantiodifferentiation of the antitumor alkaloid crispine A using the NMR chiral solvating agents (R)- and (S)-binol. *Tetrahedron: Asymmetry*, 25:224–228, 2014.

[136] K. C. Nicolaou and S. A. Snyder. Chasing molecules that were never there: Misassigned natural products and the role of chemical synthesis in modern structure elucidation. *Angew. Chem., Int. Ed.*, 44:1012–1044, 2005.

[137] M. E. Maier. Structural revisions of natural products by total synthesis. *Nat. Prod. Rep.*, 26:1105–1124, 2009.

[138] T. L. Suyama, W. H. Gerwick, and K. L. McPhail. Survey of marine natural product structure revisions: A synergy of spectroscopy and chemical synthesis. *Bioorg. Med. Chem.*, 19:6675–6701, 2011.

[139] L. A. Nafie and R. K. Dukor. Vibrational optical activity. *Chiral Analysis, 2nd ed*, Elsevier:Ch5: 201–247, 2018.

[140] S. Ostovar pour, L. D. Barron, S. T. Mutter, and E. W. Blanch. Raman optical activity. *Chiral Analysis, 2nd ed*, Elsevier:Ch6: 249–291, 2018.

[141] P. L. Polavarapu. Why is it important to simultaneously use more than one chiroptical spectroscopic method for determining the structures of chiral molecules? *Chirality*, 20:664–672, 2008.

[142] A. D. Buckingham, P. W. Fowler, and P. A. Galwas. Velocity-dependent property surfaces and the theory of vibrational circular dichroism. *Chem. Phys.*, 112:1–14, 1987.

[143] P. A. Galwas. On the distribution of optical polarization in molecules. *Ph.D. Thesis*, Cambridge University: Cambridge, UK, 1983.

[144] P. J. Stephens. Theory of vibrational circular dichroism. *J. Phys. Chem.*, 89:748–752, 1985.

[145] DALTON:a molecular electronic structure program. *http://daltonprogram.org*, 2015.

[146] PSI4: A open-source suite of ab initio quantum chemistry programs. *www.psicode.org/*, 2014.

[147] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.*, 132(No. 154104), 2010.

[148] A. Padwa, T. M. Heidelbaugh, J. T. Kuethe, and M. S. McClure. Heterocyclic synthesis via the tandem thionium/N-acyliminium ion cascade. *J. Org. Chem.*, 63:6778–6779, 1998.

[149] T. Luker, W.-J. Koot, H. Hiemstra, and W. N. Michael Speckamp. Additions to (R)-1-acetyl-5-isopropoxy-3-pyrrolin-2-one and subsequent N-acyliminium ion generation: Synthesis of enantiopure 1-azabicycles and preparation of an intermediate for a projected synthesis of roseophilin. *J. Org. Chem.*, 63:220–221, 1998.

[150] B. P. Wijnberg and W. N. Speckamp. Diastereoselective cyclisations of chiral α-acyliminium ions. *Tetrahedron Lett.*, 21:1987–1990, 1980.

[151] V. Prelog and G. Helmchen. Basic principles of the CIP-system and proposals for a revision. *Angew. Chem., Int. Ed.*, 21:567– 583, 1982.

[152] S. R. Domingos, C. Peŕez, and M. Schnell. Sensing chirality with rotational spectroscopy. *Annu. Rev. Phys. Chem.*, 69:499–519, 2018.

[153] B. H. Pate, L. Evangelisti, W. Caminati, Y. Xu, J. Thomas, D. Patterson, C. Perez, and M. Schnell. Quantitative chiral analysis by molecular rotational spectroscopy. *Chiral Analysis, 2nd ed.*, pages Ch 17: 679–729, 2018.

[154] P. L. Polavarapu. Determination of the structures of chiral natural products using vibrational circular dichroism. in comprehensive chiroptical spectroscopy. *John Wiley & sons: Hoboken, New Jersey*, Vol. 2., 2012.

