Conformational Study of Amino Acid Type Molecules in Solution Using Raman Optical Activity

by

Joseph Cheramy

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Chemistry
University of Alberta

Abstract

Raman and Raman optical activity (ROA) spectroscopy are a set of especially effective spectroscopic practices for measuring the vibrational transitions of molecules in solution. They are particularly well suited for measurements of aqueous solutions because of the relatively little Raman scattering produced by water. This is useful for biologically relevant molecules, as such systems exist by and large in aqueous environments.

Through the use of quantum mechanical calculations, Raman and ROA spectra can be determined for given molecules, and even molecules solvated explicitly. This explicit solvation is key in understanding the conformational changes a molecule undergoes when dissolved in a solvent. Certain solute-solvent complexes form and it is proposed that the lifetimes of some are longer than others. These longer-lived complexes likely contribute more to the observed experimental Raman and ROA spectral features.

The following chapters detail work to determine what the solute-solvent complexes are for the amino acid type molecules N-acetyl-L-cysteine and serine through combined theoretical and experimental method. Sets of solute-solvent complexes are constructed and their respective spectra are fit and compared to those produced experimentally. In this way, the conformational changes and interactions between the solvent molecules and target molecule can be determined in the hope of better understanding how the molecules behave in solution.

Preface

This thesis is based on research I performed at the University of Alberta between September 2014 and January 2017. I was responsible for data collection, analysis, and manuscript writing for Chapters 2 and 3. Dr. M. Reza Poopari was responsible for some data collection, theoretical calculation, and guidance for the work presented in Chapter 2. Professor Dr. Yunjie Xu was the supervisory author and involved in concept formation, data analysis, and manuscript composition for all the presented work.

Acknowledgments

I would first like to thank my advisor Prof. Yunjie Xu. It is because of her that I was able to do graduate research at all. Her guidance was always great and I learned more than I could have imagined from her. I'm not sure how she has time to do everything that she does. Secondly, I would like to thank all the members of the Xu and Jäger groups: Zahra, Reza, Angelo, Jason, Amin, Supriya, Javix, Chrissy, Jiao, Elijah, Prasanta, Nate, Leo, Yanqing, David, and everyone else. Everyone was always very helpful and ready to discuss any problems. I would also like to thank Prof. Wolfgang Jäger for his always appreciated and insightful questions and comments.

I am glad to have had Dr. Anna Jordan as my lab coordinator for Chemistry 37X. I always enjoy our talks and I appreciate all the support you have shown me. I'll still come visit as long as I'm around.

A thanks to the members of the mass spectroscopy lab, analytical lab, and NMR lab. Everyone in these labs was always very helpful and their expertise appreciated. Thanks to the members of the CGSS for letting me have a bit of fun outside of my work.

Finally, I would of course like to thank my friends and family. You all mean a lot to me, even though I don't show it all the time. My parents and brother have been very supportive and I guess need to take some responsibility for this work.

This research was funded by the University of Alberta, the Natural Sciences and Engineering Research Council of Canada, the Canadian Foundation for Innovation, and the Alberta Enterprise and Advanced Education. I also gratefully acknowledge access to the computing facilities of the Western Canada Research Grid (Westgrid) and Compute/Calcul Canada. I acknowledge the Government of Alberta for a Queen Elizabeth II Graduate Scholarship.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Overview

The need to obtain structural information such as absolute configuration and conformation of a targeted chiral compound directly in solution has motivated countless studies and innovations in the past. Such structural information is essential in the understanding of many biological and chemical processes. In particular, the knowledge of the absolute configuration of a chiral molecule is of great importance in stereoselective syntheses,[1,2] reaction mechanisms,[3] biological enzymatic reactions, and in drug design.[4] Currently, there are a few spectroscopic techniques that can be used in identifying the chirality of a given molecule, each with their own advantages and disadvantages. X-ray crystallography has been widely used, but requires a suitable crystal. Furthermore. researchers increasingly recognize that the configuration adopted by a chiral molecular system in solution often differs from that in the solid phase.[5,6] NMR is an extremely popular and a powerful technique, but it requires often an expensive chiral shift agent and additional syntheses.

Linear absorption spectroscopies in the infrared (IR), visible, and ultra-violet regions have long been used to characterize molecules. Chiefly, IR and Raman spectroscopy can be used to determine certain molecular structural details, including the presence of functional groups and the nature of certain bonds. This is useful information, however the introduction of circularly polarized light to these specific techniques has allowed researchers to gain much more insight into the structural information of chiral molecules in solution, in film, and on metal nanoparticles.[7] Vibrational circular dichroism (VCD) measures the difference in the absorbance of left and right circularly polarized light in the mid-infrared (fingerprint) region by a chiral molecule during a vibrational transition. Researchers have been able to use VCD spectroscopy in conjunction with quantum mechanical modelling to probe properties of chiral molecules with great success.[8–10] Put simply, VCD spectroscopy has been proven to facilitate the determination of absolute configurations of chiral molecules in solution.

Raman and Raman optical activity (ROA) spectroscopy are related to the scattering phenomenon of electromagnetic radiation by molecules and are complementary to infrared and VCD spectroscopy, respectively.[11] The main advantages of ROA spectroscopy come from the fact that water absorbs infrared light over a large frequency domain in the infrared region but

does not significantly scatter light in this same region. Thus, water is routinely used as a solvent for ROA studies. Water is also a natural environment for biological samples.[12,13] Furthermore, ROA spectrometers are capable of extending the measured spectral coverage down to ~ 200 cm^{-1}, which is normally not accessible using an Fourier transform IR instrument due to the limitations imposed by the optical components such as the sample cell windows, lenses, and IR detectors.

Generally, VCD and ROA experiments provide complementary information. Therefore, a combination of VCD and ROA spectroscopy is desirable for systematic studies of chiral molecules in solution, especially for the investigation of solvent effects and for establishing a suitable model to account for them. Such combined studies have rarely been reported directly in the past.[14,15] On the other hand, these combined studies can provide further insight and more stringent tests for any solvation models proposed.

ROA phenomenon was first fully theorized in 1971 by Barron and Buckingham,[16] and the first measurements made in 1973.[17] In the early days, it was difficult to obtain good quality ROA data. Improvements were made to the instrument layout and configurations by Nafie *et al*. in 1987.[18,19] Significant innovations in the instrumentation were made by Hug and co-workers[12] where they incorporated the latest technologies available from other research areas such as back-thinned charge-coupled device (CCD) and holographic notch filters. Later one, the same research group introduced the concept of "virtual enantiomer" and realized it experimentally.[13] This latest version of the ROA instrument is the blueprint for the ROA spectrometer commercialized by BioTools nowadays.

At the same time, *ab initio* calculations for ROA tensors were only first established in 1990 by Polavarapu.[20] After this, many developments and optimizations were made with respect to the simulation of Raman and ROA experiments.[21–28] Still, there is much work to be done in experiment and theoretical calculations for ROA, much like what has been done for VCD.

In this regard, systems previously studied using VCD spectroscopy act as an excellent starting point. After a brief overview of the theory behind the ROA phenomenon and the instrumental setup (including some operational tips/observations), this text details the study of two amino acid type systems, namely serine and N-acetyl-L-cysteine, which were previously investigated using VCD spectroscopy.[29–31] The present studies benefit from the configurational information gained from the VCD studies and extends the work to take advantage of the

capabilities of ROA spectroscopy. The current work is accomplished with a combination of experimental measurements, quantum mechanical calculations, and data analysis. Each of these aspects and the way in which they synergize are detailed in Chapters 2 and 3.

1.2 Theory of ROA

Raman scattering is the inelastic scattering of monochromatic light by a molecule. For this phenomenon to occur, there must be a change in the polarizability (α) of the molecule for a given normal mode (Q_k) due to the electric field of the incident photon (E) (see Equation 1). Photons that are scattered inelastically either gain energy (Anti-Stokes) or lose energy (Stokes) when the molecule is elevated to a virtual excited state, depending on the initial and final vibrational state of the molecule. This is a weak process and generally occurs once for every 10^{10} scattered photons.

$$\mu_{ind} = \alpha_0 E_0 \left(\frac{\partial \alpha}{\partial Q_k} \right) Q_k E \tag{1}$$

To model Raman intensity, one usually only needs to consider the oscillating electric dipole (μ) of the molecule induced by the electric dipole of the exciting light. The optical tensor used to calculate these intensities is:

$$\alpha = \frac{2}{\hbar} \sum_n \frac{\omega_{gn}}{\omega_{ng}^2 - \omega^2} Re(\mu_{gn}\mu_{ng}) \tag{2}$$

where ω is the angular frequency.

ROA is defined by the difference in the intensity of scattered right and left circularly polarized light by a molecule. To model ROA intensity, which is usually about 10^3 times weaker than Raman, one needs to also consider the electric dipole-magnetic dipole optical tensor (G') and therefore the magnetic dipole (m), and the electric dipole-electric quadrupole tensor (A) and therefore the electric quadrupole (Θ):

$$A_{\alpha,\beta_\gamma} = \frac{2}{\hbar} \sum_n \frac{\omega_{gn}}{\omega_{ng}^2 - \omega^2} Re(\mu_{gn}\Theta_{\beta_\gamma,ng}) \tag{3}$$

$$G'_{\alpha,\beta} = -\frac{2}{\hbar}\sum_{n}\frac{\omega_{gn}}{\omega_{ng}^2 - \omega^2} Im(\mu_{gn} m_{\beta,ng}) \quad (4)$$

Thus, considerable work needed to be done to implement these properties to various computational methods to calculate the ROA tensors for molecules.

There are four different instrumental setup schemes possible for measuring ROA signal. In the commercial instrument we use, Raman and ROA signals are measured using a scattered circular polarization (SCP) scheme. For this type of instrument, the incident light is not circularly polarized specifically and the intensity of the scattered right and left circularly polarized light from the sample are summed to give the Raman intensity and subtracted to give the ROA intensity. This is the most common type of instrument setup due to the theoretical ratio of ROA intensity to Raman intensity (circular intensity difference (CID)) for anisotropic scattering being 0 in the forward scattering direction and maximized in the backwards scattering direction. In essence, the ROA scattering intensity is higher in a backscattering configuration relative to Raman scattering intensity than in a forward scattering configuration.

1.3 Experimental Setup and Instrumentation

Phenomenal work has been done in the development of the modern ROA instrument. The main developments have been detailed in published work,[12,13,18,32] and many more minor but incredibly helpful details are described in dissertations.[33–36] In particular the PhD thesis of Haesler contains the best overall description I have found.[37] I will begin by giving a brief summary of the main components contained in the commercial ROA instrument sold by BioTools (Florida, USA), the ChiralRAMAN-2X spectrometer (Figure 1) (which contains components that differ from previously described instruments). Following this summary, I will detail a few observations I have made during my time testing and using the instrument, including some tips for its use.

Figure 1. Exterior of ChiralRAMAN-2X instrument with a view of the interior inset.

The instrument is set up in a backscattering geometry to maximize the ROA signal produced by the sample (as diagrammed in Figure 2). To simplify things, it is easier to go through each component individually, in the order in which light passes through the system. It is first important to know that waveplates are used to modify the polarity of light passing through them. This is done by modifying the thickness and material used depending on the wavelength of light being passed to change the refractive index of the material in the x and y axes. A λ/2 waveplate retards the axis of light that is parallel to the optical axis of the waveplate. In this way, the phase of one axis of the light is offset by one half the wavelength. As a result, linearly polarized light of the correct wavelength will be converted to the opposite plane of polarization after it has passed through the waveplate. Similarly, circularly polarized light will be converted to the opposite hand. A λ/4 waveplate has the same properties in terms of construction, but the axes of the light are instead offset by one quarter the wavelength. As such, it is possible to create circularly polarized light from linearly polarized light, and linearly polarized light from circularly polarized light. Both types of waveplate are important in the function of an ROA spectrometer.

The incident light produced by the laser is inherently polarized linearly. However, after passing through several optical elements, whatever pure linear polarization that existed can be lost, including the creating of some elliptically or circularly polarization. To account for this, the instrument contains λ/2 waveplate rotators that serve the purpose of creating all orientations of linearly (and elliptically/circularly) polarized incident light. These rotators remain in the beam path during the entire experiment. On the other hand, two additional λ/2 waveplates are used as circularity converters. The first circularity converter (CC1) converts any circular or elliptical polarization that may exist to the opposite hand. The second (CC2) creates the "virtual enantiomer", as described by Hug, of the light scattered by the sample by creating circularly polarized light of the opposite hand.[13] Finally, the liquid crystal retarder (LCR) and beam splitter pass the right and left circularly polarized scattered light to fiber optic arrays that then pass light to the spectrograph. The LCR is a λ/4 waveplate that is used to alternately pass the right or left circularly polarized elements of the scattered light as p- and s-linearly polarized light to the beam splitter, which is then able to separate the two polarizations and pass each to a branch of the fiber optic array. This step is necessary to correct for any differences in the two arms of the array.

Data is collected in 16 scans per measurement cycle. This procedure allows for the correction of offset issues. Four scans are completed with neither circularity converter in the light path, two for each LCR state (to alternate the light passed to each branch of the fiber optic array). Then another four scans are completed with CC1 in the light path, followed by four with CC1 and CC2 in place, and finally four with only CC2 in place. These scans are then combined to get either a Raman spectrum or ROA spectrum and most offset issues created by imperfections of the optics are reduced.

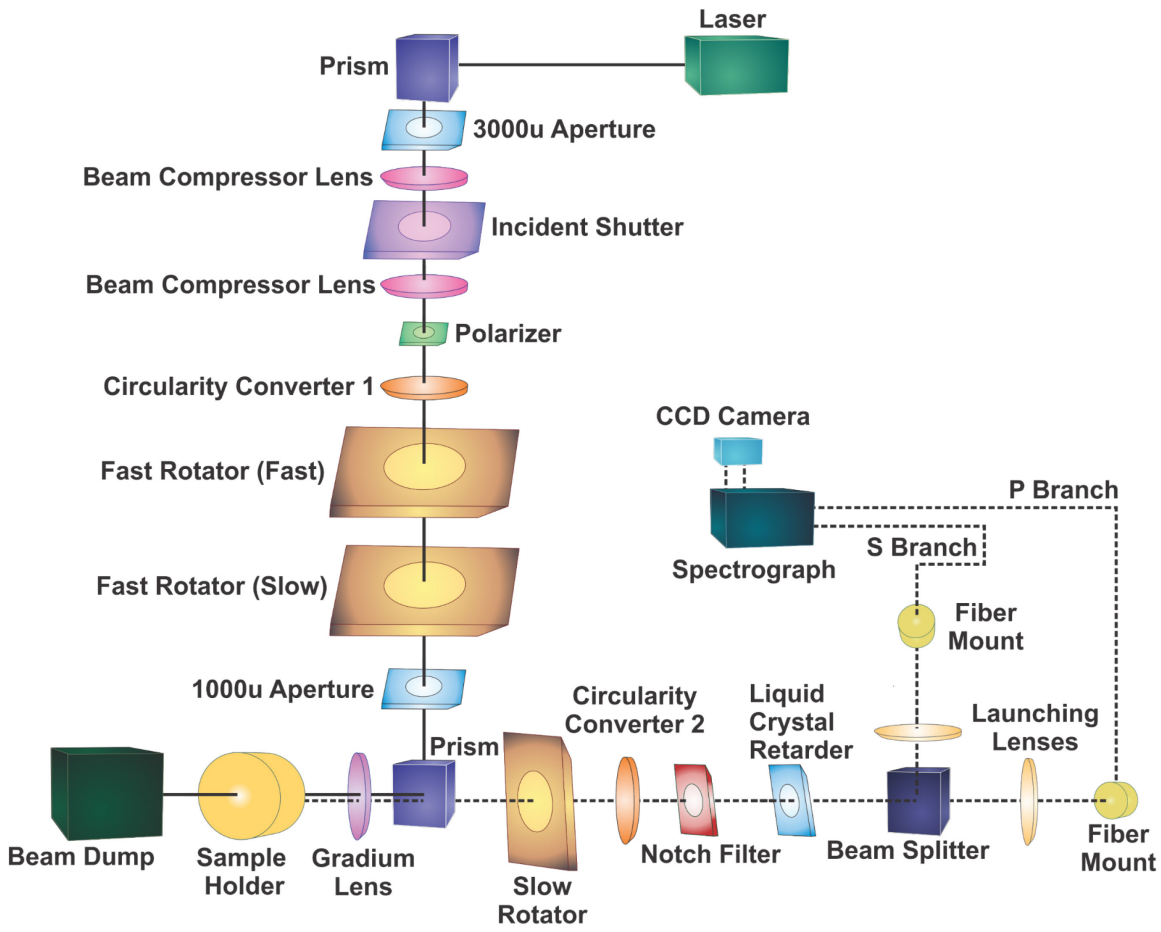

Figure 2. Schematic of the components contained in the ROA instrument. See text for more details on each component. (Figure adapted from conference presentation.)[38]

- **Laser**: A frequency doubled OPUS 532 laser manufactured by Laser Quantum. It produces continuous wave coherent light with a wavelength of 523 nm, power between ~ 0 and 2 W, and horizontal polarization. The laser system consists of the physical laser (Figure 4) and a control module (Figure 3). The laser power and on/off state are set from the control module. The laser and power supply temperatures can also be monitored here. It is possible to connect the module to a computer using the serial port on the back. This allows for control and monitoring through the provided software.

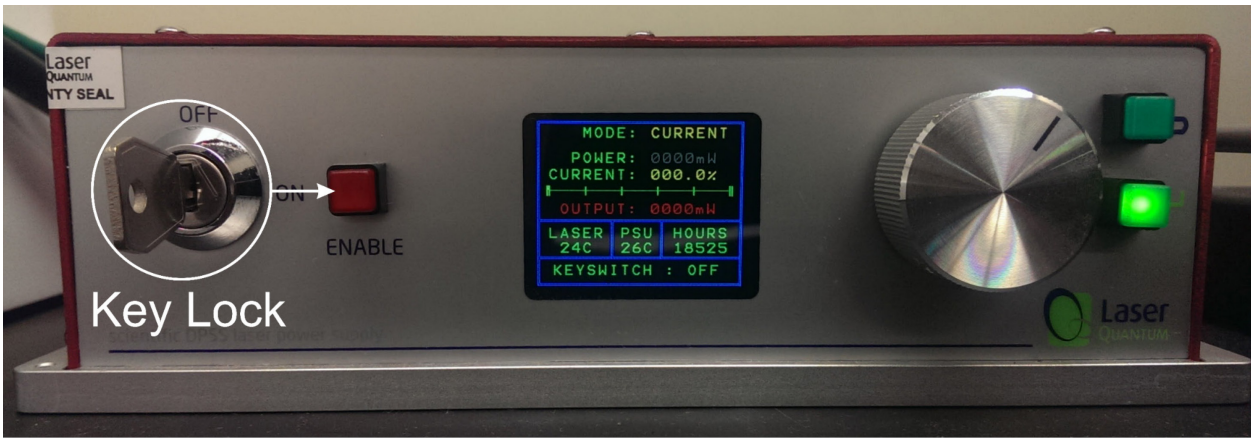

Figure 3. Laser control module.

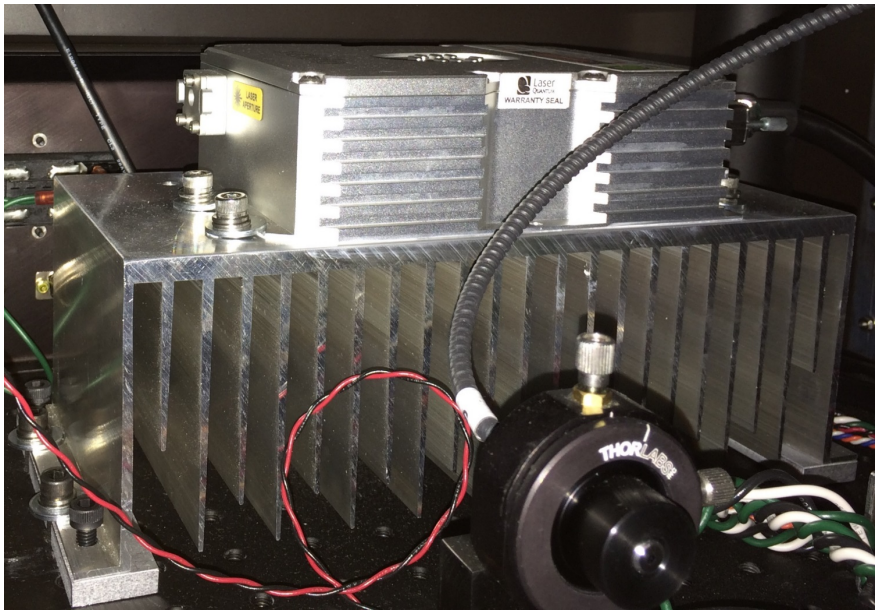

Figure 4. Laser Quantum OPUS 532 Laser inside of the ROA spectrometer, mounted to a large passive heat sink.

- **Incident Shutter**: Used to prevent incident laser light from reaching the sample. Useful to limit exposure time of less stable samples to laser.
- **Circularity Converter 1:** A λ/2 waveplate used to correct for unwanted circular polarization of the incident light. Moved in and out of the optical path.
- **Fast Linear Rotators (Fast/Slow):** A set of counter-rotating λ/2 waveplates with the purpose of creating all orientations of linearly polarized light to correct for linear polarization created by other optical elements. (Seen in Figure 5,along with above components.)

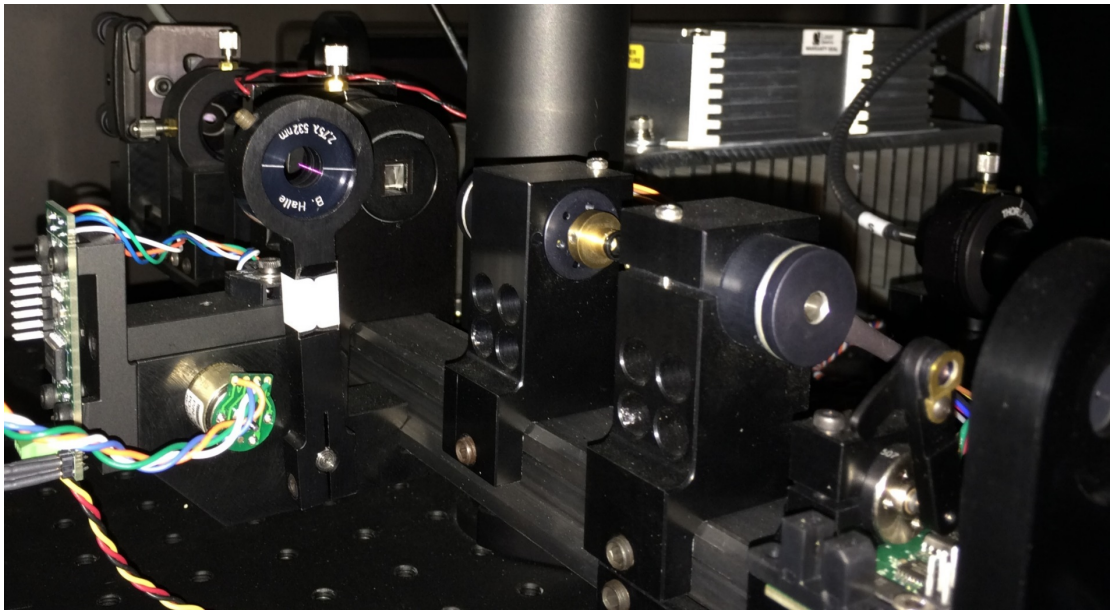

Figure 5. Components of the ROA instrument that pertain to the modulation and direction of the incident laser light. See Figure 2 for labelling of each component.

- **Prism:** A Risley prism used to divert the perpendicular incident light towards the sample cell and allow parallel scattering light to pass (Figure 6).

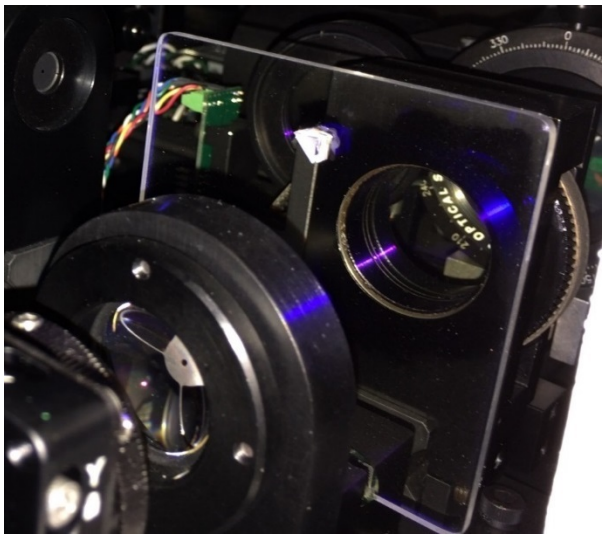

Figure 6. Prism shown attached to a transparent holding plate.