[155] A. Klamt and G. Schuurmann. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc. Perkin Trans. 2*, pages 799–805, 1993.

[156] A.S.Perera, J.Thomas, M.R.Poopari, and Y.Xu. The clusters-in-a-liquid approach for solvation: new insights from the conformer specific gas phase spectroscopy and vibrational optical activity spectroscopy. *Front. Chem. 4*, 9:1–17, 2016.

[157] T.Giovannini, G.DelFrate, P.Lafiosca, and C.Cappelli. Effective computational route towards vibrational optical activity spectra of chiral molecules in aqueous solution. *Phys. Chem. Chem. Phys.*, 20:9181–9197, 2018.

[158] T. Giovannini, M. Olszòwka, and C. Cappelli. Effective fully polarizable QM/MM approach to model vibrational circular dichroism spectra of systems in aqueous solution. *J. Chem. Theory Comput. 12*, pages 5483–5492, 2016.

[159] V. Andrushchenko, L. Benda, O. Páv, M. Dračínský, and P. Bouř. Vibrational properties of the phosphate group investigated by molecular dynamics and density functional theory. *J. Phys. Chem. B 119*, pages 10682–10692.

[160] V. Andrushchenko, D. Tsankov, M. Krasteva, H. Wieser, and P. Bouř. Spectroscopic detection of DNA quadruplexes by vibrational circular dichroism. *J. Am. Chem. Soc. 133*, pages 15055–15064, 2011.

[161] P.Bouř, V.Andrushchenko, M.Kabeláč, V.Maharaj, and H.Wieser. Simulations of structure and vibrational spectra of deoxyoctanucleotides. *J. Phys. Chem. B 109*, pages 20579–20587, 2005.

[162] J.L. Johnson and P.L. Polavarapu. Chiral molecular structures of substituted indans: ring puckering, rotatable substituents, and vibrational circular dichroism. *ACS Omega 4*, pages 4963–4976, 2019.

[163] K. Le Barbu-Debus, A. Scherrer, A. Bouchet, D. Sebastiani, R. Vuilleumier, and A. Zehnacker. Effect of puckering motion and hydrogen bond formation on the vibrational circular dichroism spectrum of a flexible molecule: the case of (S)-1-indanol. *Phys. Chem. Chem. Phys. 20*, pages 14635–14646, 2018.

[164] P.L. Polavarapu, E.A. Donahue, K.C. Hammer, V. Raghavan, G. Shanmugam, I. Ibnusaud, D.S. Nair, C. Gopinath, and D. Habel. Chiroptical spectroscopy of natural products: avoiding the aggregation effects of chiral carboxylic acids. *J. Nat. Prod*, 75:1441–1450, 2012.

[165] S.Ghidinelli, S., Abbate, J.Koshoubu, Y.Araki, T.Wada, and G.Longhi. Solvent effects and aggregation phenomena studied by vibrational optical activity and molecular dynamics: the case of pantolactone. *J. Phys. Chem. B*, 124:4512–4526, 2020.

[166] K. Bünnemann and C. Merten. Solvation of a chiral carboxylic acid: effects of hydrogen bonding on the IR and VCD spectra of α-methoxyphenylacetic acid. *Phys. Chem. Chem. Phys.*, 19:18948–18956, 2017.

[167] L.Weirich and C.Merten. Solvation and self-aggregation of chiral alcohols:how hydrogen bonding affects their VCD spectral signatures. *Phys. Chem. Chem. Phys.*, 21:13494–13503, 2019.

[168] L.Weirich, K.Blanke, and C.Merten. More complex, less complicated? explicit solvation of hydroxyl groups for the analysis of VCD spectra. *Phys. Chem. Chem. Phys.*, 22:12515–12523, 2020.