- **Sample Cell:** A mounting bracket that holds a fused quartz (silica) cuvette (or other container). The position can be manipulated in all three axes to maximize Raman signal. A quartz cuvette is used due to its wide spectral window in the infrared (Figure 7).

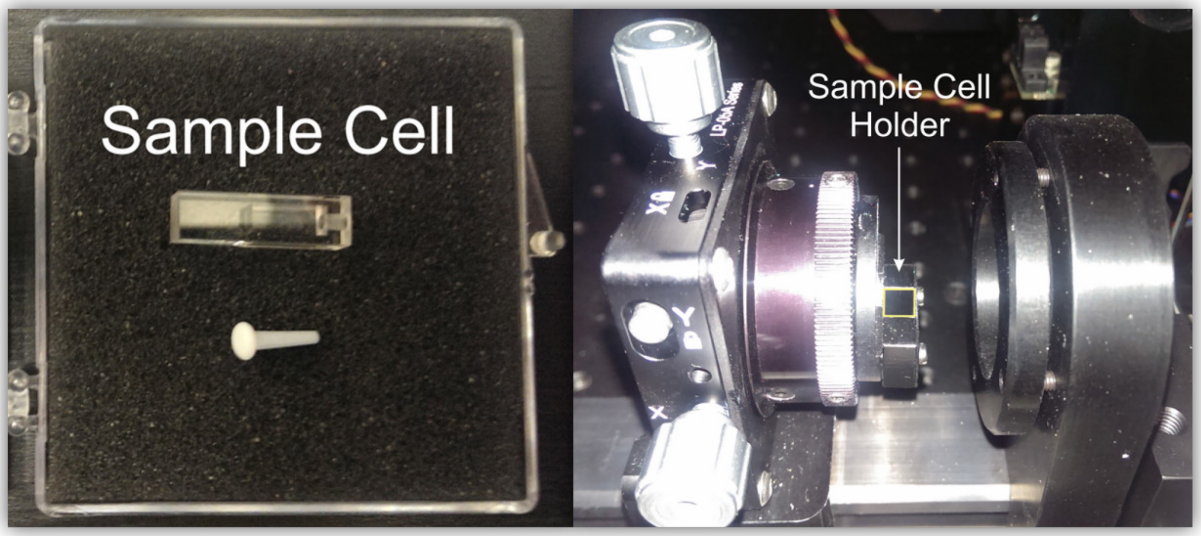

Figure 7. Quartz sample cell (left) and sample cell holder in instrument (right).

- **Slow Linear Rotator:** A rotating λ/2 waveplate that is used to correct for small amounts of linear polarization created by imperfections in the many preceding optical elements.
- **Circularity Converter 2:** A λ/2 waveplate placed in line with the scattered light to create the "virtual enantiomer" of the sample. Moved in and out of the optical path.
- **Notch Filter:** A holographic notch filter manufactured by Kaiser Optical Systems, Inc. (Michigan, USA). The notch filter is constructed from a dichromated gelatin film between two glass plates to create an interference pattern. A notch filter is necessary in ROA experiment to eliminate the Rayleigh scattering of the molecule (which is relatively strong). The advantage of this type of filter is a narrow notch bandwidth and a sharp cutoff.[39]
- **Liquid Crystal Retarder:** A λ/4 waveplate that converts the right and left circularly polarized light into either p- or s-linearly polarized light depending on the orientation of the fast and slow axes of the crystal. This orientation is alternated by electronic switching.
- **Beam Splitter:** Separates the p- and s-linearly polarized light from the liquid crystal retarder and passes each to a separate fiber optic branch.
- **Fiber Optic Branches:** A pair of fiber optic bundles that collect the scattered light on one end (Figure 8). The bundles are spread in a parabolic shape at the other end and lined up with the slit of the spectrograph.

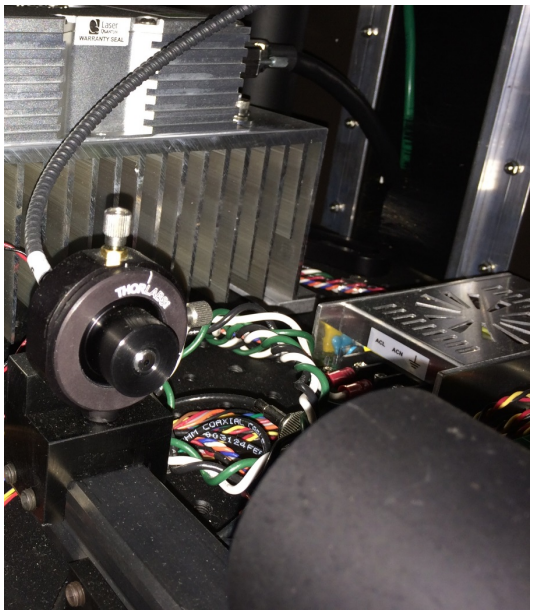

Figure 8. Launching lens, fiber mount, and fiber optic branch leading to the spectrograph.

- **Spectrograph:** A Kaiser Optical Systems Inc. HoloSpec f/1.8 spectrograph (Michigan, USA) (Figure 9). The spectrograph disperses the collected scattered light from the fiber optics and disperses it onto the CCD using a volume-phase holographic transmission grating. This is constructed from dichromatic gelatin, in a similar way to the holographic notch filter.

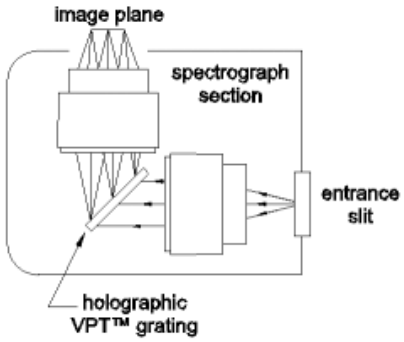

Figure 9. HoloSpec f/1.8 spectrograph. Diagram from Kaiser Optical Systems webpage.[40]

- **CCD Camera:** A back-thinned CCD camera with a 1024 x 256 pixel array (Figure 10**Figure 11**). The array is separated into two halves horizontally, each having light from one optical fiber optic branch dispersed onto it.

Figure 10. Fiber optic cable entering the spectrograph and CCD camera enclosure.

There are several issues that arose during my time using the instrument. To begin, the way in which the instrument and computer interface with each other should be known. There are two connections between the computer and instrument and therefore two possible points of failure. The CCD camera connects to the computer via an Ethernet cable and is powered via an external power supply that plugs in to the back of the instrument (Figure 11). The instrument itself connects to the computer via serial port, which is converted to USB using a serial to USB adaptor (Figure 12).

Figure 11. (Left) CCD camera bolted on to the spectrograph enclosure. Ethernet cable (green) to communicate with the computer and power cable (beige) also shown. (Right) Power for instrument (black power cable), power for CCD camera (beige, metal fitting), and serial data port.

The connection between the computer and CCD camera is very stable in general and the only problem occurs when the ROA software is manually closed. In this case, the computer and instrument must be shut off and removed from their power source. This can be accomplished by

cycling the power bar after shutting both off. Not doing so will result in no connection, which is evident by the software displaying a flat line across the Raman and ROA spectra should a scan be started.

The connection between the computer and instrument is a bit more problematic. The driver for the serial to USB adaptor often causes the computer to crash. It has been suggested that updating the version of the driver (USA19Hx64 ver. 3.7.0.5 to ver. 3.7.1) may fix the problem. However, the driver refuses to update using any installation method. It is possible that another serial to USB adaptor could be used to fix this problem.

Figure 12. Keyspan serial to USB adaptor. Connects serial port of instrument to computer.

Should the instrument or computer crash for any reason during measurement, it is possible to resume the experiment by restoring all connections and loading the last saved file for the measurement in the ROA software, then choosing not to clear all data before starting the new measurement. This loads the previous data and continues to add subsequent measurement cycles to it.

In terms of instrument setup and optimization, there are a few points to go over. The horizontal pixels of the CCD camera can be calibrated to the correct wavenumber reading using a neon lamp (Figure 13). This procedure is described in the manual, but it is helpful to secure the lamp near the sample compartment using tape, as there is no built in mounting option.

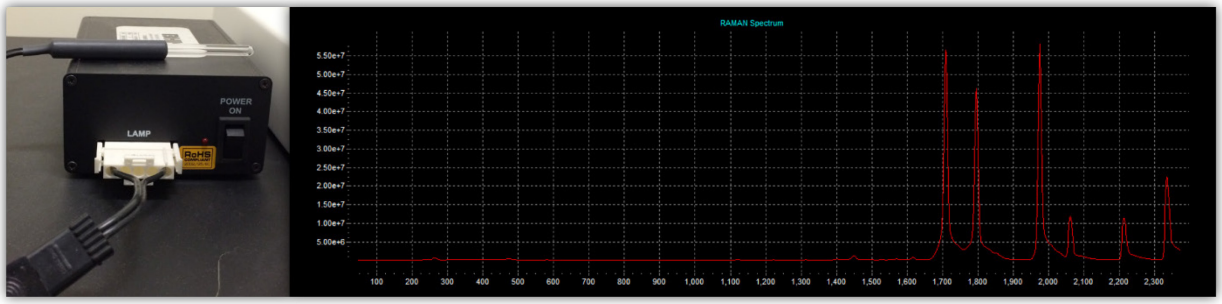

Figure 13. (Left) Neon discharge lamp used for calibrating the CCD camera. (Right) Example of the Raman spectrum produced by the neon lamp.

The focus of the laser on the sample cell is also important for effective measurements. The position of the cell can be adjusted in all three axes. The x and y axes are adjusted using thumb screws on the cell mount. This simply adjusts the section of the cell on which the incident light hits. The sample cell position in the z-axis is more important. If the laser is focused on the front of the cell, the quartz material of the cell will create Raman scattering that will show up in the spectrum. An example of this is shown in Figure 14.

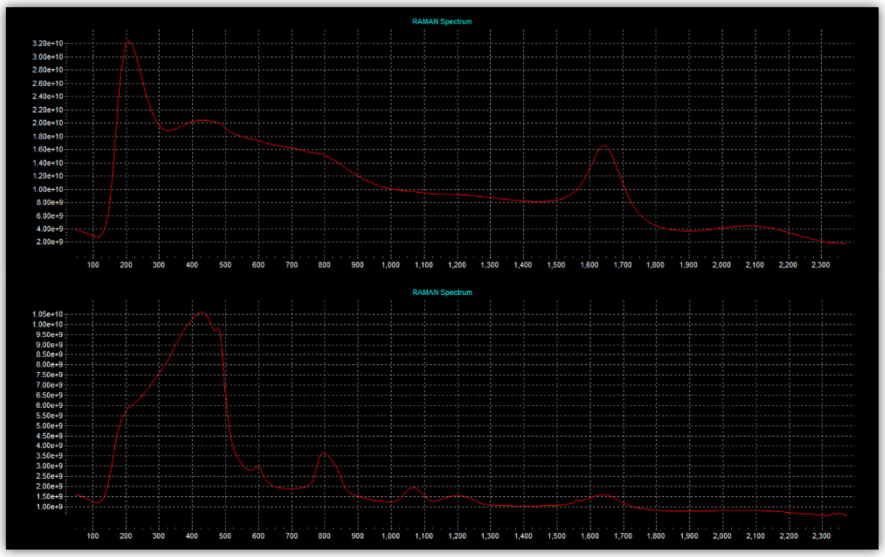

Figure 14. (Top) Raman spectrum of water. (Bottom) Raman spectrum produced when laser improperly focused on face of sample cell.

A final step before carrying out a measurement is to calibrate the binning pattern for the sample. This is done by opening the “Image Alignment and Binning” window, starting acquisition with the laser on, and selecting “Calculate Binning”. This step optimizes the binning pattern of the CCD for the sample, reducing noise and read-out time (related to the saturation of

the pixels). An example of a binning pattern is shown in Figure 15. The instrumental setup and calibration can be tested using a standard sample, often (+)-α-pinene. The Raman and ROA spectra of (+)-α-pinene are shown in Figure 16.

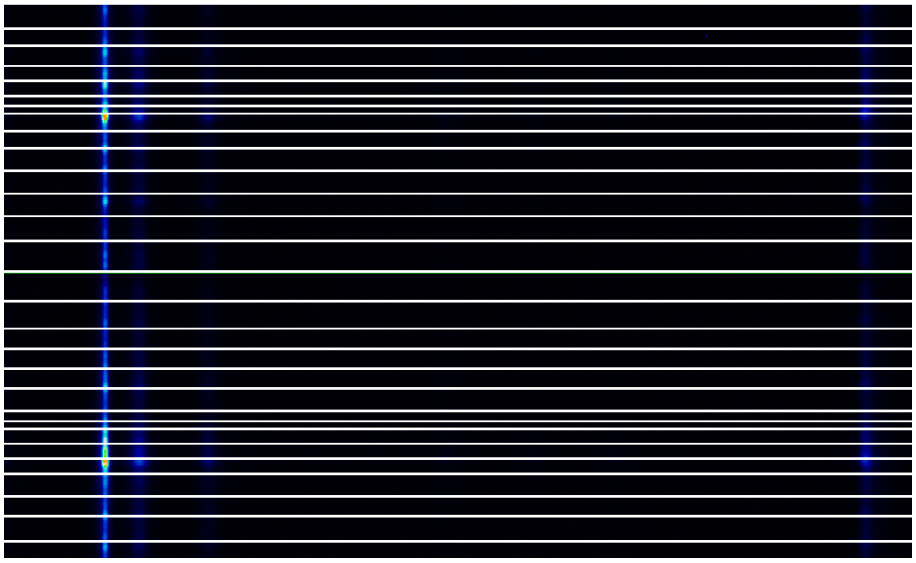

Figure 15. Binning pattern for $CDCl_3$. Each fiber optic branch is dispersed over one half (top and bottom) of the CCD. The pattern bins few pixels near the centre of the CCD where the light is most intense.

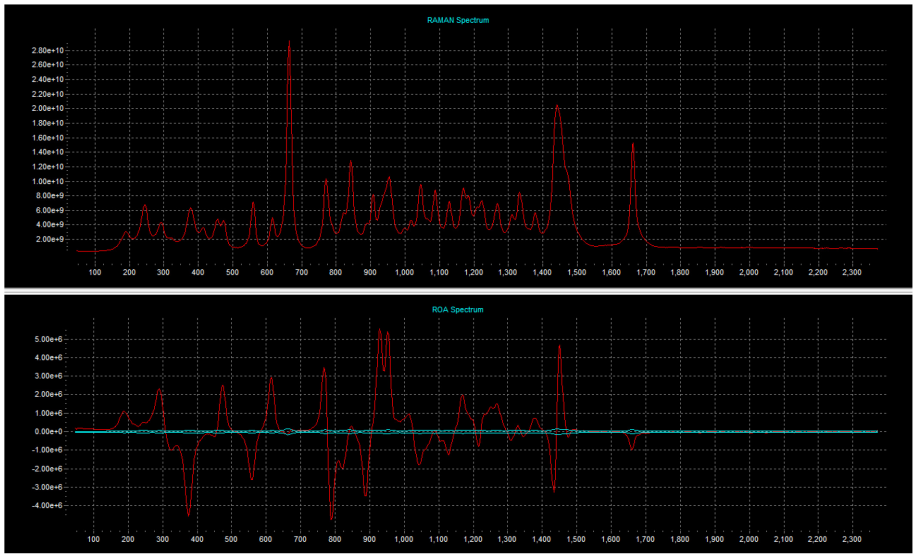

Figure 16. Raman and ROA spectra of (+)-α-pinene as displayed in the ROA software.

A final note on ROA measurements relates to the fluorescence of certain samples. Fluorescence is observed when either the sample or some other species in the solution is electronically resonant with the incident laser light. This is very common for organic samples or molecules that contain conjugated systems. Two examples of this are displayed in Figure 17. It is

often very obvious when the sample is fluorescent by observing an overall elevated baseline that can take a parabolic shape. Sometimes, this maybe to due to a fluorescent impurity. One solution is to burn off the fluorescent impurity in the sample using the instrument's laser. This is done by opening the incident timing shutter (idle) in the engineering controls (by sending the parameters to the instrument in the same control menu) and turning the power of the laser higher than is usual for a measurement. This prevents the strong fluorescent light from reaching the CCD, while allowing the sample to be illuminated. After some time (depending on the sample), the fluorescent contaminant will have burned off, allowing the measurement of the sample. This technique also works in rare instances for the measurement of molecules that are themselves fluorescent.

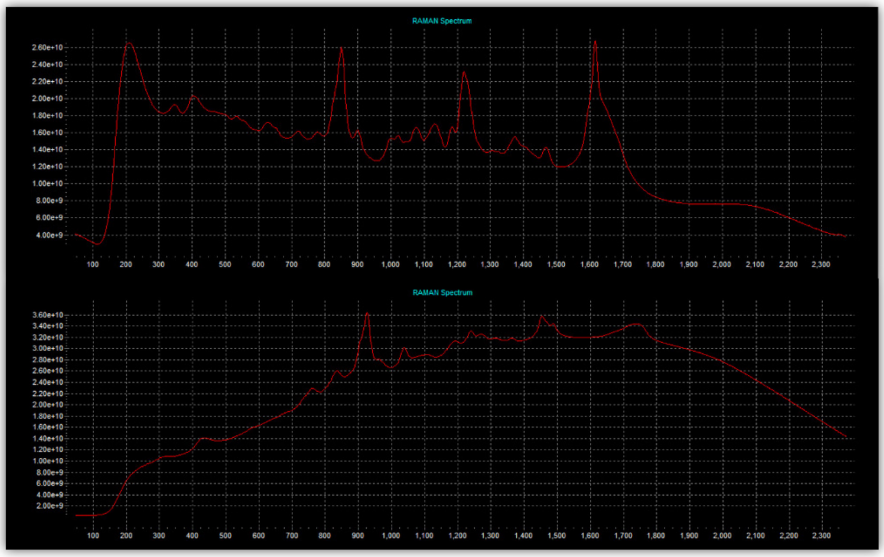

Figure 17. Examples of fluorescent samples. Note the elevated and sometimes parabolic baseline.

Chapter 2: Long-Lived Solute-Solvent Complexes of N-Acetyl-L-Cysteine as Determined by Raman Optical Activity Spectroscopy

2.1 Introduction

As biological environments are predominantly aqueous, the study of the absolute configuration and conformation of chiral compounds such as amino acids, proteins, and drug molecules in water is of particular relevance and interest for several reasons. The interaction between chiral solute and solvent molecules can induce drastic conformational changes of the former. As a result of such conformational changes, the ability of two solute species to bind to one another can be compromised.[41,42] In general, protic solvents such as water and methanol interact with soluble molecules through strong hydrogen-bonding. This hydrogen-bonding plays a dominant role (among other factors such as other dipole-dipole and Van der Waals interactions) in modifying the structural properties of the solute molecules.[43,44] To this end, N-acetyl-L-cysteine, an amino acid derivative, has several hydrogen-bonding sites to which protic solvent molecules can attach, and is therefore a good model system for developing methods to account for changes in such solutions.[29,31] Raman optical activity (ROA) is defined as the difference in Raman scattering intensity of right versus left circularly polarized light produced by a sample and is very sensitive to both molecular chirality and conformation. For this reason, ROA is an ideal technique for this type of study. ROA is capable of measuring vibrational transitions in a wide spectral range, typically all the way down to the ~ 200 cm^{-1}. Complementary information over the more established vibrational circular dichroism (VCD) technique, which has a more limited spectral range due to optics, can be obtained. Since water is a weak Raman scatterer, ROA also has the advantage of working well with water as a solvent for measurements, whereas large absorption of water in this region is traditionally problematic for absorption based IR and VCD spectroscopy.[12] Finally, since Raman and ROA spectra are able to be measured to much lower frequency than vibrational absorption and VCD spectra for the same molecules, additional vibrational modes are able to be analyzed.[45–47]

In the "clusters-in-a-liquid" model,[48] it is hypothesized that the main species responsible for the observed chiroptical features in aqueous solution are not chiral molecules themselves, but rather the long-lived chiral molecule—solvent molecule clusters. The challenge is determining not only the conformation these molecules adopt in solution, but also to what extent these chiral

molecule—solvent complexes actually exist, i.e., number of solvent molecules and locations/orientations.[49]

When considering all of these factors, it becomes difficult to compare the theoretical results with the experimental ones due to the sheer number of possibilities. It is becoming increasingly common to take advantage of molecular dynamics simulations to capture many possible solvation possibilities,[15] which is a useful approach.[23,50,51] Traditionally, averaging the different conformers through the use of a Boltzmann distribution is performed. However, for these systems, this procedure tends to yield inferior results due to the complex nature of a solution. Instead, using different combinations of conformers/complexes and fitting the calculated results to the experimental ones leads to better agreement.[52] An empirical approach may be best to capture the long-lived clusters in aqueous solutions. In this way, analysis of the spectral features and agreement between experimental and calculated results through visual inspection and experience can be aided by a numerical method of determining the contribution of each conformer. Additionally, the effects of a similar, but larger, protic solvent are examined through repeating this work using methanol as a solvent.

2.2 Experimental

N–Acetyl–L–Cysteine and methanol were obtained from Sigma-Aldrich and used without further purification. Sample solutions were prepared by dissolving NALC in either methanol (1.6 M) or double distilled water (0.67 M).Raman and ROA spectra were measured using a BioTools ChiralRAMAN-2X spectrometer (532 nm excitation laser).[11,13] Measurements for solutions in both solvents were done using an illumination time of 1.029 seconds (default setting) in a quartz cell. The water solution was measured for 1742 complete cycles (32 scans per cycle) at a measured laser power of ~ 350 mW at the sample. The methanol solution was measured for 3496 complete cycles at a measured laser power of ~ 60 mW at the sample. Raman spectra were corrected by solvent subtraction and baseline correction. ROA spectra were corrected by solvent subtraction.

2.3 Theoretical

Starting conformers for this work were based on conformational searches done in a previous study.[29] All quantum mechanical calculations were performed using the Gaussian 09 software package (Versions D01 and E01).[53] Optimization of geometries and Hessian

calculations were completed using DFT[54] at the B3LYP[55,56]/cc-pVTZ[57] level of theory, using the integral equation formalism (IEF) version of the polarization continuum model (PCM) to account for the bulk solvent environment (ε = 78.3553 for water and ε = 32.613 for methanol),[58,59] and using an ultrafine integration grid and default convergence criteria. ROA calculations were performed using the same optimized geometries and with the aug-cc-pVDZ[60] basis set where 532 nm was set as the excitation wavelength.[27]

A Lorentzian line shape with a half-width at half-height (HWHH) of 10 cm^{-1} was used for the simulation of Raman and ROA spectra. It was pointed out that the simulated Raman and ROA spectra simulated using GaussView have not properly taken into account the dependence of the scattered intensity on the frequency dependent Boltzmann type factor.[28,61] This is particularly severe at the low frequency region where intensities of bands are very much underestimated.[62,63] No frequency scaling factors were used.

We have adopted a spectral fitting procedure proposed by a number of research groups in dealing with the percentage weight factors for the explicit solvated chiral solute clusters.[52,64,65] Calculated spectra were fit to the experimental results for each set of explicit solvent molecules (with zero to six solvent molecules). Equation 5 outlines the fitting procedure.

$$\sigma = \int_{\nu_i}^{\nu_f} \left[R_{exp} - \sum_n (c_n R_{n,calc}) \right]^4 dR + \int_{\nu_i}^{\nu_f} \left[ROA_{exp} - \sum_n (c_n ROA_{n,calc}) \right]^4 dROA \qquad (5)$$

In short, the sum of the residuals to the fourth power (σ) (minimizes the effects of weaker bands)[52] were minimized for the difference between the experimental (R_{exp}/ROA_{exp}) and sum of the normalized calculated Raman ($R_{n,calc}$) and ROA ($ROA_{n,calc}$) spectra (multiplied by a constant, c_n). In the case of the methanol solution spectra, the residual values between 964 – 1140 cm^{-1} and 1397 – 1533 cm^{-1} were omitted from this procedure because of the strong solvent interference in these two regions.

2.4 Results and Discussion

The collected experimental Raman and ROA spectra for NALC in both water and methanol are shown in Figure 18. The two sets of data are for the most part the same, save for the regions in which methanol produces very strong Raman scattering. In fact, this strong scattering by the solvent results in the saturation of the instrument's CCD at a much lower incident laser power than the water based solution. As such, the methanol solution was measured

for a much longer amount time and yet the signal to noise ratio is still less than that of the water solution. Despite this issue and somewhat surprisingly, the ROA spectra in the studied region are the same between the two solutions after solvent subtraction (raw spectra shown in Figure 32), demonstrating that at a basic level the solvent-solute interactions are alike for both solvents. Water and methanol cause comparable conformational changes in the NALC molecule itself. This being said, certain NALC-solvent complexes likely contribute disproportionately to the observed spectra and are therefore longer-lived than others.

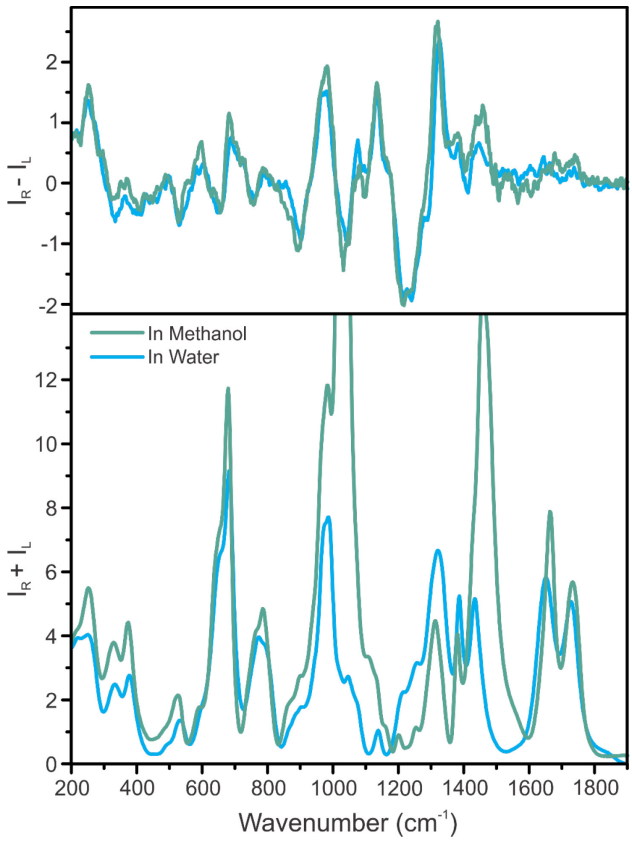

Figure 18. Experimental Raman (bottom) and ROA (top) spectra of the NALC molecule dissolved in water (blue) and methanol (green).

To prevent interferences due to the addition of buffer and/or acid and bases, none were added to the samples. The resultant pH of the measured samples was therefore about 2.5, giving mainly the neutral form NALC in solution. Based on the pK_a value and the simple relationship: $pH=pKa+\log\{[A^-]/[HA]\}$, one may estimate that the relative abundance of the neutral (COOH) relative to the COO^- form is about 82% to 18%. On the other hand, in the previous IR and VCD study, the IR spectra measured at pH=2.5 and pH=0.65 look identical and the anti-symmetric stretching band of COO^- at 1599 cm^{-1} is not at all visible at pH= 2.5,[29] suggesting that the

abundance of COO^- is much lower than the simple estimation implied. For this reason, in the current study, only the neutral form is modelled herein.

A starting point for most combined theoretical and experimental vibrational spectroscopic studies involves modeling of several of the target molecule's monomeric conformations. For solution based studies, taking advantage of the PCM is sometimes sufficient in such situations for aprotic solvents (and sometimes protic ones). For the study of NALC using ROA, this not the case. Figure 33 shows the experimental and calculated Raman and ROA spectra for the three starting NALC monomer conformations in water using the implicit PCM solvation model. It is immediately apparent that the modeling fails to accurately represent what is occurring in the experiment. The three dominant NALC conformers can be seen in Figure 19.

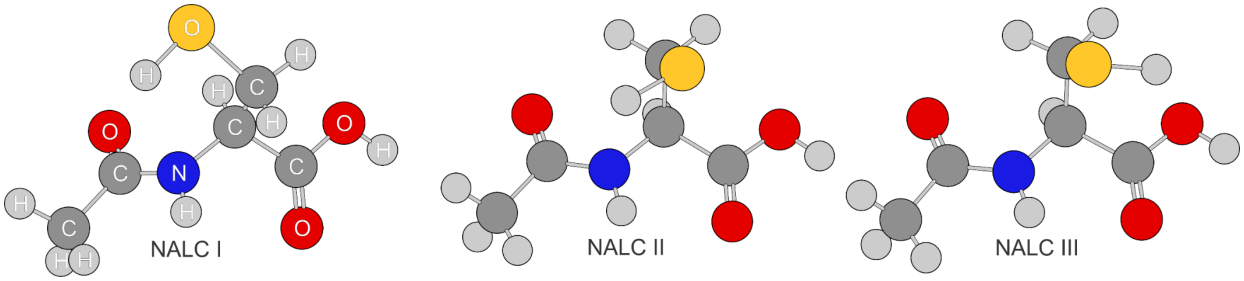

Figure 19. The three most stable monomeric conformations of NALC calculated at the B3LYP/cc-pVTZ (PCM:Water) level of theory.

As a next step, explicit solvent molecules were added to the model in an attempt to improve the results. Free energy considerations were looked at in more detail in a previous study,[31] but due to some differences in calculation aspects (basis set selection), the relative energies for this study are found in Table 1. The systematic addition of increasing amounts of explicit molecules should eventually yield some conformational arrangement(s) whose Raman and ROA spectra closely replicate the key features of the experimental Raman and ROA spectra, respectively, and therefore reveal the key conformation(s) in the experimental solution. No single NALC-solvent complex comes close to replicating a majority of the spectral features. Examples of this are given in Figure 34, where single conformer Raman and ROA spectra are given for the NALC-$(water)_4$ complex in addition to the Boltzmann averaged spectra. Using the Boltzmann population at room temperature based on the free energy, comparatively reasonable results can be achieved. The use of a fitting procedure for the same sets of conformers (based on a number of explicit solvent molecules) further improves the agreement to the experimental data. This fitting procedure emphasizes that the predicted spectral features for Raman and ROA are

generally quite reliably, whereas the related free energies obtained are less so at this level of theory for such species in solutions. In addition, there is no easy way to predict the relative abundances of solvated clusters containing different number of solvent molecules.

Table 1. List of the relative free energy (ΔG), Boltzmann population (%), and empirical weight (%) for each conformer in water and methanol, explicitly solvated by a number of solvent molecules, computed at the B3LYP/cc-pVTZ (PCM:Water/Methanol) level of theory.

Conformer of NALC-Water Complex	Relative ΔG (kJ/mol)	Boltzmann Population	Empirical Weight
I-0Water	0	40.14%	37.30%
II-0Water	0.53	32.37%	0.33%
III-0Water	0.94	27.49%	62.37%
I-4Water-a	1.79	20.72%	39.57%
I-4Water-b	7.90	1.76%	3.00%
II-4Water-a	0	42.69%	0.00%
IIII-4Water-a	0.73	31.80%	37.67%
IIII-4Water-b	6.55	3.03%	19.76%
I-6Water-a	0	91.45%	50.99%
III-6Water-a	5.87	8.55%	49.01%

Conformer of NALC-MeOH Complex	Relative ΔG (kJ/mol)	Boltzmann Population	Empirical Weight
I-0MeOH	0.74	33.46%	59.19%
II-0MeOH	0	45.02%	0.00%
III-0MeOH	1.83	21.51%	40.81%
I-3MeOH-a	2.91	7.76%	40.57%
I-3MeOH-b	0.78	18.36%	16.69%
II-3MeOH-a	2.02	11.15%	0.00%
II-3MeOH-b	0	25.15%	0.00%
III-3MeOH-a	1.15	15.80%	0.00%
III-3MeOH-b	0.36	21.78%	42.74%
I-4MeOH-a	4.56	3.67%	0.00%
I-4MeOH-b	2.80	7.49%	50.02%
II-4MeOH-a	1.53	12.50%	0.00%
II-4MeOH-b	0.55	18.54%	0.41%
II-4MeOH-c	0	23.14%	35.93%
III-4MeOH-a	2.84	7.37%	0.00%
III-4MeOH-b	1.87	10.88%	2.55%
III-4MeOH-c	0.85	16.41%	11.09%

After examination, employing empirical fitting for each set of NALC-solvent complexes gave the best results. Different versions of a similar approach have been performed in other studies with some success.[52,64,66] The results herein are perhaps the most convincing example for the use of empirical fitting for solution based spectroscopic studies. All of the empirically fit calculated spectra for the water complexes are shown in Figure 35 and those for just the four and six water complexes are shown in Figure 20. The key section of the spectrum is the 800 to 1400 cm^{-1} region. The vibrations contained in this range can predominantly be assigned to variations of S-H bending, C-H bending, C-H_2 bending, C-H_3 and N-H bending modes. The two strong bands at higher frequencies between 1600 and 1800 cm^{-1} are assigned to the C=O stretching mode of the carbonyl group (lower frequency) and C=O stretching mode of the carboxyl group (higher frequency). While the bands below about 800 cm^{-1} certainly contain valuable information and an agreement between the experiment and theory would be ideal, they mostly consist of molecular breathing vibrations and solvent (O-H) bending modes and are therefore difficult to analyze in detail. Regardless of this, the agreement between the experiment and fitted calculated spectra in this region is surprising. Selected example vibrational modes are detailed in Figure 21, with a focus placed on those that are most greatly affected by the explicit solvent molecules. Specifically, these selected vibrational modes have a noticeable difference between the monomer and explicitly solvated models in terms of intensity and/or character. For example, the bands labeled 9 to 12 in Figure 20 are significantly indirectly affected by the explicit water molecules. These correspond to the vibrational modes below in Figure 21. Band 9 is assigned to antisymmetric C-H_3 wagging of the methyl group and S-H bending (and N-H bending). Band 10 represents another mode of antisymmetric C-H_3 wagging. Band 11 can primarily be assigned to S-H bending, with contributions from C-H bending, C-H_2 bending, and C-C stretching. Finally, band 12 is assigned to N-H bending, with smaller contributions from symmetric C-H_3 wagging and C-H bending. It is interesting to note that while the observed motions are very similar between monomer and explicitly solvated complexes, the spectroscopic features are modified by small geometric changes caused by the solvation rather than the direct changes due to hydrogen-bonding of hydroxyl and carbonyl groups. A key example is the vibrational mode at 996 cm^{-1} (band 8). For the monomer, this band has a negative ROA intensity, while it has positive intensity for the water complex.

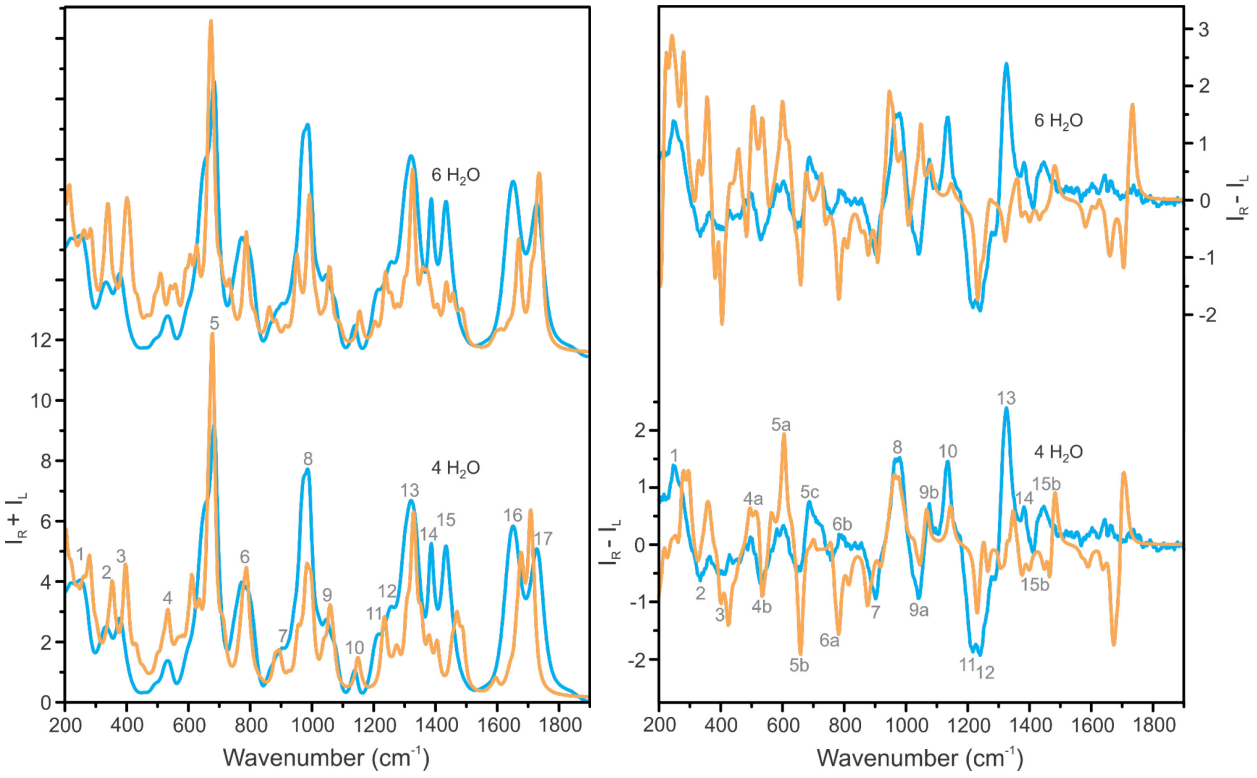

Figure 20. Experimental (blue) and empirically fit (orange) Raman (left) and ROA (right) spectra of the NALC-4 Water complexes and the NALC-6 Water complexes.

Often the vibrational modes of the explicit solvent molecules are frozen when calculating such complexes, resulting in the calculated vibrational bands only containing information about the conformation of the main molecule. Here, the vibrations of the solvent molecules are considered, as the experimental results support that there is evidence of some contribution from these complexed explicit solvent molecules in the physical experiment. This is especially apparent when examining the agreement between the empirically fit data for the four water complexes with that of the experiment and the calculated monomeric results. The agreement between this explicitly solvated model and experiment has far better agreement in the < 700 cm^{-1} regions, where the vibrational modes of the explicit solvent molecules primarily reside. This supports the theory that the solvent does not only force small conformational changes in the solute molecule, but that the long-lived complexes as a whole contribute to the measured Raman and ROA spectra.

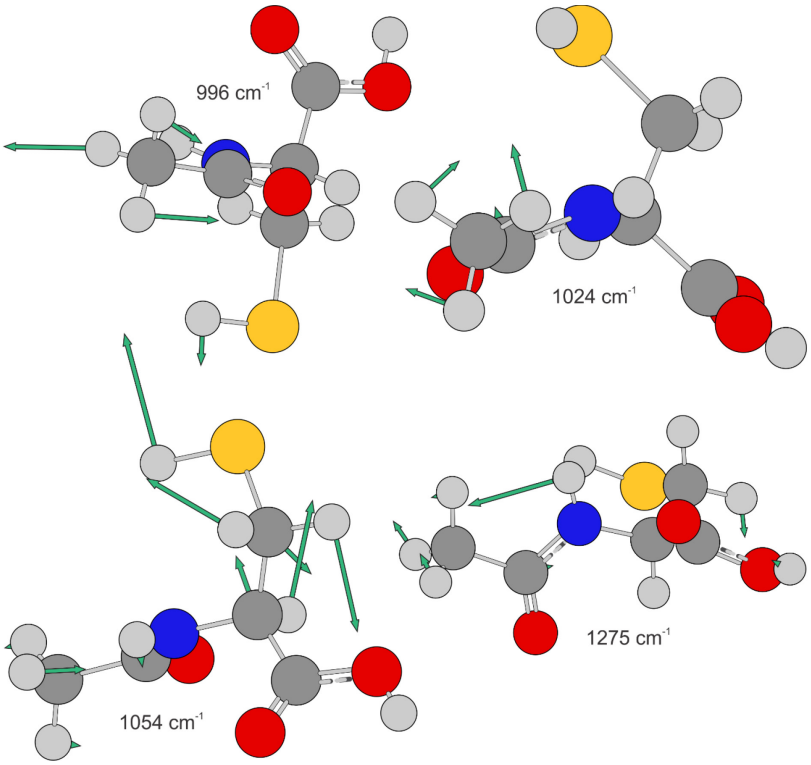

Figure 21. Four examples of vibrational modes (assigned as bands 9 to 12 in Figure 20, in order of increasing frequency) for the NALC monomer (conformer 1) (thus non-matching frequencies in this Figure and Figure 20) that have intensity differences or notable spectral features between the monomeric and explicitly solvated models.

Using this information and the empirical fitting results, the predominant NALC-solvent complex conformations can be determined for both the water and methanol solutions. In selecting which set of complexes (based on number of solvent molecules) best matches the experimental Raman and ROA spectra, a combination of visual inspection and the lowest sum of residual values was used.

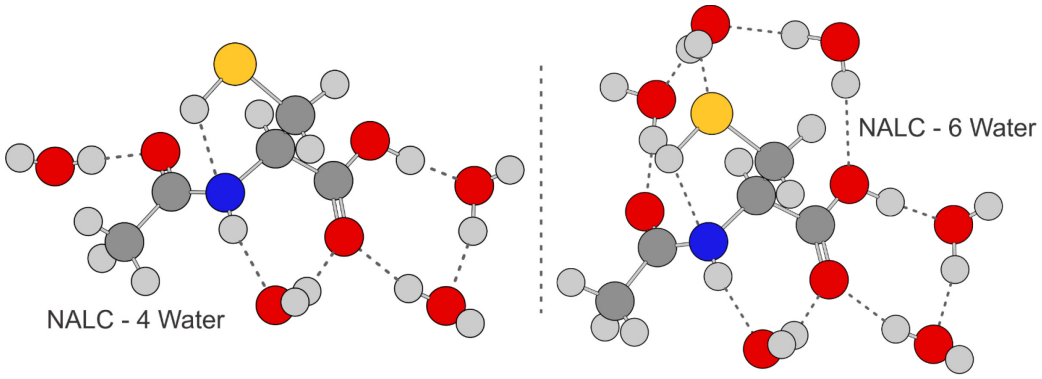

Figure 22. An example of a NALC-4 Water and NALC-6 Water complex calculated at the B3LYP/cc-pVTZ (PCM:Water) level of theory.

Figure 22 shows an example of the four and six water complexes with NALC. In general, more energetically stable structures are achieved for both explicit methanol and water molecules when two solvent molecules form an eight-membered ring with the carboxyl group. The C=O stretching frequencies become more red-shifted as the size of the complex increases and as a result, the agreement with experiment improves. The experimental ROA does not discernibly show these same bands. This is an interesting phenomenon that appears to be the case for many measured systems measured by the authors. All of the considered conformers for NALC show strong bands for this region in the ROA spectrum. This effect is not well understood. One possible explanation is that the experimental bands are farther apart in frequency than in the calculations. Thus, there is no exciton coupling between the modes and the ROA intensity is low in the experiment as a result.

As can be seen, the agreement between these empirically fit Raman and ROA spectra for the four and six water complexes and the experimental results is quite convincing. The key spectral region of bands 7 to 10 is well reproduced for both complexes. The experiment is also remarkably well reproduced below 900 cm^{-1} in the Raman spectrum and the spectral features are similarly in good agreement with the ROA spectrum, save for intensity differences. Finally, the bands from 11 to 15 are similar between these two sets of calculated complexes, and while the agreement with experiment is somewhat lacking for the Raman spectra, the correct sign is predicted for the ROA spectra. It is clear that the four water complex serves as the best model for replicating the experiment using this visual inspection, which is backed up by a sum of residuals squared value that is roughly 860 intensity units less than the six water complex.

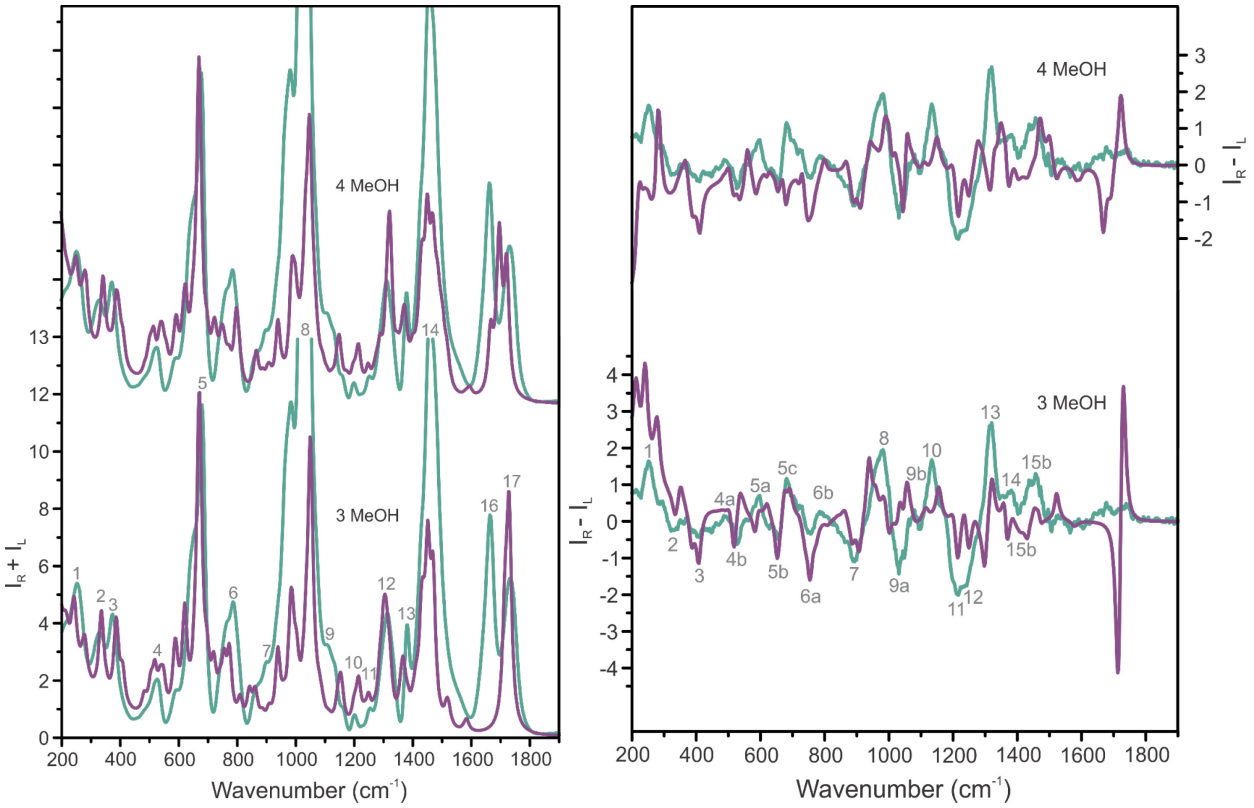

Figure 23. Experimental (green) and empirically fit (purple) Raman (left) and ROA (right) spectra of the NALC-3 MeOH complexes and the NALC-4 MeOH complexes.

The same is true for the three and four methanol complexes shown in Figure 23 (the spectra for all the empirically fit methanol complexes are shown in Figure 36). The empirically fit set of four methanol complexes has an overall sum of residuals squared value that is almost 18,000 intensity units less than that of the set of three methanol complexes. In general, visual inspection of the spectra agrees with this metric. The spectra for the set of three methanol complexes is deficient in a few places, specifically bands 9a and 9b in the ROA spectrum and the shift of the carbonyl and carboxyl bands together to overlap and produce a single band (17) in the Raman spectrum. While the set of four methanol complexes falls behind the set of three methanol with respect to the bands 5a-c and 13 in the ROA spectrum. An additional observation of this approach is that the bands that are saturated in the experiment due to the large Raman scattering of the solvent methanol (bands 8 and 14) are in fact better reproduced as the explicit solvation shell increases in size. The five methanol complex shows this effect nicely (Figure 36).

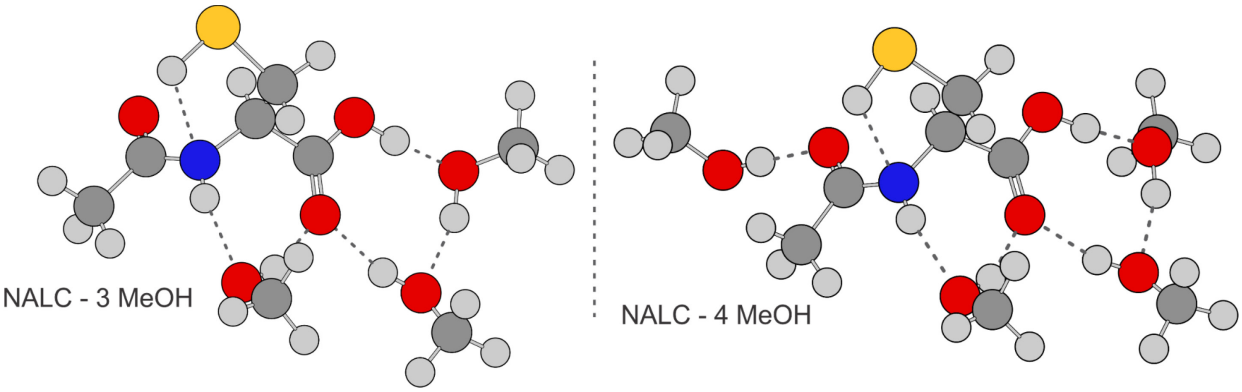

Figure 24. An example of a NALC-3 MeOH and NALC-6 MeOH complex calculated at the B3LYP/cc-pVTZ (PCM:Methanol) level of theory.