[169] L.Weirich, J. Magalhães de Oliveira, and C. Merten. How many solvent molecules are required to solvate chiral 1,2-diols with hydrogen bonding solvents? a VCD spectroscopic study. *Phys. Chem. Chem. Phys*, 22:1525–1533, 2020.

[170] V. Andrushchenko, H. Wieser, and P. Bouř. Rna structural forms studied by vibrational circular dichroism: ab initio interpretation of the spectra. *J. Phys. Chem. B*, 108:3899–3911, 2004.

[171] C.L. Covington, F.M.S. Junior, J.H.S. Silva, R.M. Kuster, M.B. de Amorim, and P.L. Polavarapu. Atropoisomerism in biflavones: the absolute configuration of (-)- agathisflavone via chiroptical spectroscopy. *J. Nat. Prod.*, 79:2530–2537, 2016.

[172] D.P. Demarque, D.R. Pinho, N.P. Lopes, and C. Merten. Revisiting empirical rules for the determination of the absolute configuration of cascarosides and other (ox-) anthrones. *Chirality*, 30:432–438, 2018.

[173] T.Itoh, K.Nagata, M.Yokoya, M.Miyazaki, K.Kameoka, S.Nakamura, and A.Ohsawa. The synthesis of isoquinoline alkaloid and its related compounds using alanine derivatives as chiral auxiliaries. *Chem. Pharm. Bull.*, 51:951–955, 2003.

[174] Y.S. Lee, D.W. Kang, S.J. Lee, and H. Park. Asymmetric synthesis of both enantiomers of pyrrolidinoisoquinoline derivatives from l-malic acid and l-tartaric acid. *J. Org. Chem.*, 60:7149–7152, 1995.

[175] S. Okamoto, X. Teng, S. Fujii, Y. Takayama, and F. Sato. An allyltitanium derived from acrolein 1,2-dicyclohexylethylene acetal and (2-propene)ti(o-i-pr)2 as a chiral propionaldehyde homoenolate equivalent that reacts with imines with excellent stereoselectivity. an efficient and practical access to optically active -amino carbonyl compounds. *Journal of the American Chemical Society*, 123:3462–3471, 2001.

[176] W.H. Pearson and W. k. Fang. Synthesis of benzo-fused 1-azabicyclo[m.n.0]alkanes via the schmidt reaction: a formal synthesis of gephyrotoxin. *J. Org. Chem.*, 65:7158–7174, 2000.

[177] D. Habel, D.S. Nair, Z. Kallingathodi, C. Mohan, S.M. Pillai, R.R. Nair, G. Thomas, S. Haleema, C. Gopinath, R.V. Abdul, M. Fritz, A.R. Puente, J.L. Johnson, P.L. Polavarapu, and I. Ibnusaud. Natural product-derived chiral pyrrolidine-2,5-diones, their molecular structures and conversion to pharmacologically important skeletons. *J. Nat. Prod.*, 83:2178–2190, 2020.

[178] D.A. Case, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, I. T. E. Cheatham, V.W.D. Cruzeiro, T.A. Darden, R. E.Duke, D. Ghoreishi, M. K.Gilson, H. Gohlke, A. W.Goetz, D. Greene, R. Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D.J. Mermelstein, K.M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, J. Smith, R. SalomonFerrer, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, L. Xiao, D.M. York, and P.A. Kollman. Amber. University of California, San Francisco, California, 2018.

[179] J. Wang, W. Wang, P.A. Kollman, and D.A. Case. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.*, 25:247–260, 2006.

[180] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, and D.A. Case. Development and testing of a general amber force field. *J. Comput. Chem.*, 25:1157–1174, 2004.

[181] A. Jakalian, B.L. Bush, D.B. Jack, and C.I. Bayly. Fast, efficient generation of high-quality atomic charges. am1-bcc model: I. method. *J. Comp. Chem*, 21:132–146, 2000.