The fact that the experimental ROA spectra in water and methanol are essentially identical reveals some interesting information about the systems. It has been established that a set of four explicit solvent molecules provides the best agreement theoretically with the experiment for both solutions. In examining the four solvent complex with the highest empirical weight for each system, the same hydrogen-bonding motif is shown in both. The carboxyl group forms an eight-member ring with two solvent molecules, with an additional solvent molecules forming a hydrogen-bond with the opposite lone pair of the oxygen atom in the carbonyl group. Likewise, one lone pair of the carbonyl component of the amide group forms a hydrogen-bond with a solvent molecule, while the other forms a bond with the hydrogen of the thiol group. It can therefore be concluded that it is likely that very similar NALC-solvent complexes exist in both water and methanol solutions because of their identical experimental ROA spectra. The calculated spectra for these complexes back this up by demonstrating that certain hydrogen-bonding and solvent arrangements better replicate the experiment, and that these two factors are therefore important in the modelling the systems.

2.5 Conclusion

Experimental Raman and ROA spectra were collected for the NALC molecule in solutions of water and methanol. To determine the interactions between NALC and solvent molecules, and therefore determine the longest-lived complexes that exist each solution, DFT calculations we performed to simulate NALC-solvent complexes and their respective Raman and ROA spectra. This was accomplished by starting with previously determined clusters found using molecular dynamics calculations, and then focusing on the more stable complexes by

performing optimization and frequency calculations at the B3LYP/cc-PVTZ (PCM) level of theory. As a result, several NALC-water complexes (with one to six explicit solvent molecules) and NALC-MeOH complexes (with one to five explicit solvent molecules) were determined and their Raman and ROA optical tensors calculated at the BYLYP/aug-cc-pVDZ (PCM) level of theory. Monomeric NALC and any single NALC-solvent complex provided poor agreement with the experimental results. Therefore, empirical fitting was employed in an effort to highlight the importance of replicating the spectral features generated in the experiment in way that incorporates automation and visual inspection. To that effect, this approach has been successful, it has ultimately been determined that for both NALC in water and in methanol, the longest-lived species are NALC-4 solvent complexes. In this case, ROA spectroscopy serves as a very useful technique in determining the structure of this chiral molecule in protic solvents.

2.6 Supporting Information

Supporting information, including raw experimental spectra, additional calculated spectra, comparisons between NALC in water and methanol, a full table of conformational energies, and monomeric Cartesian coordinates can be found in Appendix A.

Chapter 3: Raman Optical Activity Spectroscopy of Serine-Water Complexes

3.1 Introduction

Serine is an amino acid with many functions in biological systems. It occurs naturally and is important in enzymatic function. The levorotary enantiomer is a precursor for proteins and is involved in the synthesis of many other biologically important molecules.[67] Furthermore, there are several medical applications for serine including serine deficiency (leads to enzyme defects),[68] treatment of schizophrenia and neurodegenerative diseases,[69,70] and as a biomarker for Alzheimer's disease.[71] An example of the interesting variability of serine is that it has been shown to preferentially form an octameric cluster under certain conditions over all smaller clusters.[72–74]

It can be concluded that knowledge of the conformations of serine would be very beneficial. This is especially true for its conformations in an aqueous environment, as the biological systems in which serine takes part are of an aqueous nature. Serine has been studied extensively using infrared (IR), Raman, and vibrational circular dichroism (VCD) techniques.[75–79] Several theoretically based studies have also been performed.[80–82] The findings in these studies are helpful as a starting point in determining the general conformation of the molecules, but do not address the potential interactions that it can experience in solution. Zhu, *et al.* performed such a study using different concentration and pH parameters. Their combined IR, VCD, and quantum mechanical calculation based study found that it was advantageous to include explicit solvent molecules for the simulation of their experiments, specifically serine complexed with six water molecules.[30] This concept was later termed "clusters-in-a-liquid", and promotes the idea that the spectral features observed from measurement of a molecule in solution are due to the solute-solvent complexes.[48]

Raman and Raman optical activity (ROA) spectroscopy are complementary to IR and VCD spectroscopy. Where IR and VCD spectroscopy take advantage of the vibrational absorption of a molecule to gather structural information, Raman and ROA take advantage of the inelastic scattering of light from the molecule to gain complementary information. In short, ROA is the differential scattering of right and left circularly polarized light. The latter two techniques are relevant to the study of serine in solution because of the ease of using water as a solvent for

measurements. While water has strong absorption of light in the IR fingerprint regions, it does not exhibit significant Raman scattering and permits the measurement of a larger spectral range.

This study is complementary to the previous VCD study, instead using ROA in an attempt to extract additional details about the conformation of serine and its water clusters. In addition to serine complexed with six explicit water molecules, complexes with four explicit water molecules were studied, as well as the monomeric serine conformers with their geometries adopted from the complexes, but with the solvent molecules removed. We used an empirical fitting procedure (detailed below) to simulate the final Raman and ROA spectra and compare them to the experimental results. We also examined two different approaches to improve the calculations of hydrogen-bonding and van der Waals interactions between the serine and water molecules. First, the X3LYP functional was used, and second, Grimme's dispersion correction (GD3).[83,84]

3.2 Experimental

D- and L-Serine were obtained from Sigma-Aldrich and used without further purification. The samples were prepared as solutions by dissolving either enantiomer in double distilled water with a concentration of 1.9 M. The sample solutions were placed in a quartz cell and measured using a BioTools ChiralRAMAN-2X spectrometer equipped with a 532 nm excitation laser.[13,11] An illumination time of 1.25 seconds was used with measurement times of 18 hours for D-Serine and 31 hours for L-Serine. Laser power was measured to be 188 mW at the sample. Raman spectra were corrected by solvent subtraction and baseline correct. ROA spectra were corrected by subtraction of the opposite enantiomer and dividing by two.

3.3 Theoretical

Initial geometries of the SER6W complexes were taken from the previous work.[30] SER4W and SER0W conformers were created by removing selected/all explicit solvent molecules from SER6W. We note that the previous calculations did not include the implicit solvent surrounding, while in the current calculations, we used the integral equation formalism (IEF) version of the polarization continuum model (PCM) to account for the bulk solvent environment ($\varepsilon = 78.3553$ for water), with an ultrafine integration grid and default convergence criteria.[58,59] All quantum mechanical calculations were performed using the Gaussian 09 software package (Versions D01 and E01).[53] Optimization of geometries, frequency calculations,

and ROA calculations were completed using DFT[54] at the B3LYP[55,56]/aug-cc-pVTZ[57] level of theory. This functional and basis set have been benchmarked as an effective level of theory for ROA calculations.[27] The excitation wavelength was set to 532 nm for ROA calculations.[27] Additional calculations were performed for the SER6W complexes using the same methodology, but instead using either the X3LYP[83] functional or by implementing Grimme's dispersion correction (GD3).[84]

A Lorentzian line shape with a half-width at half-height (HWHH) of 10 cm^{-1} was used for the simulation of Raman and ROA spectra. It was pointed out that the simulated Raman and ROA spectra simulated using GaussView have not properly taken into account the dependence of the scattered intensity on the frequency dependent Boltzmann type factor.[61,28] This is particularly severe at the low frequency region where intensities of bands are very much underestimated.[62,63] No frequency scaling factors were used.

We have adopted a spectral fitting procedure proposed by a number of research groups in dealing with the percentage weight factors for the explicit solvated chiral solute clusters.[52,64,65] Calculated spectra were fit to the experimental results for each set of explicit solvent molecules (with zero, four, or six solvent molecules). Equation 5 outlines the fitting procedure.

$$\sigma = \int_{\nu_i}^{\nu_f} \left[R_{exp} - \sum_n (c_n R_{n,calc}) \right]^4 dR + \int_{\nu_i}^{\nu_f} \left[ROA_{exp} - \sum_n (c_n ROA_{n,calc}) \right]^4 dROA \qquad (5)$$

In short, the sum of the residuals to the fourth power (σ) (minimizes the effects of weaker bands)[52] were minimized for the difference between the experimental (R_{exp}/ROA_{exp}) and sum of the normalized calculated Raman ($R_{n,calc}$) and ROA ($ROA_{n,calc}$) spectra (multiplied by a constant, c_n).

3.4 Results and Discussion

The zwitterionic form (having both a positively charged NH_3^+ group and negative COO^- group, and therefore a net neutral charge) of serine was used for all calculations, as the pH of serine has been experimentally determined to be 5.68 and the pK_a of the COOH and NH_3^+ groups to be 2.21 and 9.15, respectively.[30,85] At a pH of around 5.7, serine should predominantly exist in the zwitterionic form for pure solutions. The five SER6W complexes found in a previous study[30] were re-optimized at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory and the Raman and ROA optical tensors were calculated using this same methodology. These complexes can be seen

in Figure 25. In general, the COO^- functional is solvated by two water molecules, and the NH_3^+ functional by two additional water molecules. The last two water molecules in the complex are placed in one of two possible positions. One is between the O-H functional and either of the other functional groups or another water molecule. A second is between the COO^- and NH_3^+ functional groups. An example of this hydrogen-bonding network is displayed in Figure 26 for the SER6W2 complex, which is the lowest energy conformer.

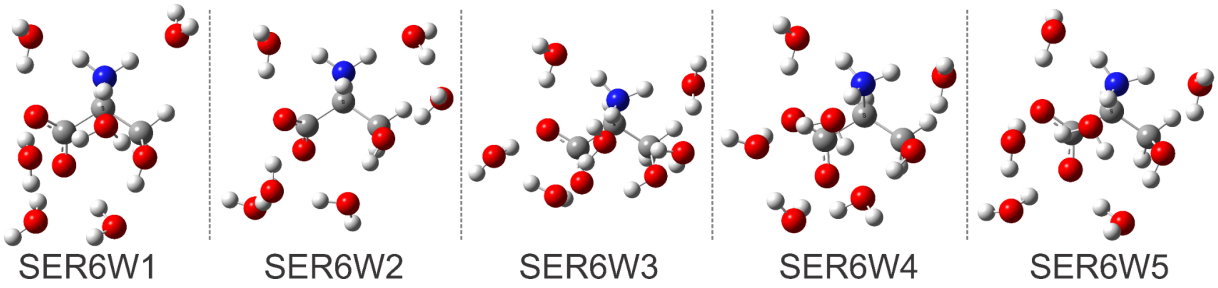

Figure 25. Six-water serine complexes calculated at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory.

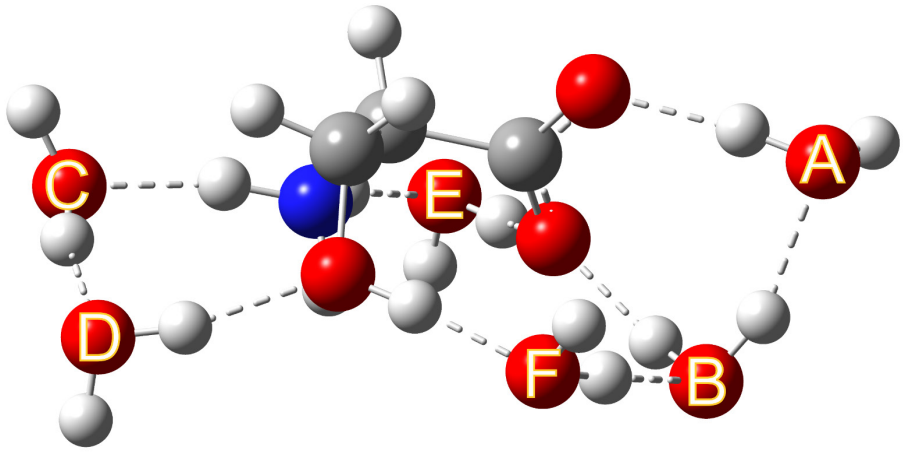

Figure 26. Example of the hydrogen-bonding network in the SER6W2 complex. Potential hydrogen-bonds shown as dashed lines.

The hydrogen-bonding network for SER6W2 consists of an eight-member ring involving the COO^- functional group and two water molecules (A and B), a seven-member ring involving the NH_3^+ functional group and two water molecules (C and D), one water molecule (E) inserted between the COO^- and NH_3^+ functional groups, and one final water molecule (F) inserted between the O-H functional group and the ring structure the COO^- ring structure. Such

cooperativity increases stability and lowers the energy relative to other hydration schemes without such arrangements.

The experimental Raman and ROA spectra for L-serine are shown in Figure 27, along with those calculated for the empirically fit six-water complexes (SER6W). An empirical fitting approach was implemented here to capture the spectral features of the long-lived clusters in the solution using several possible water-serine complexes simultaneously, rather than relying on their calculated free energy estimates. The major points of interest with respect to the comparison between the two are the 1200 to 1600 cm^{-1}, 800 to 1200 cm^{-1}, and < 800 cm^{-1} regions. The region > 1600 cm^{-1} contains vibrational bands related to O-H bending of the water molecules, N-H_3 bending, and to a lesser extent, (O)-C-O stretching. The agreement between the experimental and calculated Raman spectra here is quite good, although there is some additional broadening due to the bulk solvent in the experiment. The large negative band at 1650 cm^{-1} in the calculated ROA spectrum represents the H-O-H bending motion of the water molecule joining the COO^- and NH_3^+ functional groups. This band, along with any other significant ROA signal, are not observed in the experiment in the same range.

Moving to the next region (1200 to 1600 cm^{-1}), the calculated Raman spectrum again gives good agreement with the experiment. The three most significant bands between 1400 and 1600 cm^{-1} are assigned to C-H_2 symmetric bending (~ 1466 cm^{-1} – all specific wavenumber refer to the SER6W2 conformer), a combination of C-H_2 wagging, C-H bending, N-H_3 symmetric bending, and O-C-O symmetric stretching (~ 1425 cm^{-1}), and a combination of the symmetric motion of C-H bending, N-H_3 bending, and C-H_2 wagging (~ 1361 cm^{-1}). Relative intensities and transition frequencies are well replicated here. The ROA spectra show a slightly different story. If a small blue shift is applied to the calculated spectrum, sign agreement for the bands improves, however certain experimental features are not replicated. Specifically, the negative experimental band at 1250 cm^{-1} is not present in the calculated spectrum. Two additional bands of note near this region are the O-H bending, C-H_2 bending, and C-H bending mode at 1259 cm^{-1} and the N-H_3 asymmetric bending and C-H bending mode at 1191 cm^{-1}. The next region (800 to 1200 cm^{-1}) shows excellent agreement between experimental and calculated spectra for both Raman and ROA, with relative intensities and band position being predicted correctly for the most part. Three major contributions here come from the symmetric and antisymmetric combinations of C-O(H) stretching, O-H bending, N-H_3 bending, C-H_2 wagging, and C-H bending (1047 and 1079

cm^{-1}, respectively – negative bands), and the combination of O-H bending of serine and water molecules at 955 cm^{-1} (positive band).

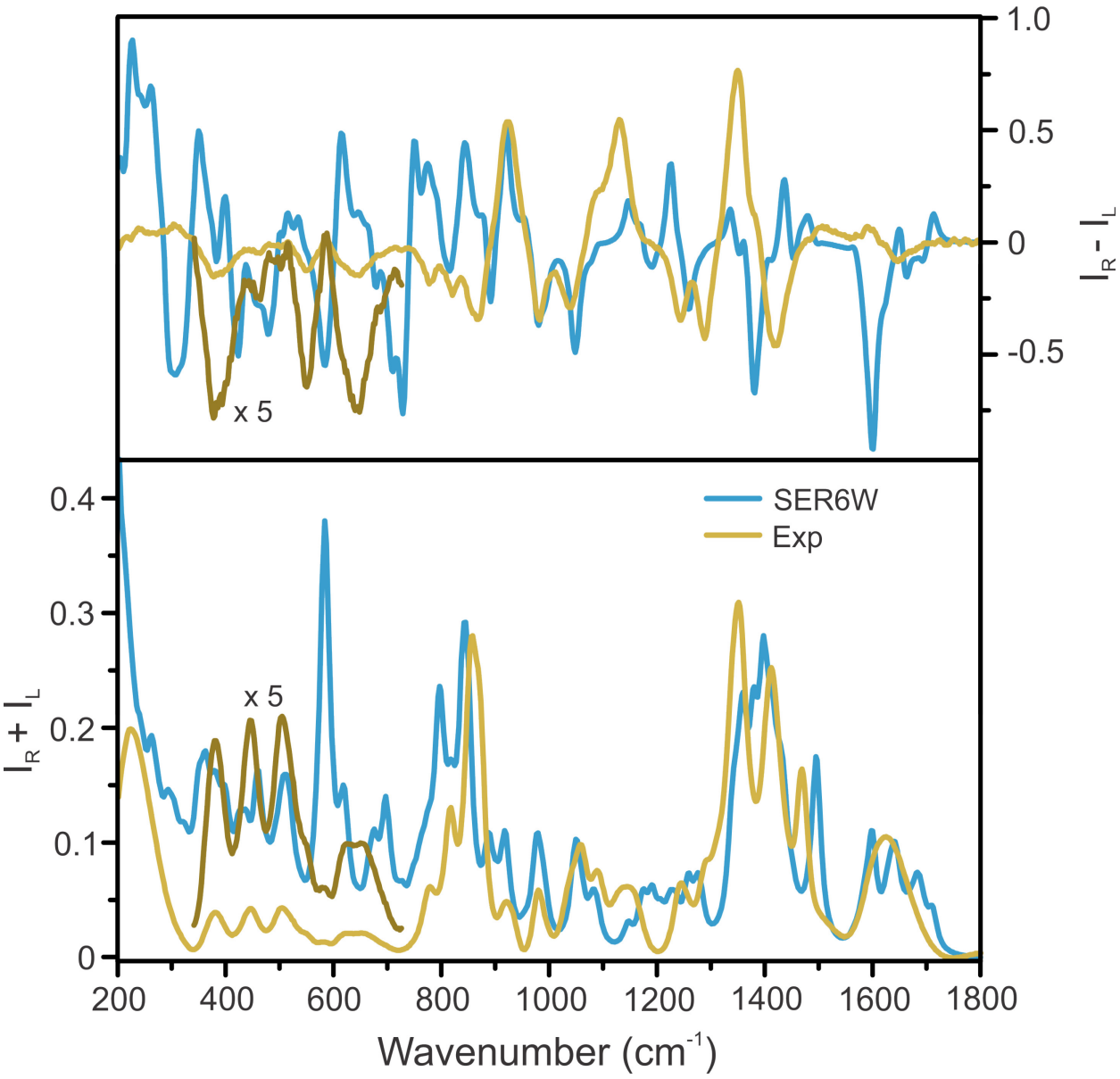

Figure 27. Experimental (yellow) and empirically fit (blue) Raman (bottom) and ROA (top) spectra of L-serine and SER6W. The darker yellow trace represents a five times expansion of the experimental spectrum for the same region.

The region below 800 cm^{-1} is dominated by breathing motions of the serine molecule, solvent vibrations, and combinations of the two. The experimental bands are relatively weak compared to the rest of the spectrum in this range. As such, there is poor agreement between experimental and calculated results. Amplification of the experimental data in this range allows

some similarity to be seen in the Raman spectra, however the calculated band at 600 cm^{-1} (serine breathing and solvent vibrations) is not seen at all in the experiment. Furthermore, the agreement is increasingly poor here for the ROA spectra. There are a number of possible reasons for this deficiency, including anharmonicity being rather significant at these lower frequencies. Anharmonic frequency calculations often correct for errors in transition frequencies. While this inherently does not change a vibrational spectrum (although the corrections are important), a shift promotes a noticeable change in the band shape for close transitions due to the line broadening imposed when generating the calculated spectrum.[86] For example, should the transitions be shifted close together due to the correction, a single narrower band will be observed than what would appear as two more distinct bands absent this correction.

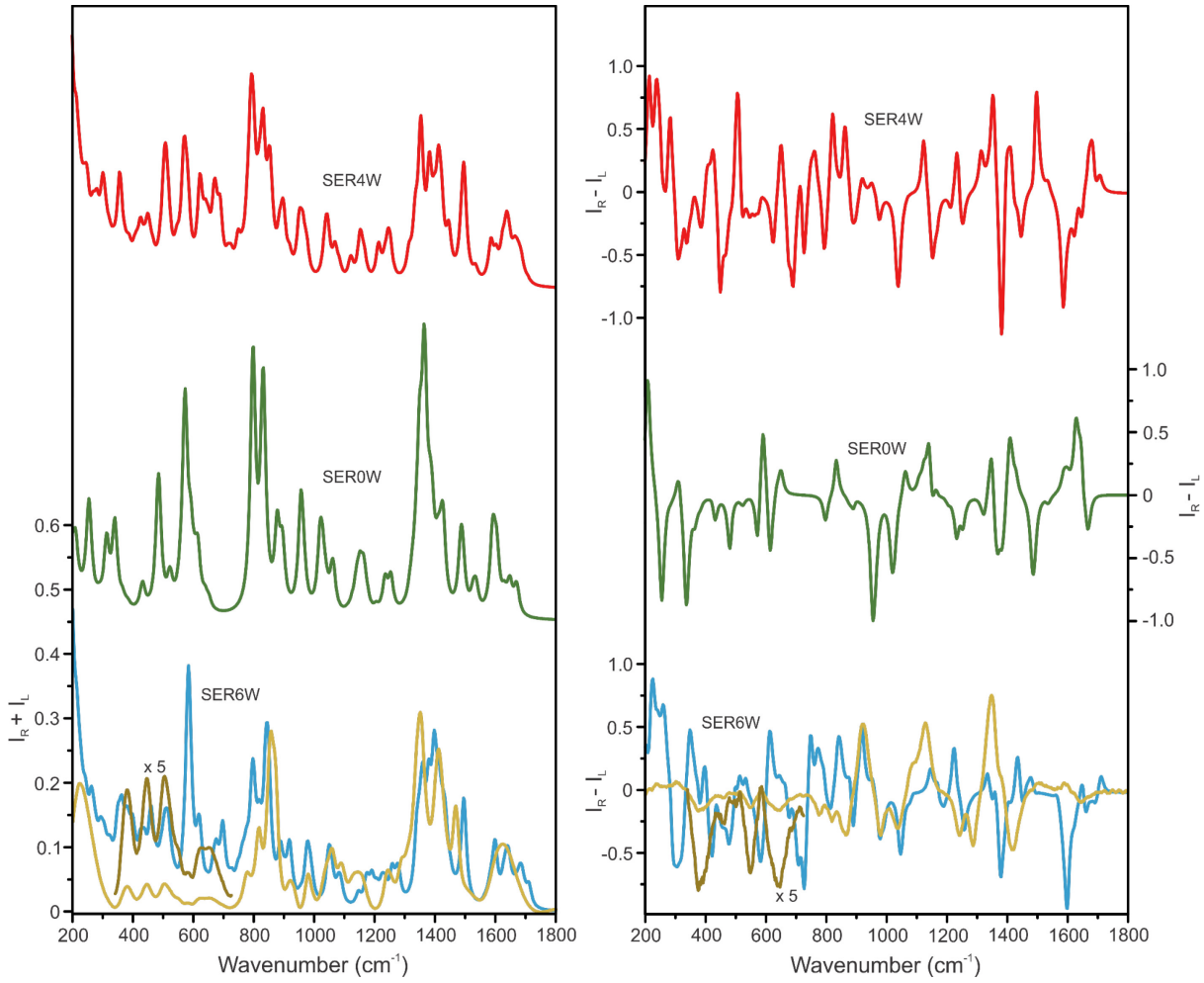

Figure 28. Experimental (yellow) and empirically fit SER0W (green), SER4W (red), and SER6W (blue) Raman (left) and ROA (right) spectra for L-serine. The darker yellow trace represents a five times expansion of the experimental spectrum for the same region.

To account for the possibility of lower order clusters and monomeric serine in solution, explicit water molecules were systematically or completely removed from these SER6W complexes to create four-water complexes and serine monomers with the same conformation as the respective parent SER6W complexes. The empirically fit calculated spectra for SER0W and SER4W are shown along with previous data in Figure 28. These specific complexes are displayed in Figure 29 and Figure 30. In terms of configuration, there are only minor changes in geometry when comparing the two sets of complexes and the monomer. One notable exception is the SER4W4 complex. The NH_3^+ functional group no longer has two water molecules to bond with and therefore rotates to stabilize instead with the COO^- functional group. In terms of Raman and ROA spectra, there are some similarities between the three sets of fit spectra. Overall, the

SER6W spectra provide the best agreement with the experiment. Chiefly, the sets of bands centred at 900 cm^{-1} and 1500 cm^{-1} have better relative intensities, and the 900 to 1200 cm^{-1} has considerably better agreement with the experiment, especially with respect to the ROA spectra.

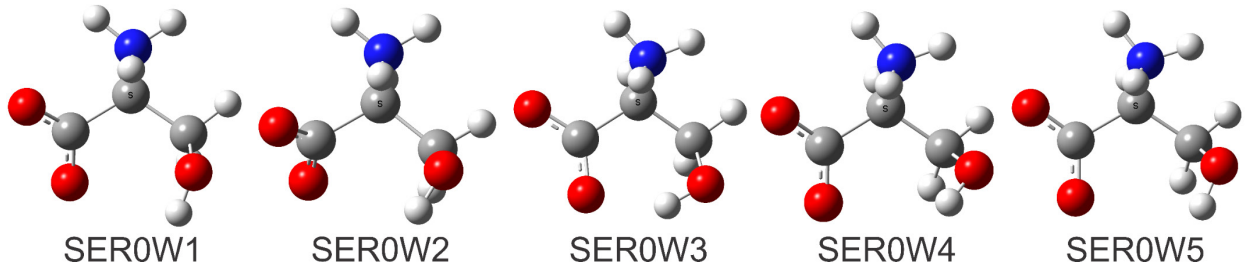

Figure 29. Serine conformers calculated at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory.

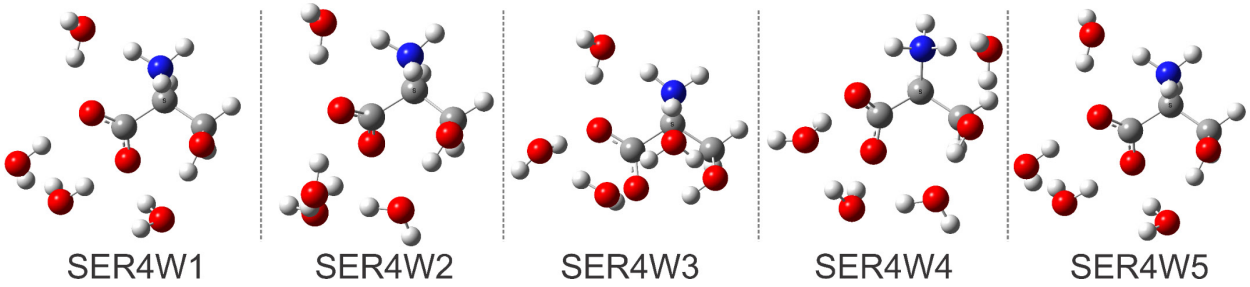

Figure 30. Four-water serine complexes calculated at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory.

Table 2 contains the relative free energies, Boltzmann populations, and empirical weights for each set of complexes and the monomers. The second conformer (SERxW2) is lowest in energy for the six-water complexes and monomer, while the fifth conformer is lowest for the four-water complex. Comparing the Boltzmann populations with the empirical weights, the conformational distribution is more spread out, while at the same time the lowerst energy conformers still contribute notable percentages. This is significant in the demonstration that replicating experimental spectral features is key in determining the long-lived clusters in such solutions and that multiple clusters contribute to what is represented in experimental measurements. Relying on unreliable energy estimates is less advantageous than relying on vibrational mode and spectral analysis here.

Table 2. List of the relative free energy (ΔG), Boltzmann population (%), and empirical weight (%) for all serine conformers and complexes, computed at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory.

Conformer	Relative ΔG (kJ/mol)	Boltzmann Population	Empirical Weight
SER0W1	5.99	6.50%	44.86%
SER0W2	0.00	73.06%	31.64%
SER0W3	3.17	20.31%	8.05%
SER0W4	19.17	0.03%	15.45%
SER0W5	16.60	0.09%	0.00%
SER4W1	18.09	0.05%	37.92%
SER4W2	6.66	4.74%	34.20%
SER4W3	3.18	19.34%	0.00%
SER4W4	6.00	6.18%	0.00%
SER4W5	0.00	69.70%	27.88%
SER6W1	8.87	2.66%	30.61%
SER6W2	0.00	95.41%	29.76%
SER6W3	10.28	1.50%	17.82%
SER6W4	17.83	0.07%	20.75%
SER6W5	13.88	0.35%	1.06%

As a final step, different computational approaches were undertaken to account for important hydrogen-bonding and van der Waals interactions. These types of interactions are abundant in these six-water complexes and it is logical to expect some potential improvement in the simulation of Raman and ROA spectra when they are accounted for. Figure 31 contains the experimental Raman and ROA spectra, along with those for empirically fit SER6W (calculations performed using the B3LYP functional), SER6W-X3LYP (calculation performed using the X3LYP functional rather than B3LYP), and SER6W-GD3 (calculation performed using the B3LYP functional and GD3 correction) calculated spectra.

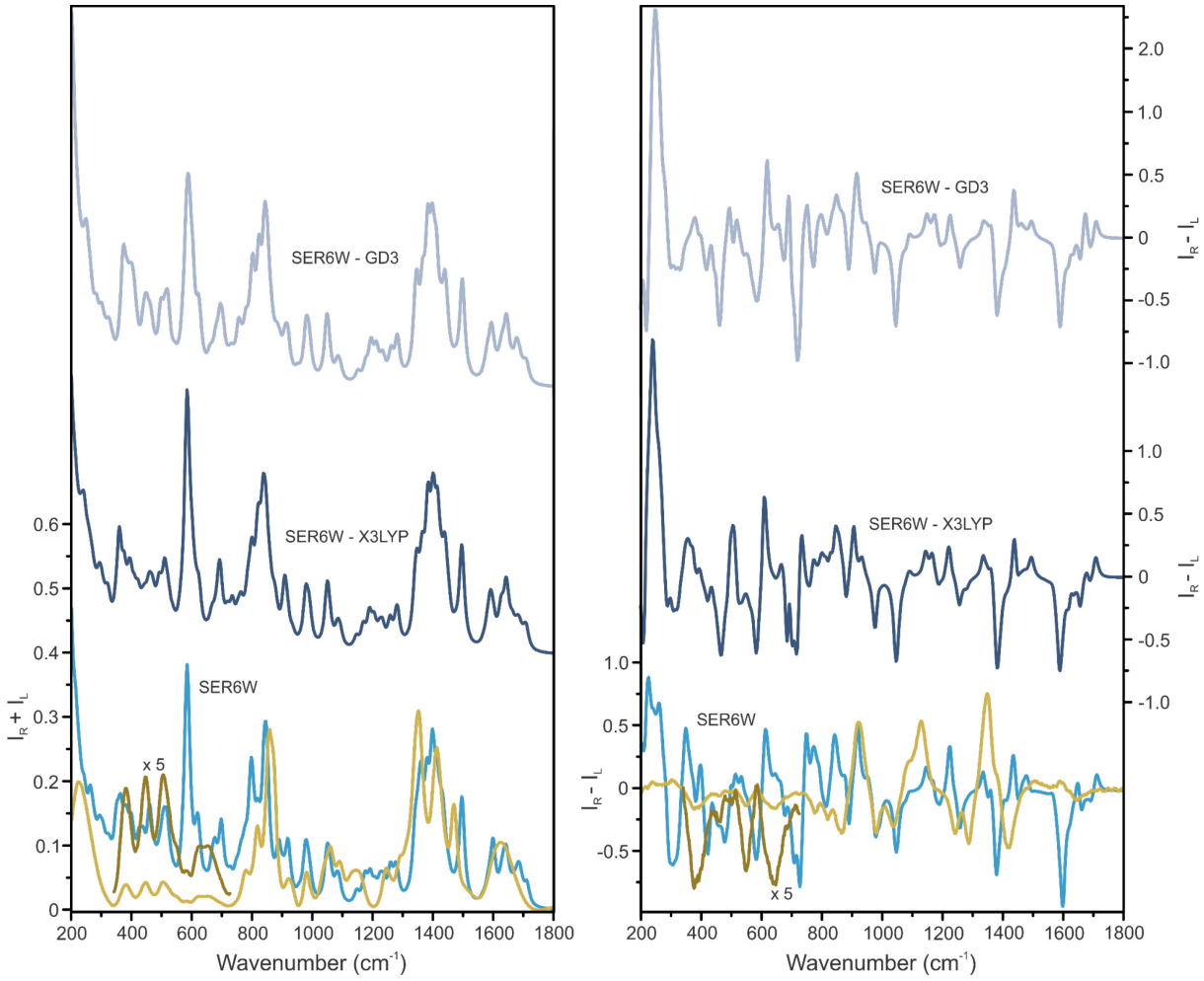

Figure 31. Experimental (yellow) and empirically fit SER6W (blue), SER6W-X3LYP (dark blue), and SER6W-GD3 (blueish grey) Raman (left) and ROA (right) spectra for L-serine. The darker yellow trace represents a five times expansion of the experimental spectrum for the same region.

X3LYP is a functional specifically designed to improve calculations of properties for hydrogen-bonded complexes, including dipole moments and polarizabilities, which are related to the calculation of Raman and ROA tensors.[83] Likewise, dispersion corrections implemented using correction parameters (GD3) should also improve the simulation of these larger serine-water complexes.[84] In reality, there is very little change in the geometry of the complexes and in the simulated spectra. Energetically, one change is that there is a more even energy distribution for the SER6W-GD3 set of complexes than the other two sets (relative free energies and empirical weights shown in Table 3). Conversely, the distribution is very similar between the three sets in terms of empirical weights. Inspection of the Raman and ROA spectra reveals only

very small differences between each, and none offer a noticeable improvement in terms of agreement with the experimental results.

Table 3. List of the relative free energy (ΔG), Boltzmann population (%), and empirical weight (%) for the six-water serine complexes, computed at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory.

Conformer	Relative ΔG (kJ/mol)	Boltzmann Population	Empirical Weight
SER6W1	8.87	2.66%	30.61%
SER6W2	0.00	95.41%	29.76%
SER6W3	10.28	1.50%	17.82%
SER6W4	17.83	0.07%	20.75%
SER6W5	13.88	0.35%	1.06%
SER6W1-X3LYP	10.32	1.49%	34.69%
SER6W2-X3LYP	0.00	95.81%	24.49%
SER6W3-X3LYP	9.49	2.08%	34.87%
SER6W4-X3LYP	14.48	0.28%	5.95%
SER6W5-X3LYP	13.92	0.35%	0.00%
SER6W1-GD3	3.32	15.04%	31.04%
SER6W2-GD3	0.00	57.52%	25.59%
SER6W3-GD3	4.17	10.68%	43.37%
SER6W4-GD3	9.88	1.07%	0.00%
SER6W5-GD3	3.22	15.69%	0.00%

3.5 Conclusion

Based on the accurate replication of the features of the experimental Raman and ROA spectra of L-serine in water using quantum mechanical calculations, the distribution of water-serine complexes has been estimated for conformers solvated with six explicit water molecules. This was ultimately accomplished by employing an empirical fitting procedure for a set of five complexes. In this way, it is proposed that the long-lived clusters responsible for the observed experimental spectra can be identified. In the case of the SER6W complexes, empirical fitting reveals a somewhat balanced distribution over four complexes, with one conformer barely contributing, whereas the Boltzmann population suggests that a single conformer contributes around 95% based on free energies. The empirically fit spectra for this set of complexes gives reasonable agreement with the experiment.

Two methods of estimating hydrogen-bonding and van der Waals interactions in these complexes were used, the X3LYP functional and GD3 corrections. In both cases, the empirically

fit spectra Raman and ROA spectra offer no benefit over the B3LYP functional by itself. Additional corrections should be considered, including Becke-Johnson damping for dispersion corrections.[87] In conclusion, empirical fitting proves to be an effective technique in accurately predicting experimental Raman and ROA spectra for this amino acid.

3.6 Supporting Information

Supporting information, including monomeric Cartesian coordinates can be found in Appendix B.

Chapter 4: Conclusion and Future Applications

The thesis begins by describing the incredibly versatile and powerful techniques of Raman and ROA spectroscopies. Hopefully this was conveyed in the introduction and in the many excellent spectroscopic studies cited herein. Next, a breakdown of the ROA instrument itself and a brief overview of the theory involved in Raman and ROA spectroscopy were covered. To summarize, traditional vibrational spectroscopies only consider the electric dipole moment of the studied molecules when estimating transition frequencies and intensities. For ROA, the magnetic dipole (which is also needed for VCD) and electric quadrupole moments are necessarily considered. Modern quantum mechanical software now implements methods for routine calculation of ROA tensors, making such calculations rather accessible.[53]

Following this summary, a combined experimental and theoretical study of the N-acetyl-L-cysteine molecule in water and methanol is detailed. For this study, the Raman and ROA spectra of several NALC-water complexes were calculated and compared to the experimental results using a combination of visual inspection and empirical fitting of the data sets. In this way, the important experimental spectral features are emphasized when determining which theoretical complexes best replicate the experiment. It was determined that the NALC-4 solvent complexes were likely the most abundant for these solutions and contribute most strongly to the experimental spectra.

Likewise, a similar study of serine in water was carried out in Chapter 4. In this case, a set of serine-6 water complexes was found to best exemplify the experimentally measured Raman and ROA spectroscopy. As an added element, the X3LYP function and dispersion corrections were used to account for the propensity of hydrogen-bonding and dispersion interactions in these larger clusters. However, these results in both cases were comparable to the more standard B3LYP functional with no dispersion corrections.

For both the study of NALC and serine, the creation of many solvent-solute clusters and the subsequent empirical fitting of their Raman and ROA spectra to experimental data proved to be an effective approach. In this way, the longest-lived clusters in these solutions are simulated based on the produced spectral characteristics.

In future work, it seems logical to extends similar studies even further in a few ways. First, the consideration of many more conformers and complexes found using molecular dynamics is possible. The computational expense of performing DFT calculations for additional

species is not prohibitive. For example, geometry optimization, frequency, and Raman and ROA optical tensor calculations for a six water NALC complex at the level of theory used in Chapter 3 takes around 15 days of CPU core time (Intel Xeon at 2.6 GHz). Second, the application of other computational methods that are advantageous for such hydrogen-bonding complexes could be beneficial, such as damping corrections for GD3 dispersion corrections.[87] Third, anharmonic frequency calculations could improve agreement between experimental and calculated frequencies and Raman and ROA intensities. It is possible to perform anharmonic Raman calculations currently using available software, however anharmonic ROA calculations are more challenging and are being developed by research groups.[28,88,89] The fourth and last extension of this work is into establishing a reliable way to describe the geometric alterations of proteins and enzymes in a solution when they interact with a drug or other molecules using ROA spectroscopy. This is complicated due to the large size of proteins and enzymes and the experimental measurement of solutions containing many species. Despite these challenges, there is published work on extending present knowledge and methods towards this subject.[64,90–92]

Appendix A: Long-Lived Solute-Solvent Complexes of N-Acetyl-L-Cysteine as Determined by Raman Optical Activity Spectroscopy

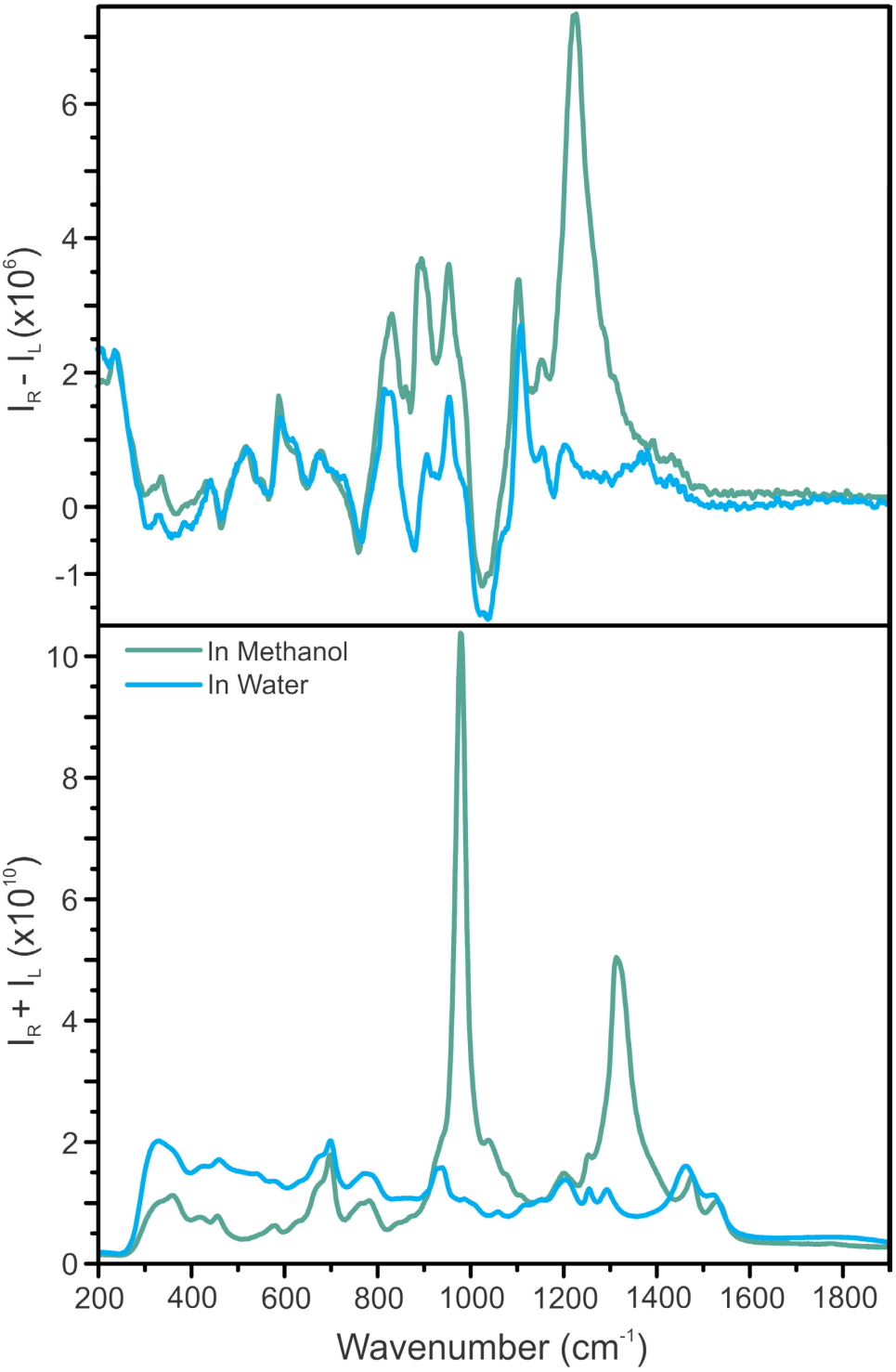

Figure 32. Raw experimental Raman (bottom) and ROA (top) spectra of the NALC molecule dissolved in water (blue) and methanol (green).

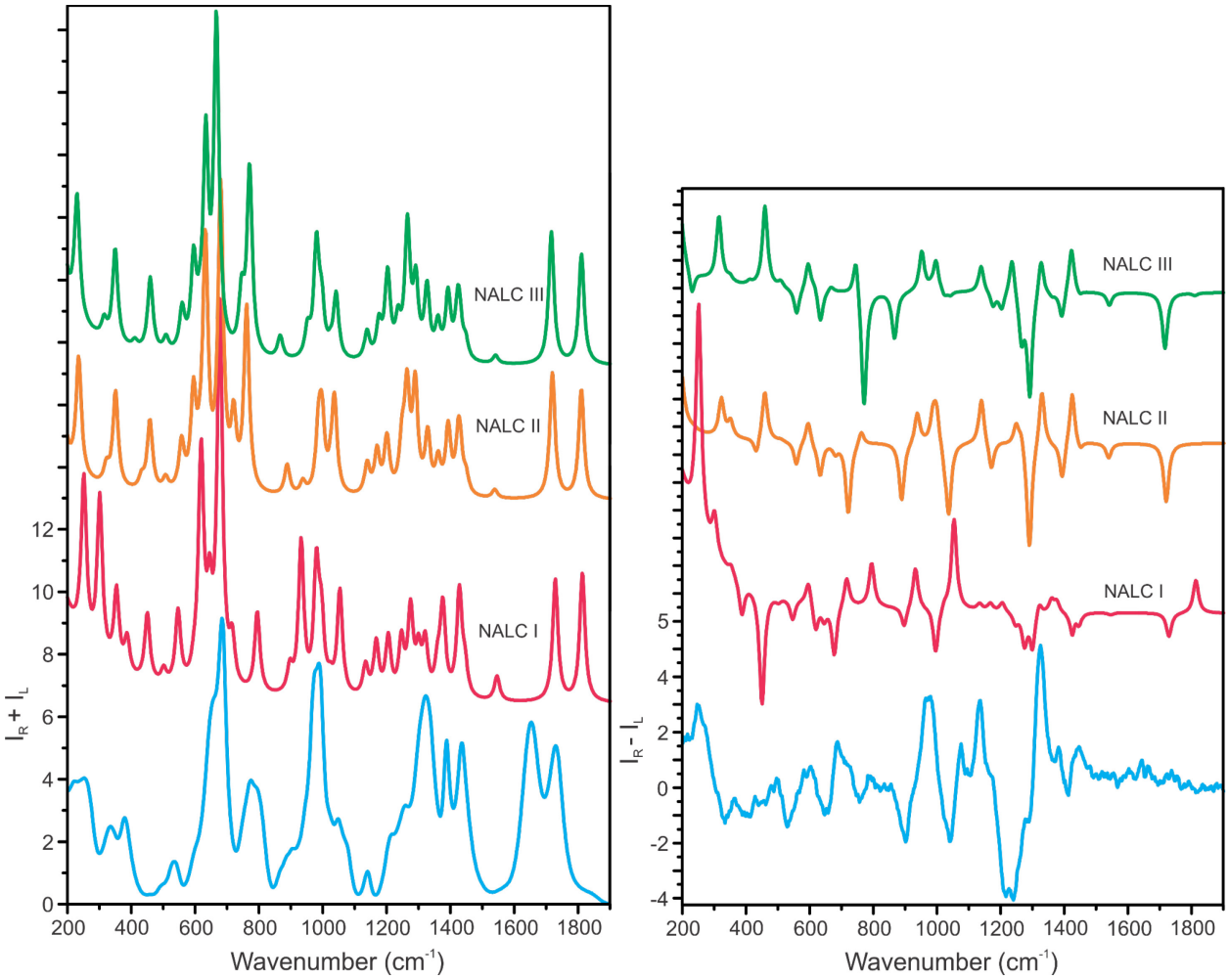

Figure 33. Experimental NALC in water (blue), individual monomer conformations (various colours) Raman (left) and ROA (right) spectra of NALC (with PCM:Water).

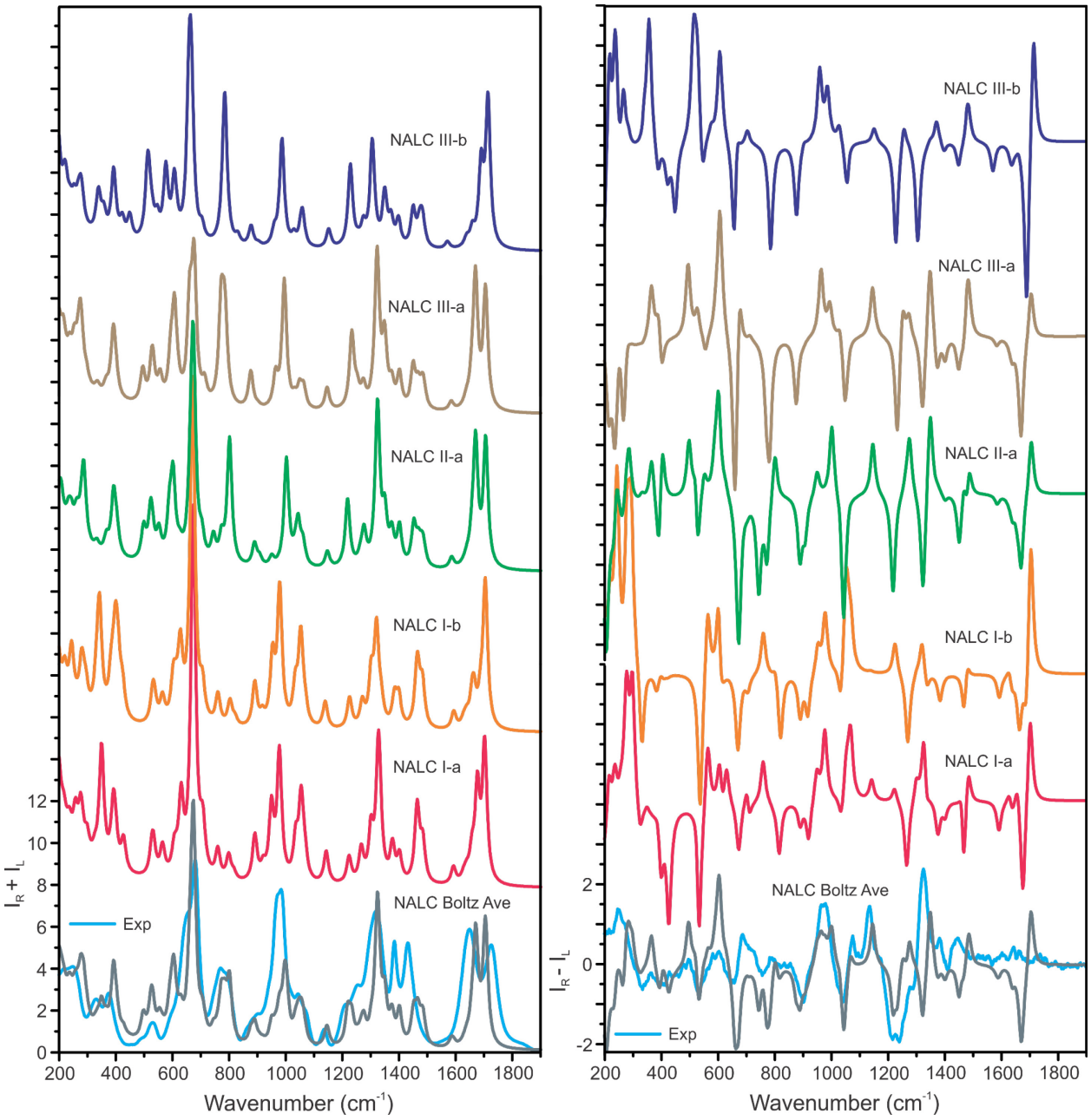

Figure 34. Experimental NALC in water (blue), individual NALC-4 Water complexes (various colours), and Boltzmann population weighted (grey) Raman (left) and ROA (right) spectra.