[182] A. Jakalian, D.B. Jack, and C.I. Bayly. Fast, efficient generation of high-quality atomic charges. am1-bcc model: II. parameterization and validation. *J. Comput. Chem.*, 23:1623–1641, 2002.

[183] D.R.Roe and T.E.Cheatham. Ptraj and cpptraj:software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.*, 9:3084–3095, 2013.

[184] P.L. Polavarapu. Determination of the structures of chiral natural products using vibrational circular dichroism. *Comprehensive Chiroptical Spectroscopy*, John Wiley & Sons, New York:387–420, 2012.

[185] N. Berova, K. Nakanishi, and R. W Woody. Circular dichroism: Principles and applications. Wiley-VCH:877, 2000.

[186] S. D. Allen and O Schnepp. Circular dichroism of an optically active benzene chromophore-1-methylindan. *J. Chem. Phys*, 59:4547–4556, 1973.

[187] H. E. Smith, B. G. Padilla, J. R. Neergaard, and F.-M. Chen. Optically active amines. 25. circular dichroism of 1-substituted indans. *J. Am. Chem. Soc*, 59:6035–6039, 1978.

[188] D. Scuderi, A. Paladini, M. Satta, D. Catone, S. Piccirillo, M. Speranza, and A. G Guidoni. Chiral aggregates of indan-1-ol with secondary alcohols and water: Laser spectroscopy in supersonic beams. *Phys. Chem. Chem. Phys*, pages 4999–5003, 2002.

[189] K. L. Barbu-Debus, F. Lahmani, A. Zehnacker-Rentien, and N Guchhait. Laser-induced fluorescence and single vibronic level emission spectroscopy of chiral (R)-1-aminoindan and some of its clusters in a supersonic jet. *Phys. Chem. Chem. Phys*, 8:1001–1006, 2006.

[190] K. Le Barbu, V. Brenner, P. Millie, F. Lahmani, and A Zehnacker-Rentien. An experimental and theoretical study of jet-cooled complexes of chiral molecules: The role of dispersive forces in chiral discrimination. *J. Phys. Chem. A*, 102:128–137, 1998.

[191] K. Le Barbu-Debusa, F. Lahmani, A. Zehnacker-Rentien, N. Guchhait, S. S. Panja, and T Chakraborty. Fluorescence spectroscopy of jet-cooled chiral (±)-indan-1-ol and its cluster with (±)-methyland ethyl-lactate. *J. Chem. Phys*, 125:174305, 2006.

[192] T. Isozaki, H. Iga, T. Suzuki, and T Ichimura. Low-frequency vibrations specific for conformers of 1-aminoindan studied by uv-uv hole-burning spectroscopy. *J. Chem. Phys*, 126:214304, 2007.

[193] A. Bouchet, J. Altnoder, M. Broquier, and A Zehnacker. Ir-uv spectroscopy of jet-cooled 1-indanol: Restriction of the conformational space by hydration. *J. Mol. Struc*, 1076:344–351, 2014.

[194] L. P. Fontana, T. Chandramouly, H. E. Smith, and P. L Polavarapu. Vibrational circular dichroism and absolute configuration of 1-substituted indans. *J. Org. Chem*, 53:3381–3383, 1988.

[195] P. A Galwas. *On the Distribution of Optical Polarization in Molecules. Ph.D. Thesis*, Cambridge University: Cambridge, UK, 1983.

[196] J. R. Cheeseman, M. J. Frisch, F. J. Devlin, and P. J Stephens. Ab initio calculation of atomic axial tensors and vibrational rotational strengths using density functional theory. *Chem. Phys. Lett*, 252:211–220, 1996.

[197] F. Egidi, T. Giovannini, G. Del Frate, P. M. Lemler, P. H. Vaccaro, and C Cappelli. A combined experimental and theoretical study of optical rotatory dispersion for (R)-glycidyl methyl ether in aqueous solution. *Phys. Chem. Chem. Phys*, pages 3644–3655, 2019.