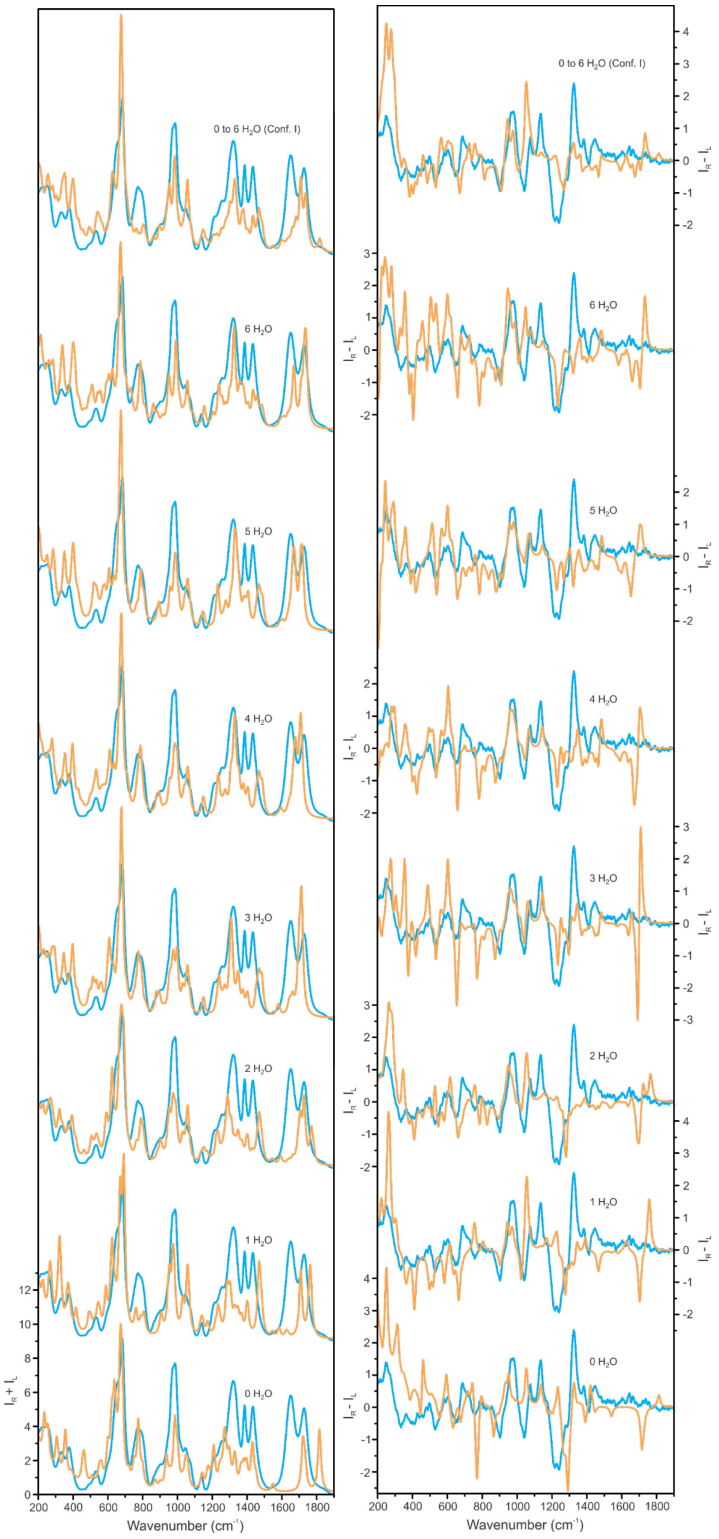

Figure 35. Experimental (blue) and empirically fit (orange) Raman (left) and ROA (right) spectra of the NALC Water complexes for zero to six explicit solvent molecules, as well as an average of the the zero to six explicit solvent complexes just of conformer I of NALC.

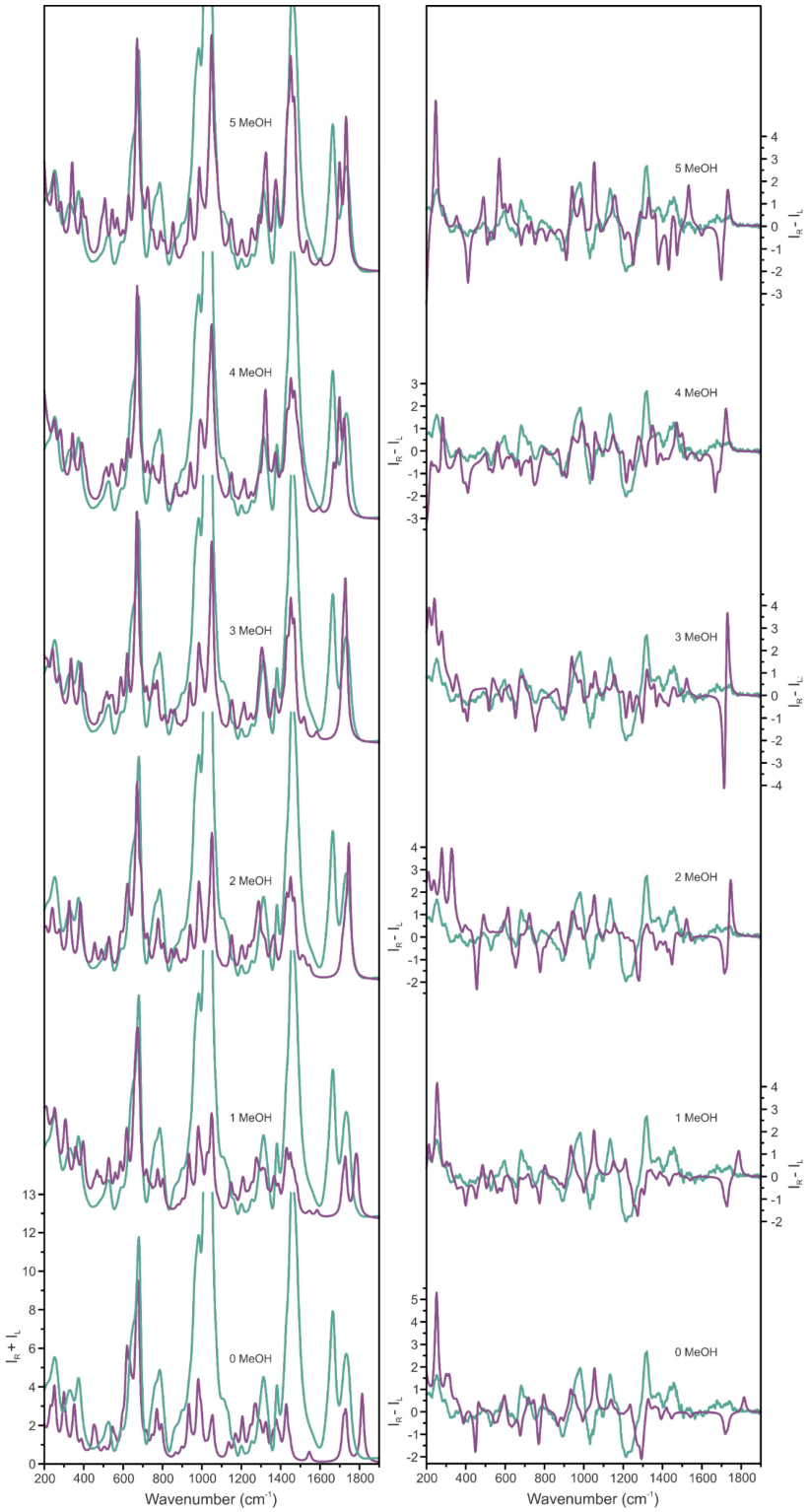

Figure 36. Experimental (green) and empirically fit (purple) Raman (left) and ROA (right) spectra of the NALC MeOH complexes for zero to five explicit solvent molecules.

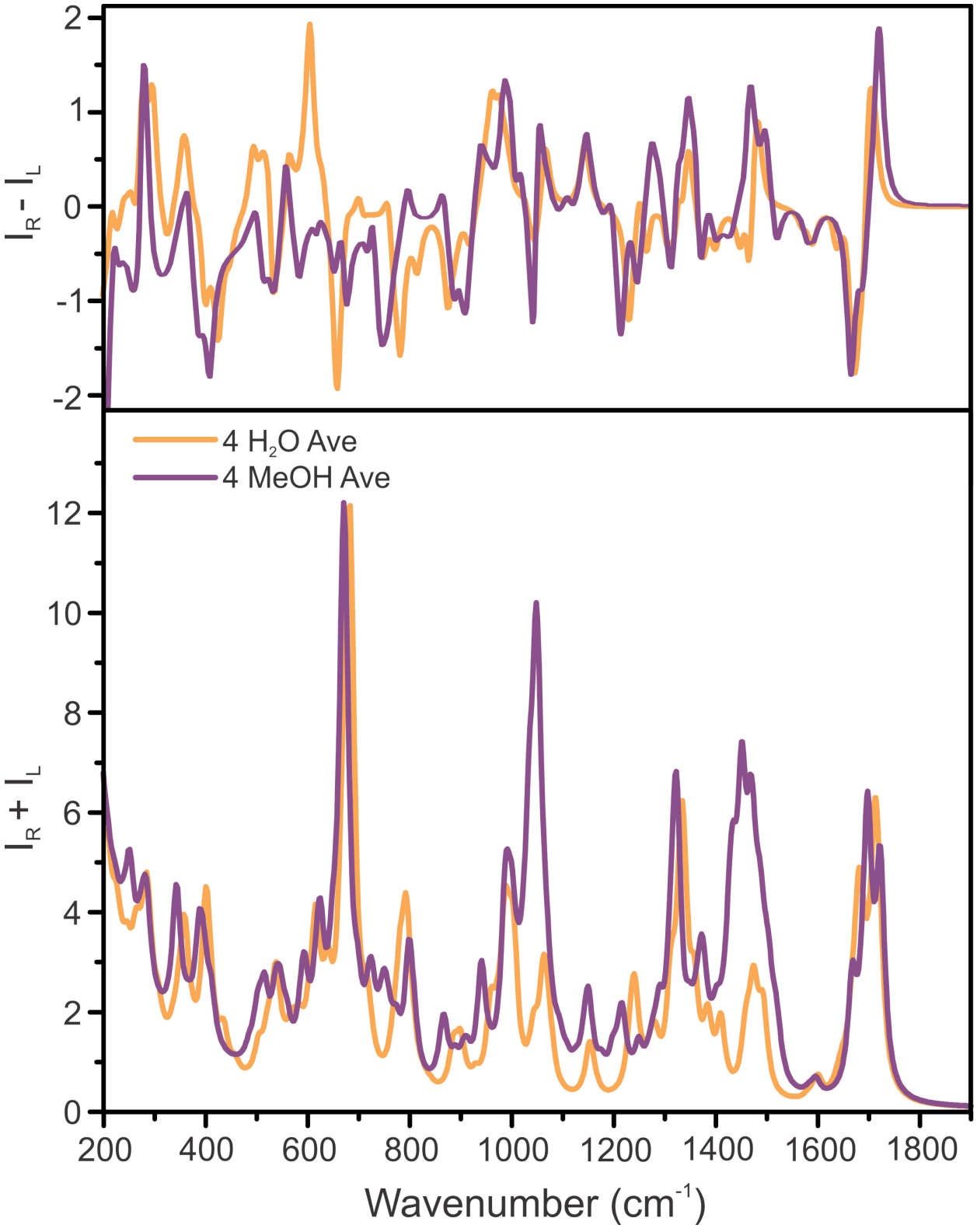

Figure 37. Emperically fit (orange) Raman (bottom) and ROA (top) spectra of the NALC-4 Water complexes and the emperically fit NALC-4 MeOH complexes (purple).

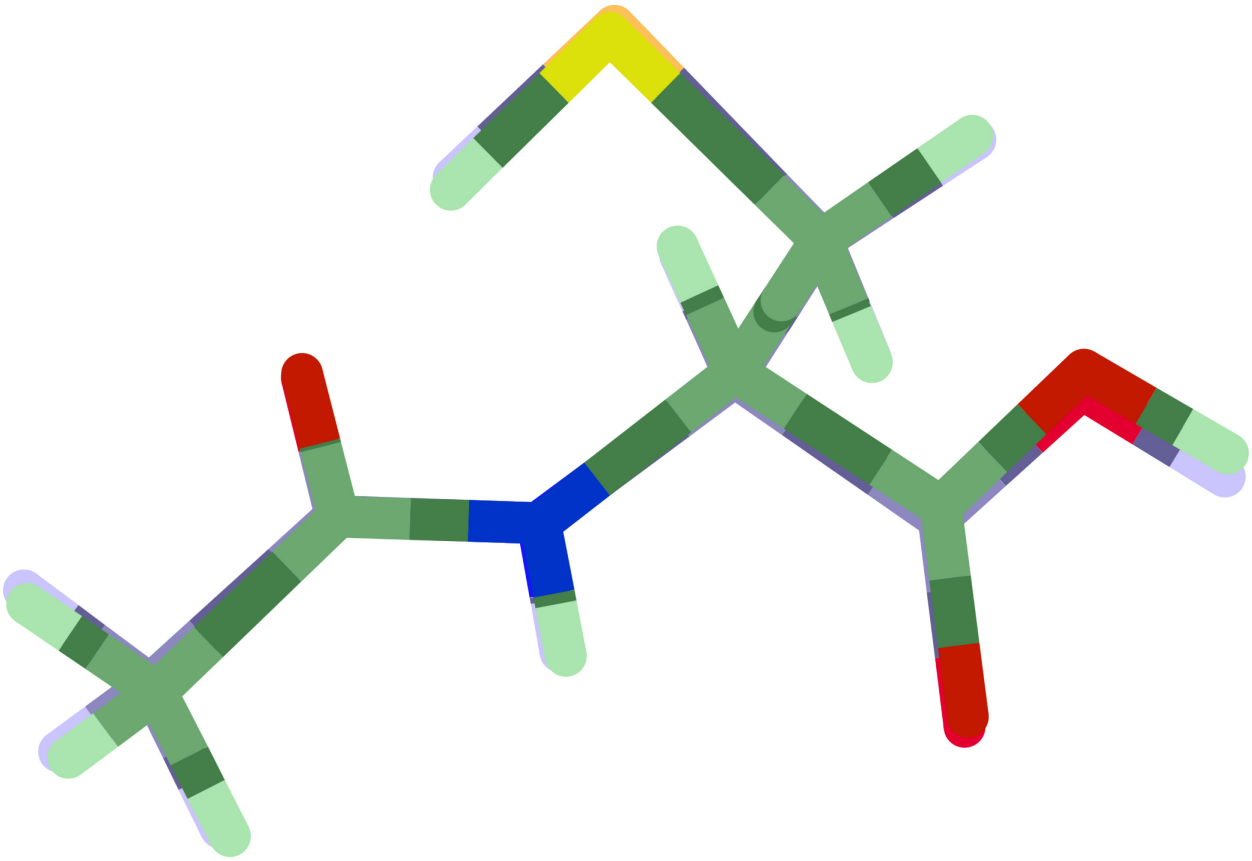

Figure 38. Stick figure overlay of the NALC-I-4 MeOH-a conformer (tinted green/front) and the NALC-4 H2O-1 conformer (tinted blue/back).

Table 4. List of the relative free energy (ΔG), Boltzmann population (%), and empirical weight (%) for each conformer in water and methanol, separated by a number of solvent molecules, computed at the B3LYP/cc-pVTZ (PCM:Water) level of theory.

Conformer of NALC-Water Complex	Relative ΔG (kJ/mol)	Boltzmann Population	Empirical Weight
I-0Water	0	40.14%	37.30%
II-0Water	0.53	32.37%	0.33%
III-0Water	0.94	27.49%	62.37%
I-1Water-a	0.38	46.07%	37.49%
I-1Water-b	0	53.72%	0.00%
I-1Water-c	13.70	0.21%	62.51%
I-2Water-a	1.70	18.78%	19.00%
I-2Water-b	2.26	14.97%	0.00%

I-2Water-c	0	37.31%	14.71%
I-2Water-d	26.15	0.00%	0.00%
I-2Water-e	25.27	0.00%	37.90%
I-2Water-f	3.85	7.89%	0.00%
II-2Water-a	2.98	11.20%	0.00%
III-2Water-a	3.30	9.86%	28.40%
I-3Water-a	1.42	29.79%	42.03%
II-3Water-a	0	52.80%	0.00%
III-3Water-a	2.75	17.41%	57.97%
I-4Water-a	1.79	20.72%	39.57%
I-4Water-b	7.90	1.76%	3.00%
II-4Water-a	0	42.69%	0.00%
IIII-4Water-a	0.73	31.80%	37.67%
IIII-4Water-b	6.55	3.03%	19.76%
I-5Water-a	5.34	6.94%	37.31%
I-5Water-b	0	59.88%	9.34%
II-5Water-a	7.21	3.26%	14.89%
III-5Water-a	1.72	29.91%	38.47%
I-6Water-a	0	91.45%	50.99%
III-6Water-a	5.87	8.55%	49.01%

Conformer of NALC-MeOH Complex	**Relative ΔG (kJ/mol)**	**Boltzmann Population**	**Empirical Weight**
I-0MeOH	0.74	33.46%	59.19%
II-0MeOH	0	45.02%	0.00%
III-0MeOH	1.83	21.51%	40.81%
I-1MeOH-a	3.18	7.25%	22.58%
I-1MeOH-b	2.66	8.92%	0.00%
I-1MeOH-c	17.46	0.02%	45.06%
II-1MeOH-a	0	26.14%	8.02%
II-1MeOH-b	0.40	22.25%	0.00%

III-1MeOH-a	1.45	14.55%	16.32%
III-1MeOH-b	0.56	20.86%	8.02%
I-2MeOH-a	3.28	9.27%	22.10%
I-2MeOH-b	2.26	14.01%	38.00%
II-2MeOH-a	2.14	14.71%	0.00%
II-2MeOH-b	0	34.84%	0.00%
III-2MeOH-a	2.80	11.27%	31.04%
III-2MeOH-b	1.94	15.91%	8.87%
I-3MeOH-a	2.91	7.76%	40.57%
I-3MeOH-b	0.78	18.36%	16.69%
II-3MeOH-a	2.02	11.15%	0.00%
II-3MeOH-b	0	25.15%	0.00%
III-3MeOH-a	1.15	15.80%	0.00%
III-3MeOH-b	0.36	21.78%	42.74%
I-4MeOH-a	4.56	3.67%	0.00%
I-4MeOH-b	2.80	7.49%	50.02%
II-4MeOH-a	1.53	12.50%	0.00%
II-4MeOH-b	0.55	18.54%	0.41%
II-4MeOH-c	0	23.14%	35.93%
III-4MeOH-a	2.84	7.37%	0.00%
III-4MeOH-b	1.87	10.88%	2.55%
III-4MeOH-c	0.85	16.41%	11.09%
I-5MeOH-a	2.58	15.97%	79.05%
I-5MeOH-b	1.81	21.78%	0.00%
II-5MeOH-a	2.43	16.96%	20.95%
II-5MeOH-b	0	45.29%	0.00%

$$\frac{d\sigma}{d\Omega} = \frac{\pi^2}{\epsilon_o^2}(\tilde{\nu}_{in} - \tilde{\nu}_p)^4 \frac{h}{8\pi^2 c\tilde{\nu}_p} \frac{45a_p'^2 + 7\gamma_p'^2}{45} \frac{1}{1 - exp[-\tilde{\nu}_p hc/k_B T]} \quad (6)$$

Equation 6. Differential Raman scattering cross-section used to calculate Raman intensity from theoretically calculated transitions. See Reference 63 for further details.

$$BP_i = \frac{e^{-\Delta G_i/k_B T}}{\sum_{j=1}^{n} e^{-\Delta G_j/k_B T}} \quad (7)$$

Equation 7. Standard equation used to calculate Boltzmann populations using the determined free energy at room temperature for a given set of conformers.

Cartesian coordinates for the NALC I conformer calculated at the B3LYP/cc-pVTZ (PCM:Water) level of theory using ultra tight integrals, ultra fine grid, and tight SCF convergence criteria.

0 1

N	0.77058800	-0.28564900	-0.45668700
H	0.82850000	-0.48578200	-1.44180100
C	-0.54218900	-0.18956400	0.13200300
H	-0.40263700	-0.03916900	1.20078900
C	-1.36055800	0.98119900	-0.44231800
H	-2.35260300	0.98380900	0.00207400
O	-2.42788700	-1.51170900	0.69136800
H	-2.90181000	-2.34135500	0.52279700
C	-1.31459600	-1.49656800	-0.05796000
O	-0.98845100	-2.39591300	-0.79084100
C	1.91179700	-0.34793600	0.28272200
O	1.91520300	-0.24451500	1.50430200
C	3.18740000	-0.54068900	-0.50362700
H	3.65405800	-1.47304100	-0.18646300
H	3.03328600	-0.56821800	-1.58011600

H	3.87302900	0.27015500	-0.26102000
H	-1.47513300	0.87777300	-1.51974000
S	-0.66817200	2.63323200	-0.06739500
H	0.45496200	2.51606700	-0.79787700

Cartesian coordinates for the NALC II conformer calculated at the B3LYP/cc-pVTZ (PCM:Water) level of theory using ultra tight integrals, ultra fine grid, and tight SCF convergence criteria.

0 1

N	0.91541400	-0.34074300	-0.29251800
H	0.77953200	-0.68848700	-1.22832600
C	-0.24245900	-0.33558600	0.57350300
H	0.03863300	-0.77312700	1.53554800
C	-0.77125600	1.07774900	0.87685300
H	0.02476900	1.65705800	1.33442400
O	-2.38294500	-1.32278100	0.75274900
H	-3.03846200	-1.89135100	0.31903400
C	-1.30922800	-1.22878200	-0.04007400
O	-1.20772800	-1.78910700	-1.10403400
C	2.16879200	-0.11414000	0.17276100
O	2.38420400	0.19195600	1.34238500
C	3.28404500	-0.25370500	-0.83675600
H	3.96933700	-1.02858500	-0.49393800
H	2.93427400	-0.50369900	-1.83597000
H	3.83919400	0.68278100	-0.87673000
H	-1.59779500	1.01633700	1.58017700
S	-1.43633100	1.98573100	-0.57064000
H	-0.27646800	2.10446900	-1.24148800

Cartesian coordinates for the NALC III conformer calculated at the B3LYP/cc-pVTZ (PCM:Water) level of theory using ultra tight integrals, ultra fine grid, and tight SCF convergence criteria.

0 1

N	0.98241700	-0.34047800	-0.30587800
H	0.87270500	-0.66470600	-1.25288000
C	-0.18630200	-0.38430500	0.54154600
H	0.10072800	-0.80092300	1.51104700
C	-0.78696800	1.00524600	0.83249700
H	0.00059400	1.62713700	1.24954200
O	-2.27894500	-1.47070100	0.70373300
H	-2.90244800	-2.07141200	0.26586900
C	-1.20048300	-1.33151100	-0.07739900
O	-1.06476800	-1.88604500	-1.14040500
C	2.21214300	-0.03432300	0.17198000
O	2.39514800	0.28441800	1.34447000
C	3.34523900	-0.10262500	-0.82494500
H	4.08842000	-0.81084800	-0.45974400
H	3.02708800	-0.40100800	-1.82151700
H	3.82194100	0.87552600	-0.88086600
H	-1.58157200	0.92555500	1.56811500
S	-1.40836800	1.91749100	-0.63277300
H	-2.59810900	1.29744600	-0.73390200

Appendix B: Raman Optical Activity Spectroscopy of Serine-Water Complexes

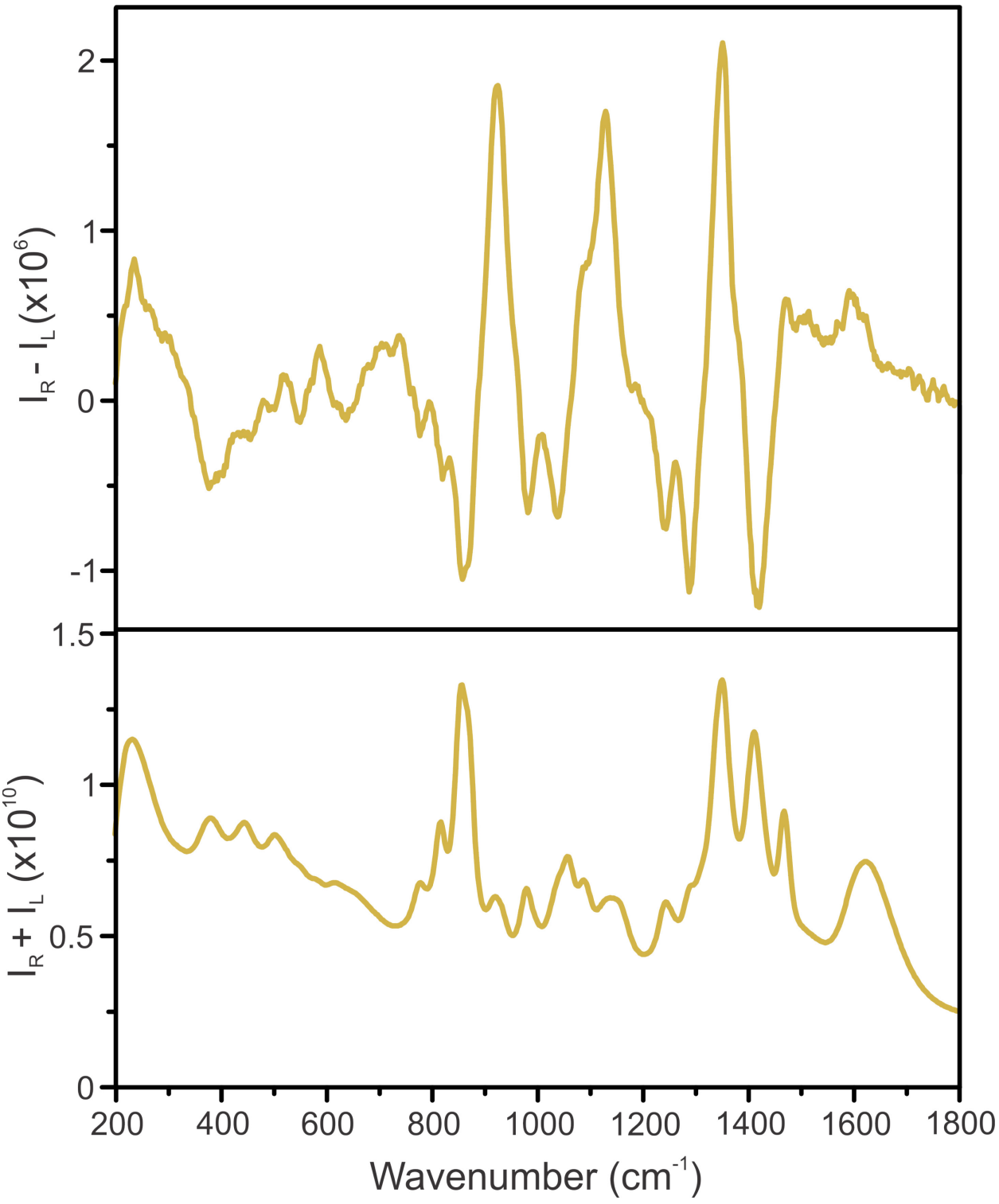

Figure 39. Raw experimental Raman (bottom) and ROA (top) spectra of L-serine in water.

Cartesian coordinates for the SER0W1 conformer calculated at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory using ultra tight integrals, ultra fine grid, and tight SCF convergence criteria.

0 1

C	-1.40207000	-0.36342400	0.64504600
O	-2.05846300	-0.48797300	-0.60986600
C	-0.08725400	0.41107400	0.56796100
N	-0.26393300	1.65046200	-0.26697200
C	1.15979500	-0.35072800	0.04223300
O	1.07047800	-1.58549300	-0.15732700
O	2.19123600	0.34793000	-0.12593000
H	-1.21896200	2.03215500	-0.17897900
H	0.16098900	0.74230400	1.57737700
H	-1.20355900	-1.33793300	1.09030500
H	-2.08096800	0.17386400	1.30840700
H	-1.77311500	-1.32092500	-1.04394100
H	-0.12404500	1.42113700	-1.27544100
H	0.43834800	2.35892900	0.00461600

Cartesian coordinates for the SER0W2 conformer calculated at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory using ultra tight integrals, ultra fine grid, and tight SCF convergence criteria.

0 1

C	1.25156100	-0.78449000	-0.43561000
C	0.15536400	0.25253700	-0.71218400
C	-1.17223500	-0.11757600	-0.01211800
O	-1.60059100	0.66772200	0.87842300
O	1.73756300	-0.67852400	0.89582700
N	0.64295500	1.58978700	-0.25259000

O	-1.67862200	-1.19205700	-0.38816200
H	1.54207400	1.85432200	-0.71069900
H	-0.01990400	0.31194100	-1.78459900
H	0.84730400	-1.77411700	-0.64264000
H	2.09370900	-0.60822000	-1.10439000
H	1.21423500	-1.24480500	1.51672700
H	0.82185700	1.52841800	0.74978300
H	-0.07489500	2.32400900	-0.38528700

Cartesian coordinates for the SER0W3 conformer calculated at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory using ultra tight integrals, ultra fine grid, and tight SCF convergence criteria.

0 1

C	-1.37690100	-0.48250800	0.57960700
C	-0.10123700	0.36383200	0.60405700
C	1.16651800	-0.33170200	0.03634200
O	2.12097400	0.43594800	-0.28619800
O	-2.04688100	-0.48174600	-0.67437100
N	-0.31577100	1.66528400	-0.11728600
O	1.15143500	-1.57164000	-0.04452800
H	-1.27341100	2.02186400	0.02342600
H	0.11467100	0.60786900	1.64546400
H	-1.11308400	-1.49126800	0.89346100
H	-2.08880100	-0.07400600	1.29609300
H	-1.55314100	-0.97278200	-1.36115900
H	-0.19314100	1.53117500	-1.14554900
H	0.38280100	2.36192500	0.19000600

Cartesian coordinates for the SER0W4 conformer calculated at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory using ultra tight integrals, ultra fine grid, and tight SCF convergence criteria.

0 1

C	-1.46747300	-0.26568400	0.50632500
C	-0.11947200	0.45271700	0.59714000
C	1.09084900	-0.36341900	0.07345900
O	2.02413500	0.27463800	-0.47609100
O	-1.81563300	-0.67487400	-0.81360500
N	-0.19078600	1.77931600	-0.10041600
O	1.05549400	-1.60286400	0.25144100
H	-0.97915500	2.34007400	0.26752300
H	0.06710400	0.66462700	1.65165400
H	-1.43990600	-1.12128300	1.17980200
H	-2.24574400	0.41176400	0.85953800
H	-1.56299300	-1.61682100	-0.91833900
H	-0.32344200	1.63531600	-1.12376300
H	0.68424900	2.31422700	0.03098800

Cartesian coordinates for the SER0W5 conformer calculated at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory using ultra tight integrals, ultra fine grid, and tight SCF convergence criteria.

0 1

C	-1.45842100	0.00635900	0.66106100
C	-0.05913000	0.59243200	0.50707600
C	1.02585900	-0.44404800	0.09658600
O	2.21631000	-0.03815400	0.12931800
O	-1.89627300	-0.72346800	-0.48137600
N	-0.01817000	1.71236900	-0.50226100
O	0.63661600	-1.57644800	-0.27321700
H	-0.87743500	2.29694000	-0.47207300
H	0.22835800	1.03298600	1.46051600

H	-1.44882900	-0.63819800	1.54363800
H	-2.18127000	0.80197700	0.83536600
H	-1.17287700	-1.35875900	-0.64635100
H	0.05845300	1.35788200	-1.47493800
H	0.81772100	2.29669400	-0.31648100

Cartesian coordinates for the SER4W1 conformer calculated at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory using ultra tight integrals, ultra fine grid, and tight SCF convergence criteria.

0 1

C	-2.05967200	-1.35170100	0.63240200
O	-2.10070300	-1.74809300	-0.72895900
C	-1.48296700	0.05701700	0.77183500
N	-2.30859400	0.97764200	-0.08071100
C	0.01488500	0.15622300	0.37959800
O	0.73873200	-0.76975400	0.82194600
O	0.37622900	1.13278700	-0.31335400
O	-1.03424200	3.46080700	-0.22363500
O	0.41777400	-2.90168500	-0.94367400
O	3.49020500	-0.58571900	1.02959100
O	3.11725000	1.23609200	-1.15029500
H	-3.26944300	1.02063000	0.25597500
H	-1.58037900	0.39806300	1.80237200
H	-1.46360000	-2.02998700	1.24099800
H	-3.08236400	-1.36334900	1.01355300
H	-1.25527900	-2.19511900	-0.94085500
H	-2.33261300	0.60195800	-1.03054100
H	-1.92722100	1.94989900	-0.11031000
H	3.46542700	0.64978200	-0.45793400
H	2.16676700	1.24880300	-0.94276500
H	2.51093700	-0.64798600	0.98433000
H	3.66995800	-0.08142500	1.82932400

H	-0.96152200	3.94418000	0.60636000
H	-0.27654600	2.84319200	-0.24110300
H	0.73905600	-2.22882800	-0.31288200
H	0.88153300	-2.71801700	-1.76751100

Cartesian coordinates for the SER4W2 conformer calculated at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory using ultra tight integrals, ultra fine grid, and tight SCF convergence criteria.

0 1

C	1.08348200	1.90306300	0.90951800
C	1.44835700	0.41042900	0.93482600
C	0.21660600	-0.48219500	0.66607800
O	0.18663500	-1.14464000	-0.40712100
O	0.81795800	2.33761900	-0.41226000
N	2.51171200	0.19407000	-0.10361200
O	-0.66049900	-0.43057400	1.55027300
O	-1.75684600	2.00752300	-1.28068400
O	-2.34448200	-0.66731600	-1.47099100
O	2.48813000	-2.54769200	-0.81350200
O	-3.23733200	-1.19965000	1.14477000
H	3.40559600	0.57590700	0.20163800
H	1.86731200	0.14535900	1.90252100
H	0.23760000	2.06080600	1.57770300
H	1.92765100	2.48601800	1.27973500
H	-0.12507600	2.18886600	-0.64904400
H	2.22745500	0.69749500	-0.94838300
H	2.64170500	-0.80834200	-0.34922300
H	2.74220200	-2.77322000	-1.71480000
H	1.54079500	-2.30494900	-0.84603800
H	-1.44568600	-0.96165100	-1.21377100
H	-2.87260500	-0.89346800	-0.68058300
H	-3.38810100	-2.11599000	1.39721200
H	-2.29052100	-1.01111100	1.35401400

H	-2.44207300	2.45940900	-0.77798200
H	-2.04742700	1.06643500	-1.35412600

Cartesian coordinates for the SER4W3 conformer calculated at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory using ultra tight integrals, ultra fine grid, and tight SCF convergence criteria.

0 1

C	-1.60293200	0.70498700	-1.18658700
C	-1.18723700	-0.76533400	-1.10790200
C	0.32246900	-0.93907800	-0.79724400
O	0.59823800	-1.63337200	0.21675100
O	-1.43350900	1.39954400	0.03848900
N	-1.95622000	-1.51995300	-0.05499700
O	1.10195700	-0.36886300	-1.58299500
O	2.96782900	-1.27101200	1.54539100
O	-3.24660200	0.27843200	1.70054800
O	3.03466700	0.94794600	-0.18163600
O	0.95393600	2.58088200	0.45343000
H	-2.46991700	-0.90514800	0.61350800
H	-1.39039300	-1.23385200	-2.06909200
H	-1.01969000	1.16290400	-1.98500500
H	-2.65763100	0.76973500	-1.45685800
H	-0.52453800	1.77640700	0.12582100
H	-1.22408600	-2.02342000	0.47435900
H	-2.61429700	-2.18456100	-0.45199500
H	3.60529400	-1.92279300	1.23676200
H	2.13382700	-1.48090600	1.06532800
H	3.18825500	0.29452600	0.52855500
H	2.46019300	0.44465700	-0.80070700
H	-2.99609000	0.15003300	2.62155000
H	-2.60948200	0.92601500	1.33732300
H	1.05270500	3.36632000	-0.09391800
H	1.75345900	2.02785700	0.26992300

Cartesian coordinates for the SER4W4 conformer calculated at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory using ultra tight integrals, ultra fine grid, and tight SCF convergence criteria.

0 1

C	-2.06810300	-1.34381000	0.63306600
C	-1.46932000	0.05330700	0.79686700
C	0.02538600	0.14209100	0.38994700
O	0.39371700	1.13415500	-0.27600400
O	-2.13556200	-1.70585100	-0.73693400
N	-2.29398900	1.00210200	-0.02550400
O	0.74059500	-0.80695900	0.79683300
O	-0.98738400	3.48032900	-0.08729100
O	0.36033400	-2.89089900	-1.00939900
O	3.49106100	-0.64442800	1.01547800
O	3.13469900	1.24255500	-1.11130500
H	-3.24022800	1.07565700	0.34542500
H	-1.54932500	0.37254700	1.83576900
H	-1.47250000	-2.04490600	1.21561600
H	-3.08497300	-1.35072900	1.02921900
H	-1.29955400	-2.15898800	-0.97260900
H	-2.36309900	0.62523900	-0.97302100
H	-1.88637300	1.96276100	-0.07267800
H	-0.24075800	2.86445900	-0.22045800
H	-1.07683000	3.96763900	-0.91337400
H	3.47854000	0.63530200	-0.43499300
H	2.18400100	1.25469100	-0.90459900
H	3.67042000	-0.16401600	1.82988400
H	2.51152100	-0.69929600	0.96457900
H	0.69939900	-2.23751600	-0.36720500
H	0.82023000	-2.69830600	-1.83333900

Cartesian coordinates for the SER4W5 conformer calculated at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory using ultra tight integrals, ultra fine grid, and tight SCF convergence criteria.

0 1

C	2.54874300	0.27041100	-0.64085100
C	1.21548800	0.99918400	-0.49597000
C	-0.00581700	0.21240500	-1.06216500
O	-1.11701900	0.80035600	-0.99391300
O	2.52151600	-1.06327300	-0.14356200
N	0.92777300	1.32804700	0.94663900
O	0.20151400	-0.93356900	-1.51648400
O	-1.35784000	2.93298300	0.66668400
O	-3.30229500	-0.67071500	-0.19325500
O	0.62940900	-1.19433000	2.10059400
O	-1.24841000	-2.44646100	0.47441500
H	1.71457700	1.81912100	1.36728100
H	1.27950500	1.95090000	-1.02212100
H	2.81575900	0.28060000	-1.70036000
H	3.33055900	0.80287000	-0.09885100
H	1.82290700	-1.48785900	-0.67388200
H	0.75612000	0.44929000	1.49138700
H	0.09251000	1.94700800	1.02029500
H	-1.48276100	2.26226200	-0.03927000
H	-2.07610600	2.79068300	1.29223100
H	-0.77099100	-2.18967000	-0.33276800
H	-2.07921500	-1.93532600	0.39006400
H	-3.80156900	-1.03177000	-0.93283400
H	-2.59128400	-0.12852500	-0.59974500
H	1.42958100	-1.45797400	1.62044900
H	-0.09948500	-1.65986000	1.62973900

Cartesian coordinates for the SER6W1 conformer calculated at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory using ultra tight integrals, ultra fine grid, and tight SCF convergence criteria.

0 1

C	1.62831600	-1.08666100	-1.20478700
O	1.68988700	-1.79496000	0.02634700
C	0.97295200	0.28808300	-1.08187800
N	1.50506500	1.00978800	0.12659700
C	-0.57837600	0.33672800	-1.02823500
O	-1.22067400	-0.72139500	-1.22892300
O	-1.07811400	1.46775800	-0.80172900
O	0.58557400	3.61363500	-0.51551500
O	3.89915000	-0.39117800	0.92726900
O	-0.66680700	-3.14959000	-0.07814300
O	-3.72251000	-0.89903500	-0.05546900
O	-0.10111800	0.24919300	2.31564100
O	-2.58569600	1.11771700	1.52988400
H	2.48862700	0.75863800	0.31379200
H	1.26631100	0.87338500	-1.95458700
H	1.09688400	-1.65599400	-1.96686900
H	2.65516500	-0.95041400	-1.54632200
H	0.88992800	-2.35651300	0.11786900
H	0.95886900	0.73493300	0.97234700
H	1.40796600	2.03075600	-0.00409700
H	-3.12347100	0.37822000	1.18347600
H	-2.17778900	1.45875000	0.71129100
H	-2.91388300	-0.79601100	-0.60305600
H	-4.43779800	-0.52766900	-0.58171800
H	0.88462600	3.98748500	-1.35072900
H	-0.14048700	2.99314000	-0.74391100

H	3.26811500	-1.13161800	0.89440800
H	4.10616200	-0.26738100	1.85937000
H	-0.10081500	-0.70878500	2.41087000
H	-1.03639900	0.49760900	2.12166100
H	-0.99706500	-2.34463200	-0.52893600
H	-1.16527400	-3.20848500	0.74347500

Cartesian coordinates for the SER6W2 conformer calculated at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory using ultra tight integrals, ultra fine grid, and tight SCF convergence criteria.

0 1

C	-0.99056800	-0.78889800	1.22268000
C	-0.55692000	0.65843900	0.95672700
C	0.93883500	0.74698400	0.57663400
O	1.22290300	1.16635500	-0.57951400
O	-1.08864800	-1.52841300	0.01286600
N	-1.41555800	1.21639400	-0.13351900
O	1.72590600	0.35922900	1.46123100
O	1.20953300	-2.70147500	-0.72437300
O	2.96010000	-0.78268100	-1.56194800
O	-0.06334700	3.57968200	-0.95432000
O	-3.75963800	-1.90083800	-0.57228200
O	-4.13550500	0.69262200	0.26175200
O	4.29072900	-0.29163200	0.84409000
H	-2.43000500	1.15991000	0.10305100
H	-0.72573700	1.25375800	1.85194700
H	-0.28099200	-1.23739200	1.91622200
H	-1.97656400	-0.79193900	1.68674300
H	-0.22682000	-1.95716000	-0.21841500
H	-1.26833300	0.64799400	-0.96764700

H	-1.14149700	2.18458100	-0.37745400
H	-0.24455000	3.82047600	-1.86896200
H	0.60214300	2.86458500	-0.99611300
H	2.37886500	-0.01639600	-1.37896900
H	3.60706900	-0.71931600	-0.83015400
H	4.82146000	0.50806100	0.77234200
H	3.39150700	0.01710300	1.11286400
H	-4.13993800	-0.24997600	-0.02149700
H	-4.52123800	0.70770600	1.14350000
H	-2.78739900	-1.92498300	-0.45942900
H	-3.91063400	-1.99399200	-1.51855700
H	1.63707000	-3.18336300	-0.00920800
H	1.88015300	-2.04636300	-1.04189100

Cartesian coordinates for the SER6W3 conformer calculated at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory using ultra tight integrals, ultra fine grid, and tight SCF convergence criteria.

0 1

C	1.62474600	-0.88629400	-1.43928200
C	1.06650600	0.49849300	-1.10025200
C	-0.47506800	0.64119700	-1.22853500
O	-0.99124500	1.62816800	-0.62541600
O	1.53887300	-1.82116900	-0.37158900
N	1.48809100	0.91524500	0.28131300
O	-1.07318000	-0.20518900	-1.91417800
O	-2.77465200	0.86428300	1.24496200
O	0.67560400	3.63820600	0.17800500
O	3.71379200	-0.77910400	0.99836200
O	-3.07168100	-1.40451800	-0.45989000
O	-0.91404500	-2.58020000	0.66364500

O	-0.31282600	-0.25096600	2.14687700
H	2.43914200	0.58271000	0.50255000
H	1.50675000	1.21310800	-1.79753800
H	1.10914200	-1.24425900	-2.32906600
H	2.68587600	-0.79574200	-1.66913300
H	0.61782500	-2.09464500	-0.18868300
H	0.84998800	0.49640500	0.99387600
H	1.42819300	1.94236700	0.37708300
H	-3.36988900	1.53047000	1.60214500
H	-2.26632200	1.29197100	0.51654200
H	-3.24464700	-0.70603500	0.19561800
H	-2.50745100	-0.94942300	-1.12135000
H	-0.07766000	3.08146700	-0.11742000
H	0.40798400	4.00930200	1.02511900
H	3.81893000	-0.86669300	1.95135300
H	3.02688800	-1.42384600	0.74902300
H	-0.48651700	-1.15016100	1.81142200
H	-1.16849100	0.20423800	2.04287500
H	-1.04283600	-3.48980000	0.94944600
H	-1.76576300	-2.27460000	0.25914000

Cartesian coordinates for the SER6W4 conformer calculated at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory using ultra tight integrals, ultra fine grid, and tight SCF convergence criteria.

0 1

C	0.68363900	-1.86115900	-0.49531400
C	0.93158400	-0.40521900	-0.89558700
C	-0.35093800	0.44393600	-1.09375200
O	-0.29060000	1.66203500	-0.78922300
O	-0.07069300	-2.00938700	0.70463500

N	1.84354500	0.25688500	0.09511000
O	-1.35272400	-0.14666900	-1.55933000
O	2.11289600	2.94738600	-0.89665900
O	-2.59820100	-2.31784500	-0.37304200
O	0.26040400	0.18324000	2.45675200
O	-3.56916800	0.40790000	0.03619100
O	-1.75073700	1.94239500	1.60756900
O	4.15680300	-1.40729400	0.32308500
H	2.71460700	-0.29074300	0.20783500
H	1.46203300	-0.41476900	-1.84965700
H	0.18653000	-2.35964300	-1.32644300
H	1.64771800	-2.34776200	-0.34310000
H	-1.00936900	-2.15425700	0.45976700
H	1.36768700	0.33415500	1.01877700
H	2.09195900	1.21258300	-0.21151000
H	1.16025400	2.72060200	-0.92218200
H	2.20414700	3.59828300	-0.19291000
H	-2.52162700	1.44282700	1.28258200
H	-1.24532400	2.08643700	0.78685500
H	-4.35901000	0.81958000	-0.32901100
H	-2.90149700	0.41021700	-0.68180100
H	-2.20920400	-1.73888900	-1.05472500
H	-3.21474900	-1.71245000	0.06541100
H	4.87860000	-1.21701600	-0.28647300
H	4.56833900	-1.45457400	1.19325800
H	-0.49515100	0.78107100	2.27108100
H	-0.00030900	-0.67328600	2.07457900

Cartesian coordinates for the SER6W5 conformer calculated at the B3LYP/aug-cc-pVTZ (PCM:Water) level of theory using ultra tight integrals, ultra fine grid, and tight SCF convergence criteria.