[198] S. Ghidinelli, S. Abbate, S. E. Boiadjiev, D. A. Lightner, and G Longhi. l-stercobilin-hcl and d-urobilin-hcl. analysis of their chiroptical and conformational properties by VCD, ECD, and CPL experiments and MD and DFT calculations. *J. Phys. Chem. B*, 122:12351–12362, 2018.

[199] E. Debie, E. De Gussem, R. K. Dukor, W. Herrebout, L. A. Nafie, and P Bultinck. A confidence level algorithm for the determination of absolute configuration using vibrational circular dichroism or Raman optical activity. *ChemPhysChem*, 12:1542–1549, 2011.

[200] A Rohatgi. WebPlotDigitizer. *https://automeris.io/WebPlotDigitizer*, 2018.

[201] A. Scherrer, R. Vuilleumier, and D Sebastiani. Vibrational circular dichroism from ab initio molecular dynamics and nuclear velocity perturbation theory in the liquid phase. *J. Chem. Phys*, 145:084101, 2016.

[202] A. Scherrer, R. Vuilleumier, and D Sebastiani. Nuclear velocity perturbation theory of vibrational circular dichroism. *J. Chem. Theory Comput*, 9:5305–5312, 2013.

[203] A. Scherrer, F. Agostini, D. Sebastiani, E. K. U. Gross, and R Vuilleumier. Nuclear velocity perturbation theory for vibrational circular dichroism: An approach based on the exact factorization of the electron-nuclear wave function. *J. Chem. Phys*, 143:074106, 2015.

[204] L. A Nafie. Adiabatic molecular properties beyond the Born-Oppenheimer approximation. complete adiabatic wave functions and vibrationally induced electronic current density. *J. Chem. Phys*, 79:4950–4957, 1983.

[205] Y. Xia, M. A. J. Koenis, J. F. Collados, P. Ortiz, S. R. Harutyunyan, L. Visscher, W. J. Buma, and V. P Nicu. Regional susceptibility in VCD spectra to dynamic molecular motions: The case of a benzyl α-hydroxysilane. *ChemPhysChem*, 19:561–565, 2018.

[206] J.L. Johnson, G. Zajac, M. Baranska, and P.L. Polavarapu. Vibrational Raman optical activity of camphor: The importance of electric-dipole—electric-quadrupole polarizability contribution. *J. Raman Spectrosc*, 51:669–679, 2020.

[207] L.D. Barron and B.P. Clark. Raman optical activity of menthol and related molecules. *Chem. Soc. Perkins II*, pages 1164–1170, 1979.

[208] S. Yamamato and P. Bour. Transition polarizability model of induced resonance Raman optical activity. *J. Comp. Chem*, 34:2152, 2013.

[209] L. D. Barron, J. R. Escribano, and J. F. Torrance. Polarized Raman optical activity and the bond polarizability model. *Mol. Phys*, 57:653–660, 1986.

[210] L.D. Barron and P.L. Polavarapu. Raman optical activity in ring deformations of methyloxirane, methylthiirane and trans-2,3-dimethyloxirane. *Mol. Phys*, 65:659–667, 1988.

[211] T. M. Black, P. K. Bose, P. L. Polavarapu, L. D. Barron, and L. Hecht. Vibrational optical activity in trans-2,3-dimethyloxirane. *J. Am. Chem. Soc*, 112(4):1479–1489, 1990.

[212] L. D. Barron and J. Vrbancich. Methyl torsion Raman optical activity in (2S,3S)-(-)2,3-epoxybutane. *Mol.Phys*, 48(4):833–845, 1983.

[213] P. L. Polavarapu, L. Hecht, and L. D. Barron. Vibrational Raman optical activity in substituted oxiranes. *J. Phys. Chem*, 97(9):1793–1799, 1993.

[214] L. D. Barron, L. Hecht, and S. M. Blyth. Polarized Raman optical activity of menthol and related molecule. *Spectrochim. Acta A*, 45:375–379, 1989.

[215] D. Che and L. A. Nafie. Isolation of Raman optical activity invariants. *Chem. Phys. Lett*, 189(1):35–42, 1992.

[216] L. D. Barron and A. D. Buckingham. Simple two-group model for rayleigh and Raman optical activity. *J. Am. Chem. Soc*, 96(15):4769–4773, 1974.

[217] J. R. Escribano and L. D. Barron. Valence optical theory of vibrational circular dichroism and Raman optical activity. *Mol. Phys*, 65:327–344, 1988.

[218] L. Hecht. The isolation of Raman optical activity invariants. *Chem. Phys. Lett*, 195:518–522, 1992.

[219] G. Yu, T. B. Freedman, and L. A. Nafie. Dual circular polarization Raman optical activity of related terpene molecules: Comparison of backscattering DCPI and right-angle ICP spectra. *J. Raman Spectrosc*, 26:733–743, 1995.

[220] L. D. Barron and J. R. Escribano. Polarized Raman optical activity of β-pinene: electric quadrupole contributions. *Chem. Phys. Lett*, 126(5):461–464, 1986.

[221] L. D. Barron, L. Hecht, A. R. Gargaro, and W. Hug. Vibrational Raman optical activity in forward scattering: Trans-pinane and β-pinene. *J. Raman Spectrosc*, 21:375–379, 1990.

[222] V. Liégeois, O. Quinet, and B.Champagne. Vibrational Raman optical activity as a mean for revealing the helicity of oligosilanes: A quantum chemical investigation. *J. Chem. Phys*, 122(21):214304, 2005.

[223] G. Zuber, M. Goldsmith, D. N. Beratan, and P. Wipf. Towards Raman optical activity calculations of large molecules. *ChemPhysChem*, 6:595–597, 2005.

[224] G. Zuber and W. Hug. Rarefied basis sets for the calculation of optical tensors. 1. the importance of gradients on hydrogen atoms for the Raman scattering tensor. *J. Phys. Chem. A*, 108(11):2108–2118, 2004.

[225] A. J. Thorvaldsen, B. Gao, K. Ruud, M. Fedorovsky, G. Zuber, and W. Hug. Efficient calculation of ROA tensors with analytical gradients and fragmentation. *Chirality*, 24:1018–1030, 2012.

[226] S. Luber, C. Herrmann, and M. Reiher. Relevance of the electric-dipole-electric-quadrupole contribution to Raman optical activity spectra. *J. Phys. Chem. B*, 112(7):2218–2232, 2008.

[227] P. L. Polavarapu, C. L. Covington, K. Chruszcz-Lipska, G. Zajac, and M. Baranska. Vibrational Raman optical activity of bicyclic terpenes: Comparison between experimental and calculated vibrational Raman, Raman optical activity and dimensionless circular intensity difference spectra and their similarity analysis. *J. Raman Spectrosc*, 48:305, 2017.

[228] M. McCarthy and K.L.K. Lee. Molecule identification with rotational spectroscopy and probabilistic deep learning. *J. Phys. Chem. A*, 124:3002–3017, 2020.

[229] Z. Wang, X. Feng, J. Liu, M. Lu, and M. Li. Functional groups prediction from infrared spectra based on computer-assist approaches. *Microchem. J*, 159:105395, 2020.

[230] C. G. Jones, M. W. Martynowycz, J. Hattne, T. J. Fulton, B. M. Stoltz, J. A. Rodriguez, H. M. Nelson, and T. Gonen. The cryoem method microed as a powerful tool for small molecule structure determination. *ACS Cent. Sci*, 4:1587–1592, 2018.

[231] F. Wang and P. L. Polavarapu. Conformational stability of (+)-epichlorohydrin. *J. Phys. Chem. A*, 104:6189–6196, 2000.