0 1

C	-1.48392100	0.42496700	1.83984300
C	-0.39882500	1.08807800	0.99838100
C	1.00065800	0.41953200	1.11874300
O	1.95504100	1.03781400	0.58034300
O	-1.59821200	-0.97712800	1.61491800
N	-0.75659300	1.12554000	-0.46630200
O	1.06595800	-0.69076100	1.69630800
O	1.20724500	3.04087100	-1.12852500
O	3.86568700	-0.49220500	-0.67941600
O	-3.53568200	0.96151300	-0.90947700
O	-2.87766900	-1.70671200	-0.76884800
O	1.86235300	-2.40902300	-0.43007400
O	-0.22881300	-1.36932600	-1.79641900
H	-1.77172600	1.29288300	-0.61620900
H	-0.31334300	2.12609600	1.31618100
H	-1.25297500	0.62440200	2.88930800
H	-2.45334800	0.86968800	1.61999200
H	-0.68787500	-1.30994300	1.73588800
H	-0.53549700	0.23110400	-0.94456400
H	-0.20050200	1.87039800	-0.92522700
H	1.68215800	2.43090200	-0.52307800
H	1.64562800	2.95649700	-1.98154900
H	1.61786800	-2.03740600	0.43675400
H	2.64925700	-1.87216900	-0.66255500
H	4.59886200	-0.70445900	-0.09290000
H	3.27596200	0.09472700	-0.15538900
H	-3.45562100	-0.01729500	-0.98078600
H	-3.88341400	1.26496700	-1.75415100
H	0.54962300	-1.81461900	-1.37043300

H	-0.01226800	-1.27872900	-2.73019600
H	-2.52378300	-1.58856400	0.13418400
H	-2.06759100	-1.77307300	-1.30343600

Appendix C: Least-Square Fitting Procedure

```
import numpy
import pandas
import matplotlib.pyplot as plt
import pylab
import math
from lmfit import minimize, Minimizer, Parameters, Parameter,
report_fit

# Set variables

# Constants
pi = 3.14159265358979
planck_c = 6.626069570000000e-34
speed_light = 29979245800.0000
hartree = 4.359744340000000e-18
boltzmann_c = 1.3806488e-23
T25 = 298.15

# Unit conversion factors
ang_per_bohr = 0.52917721092e0
kg_per_amu = 1.660538921e-27
ang_to_cm = 1.0e-8
cm_to_ang = 1.0e+8
m2_to_ang = 1.0e+10
au_to_wn = hartree / (planck_c * speed_light)
hbar = planck_c * m2_to_ang**2 / (2 * pi * kg_per_amu)
fact_g = 2 * pi * speed_light / hbar
massweighted_to_normco = 1.0 / (fact_g**0.5 * ang_per_bohr)
ramact4_to_ramact6 = 2.0 / (massweighted_to_normco**2 *
ang_per_bohr**2)
fac_k = 16.0 * (pi**4)
scale_f = 1.0e-52
fac_e = (planck_c * speed_light) / (boltzmann_c * T25)

# Parameters
incident_light = 18796.99
factor = 90.0
gamma = 10
exp_scale_raman = 10 ** (-9)
exp_scale_roa = 10 ** (-6)
calc_scale_raman = 10 ** (36)
calc_scale_roa = 10 ** (39)
# Need to set to number of conformers.
n = 5

excel_file = r'/Directory/File.xlsx'
out1_input = pandas.read_excel(excel_file, sheetname = 'Calc_Freq',
parse_cols = 'C:E', header = 1)
out1_input.dropna(inplace = True)
```

```
out2_input = pandas.read_excel(excel_file, sheetname = 'Calc_Freq',
parse_cols = 'K:M', header = 1)
out2_input.dropna(inplace = True)
out3_input = pandas.read_excel(excel_file, sheetname = 'Calc_Freq',
parse_cols = 'S:U', header = 1)
out3_input.dropna(inplace = True)
out4_input = pandas.read_excel(excel_file, sheetname = 'Calc_Freq',
parse_cols = 'AA:AC', header = 1)
out4_input.dropna(inplace = True)
out5_input = pandas.read_excel(excel_file, sheetname = 'Calc_Freq',
parse_cols = 'AI:AK', header = 1)
out5_input.dropna(inplace = True)
exp_input = pandas.read_excel(excel_file, sheetname = 'Exp',
parse_cols = 'A:C', header = 0)

exp_input['Raman'] = exp_input['Raman'] * exp_scale_raman
exp_input['ROA'] = exp_input['ROA'] * exp_scale_roa

exp_x_vals = exp_input['Frequencies'].values.tolist()
exp_ram_vals = exp_input['Raman'].values.tolist()
exp_roa_vals = exp_input['ROA'].values.tolist()

def genA6(input_series1):
    input_series1['Raman'] = input_series1['Raman'] /
(ramact4_to_ramact6 * input_series1['Frequencies'] * 1)
    input_series1['ROA'] = input_series1['ROA'] / (ramact4_to_ramact6
* input_series1['Frequencies'] * 10000)
    return input_series1

def genSigma(input_series2):
    input_series2['Raman'] = scale_f * fac_k * ((incident_light -
input_series2['Frequencies']) ** 4) \
                             * (input_series2['Raman']/factor) * (1.0 /
(1.0 - numpy.exp(-1 * fac_e * input_series2['Frequencies'])))
    input_series2['ROA'] = scale_f * fac_k * ((incident_light -
input_series2['Frequencies']) ** 4) * (input_series2['ROA']/factor) \
                             * (1.0 / (1.0 - numpy.exp(-1 * fac_e *
input_series2['Frequencies'])))
    return input_series2

def lorentzian(exp_freqs, calc_freq, fwhm):
    return (fwhm/pi) / ((exp_freqs - calc_freq) ** 2 + fwhm ** 2)

def genLor(exp_freqs, calc_data, fwhm, data_type):
    ram_int = numpy.zeros(len(exp_freqs))
    for i, row in calc_data.iterrows():
        ram_int += lorentzian(exp_freqs, calc_data['Frequencies'][i],
fwhm) * calc_data[data_type][i]
    return ram_int

def fitfn(params, exp_ram_int, exp_roa_int, data1, data2, data3,
data4, data5):
```

```
    const_a = params['const_a']
    const_b = params['const_b']
    const_c = params['const_c']
    const_d = params['const_d']
    const_e = params['const_e']
    #fwhm_var = params['fwhm_var']
    return ((exp_ram_int - (const_a * genLor(exp_x_vals, data1, gamma,
'Raman') \
            + const_b * genLor(exp_x_vals, data2, gamma, 'Raman') \
            + const_c * genLor(exp_x_vals, data3, gamma, 'Raman') \
            + const_d * genLor(exp_x_vals, data4, gamma, 'Raman') \
            + const_e * genLor(exp_x_vals, data5, gamma, 'Raman'))) **
2) \
            + ((exp_roa_int - (const_a * genLor(exp_x_vals, data1,
gamma, 'ROA') \
            + const_b * genLor(exp_x_vals, data2, gamma, 'ROA') \
            + const_c * genLor(exp_x_vals, data3, gamma, 'ROA') \
            + const_d * genLor(exp_x_vals, data4, gamma, 'ROA') \
            + const_e * genLor(exp_x_vals, data5, gamma, 'ROA'))) **
2)

genA6(out1_input)
genA6(out2_input)
genA6(out3_input)
genA6(out4_input)
genA6(out5_input)

genSigma(out1_input)
genSigma(out2_input)
genSigma(out3_input)
genSigma(out4_input)
genSigma(out5_input)

out1_input['Raman'] = out1_input['Raman'] * calc_scale_raman
out1_input['ROA'] = out1_input['ROA'] * calc_scale_roa
out2_input['Raman'] = out2_input['Raman'] * calc_scale_raman
out2_input['ROA'] = out2_input['ROA'] * calc_scale_roa
out3_input['Raman'] = out3_input['Raman'] * calc_scale_raman
out3_input['ROA'] = out3_input['ROA'] * calc_scale_roa
out4_input['Raman'] = out4_input['Raman'] * calc_scale_raman
out4_input['ROA'] = out4_input['ROA'] * calc_scale_roa
out5_input['Raman'] = out5_input['Raman'] * calc_scale_raman
out5_input['ROA'] = out5_input['ROA'] * calc_scale_roa

#plt.plot(exp_x_vals, genLor(exp_x_vals, out1_input, gamma, 'Raman'),
'g')
#plt.plot(exp_x_vals, genLor(exp_x_vals, out2_input, gamma, 'Raman'),
'r')
#plt.plot(exp_x_vals, exp_ram_vals, 'b')
#plt.plot(exp_x_vals, genLor(exp_x_vals, out1_input, gamma, 'ROA'),
'g')
```

```
#plt.plot(exp_x_vals, genLor(exp_x_vals, out2_input, gamma, 'ROA'),
'r')
#plt.plot(exp_x_vals, exp_roa_vals, 'b')

params = Parameters()
params.add('const_a', value = (1.0/n), min = 0.0, max = 1.0, vary =
True)
params.add('const_b', value = (1.0/n), min = 0.0, max = 1.0, vary =
True)
params.add('const_c', value = (1.0/n), min = 0.0, max = 1.0, vary =
True)
params.add('const_d', value = (1.0/n), min = 0.0, max = 1.0, vary =
True)
params.add('const_e', value = (1.0/n), min = 0.0, max = 1.0, vary =
True)
#params.add('fwhm_var', value = 10, min = 6, vary = True)

minner = Minimizer(fitfn, params, fcn_args=(exp_ram_vals,
exp_roa_vals, out1_input, out2_input, out3_input, out4_input,
out5_input))
kws  = {'options': {'maxiter':10}}
fit_result = minner.minimize()

fit_raman = fit_result.params['const_a'].value * genLor(exp_x_vals,
out1_input, fit_result.params['fwhm_var'], 'Raman') \
            + fit_result.params['const_b'].value * genLor(exp_x_vals,
out2_input, fit_result.params['fwhm_var'], 'Raman') \
            + fit_result.params['const_c'].value * genLor(exp_x_vals,
out3_input, fit_result.params['fwhm_var'], 'Raman') \
            + fit_result.params['const_d'].value * genLor(exp_x_vals,
out4_input, fit_result.params['fwhm_var'], 'Raman') \
            + fit_result.params['const_e'].value * genLor(exp_x_vals,
out5_input, fit_result.params['fwhm_var'], 'Raman')
fit_roa = fit_result.params['const_a'].value * genLor(exp_x_vals,
out1_input, fit_result.params['fwhm_var'], 'ROA') \
            + fit_result.params['const_b'].value * genLor(exp_x_vals,
out2_input, fit_result.params['fwhm_var'], 'ROA') \
            + fit_result.params['const_c'].value * genLor(exp_x_vals,
out3_input, fit_result.params['fwhm_var'], 'ROA') \
            + fit_result.params['const_d'].value * genLor(exp_x_vals,
out4_input, fit_result.params['fwhm_var'], 'ROA') \
            + fit_result.params['const_e'].value * genLor(exp_x_vals,
out5_input, fit_result.params['fwhm_var'], 'ROA')

fit_result_x = pandas.DataFrame(data = exp_x_vals, columns =
['Frequencies'])
fit_result_y1 = pandas.DataFrame(data = fit_raman, columns =
['Raman'])
fit_result_y2 = pandas.DataFrame(data = fit_roa, columns = ['ROA'])
fit_result_output = pandas.concat([fit_result_x, fit_result_y1,
fit_result_y2], axis=1)
a = fit_result.params['const_a']
```

```
b = fit_result.params['const_b']
c = fit_result.params['const_c']
d = fit_result.params['const_d']
e = fit_result.params['const_e']
#l = fit_result.params['fwhm_var']
l = gamma

#pylab.plot(exp_x_vals, fit_raman, 'k+')
#pylab.plot(exp_x_vals, fit_roa, 'k+')
pylab.show()

excel_file_out = r'/Directory/File.xlsx'
fit_result_output.to_excel(excel_file_out, sheet_name = 'Fit Result',
index = False)
fit_file_stat = open('Fit_Stats.txt', 'w')
fit_file_stat.write("%s \n %s \n %s \n %s \n %s" % (a, b, c, d,
e, l))
```

References

(1) Prusov, E. V. *Stereoselective Synthesis of Drugs and Natural Products. Edited by Vasyl Andrushko and Natalia Andrushko.*; 2014; Vol. 53.

(2) Atkinson, R. S. *Stereoselective synthesis*; 1995.

(3) Reist, M.; Carrupt, P. A.; Francotte, E.; Testa, B. *Chem. Res. Toxicol.* **1998**, *11* (12), 1521–1528.

(4) Margolin, A. L. *Enzyme Microb. Technol.* **1993**, *15* (4), 266–280.

(5) Merten, C.; McDonald, R.; Xu, Y. *Inorg. Chem.* **2014**, *53* (6), 3177–3182.

(6) Sato, H.; Yamagishi, A. *Int. J. Mol. Sci.* **2013**, *14* (1), 964–978.

(7) Yang, G.; Xu, Y. *Top. Curr. Chem.* **2011**, *298*, 189–236.

(8) Debie, E.; Bultinck, P.; Herrebout, W.; van der Veken, B. *Phys. Chem. Chem. Phys.* **2008**, *10* (24), 3498–3508.

(9) Stephens, P. J.; Devlin, F. J.; Pan, J. *Chirality* **2008**, *20* (5), 643–663.

(10) He, Y.; Wang, B.; Dukor, R. K.; Nafie, L. A. *Appl. Spectrosc.* **2011**, *65* (7), 699–723.

(11) Barron, L. D.; Zhu, F.; Hecht, L.; Tranter, G. E.; Isaacs, N. W. *J. Mol. Struct.* **2007**, *834–836*, 7–16.

(12) Hug, W.; Hangartner, G. *J. Raman Spectrosc.* **1999**, *30* (June), 841–852.

(13) Hug, W. *Appl. Spectrosc.* **2003**, *57* (1), 1–13.

(14) Brotin, T.; Daugey, N.; Vanthuyne, N.; Jeanneau, E.; Ducasse, L.; Buffeteau, T. *J. Phys. Chem. B* **2015**, *119* (27), 8631–8639.

(15) Jürgensen, V. W.; Jalkanen, K. *Phys. Biol.* **2006**, *3* (1), S63-79.

(16) Barron, L. D.; Buckingham, A. D. *Mol. Phys.* **1971**, *20* (6), 1111–1119.

(17) Barron, L. D.; Bogaard, M. P.; Buckingham, a D. *J. Am. Chem. Soc.* **1973**, *95* (2), 603–605.

(18) Spencer, K. M.; Freedman, T. B.; Nafie, L. A. *Chem. Phys. Lett.* **1988**, *149* (4), 367–374.

(19) Che, D.; Hecht, L.; Nafie, L. A. *Chem. Phys. Lett.* **1991**, *180* (3), 182–190.

(20) Polavarapu, P. L. *J. Phys. Chem.* **1990**, *94* (21), 8106–8112.

(21) Thorvaldsen, A. J.; Gao, B.; Ruud, K.; Fedorovsky, M.; Zuber, G.; Hug, W. *Chirality* **2012**, *24* (12), 1018–1030.

(22) Zuber, G.; Hug, W. *J. Phys. Chem. A* **2004**, *108* (11), 2108–2118.

(23) Cheeseman, J. R.; Shaik, M. S.; Popelier, P. L.; Blanch, E. W. *J Am Chem Soc* **2011**, *133*

(13), 4991–4997.

(24) Ruud, K.; Thorvaldsen, A. J. *Chirality* **2009**, *21* (1 E), E54–E67.

(25) Hopmann, K. H.; Ruud, K.; Pecul, M.; Kudelski, A.; Dračínský, M.; Bouř, P. *J. Phys. Chem. B* **2011**, *115* (14), 4128–4137.

(26) Nafie, L. A. *Annu. Rev. Phys. Chem.* **1997**, *48* (1), 357–386.

(27) Cheeseman, J. R.; Frisch, M. J. *J. Chem. Theory Comput.* **2011**, *7* (10), 3323–3334.

(28) Daněček, P.; Kapitán, J.; Baumruk, V.; Bednárová, L.; Kopecký, V.; Bouř, P. *J. Chem. Phys.* **2007**, *126* (22), 1–13.

(29) Poopari, M. R.; Dezhahang, Z.; Yang, G.; Xu, Y. *ChemPhysChem* **2012**, *13* (9), 2310–2321.

(30) Zhu, P.; Yang, G.; Poopari, M. R.; Bie, Z.; Xu, Y. *Chemphyschem* **2012**, *13* (5), 1272–1281.

(31) Poopari, M. R.; Dezhahang, Z.; Xu, Y. *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **2015**, *136* (PA), 131–140.

(32) Hecht, L.; Barron, L. D. *J. Raman Spectrosc.* **1994**, *25* (7–8), 443–451.

(33) Syme, C. D. The Analysis of Raman Optical Acitivity Specra of Proteins. PhD Thesis. 2002.

(34) Smyth, E. Raman Optical Activity of Proteins and Glycoproteins. PhD Thesis. 2000.

(35) Maria Vargek. Instrumental Advances and New Applications of Raman Optical Activity. PhD Thesis. 1997.

(36) Markéta Pazderková. Raman optical activity of biomolecules: From simple models to complex systems. PhD Thesis. 2015.

(37) Haesler, J. Construction of a New Forawrd and Backward Scattering Raman Optical Activity Spectrometer and Graphical Analysis of Measured and Calculated Spectra for (R)-[2H1,2H2,2H3]-Neopentane. PhD Thesis. 2006.

(38) Nafie, L. A. In *III Italian Meeting on Raman Spectroscopy and Non-Linear Optical Effects*; 2014.

(39) Keys, D.; Smothers, W.; Trout, T. *Kaiser Opt. Syst. Inc* **2001**, 2.

(40) HoloSpec f/1.8 Holographic Imaging Spectrograph http://www.kosi.com/na_en/products/raman-spectroscopy/imaging-spectrographs/holospec-f1.8.php (accessed Dec 17, 2016).

(41) Hammes, G. G. *Biochemistry* **2002**, *41* (26), 8221–8228.

(42) Loftfield, R. B.; Eigner, E. a; Pastuszyn, a; Lövgren, T. N.; Jakubowski, H. *Proc. Natl. Acad. Sci. U. S. A.* **1980**, *77* (6), 3374–3378.

(43) Gutteridge, A.; Thornton, J. *FEBS Lett.* **2004**, *567* (1), 67–73.

(44) Gasior-Glogowska, M.; Malek, K.; Zajac, G.; Baranska, M. *Analyst* **2016**, *141*, 291-296.

(45) Batista, J. M.; Wang, B.; Castelli, M. V.; Blanch, E. W.; López, S. N. *Tetrahedron Lett.* **2015**, *56* (44), 6142–6144.

(46) Ashton, L.; Johannessen, C.; Goodacre, R. *Anal. Chem.* **2011**, *83* (20), 7978–7983.

(47) Merten, C.; Barron, L. D.; Hecht, L.; Johannessen, C. *Angew. Chemie - Int. Ed.* **2011**, *50* (42), 9973–9976.

(48) Perera, A. S.; Thomas, J.; Poopari, M. R.; Xu, Y. *Front. Chem.* **2016**, *4*.

(49) Mutter, S. T.; Zielinski, F.; Cheeseman, J. R.; Johannessen, C.; Popelier, P. L. A.; Blanch, E. W. *Phys. Chem. Chem. Phys.* **2015**, *17* (8), 6016–6027.

(50) Cardamone, S.; Caine, B. A.; Blanch, E.; Lizio, M. G.; Popelier, P. L. A. *Phys. Chem. Chem. Phys.* **2016**, *18* (39), 27377–27389.

(51) Kapitn, J.; Baumruk, V.; Kopeck, V.; Bou, P. *Society* **2006**, 4689–4696.

(52) Kamiński, M. I.; Kudelski, A.; Pecul, M. *J. Phys. Chem. B* **2012**, *116* (16), 4976–4990.

(53) Gaussian 09, Revision E.01, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2009.

(54) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140* (4A).

(55) Becke, A. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(56) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37* (2), 785–789.

(57) Kendall, R. a; Dunning Jr., T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96* (May 2013), 6796.

(58) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105* (8), 2999–3093.

(59) Mennucci, B.; Tomasi, J.; Cammi, R.; Cheeseman, J. R.; Frisch, M. J.; Devlin, F. J.; Gabriel, S.; Stephens, P. J. *J. Phys. Chem. A* **2002**, *106* (25), 6102–6113.

(60) Dunning Jr, T. H. *J. Chem. Phys.* **1989**, *90* (1989), 1007.

(61) Cheeseman, J. R. *Gaussian Inc.* Personal Communication*, 2016.*

(62) Parchaňský, V.; Kapitán, J.; Bouř, P. *RSC Adv.* **2014**, *4* (100), 57125–57136.

(63) Neugebauer, J.; Reiher, M.; Kind, C.; Hess, B. A. *J. Comput. Chem.* **2002**, *23* (9), 895–910.

(64) Buděšínský, M.; Šebestík, J.; Bednárová, L.; Baumruk, V.; Šafařík, M.; Bouř, P. *J. Org. Chem.* **2008**, *73* (4), 1481–1489.

(65) Melcrová, A.; Kessler, J.; Bouř, P.; Kaminský, J. *Phys. Chem. Chem. Phys.* **2016**, *18*, 2130–2142.

(66) Buděšínský, M.; Daněček, P.; Bednárová, L.; Kapitán, J.; Baumruk, V.; Bouř, P. *J. Phys. Chem. A* **2008**, *112* (37), 8633–8640.

(67) Tabatabaie, L.; Klomp, L. W.; Berger, R.; de Koning, T. J.; et al. *Mol. Genet. Metab.* **2010**, *99* (3), 256–262.

(68) de Koning, T. J. *J. Inherit. Metab. Dis.* **2006**, *29* (2–3), 347–351.

(69) Balu, D. T.; Li, Y.; Puhl, M. D.; Benneyworth, M. A.; Basu, A. C.; Takagi, S.; Bolshakov, V. Y.; Coyle, J. T. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (26), E2400-9.

(70) Dunlop, R. A.; Cox, P. A.; Banack, S. A.; Rodgers, K. J.; Zaher, H.; et al. *PLoS One* **2013**, *8* (9), e75376.

(71) Madeira, C.; Lourenco, M. V; Vargas-Lopes, C.; Suemoto, C. K.; Brandão, C. O.; Reis, T.; Leite, R. E. P.; Laks, J.; Jacob-Filho, W.; Pasqualucci, C. A.; Grinberg, L. T.; Ferreira, S. T.; Panizzutti, R. *Transl. Psychiatry* **2015**, *5* (5), e561.

(72) Cooks, R. G.; Zhang, D.; Koch, K. J.; Gozzo, F. C.; Eberlin, M. N. *Anal. Chem.* **2001**, *73*, 3646-3655.

(73) Nanita, S. C.; Cooks, R. G. *Angew. Chemie Int. Ed.* **2006**, *45* (4), 554–569.

(74) Sunahori, F. X.; Yang, G.; Kitova, E. N.; Klassen, J. S.; Xu, Y. *Phys. Chem. Chem. Phys.* **2012**, No. 207890, 1873–1886.

(75) Gargaro, A. R.; Barron, L. D.; Hecht, L. *J. Raman Spectrosc.* **1993**, *24* (October 1992), 91–96.

(76) Lorenz, U. J.; Rizzo, T. R. *J. Am. Chem. Soc.* **2012**, *134* (27), 11053–11055.

(77) Jarmelo, S.; Carey, P. R.; Fausto, R. *Vib. Spectrosc.* **2007**, *43* (1), 104–110.

(78) Zhang, P.; Polavarapu, P. L. *Appl. Spectrosc.* **2006**, *60* (4), 378–385.

(79) Diem, M.; Photos, E.; Khouri, H.; Nafie, L. A. *J. Am. Chem. Soc.* **1979**, *101* (23), 6829–6837.

(80) Gronert, S.; O'Hair, R. A. J. *J. Am. Chem. Soc.* **1995**, *117* (7), 2071–2081.

(81) Pecul, M. *Chem. Phys. Lett.* **2006**, *418* (1–3), 1–10.

(82) Pecul, M. *Chem. Phys. Lett.* **2006**, *427* (1–3), 166–176.

(83) Xu, X.; Goddard, W. A. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (9), 2673–2677.

(84) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132* (15), 154104.

(85) Woodbury, C. P. *Biochemistry for the pharmaceutical sciences*; Jones & Bartlett Learning, 2012.

(86) Barone, V.; Biczysko, M.; Bloino, J.; Barth, A.; Zscherp, C.; Barone, V.; Baiardi, A. *Phys. Chem. Chem. Phys.* **2014**, *16* (5), 1759–1787.

(87) Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32* (7), 1456–1465.

(88) Panek, P. T.; Jacob, C. R. *J. Phys. Chem. Lett.* **2016**, *7* (16), 3084–3090.

(89) Bloino, J.; Biczysko, M.; Barone, V. *J. Phys. Chem. A* **2015**, *119* (49), 11862–11874.

(90) Haraguchi, S.; Hara, M.; Shingae, T.; Kumauchi, M.; Hoff, W. D.; Unno, M. *Angew. Chemie - Int. Ed.* **2015**, *54* (39), 11555–11558.

(91) Johannessen, C.; White, P. C.; Abdali, S. *J. Phys. Chem. A* **2007**, *111* (32), 7771–7776.

(92) Ashton, L.; Pudney, P. D. A.; Blanch, E. W.; Yakubov, G. E. *Adv. Colloid Interface Sci.* **2013**, *199–200*, 66–77.