[232] B. Nieto-Ortega, J. Casado, J. T. López Navarrete, F. J. Ramírez, M. Tommasini, G. Longhi, G. Mazzeo, and S. Abbate. Understanding the origin of the VCD signals on the basis of a nonredundant coordinate definition. *J. Chem. Theory Comput*, 11:2633–2641, 2015.

[233] M. Tommasini, G. Longhi, G. Mazzeo, S. Abbate, B. Nieto-Ortega, F. J. Ramírez, J. Casado, and J. T. López Navarrete. Mode robustness in Raman optical activity. *J. Chem. Theory Comput*, 10:5520–5527, 2014.

[234] J. He, A. G. Petrovic, and P. L. Polavarapu. Determining the conformer populations of (R)-(+)-3-methylcyclopentanone using vibrational absorption, vibrational circular dichroism, and specific rotation. *J. Phys. Chem. B*, 108:20451–20457, 2004.

[235] P. L. Polavarapu, P. K. Bose, L. Hecht, and L. D. Barron. Vibrational Raman optical activity in (R)-(+)-3-methylcyclopentanone: experimental and ab initio theoretical studies. *J. Phys. Chem. A*, 97:11211–11215, 1993.

[236] S. K. Singh and R. I. Kaiser. A vacuum ultraviolet photoionization study on the isomerization, decomposition, and molecular mass growth processes in solid nitromethane (CH_3NO_2). *Chem. Phys. Lett*, 766:138343, 2021.

[237] S. K. Singh, T.-Y. Tsai, B.-J. Sun, A. H. H. Chang, A. M. Mebel, and R. I. Kaiser. Gas phase identification of the elusive N-hydroxyoxaziridine (c-H_2CON(OH)): A chiral molecule. *J. Phys. Chem. Lett*, 11:5383–5389, 2020.

[238] J. E. Peironcely, M. Rojas-Chertó, D. Fichera, T. Reijmers, L. Coulier, J.-L. Faulon, and T. Hankemeier. OMG: Open molecule generator. *J. Cheminf*, 4(21), 2012.

[239] R. Gugisch, A. Kerber, A. Kohnert, R. Laue, M. Meringer, C. Rucker, and Wassermann A. MOLGEN 5.0, a molecular structure generator. in advances in mathematical chemistry and applications. *Advances in Mathematical Chemistry and Applications*, pages Ch6: 113–138, 2015.

[240] T. A. Halgren. MMFF VI. MMFF94s option for energy minimization studies. *J. Comput. Chem*, 20:720–729, 1999.

[241] A. D. McLean and G. S. Chandler. Contracted gaussian basis sets for molecular calculations. I. second row atoms, z = 11-18. *J. Chem. Phys*, 72:5639–5648, 1980.

[242] R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople. Self-consistent molecular orbital methods. a basis set for correlated wave functions. *J. Chem. Phys*, 72:650–654, 1980.

[243] M. J. Frisch, J. A. Pople, and J. S. Binkley. Self-consistent molecular orbital methods 25. supplementary functions for gaussian basis sets. *J. Chem. Phys*, 80:3265–3269, 1984.

[244] M. Brehm and M. Thomas. Computing bulk phase Raman optical activity spectra from ab initio molecular dynamics simulations. *J. Phys. Chem. Lett*, 8:3409–3414, 2017.

[245] L. Hung and E. A. Carter. Accurate simulations of metals at the mesoscale: Explicit treatment of 1 million atoms with quantum mechanics. *Chem. Phys. Lett.*, 475(4-6):163-170, 2009.

[246] D. R. Bowler and T. Miyazaki. Calculations for millions of atoms with density functional theory: linear scaling shows its potential. *J. Phys. Condens. Matter*, 22:074207, 2010.

[247] U. Schnupf and F. A. Momany. DFT energy optimization of a large carbohydrate: Cyclomaltohexaicosaose (ca-26). *J. Phys. Chem. B*, 116(23):6618–6627, 2012.